

## Classifications in climate research

Jucundus Jacobeit

### Angaben zur Veröffentlichung / Publication details:

Jacobeit, Jucundus. 2010. "Classifications in climate research." *Physics and Chemistry of the Earth, Parts A/B/C* 35 (9-12): 411–21. <https://doi.org/10.1016/j.pce.2009.11.010>.

### Nutzungsbedingungen / Terms of use:

CC BY-NC-ND 4.0



# Classifications in climate research

J. Jacobeit

*Institute of Geography, University of Augsburg, Universitaetsstrasse 10, Augsburg, Germany*

## Keywords:

Classifications  
Circulation dynamics  
Circulation–climate relationships  
Dynamical mode analysis

In order to improve the understanding and an appropriate application of classification approaches in climate research, several steps of considerations are reproduced in this paper. At first some fundamental lines of advance in the development of classifications will be outlined, especially referring to the increasing application of multivariate statistical techniques and the inclusion of dynamical aspects related to the atmospheric circulation. Subsequently, the growing field of applications of such classifications will be addressed, focusing on atmospheric circulation dynamics and circulation–climate relationships. Finally, classifications in general will be compared with dynamical mode analyses, the other fundamental approach for studying atmospheric circulation dynamics. Main conclusions may be summarized as follows: Strict assignments of individual objects to disjunctive groups (as done by classifications) may represent a serious simplification of complex reality, but are necessary or at least favourable if direct relations of the atmospheric circulation to individual events (e.g. meteorological extremes) are required or if internal characteristics of atmospheric circulation patterns have to be determined by means of particular parameters (e.g. relative vorticity or flow intensity). On the other hand, without such requirements mode analyses may be more appropriate for analysing complex data sets since generic circulation patterns (instead of simple composite patterns) can be derived allowing to represent original pressure fields as superimpositions of such generic patterns with varying weights.

## 1. Introduction

Classifications are an important tool in both general and applied climatology (see for example Barry and Perry, 1973; El-Kadi and Smithson, 1992; Yarnal, 1993; Yarnal et al., 2001), since often a strict assignment of individual objects to disjunctive groups is necessary or at least favourable for achieving clearly structured results from complex data sets. The derivation and application of classifications have considerably increased in climate research during the last decades, but already in earlier times of scientific evolution we can find prominent and long-lived examples like the well-known global climate classification from Wladimir Köppen (1923). This paper does not intend to give a comprehensive review of classifications that have been developed in climate research until now, since recently such a paper has been published with focus on the important part of the atmospheric circulation (Huth et al., 2008) and manifold activities concerning the evaluation and comparison of circulation and weather type classifications for European regions are presently carried out in context of the EU COST Action 733 (“Harmonisation and Applications of Weather Type Classifications for European Regions”, see this volume). This paper rather intends to improve the understanding of methodical approaches in climate

research summarized as “classifications” since the increasing field of different applications includes a growing risk that classifications are not properly used especially in view of other approaches that might be more appropriate for particular investigations. In this context several aspects have to be considered: What kind of progress has taken place in the methodical domain of classifications? Which fields of research are commonly applying classification approaches? What kind of requirements might favour classifications compared to other major approaches? Might the latter even be superior to classifications under particular conditions? This paper aims to contribute some principal aspects to these questions and will proceed in the following way:

- at first some fundamental lines of advance in the development of classifications since the times of ‘classical climatology’ in the first half of the 20th century will be outlined;
- secondly, with respect to the domain of atmospheric circulation dynamics, prominent fields of research questions addressed by applying particular classifications will be summarized;
- finally, a comparison of classifications in general with dynamical mode analyses, the other fundamental approach for studying atmospheric circulation dynamics, will be carried out in order to clarify which fields of application are the appropriate ones for classifications in contrast to cases for which mode analyses should be preferred.

