Changes of the diurnal cycle structures of selected meteorological parameters of Potsdam Station for the summer season

Peter C. Werner*, Friedrich-Wilhelm Gerstengarbe

Potsdam Institute for Climate Impact Research, PO Box 60 12 03, Telegrafenberg, 14412 Potsdam, Germany

ABSTRACT: Using cluster analysis, patterns of diurnal cycles were calculated based on hourly values of temperature, cloud cover, sunshine frequency, vapour pressure and relative humidity at Potsdam Station for the period June to August 1893–2000. It was shown that the large-scale climate variations of the last century are reflected in the temporal behaviour in some of these patterns. These temporal changes occur in a highly differentiated manner over the day. Additionally, relationships between the investigated parameters and precipitation and also the appearance of extremes were considered.

KEY WORDS: Diurnal cycle · Cluster analysis · Climate change

Resale or republication not permitted without written consent of the publisher

1. INTRODUCTION

Normally the climate and its changes are defined by meteorological parameters with a temporal resolution of a day or more. Up to now, less consideration has been focused on the question of whether climate change can also be demonstrated at the hourly temporal scale, though such information is needed in various climate impact models, for instance in hydrological models (King et al. 1999) and vegetation models (Wang et al. 2002) or in the calculation of human health risks (Laschewski & Jendritzky 2002). The reason for this is the fact that time series of hourly values very seldom exist over long observational periods and without gaps. An exception is the climate series of Potsdam Station (Branicki 1963). The main meteorological parameters have been observed on an hourly basis for more than 100 yr (www.klima-potsdam.de). To answer the question of whether climate changes are reflected in the diurnal cycles of meteorological parameters, the data of this station (for more information see Section 2), comprising temperature, precipitation, sunshine frequency, cloud cover, relative humidity and vapour pressure, are taken as the basis for further investigations. For the statistical investigation we apply the

extended cluster analysis described in Section 2 and some additional tests. The results are presented and discussed in Section 3. The final conclusions are given in Section 4.

2. DATA AND METHODS

Potsdam Station is situated at 52.38° N and 13.07° E (near Berlin). For temperature, precipitation, relative humidity and vapour pressure, 24 values d^{-1} are available for 1893–2000. The observation of cloud cover exists from 1894 to the present at a 2 h resolution. The hourly sunshine frequency is available for the astronomically longest possible duration per day. Because all time series have been measured since 1893 according to the same criteria, without any shifting of the station or changes in measuring methods, inhomogeneities cannot be detected by adequate statistical tests. Local influences (i.e. growth of the urban heat island) on the data can also be excluded (Lehmann 2002). The investigations were limited to the summer months June, July and August, because at this time of year the diurnal cycles of most meteorological parameters are more distinct than in other seasons. To

describe the structural differences between the diurnal cycles of the meteorological parameters, extended non-hierarchical cluster analysis (Gerstengarbe et al. 1999) was used.

The main idea of this cluster analysis is to relate to each other an existing number, i, of elements which are each described by N parameters p, i.e.

$$e_i = f(p_{i1,...,}p_{iN}) \tag{1}$$

In non-hierarchical methods the elements are simultaneously distributed into a given number of clusters. By displacing the elements between the clusters in the case of a given quality criterion, a given initial distribution is developed step-by-step into steadily improving groupings until the optimum is reached. For more details, see Steinhausen & Langer (1977). The starting point for the representation of the following method is the non-hierarchical minimum-distance method according to Forgy (1965), and the starting condition in the above method is to have the elements equally distributed over a number of given clusters (initial partition). In the case of M given elements and K clusters, each cluster receives $L = M/K$ elements as follows:

$$
\begin{array}{lll}
e_{1,...,} & e_L \in & c_1 \\
e_{L+1,...,} & e_{2L} \in & c_2 \\
. & & . \\
. & & . \\
. & & . \\
e_{(k-1)L+1,...,} & e_{kL} \in & c_k
\end{array}
\tag{2}
$$

(The number of clusters K must be defined empirically. The number of elements depends on the data series and the problem to be investigated.)

A so-called group centroid, e_k, is then calculated for each k of the K clusters (the cluster mean value under consideration of those existing parameters that have to be normalized accordingly in the case of different scalings):

$$\overline{e}_k = \frac{1}{L} \sum_{i=(k-1)L+1}^{kL} e_i \tag{3}$$

By applying the Euclidean distance, the following objective function, $a(g)$, for each grouping step, g, can be defined:

$$a(g) = \sum_{k=1}^{K} \sum_{i \in k} |e_i - \overline{e}_k|^2 \tag{4}$$

By considering the Euclidean distance, each grouping step can be seen as a displacement of the element e_i into that cluster which contains the respective nearest centroid. The objective can thus be minimized:

$$a(g) \forall g \rightarrow \min \tag{5}$$

This procedure is repeated until a local minimum of the objective function is reached. The objective function reaches a local minimum if 2 successive grouping steps show the same result; the iteration is in this case discontinued, i.e. the optimum classification with respect to the given number of clusters has been reached.

An important disadvantage of this method is that one does not know whether an absolute or just a secondary minimum of the objective function has been obtained (Milligan & Cooper 1985, Fovell & Fovell 1993). For this reason the quality of separation is unknown, as is the objective number of clusters. The solution to this problem is described by Gerstengarbe & Werner (1997). With this extended cluster analysis algorithm, it is possible to calculate cluster patterns which are separated in a statistically safe way. In parallel, one can estimate the optimum number of clusters. The clustering is carried out with a given number of elements. Thus, one obtains the optimal number of clusters containing different quantities of elements (i.e. days of a time series). Now, it is possible that the values of the other meteorological parameters relate to the elements within the clusters. A more detailed description of the method including a climatological example can be found in Gerstengarbe et al. (1999).

With this statistical procedure it is possible to sum up diurnal cycles of similar structure into adequate groups (clusters). A change in the frequency of a cluster can be an indicator of climate change during the observational period if it is statistically significant. The statistical security of the trend development was proved using the Spearman test (Schönwiese 1992), the Mann-Kendall test (Mann 1945, Kendall & Stuart 1977) and the shifting version of the Mann-Kendall test (Sneyers 1975). Additionally, the 4-fold test (Taubenheim 1969) was used for the parameter-free estimation of the correlation coefficient.

3. RESULTS

3.1. Temperature

The cluster analysis showed 8 clusters to be the optimal number for temperature. (The same result occurred for the other meteorological parameters.) Fig. 1a shows the mean diurnal cycle of each temperature cluster. The cycles differ mainly in the level and magnitude of the range between maxima and minima. Significant trends of the temporal development of some of the cluster frequencies can be observed (Table 1). It is obvious that the 3 warmest clusters (3, 4 and 5) show statistically significant positive trends. Negative trends can be found for the colder clusters, although with statistical significance (with 95% and more) only for Cluster 7. As an example Fig. 2 shows the trend development of Clusters 3 and 7. This result corresponds with the observed temperature increase

at Potsdam Station (see www.klima-potsdam.de). A comparison between the first 30 yr of the observational period and the last 30 yr was carried out because of the significant differences between the 2 periods (see also www.klima-potsdam.de). It is apparent that the tem-