## 2. Fundamental lines of advance concerning classifications in climate research

### 2.1. Application of multivariate statistical techniques

Early classifications as for example the above-mentioned one from Köppen (1923) or many other so-called descriptive or effective climate classifications (e.g. Thornthwaite, 1933; Creutzburg, 1950; Troll, 1958 and more recent approaches of Lauer et al., 1996; Lauer and Rafiqpoor, 2002) are based on particular predefined threshold values representing the boundaries between different classes (as for example the threshold of 18 °C for the monthly mean temperature of the coldest month as boundary between tropical and extra-tropical climates in the Köppen classification). Such thresholds for climatic variables are often based on discernible changes in some important components of the global ecosystem (e.g. vegetation or macro-scale plantsocietal units). In contrast to such threshold-based classifications, the implementation of multivariate statistical techniques has allowed to classify climate without predefined thresholds by grouping individual objects (e.g. meteorological stations) into self-generating classes according to particular statistical criteria. An instructive example is the global climate classification from Gerstengarbe and Werner (2007) generated by a non-hierarchical cluster analysis with threshold-boundaries between the 32 climatic types resulting from the objective grouping (i.e. thresholds for temperature and precipitation being the secondary result of assigning stations with similar characteristics to clusters) thus avoiding any subjective predefinition (see also Gerstengarbe et al., 1999). Of course such classifications may also be used to study climate change, for example by comparing the spatial distributions of climatic types for different periods thus identifying corresponding shifts of these types as shown by Gerstengarbe and Werner (2007) for the last 15 years of the 20th century in relation to the entire century.

Further progress induced by the application of multivariate statistical techniques has taken place in the field of regionalizing climate variables or phenomena. A broad range of different techniques has been considered, from correlation-based approaches and analyses of variance up to different clustering techniques (e.g. Beck, 2000; Jacobeit et al., 1998) and various modes of principal component analyses (PCA). Six different modes are described in detail by Richman (1986) from which the S-, Q-, T- and R-modes are relevant for regionalization issues. Two examples are given in Figs. 1 and 2. The first one is based on a R-mode PCA of long-term mean values for a set of climate variables from stations in Germany published in a geostatistical textbook of Bahrenberg and Giese (1975). The R-mode PCA processes a set of different variables resolved in space but not in time and provides spatial patterns of principal components which represent groups of original variables with similar spatial variability. Taking the first two PCs from this analysis which represent mean conditions for temperature and precipitation, respectively, and superimposing their spatial score patterns leads to Fig. 1 which depicts warm/wet, warm/dry, cool/wet and cool/dry regions in Germany according to above- and below-average values of the corresponding PC scores. Thus, a climatic regionalization from a large set of variables (27 in this case) can be derived in terms of one single distribution map.

Another example does not refer to mean values, but to temporally as well as spatially resolved data for just one variable (e.g. temperature). Such data may enter a S-mode PCA with spatial units (stations or grid points) as variables and time units (depending on the temporal resolution) as cases. Grouping those spatial units with the maximum loading on the same PC leads to regions of similar temporal variability with respect to the analysed variable.