perature increase is distributed variably over the day. While the increase between 05:00 and 13:00 h is small, it increases significantly after 13:00 h and reaches its maximum at 18:00 h (Fig. 3a). The finding is new that the maximum and minimum of the diurnal cycle occur
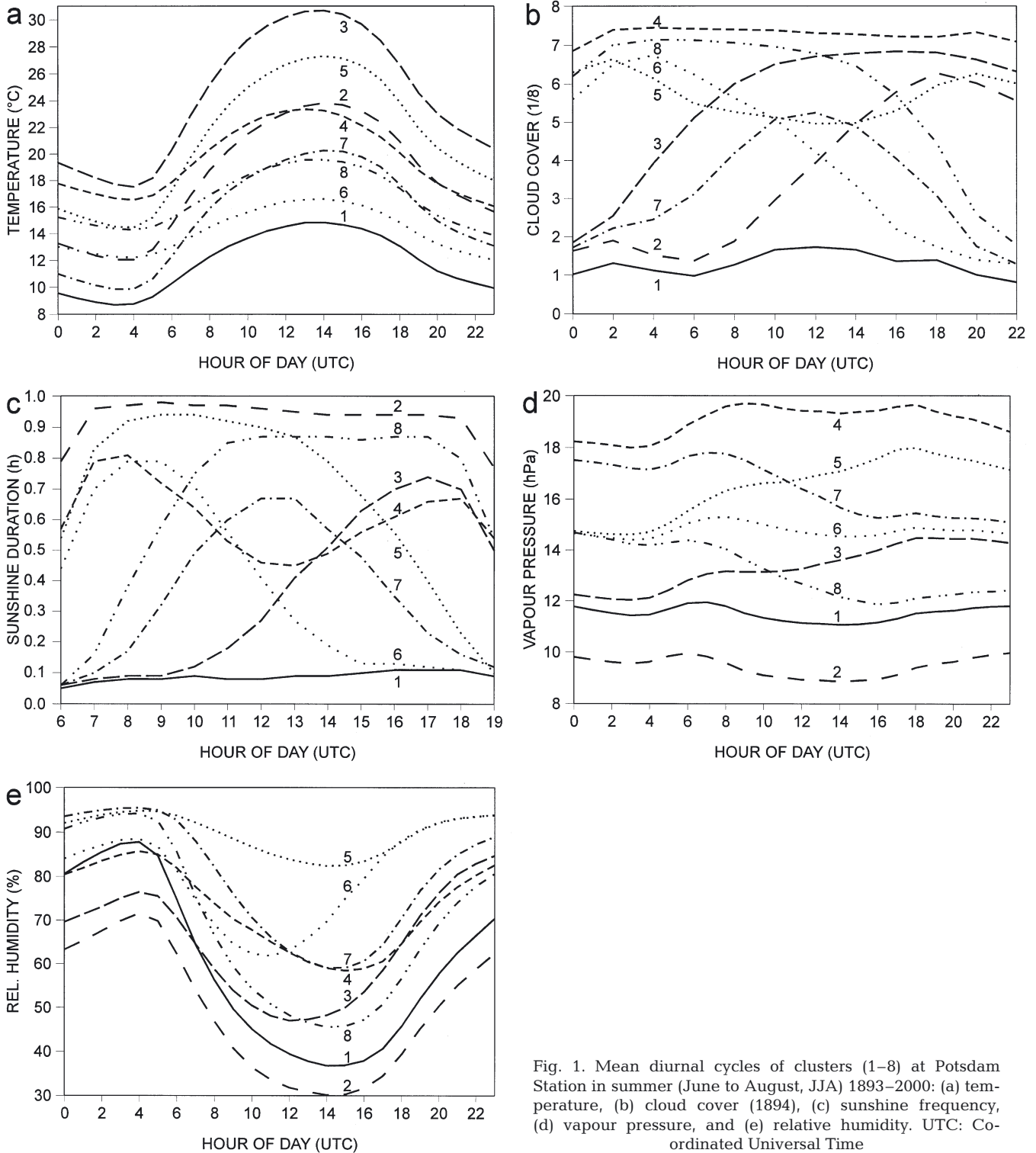

Fig. 1. Mean diurnal cycles of clusters (1–8) at Potsdam Station in summer (June to August, JJA) 1893–2000: (a) temperature, (b) cloud cover (1894), (c) sunshine frequency, (d) vapour pressure, and (e) relative humidity. UTC: Co-ordinated Universal Time

Table 1. Trend of frequency of temperature clusters (Fig. 1a), summer (June to August, JJA) 1893–2000

Cluster	Mean value (°C)	Trend of frequency[a]	Statistical significance
1	11.8	−2.48	–
2	18.2	−1.61	–
3	24.2	4.17	0.05
4	19.5	3.12	0.05
5	21.1	4.06	0.05
6	14.2	−1.45	–
7	15.3	−5.05	0.01
8	16.6	−0.77	–

[a]Related to the complete time period

simultaneously about 1 h later. In summary, one can say that the climate change may not be visible simply in a general increase of the hourly values, but in the very differentiated changes of the diurnal cycles of these values and of the position of their extremes.

3.2. Cloud cover and sunshine frequency

As for temperature, cluster analysis was carried out with 8 clusters for both cloud cover and sunshine frequency. The cycles of the single clusters within the day are reproduced in Fig. 1b,c, but considering only the time between 06:00 and 19:00 h for sunshine frequency. The cycles of the single clusters of the 2 parameters differ clearly from each other. The cycles of sunshine frequency are not independent of those of cloud cover. Table 2 shows the cluster frequencies (percent) of sunshine frequency related to those of cloud cover. As an example, Cluster 1 for cloud cover
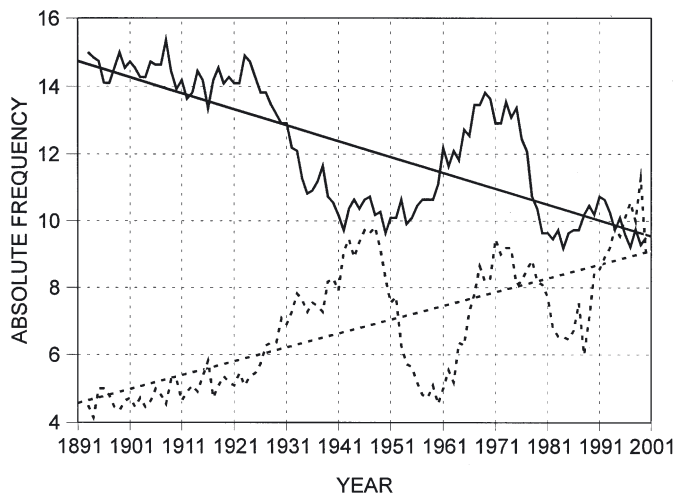

Fig. 2. Absolute frequency (smoothed lines) of Clusters 3 (dashed lines) and 7 (solid lines) of the hourly values of temperature at Potsdam Station, JJA 1893–2000

Table 2. Two-dimensional distribution of the relative frequency of the occurrence of cloud-cover clusters (Fig. 1b) and sunshine-frequency clusters (Fig. 1c), JJA 1893–2000

Cloud cover	Sunshine frequency							
	1	2	3	4	5	6	7	8
1	0.0	89.0	0.0	0.7	5.1	0.0	0.0	5.2
2	0.0	30.4	0.0	2.8	60.5	3.3	0.7	2.3
3	14.4	2.5	3.8	10.4	24.5	32.2	10.1	2.1
4	65.0	0.1	8.5	0.8	1.4	8.2	15.4	0.6
5	1.2	6.0	14.6	9.0	17.5	6.4	22.1	23.2
6	0.8	11.0	30.6	7.2	1.2	0.3	4.1	44.8
7	0.5	40.1	2.0	18.0	20.7	6.9	2.6	9.2
8	34.1	0.7	25.8	4.0	2.8	6.6	19.4	6.6

Table 3. Trends of the frequencies of selected cloud-cover (Fig. 1b) and sunshine-frequency (Fig. 1c) clusters, JJA 1893–2000

Cluster	Cloud cover		Sunshine frequency	
	Trend (d)[a]	Statistical significance	Trend (d)[a]	Statistical significance
1	−2.44	0.05	−4.23	0.05
3	5.27	0.01		
5			2.66	0.05
6	−4.80	0.01		
7	2.19	0.05		

[a]Related to the complete time period

(very low cloudiness over the day) is connected with Cluster 2 for sunshine frequency (sunshine over nearly the whole day) in 89% of all cases of Cluster 1 for cloud cover. Inverse behaviour can be obtained between Cluster 4 for cloud cover and Cluster 1 for sunshine frequency. Analogous considerations can be made between other cluster combinations.