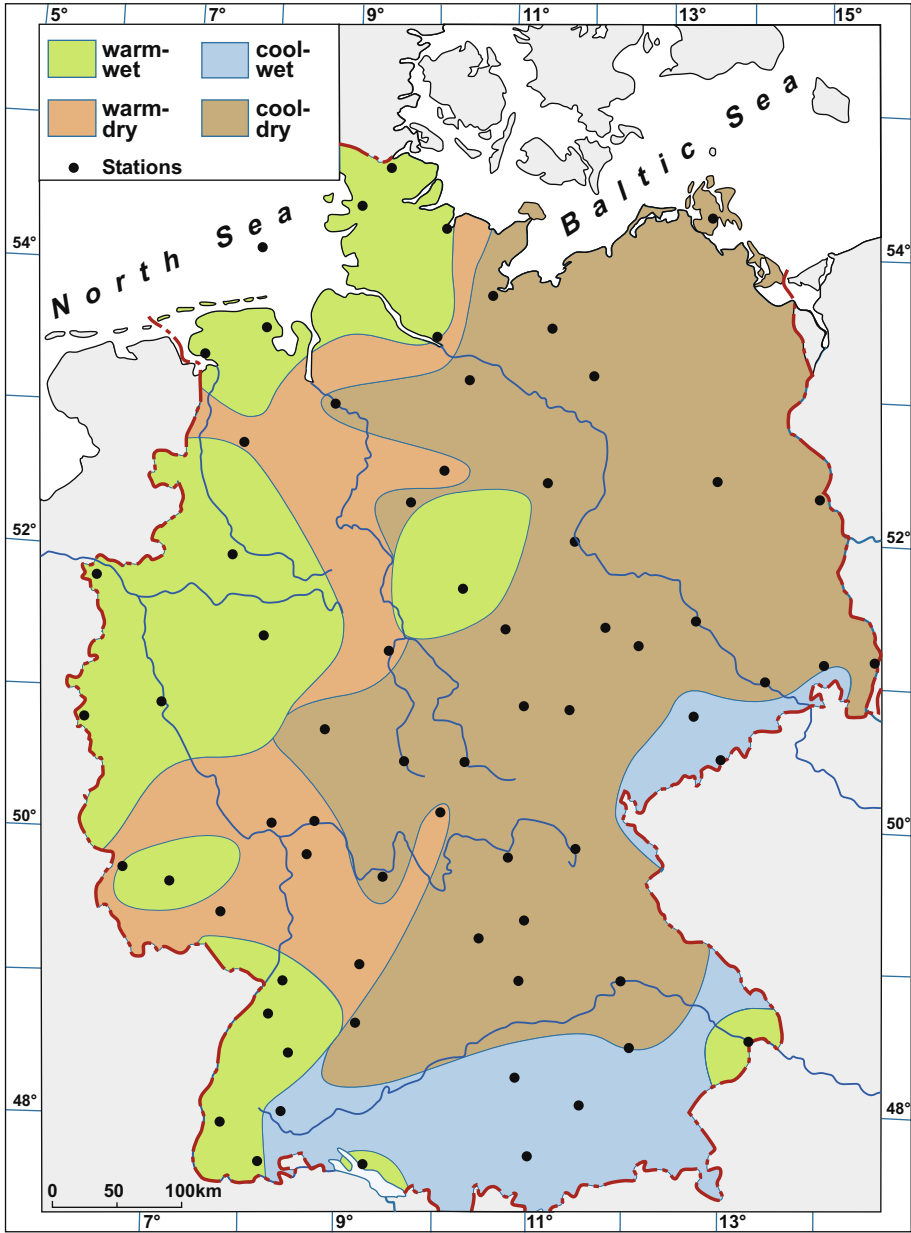
Fig. 2 gives an example referring to winter (DJ) temperatures with Mediterranean sub-regions of similar temperature variability. Such kind of regionalization is often used as starting point for further investigations on climate and circulation dynamics (e.g. Hertig and Jacobeit, 2008a,b).

Further approaches combine spatial groupings with synoptic considerations, based on a larger number of meteorological parameters (including sea level pressure, cloud cover, dew point among others) and leading to so-called spatial synoptic classifications. A prominent example is the study of Bower et al. (2007) identifying air-mass seed day criteria for centroid stations from an initial regionalization and transferring these criteria with adjustments to other stations thus obtaining generic air masses which are analysed with respect to their spatial and temporal variation across western Europe.