A totally different feature is shown in the comparison between the temporal characteristics of the clusters of the 2 parameters. Statistically significant trends of the cluster frequencies are given in Table 3. The results in Table 3 are an argument for the fact that the interannual relations between the clusters of the 2 meteorological parameters seen in Table 2 are not connected with the characteristic long-term trends of the clusters. This means that the temporal developments of the diurnal cycles of cloud cover and sunshine frequency are not coherent or only weakly so. Fig. 3b,c illustrates this. As an example, one can see that at the end of the observational period both cloud cover and sunshine frequency increase in the late afternoon. This presumed contradiction can be resolved by the possibility that in this period the amount of cirrus clouds increased over the station, but for this assumption to be found true, further investigation is necessary. On the one hand more cirrus
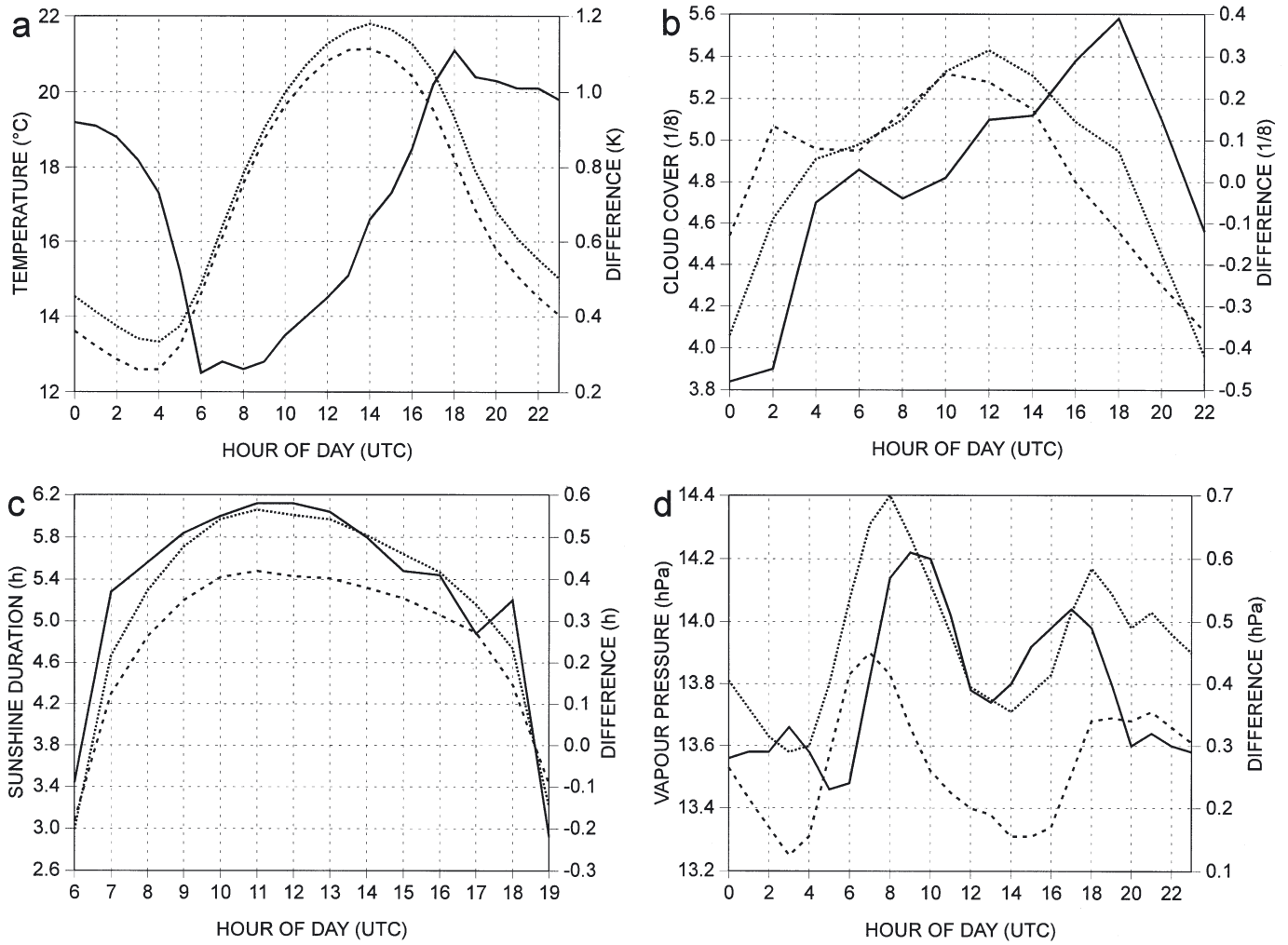
Fig. 3. Thirty-year mean of hourly values at Potsdam Station, JJA. Dashed lines: 1893–1922; dotted lines: 1971–2000; solid lines: difference between 1971–2000 and 1893–1922. (a) Temperature, (b) cloud cover (1894), (c) sunshine frequency, and (d) vapour pressure

clouds lead to more cloud cover, while on the other hand this type of cloud permits measurable sunshine.

Furthermore, the increase in sunshine frequency (Fig. 3c) causes a part of the observed warming in the summer period between 1893 and 2000 (Fig. 3a).

3.3. Vapour pressure and relative humidity

For both vapour pressure and relative humidity, 8 different clusters can be determined (Fig. 1d,e). While the diurnal cycles of vapour pressure are not unambiguously connected with other meteorological parameters, the cycle of relative humidity is strongly linked with those for temperature. This circumstance is reflected in the relative probabilities of the occurrence of relative-humidity clusters combined with the simultaneous occurrence of temperature clusters (Table 4).

Additionally, it can be pointed out that in some cases a strong connection exists between the diurnal cycles of cloud cover and relative humidity (Table 5).

Table 4. Two-dimensional distribution of the relative frequency of the occurrence of temperature clusters (Fig. 1a) and relative-humidity clusters (Fig. 1e), JJA 1893–2000

Temp.	Relative humidity							
	1	2	3	4	5	6	7	8
1	6.7	0.6	3.2	10.9	26.1	17.5	21.9	13.1
2	27.8	8.0	6.5	5.8	0.7	3.5	13.2	34.5
3	31.1	32.6	16.8	3.4	0.0	1.7	1.9	12.5
4	4.3	1.9	13.3	16.1	11.4	18.7	21.0	13.3
5	32.6	18.0	11.8	3.7	0.0	1.7	5.3	26.9
6	1.8	0.3	3.2	15.0	34.4	17.6	22.7	5.0
7	18.0	3.8	3.7	6.7	1.4	5.6	20.8	40.0
8	2.9	0.6	6.2	16.1	23.9	15.1	24.9	10.3

Table 5. Two-dimensional distribution of the relative fre-
quency of the occurrence of cloud-cover clusters (Fig. 1b) and
relative-humidity clusters (Fig. 1e), JJA 1893–2000

Cloud	Relative humidity							
cover	1	2	3	4	5	6	7	8
1	42.0	25.8	2.9	1.0	0.0	0.1	2.6	25.6
2	26.3	13.9	16.1	3.1	0.0	5.8	5.1	29.7
3	10.3	2.7	11.1	9.4	6.4	23.0	15.8	21.3
4	1.6	0.5	5.6	13.3	37.4	18.4	19.2	4.0
5	10.6	2.5	10.6	13.4	4.7	7.0	24.3	26.9
6	13.4	3.6	6.1	11.6	2.6	1.5	31.4	29.8
7	24.5	9.1	7.6	7.1	0.4	5.0	10.7	35.6
8	2.9	0.8	6.0	17.4	23.1	12.8	27.3	9.7

Table 6. Trend of the frequencies of selected vapour pressure
clusters (Fig. 1d) and relative-humidity clusters (Fig. 1e),
JJA 1893–2000

Cluster	Vapour pressure		Relative humidity	
	Trend (d)	Statistical significance	Trend (d)	Statistical significance
2			3.27	0.05
4	3.30	0.01	6.35	0.01
5	2.18	0.01		
6			−3.68	0.05
8			−5.68	0.01

A further investigation of the cluster trends leads to
the result that vapour-pressure clusters with high-
pressure values have statistically significant positive
trends (Table 6), which means that the vapour pressure
in the atmosphere increases in spite of a decrease in
precipitation (see Section 3.4). This rise is nearly

evenly distributed over the whole day, and there is a
simultaneous shifting of the daily maximum of vapour
pressure from 07:00 to 08:00 h (Fig. 3d). In contrast
to this trend development, the trends for relative
humidity are positive as well as negative (Table 6).
Therefore, the long-term changes in the mean diurnal
cycle can be neglected. The decrease in relative
humidity is on average 1.3%, occurring mainly at
night; however, the morning hours show a small
increase (not shown).

3.4. Precipitation

Because of the stochastic behaviour of precipitation
events during a day, it was impossible to arrange the
diurnal cycles into clusters. Therefore, only the mean
conditions and changes to these could be investigated.
Fig. 4a,b shows the diurnal cycles for the precipitation
sums and the frequencies of precipitation events for
the time periods 1893–1922 and 1971–2000. It is obvi-
ous that the 2 periods differ from one another. Both
precipitation sums and events decrease at the end of
the observational period. For the precipitation sums
this is valid especially for the second part of the day.
This development can be observed for the frequency
over the whole day. The intensities of precipitation
remain nearly constant (not shown) in contrast to other
investigations, where positive and negative trends of
precipitation intensities have been observed in differ-
ent parts of the world (Oki & Musiake 1994, Osborne &
Hulme 2002). The well-known relation between cloud
cover and precipitation is revealed in Fig. 5. Similar
relations can be proved for all other cloud-cover
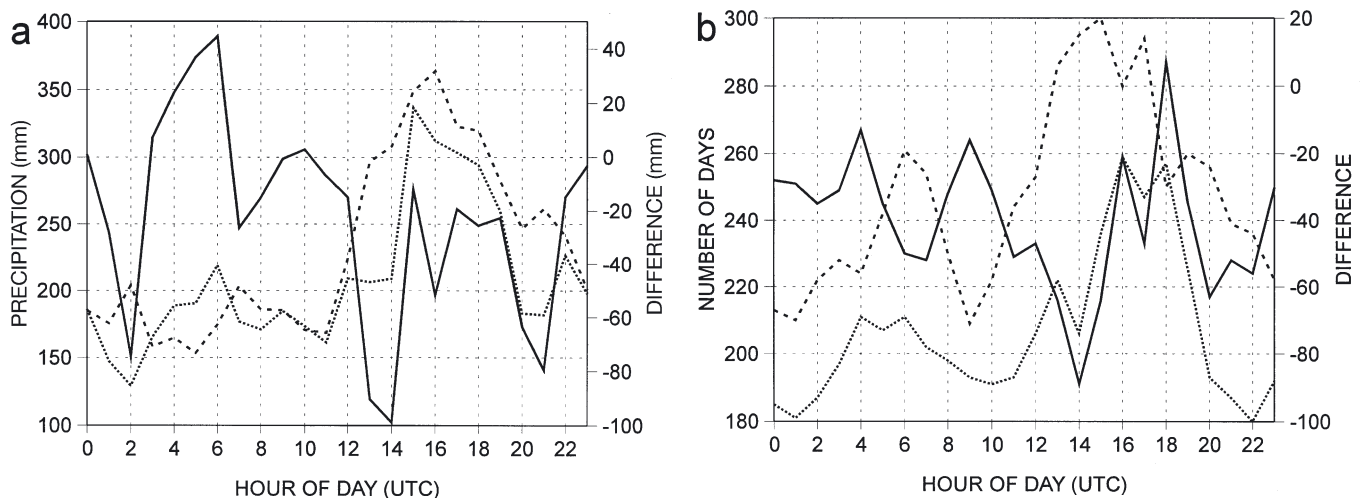clusters (not shown).