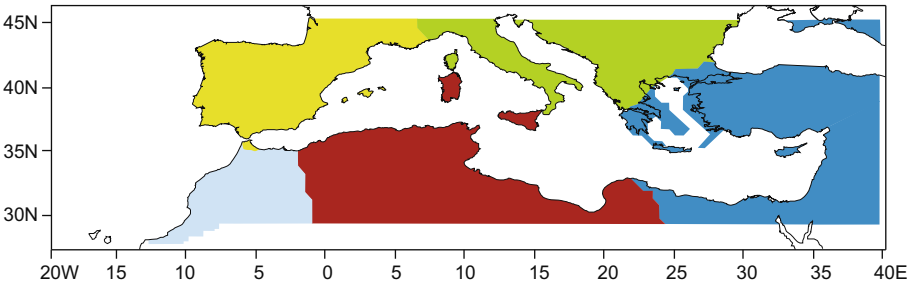
These few examples already demonstrate that multivariate statistical techniques have considerably enlarged potentials and performance of classifications in climate research. This has been completed by including the atmospheric circulation into applications and developments of classification approaches as discussed in the following sub-section.

### 2.2. Extending classification studies to the atmospheric circulation

This important line of advance in classification studies has introduced the dimensions of large-scale (so-called) weather patterns and of hemispheric, regional or meso-scale circulation types which characterize considerable parts of research in synoptic and dynamic climatology. Besides the well-known subjective or manual classifications which include a high degree of synoptic experience, but a lack of inter-subjective reproducibility – Lamb weather types for the UK (1972), Hess–Brezowsky Großwetterlagen with focus on Central Europe (1952) and many regional classifications as for example by Pédelaborde (1957) for the Paris basin, by Schüepp (1959) for the western Alpine region or by Lauscher (1972) for the eastern Alpine region – quantitative approaches gained increasing importance. Two different pathways may be identified in this respect: a first one tries to objectify manual classifications by defining numerical criteria or distance measures for the assignment of cases to predefined types as for example automated versions of the Lamb weather type classification (e.g. Jones et al., 1993; Goodess and Jones, 2002) or an objective catalogue determination of Hess–Brezowsky's Großwetterlagen (James, 2007). Similar methods have been developed by Dittmann et al. (1995) and by Beck (2000) aiming at objective assignments of individual circulation patterns to predefined circulation types. The latter approach (also described in Jacobeit et al. (1998) and Beck et al. (2007)) is based on the categories of Großwettertypes defined in terms of prototypical correlation coefficients with ideal patterns of zonal, meridional, and cyclonic flow above the study area (including central low and high pressure types and 8 directional flow types subdivided according to the sign of the vorticity coefficient). Individual pressure patterns are assigned to one of these Großwettertypes by (i) calculating correlation coefficients with the ideal flow patterns for each individual pressure pattern, (ii) removing those cases with maximum amount among the three correlation coefficients for the vorticity coefficient, being assigned to the central low or high pressure type, (iii) determining the maximum similarity (in terms of the Euclidean distance) between the actual pair of correlation coefficients with the ideal (zonal/meridional) flow patterns and the pair of prototypical ones for the Großwettertypes, (iv) determining the sub-type according to the sign of the actual vorticity coefficient. As shown in the above-mentioned publications, the composites derived from all members of the resulting sub-samples clearly reflect the characteristics of the pro-



**Fig. 1.** Regionalization of Germany into four climatic types by superimposing spatial score distributions of the two leading principal components resulting from a R-mode PCA of 27 mean climatic variables at 69 stations. This map has been derived from original maps in [Bahrenberg and Giese \(1975\)](#) for the two leading PCs representing mean conditions for temperature and precipitation, respectively. Warm/cool and wet/dry boundaries correspond to the mean values of the original PC scores.



**Fig. 2.** Mediterranean sub-regions with similar temperature variability during winter (DJ) 1948–1998 derived by a varimax-rotated S-mode PCA of gridded monthly mean temperature data from CRU (Norwich). Modified map from a Ph.D thesis ([Hertig, 2004](#)).

tototypical Großwettertypen. The added value of such objectifying studies may be seen in the definition of quantitative criteria and in the achievement of general reproducibility.

The other pathway mentioned above does not refer any more to predefined circulation types, but implements appropriate techniques which are able by themselves to generate some kind of partitions which have been called automated or objective classifications. These techniques may be grouped into methods based on correlation analyses, principal component analyses, cluster analyses, fuzzy rules and self-organizing maps. Since this paper does not intend to give a comprehensive review (like the one published by Huth et al., 2008), only some characteristic examples for these groups of statistical methods will be mentioned.