Fig. 4. Thirty-year mean of hourly values at Potsdam Station, JJA. Lines as in Fig. 3. (a) Precipitation sum and (b) number of days
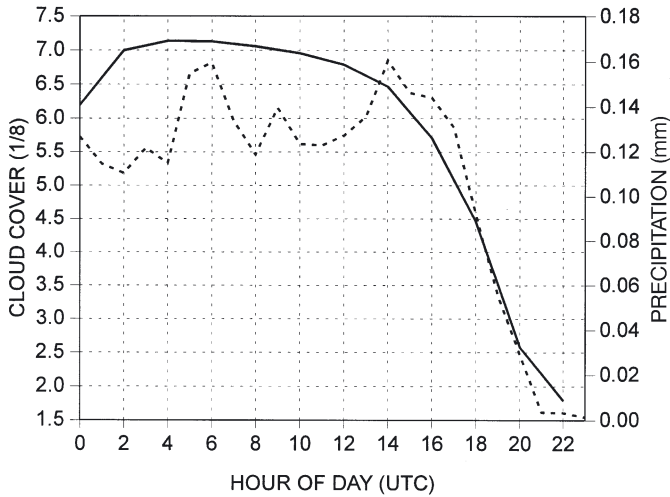with precipitation

Fig. 5. Mean diurnal cycles at Potsdam Station, JJA, Cluster 8, 1894–2000. Solid line: cloud cover; dashed line: precipitation per hour for days with cloud cover
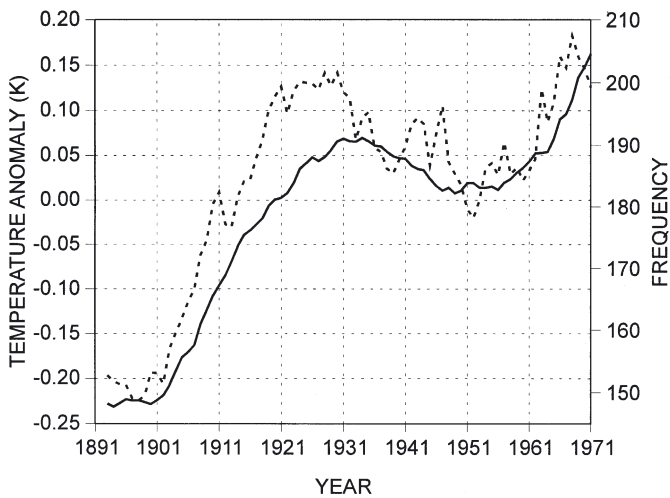

Fig. 6. Thirty-year shifting summer means, JJA 1893–2000. Solid line: anomalies of the northern-hemispheric temperature; dashed line: number of hours with ≥25°C

3.5. Extremes

A typical method for the description of climate is the investigation of threshold exceedance. An additional piece of information worth considering besides the frequency of threshold exceedance is the duration of such events. Therefore, in a further investigation, the changes of the mean number of hours per day with a temperature ≥ 25°C was estimated. The mean number of hours per summer reaches a value of 152.7 for the first 30 yr period (1893–1922) and 199.3 for the latter period (1971–2000), this is an increase of about 30%. (Similar values to those for 1971–2000 can be observed between 1921–1950 and 1931–1960 and in the last 7

Table 7. Relation between the frequencies of threshold being exceeded and not being reached (Potsdam Station), and the northern-hemispheric temperature, JJA 1893–2000

Meteorological parameter	Correlation coefficient	Statistical significance
Temperature ≥ 30.0°C	0.70	0.01
Temperature ≥ 25.0°C	0.78	0.01
Temperature < 10.0°C	−0.80	0.01
Vapour pressure ≥ 19.0 hPa	0.78	0.01
Vapour pressure < 10.0 hPa	−0.85	0.01
Precipitation ≥ 3.0 mm d^{-1}	−0.69	0.01
Relative humidity ≥ 95%	−0.85	0.01
Relative humidity < 20%	−0.82	0.01
Cloud cover ≥ $\frac{7}{8}$	0.18	–
Cloud cover ≤ $\frac{1}{8}$	−0.26	–

shifting periods.) It can be observed that this threshold is exceeded in the first period only between 07:00 and 21:00 h. In the latter period this time interval is extended up to 01:00 h. These changes become much more distinct for higher thresholds. For instance, a threshold of 30°C leads to a mean value of 20.7 h for the first period and 33.8 h for the latter. This corresponds to a value of 63.3%. In Fig. 6 the frequency of 25°C threshold extensions for 30 yr shifting intervals is presented, along with the northern-hemispheric summer temperature anomalies (Jones & Moberg 2003). It is obvious that there is good accordance between the 2 non-smoothed courses. This can be confirmed by the 4-fold test (Taubenheim 1969) with a correlation coefficient r = 0.78 and a statistical significance of 99%. Similar correlation can be found for extreme value regions of vapour pressure, precipitation and relative humidity, but not for cloud cover (Table 7). This means that the development of frequencies of extreme hourly values may reflect the climate change.

4. CONCLUSIONS

Using the extended cluster analysis it was possible to detect different structures of the diurnal cycles for some important meteorological parameters. The temporal behaviour of these structures may reflect a large-scale climate-change signal, but is very differentiated over the day. The main results are:
- simultaneous shifting of the maximum and minimum of temperature to about 1 h later;
- maximum warming in the late afternoon;
- incoherent trends of the structure of the diurnal cycles of cloud cover and sunshine frequency;
- shifting of the vapour pressure extremes from 07:00 to 08:00 h.

Additionally, a statistical relationship was found between the appearance of extreme hourly values and the course of global warming. This means that, in the future, hourly values of meteorological parameters will be of greater importance for the interpretation of climate-change signals. The results can also be used for the validation of climate-model results at this scale.

LITERATURE CITED

Branicki O (1963) Das Klima von Potsdam—Ergebnisse 60 jähriger Beobachtungen am meteorologischen Observatorium 1893–1952. Meteorol Abh XXXII/1:56

Forgy EW (1965) Cluster analysis of multivariate data: efficiency versus interpretability of classifications. Biometrics 21:768

Fovell RG, Fovell MC (1993) Climate zones of the conterminous United States defined using cluster analysis. J Clim 6:2103–2135

Gerstengarbe FW, Werner PC (1997) A method to estimate the statistical confidence of cluster separation. Theor Appl Climatol 57:103–110

Gerstengarbe FW, Werner PC, Fraedrich K (1999) Applying non-hierarchical cluster analysis algorithms to climate classification: some problems and their solution. Theor Appl Climatol 64:143–150

Jones PD, Moberg A (2003) Hemispheric and large-scale surface air temperature variations: an extensive revision and update to 2001. J Clim 16:206–223

Kendall MG, Stuart A (1977) The advanced theory of statistics, Vol 1. Charles Griffin, London

King KW, Arnold JG, Bingner RL (1999) Comparison of Green–Ampt and curve number methods on Goodwin creek watershed using SWAT. Trans ASAE 42(4): 919–925

Laschewski G, Jendritzky G (2002) Effects of the thermal environment on human health: an investigation of 30 years of daily mortality data from SW Germany. Clim Res 21:91–103

Lehmann A (2002) Die Säkulare Klimareihe von Potsdam. Klimastatusbericht 2001. Deutscher Wetterdienst, Offenbach, p 227–239

Mann HB (1945) Nonparametric test against trend. Econometrica 13:245–259

Milligan GW, Cooper MC (1985) An examination of procedures for determining the number of clusters in a data set. Psychometrika 50:159–179

Oki T, Musiake K (1994) Seasonal change of diurnal cycle of precipitation over Japan and Malaysia. J Appl Meteorol 33:1445–1463

Osborne TJ, Hulme M (2002) Evidence for trends in heavy rainfall events over the UK. Phil Trans Math Phys R Soc Lond A 360(1796):1313–1325

Schönwiese CD (1992) Praktische Statistik für Meteorologen und Geowissenschaftler. Gebr Borntraeger, Berlin

Sneyers R (1975) Statistische Analyse von Beobachtungsreihen. Technical Note 143, OMN-No. 415, WMO, Geneva, p. 191–192

Steinhausen D, Langer K (1977) Clusteranalyse—Einführung in Methoden und Verfahren der automatischen Klassifikation. Walter de Gruyter, Berlin

Taubenheim J (1969) Statistische Auswertungen geophysikalischer und meteorologischer Daten. Akad Verlagsges Geest & Portig, Leipzig

Wang SS, Grant RF, Verseghy DL, Black TA (2002) Modelling carbon dynamics of boreal forest ecosystems using the Canadian Land Surface Scheme. Clim Change 55: 451–477