Correlation-based approaches have been introduced by Lund (1963) and recently been applied to reconstructed SLP data for historical periods (Schmutz and Wanner, 1998). A prominent example for PCA-based classifications is published by Huth (1996) applying the T-mode (time units as variables and grid points as cases) to geopotential height data and assigning original fields to principal components (which define the circulation classes) by the maximum loading (positive and negative). Clustering techniques are widely used and often combined with a preceding PCA thus reducing the number of variables for clustering cases (time units in this context). An example deriving North-Atlantic European circulation patterns by means of a conventional k-means cluster analysis of monthly mean SLP grids can be found in Jacobeit et al. (1998) and Beck (2000). For more complex data sets (e.g. daily fields with a large number of grid points for an extended study period) this method has been shown to be sensitive to starting partitions (e.g. Michelangeli et al., 1995) and the ordering of objects (Philipp et al., 2007), so that more sophisticated approaches have been developed for classification purposes. A recent example is the SANDRA clustering (simulated annealing and diversified randomization, see Philipp et al., 2007) which enhances the global optimum approximation (by allowing to leave local optima) and randomizes starting partitions and ordering of objects for multiple iterations. This method may also be extended to circulation pattern sequences (similar as in Principal Sequence Pattern Analysis, (see Compagnucci et al., 2001; Jacobeit et al., 2006) thus allowing to identify dynamical developments in classified circulation patterns (Philipp, 2008).

Another approach is based on fuzzy rules (e.g. Bardossy et al., 1995, 2002) which includes multiple memberships (with different weights) of individual patterns to circulation types and thereby may be seen as an intermediate state between PCA and strictly classifying techniques (Huth et al., 2008). Further details can be found in this volume (Bardossy, 2010). Finally, as the most important example for nonlinear methods, self-organizing maps (SOM) have to be mentioned (e.g. Kohonen, 2001; Hewitson and Crane, 2002; Reusch et al., 2007). They create a spatially organized set of generalized patterns of variability, allow to represent at the same time different aspects of variability (e.g. concerning mean and standard deviation), and include intermediate patterns reflecting smooth transitions between the distinct states of the classification.

Thus, the extension of classification studies to the atmospheric circulation combined with the development and application of multivariate statistical techniques has drastically changed the state-of-the-art in climate research. However, there are still remaining some deficiencies which make appropriate analyses quite difficult. A prominent example is the SOM technique which projects multidimensional data onto a two-dimensional array of patterns with the most dissimilar pairs at the opposite ends of the main diagonals (Huth et al., 2008). This constraint of two-dimensionality implies that only two basic states with a lot of intermediate patterns can be derived as for example in the study

of Reusch et al. (2007) with a North-Atlantic monopole and the well-known NAO dipole for describing North-Atlantic circulation variability. In contrast to that, each principal component analysis is able to identify more than two significant modes of circulation variability as may be seen in studies of long-term SLP data sets for both the T-mode (e.g. Jacobeit et al., 2001) and the S-mode (e.g. Luterbacher et al., 2002). For example, both studies include a major pattern with a strong Russian High and a mid-Atlantic Low which is well-known from synoptic climatology as a prominent pattern linked with distinct anomalies in regional climates. No such pattern, however, is found in the SOM study of Reusch et al. (2007) including the intermediate patterns (which do not represent some basic state, anyway). Thus, any method which is constrained by two-dimensionality does not seem to be appropriate for a sufficient representation of large-scale circulation variability, and further progress in the development of nonlinear methods is obviously required.

### 3. Fields of research applying circulation classifications

Classifying the atmospheric circulation is not an end in itself, but stimulates further investigations in climate research which may be grouped in two categories: studies of the circulation itself on the one hand and studies of circulation–climate relationships on the other hand.

#### 3.1. Atmospheric circulation dynamics

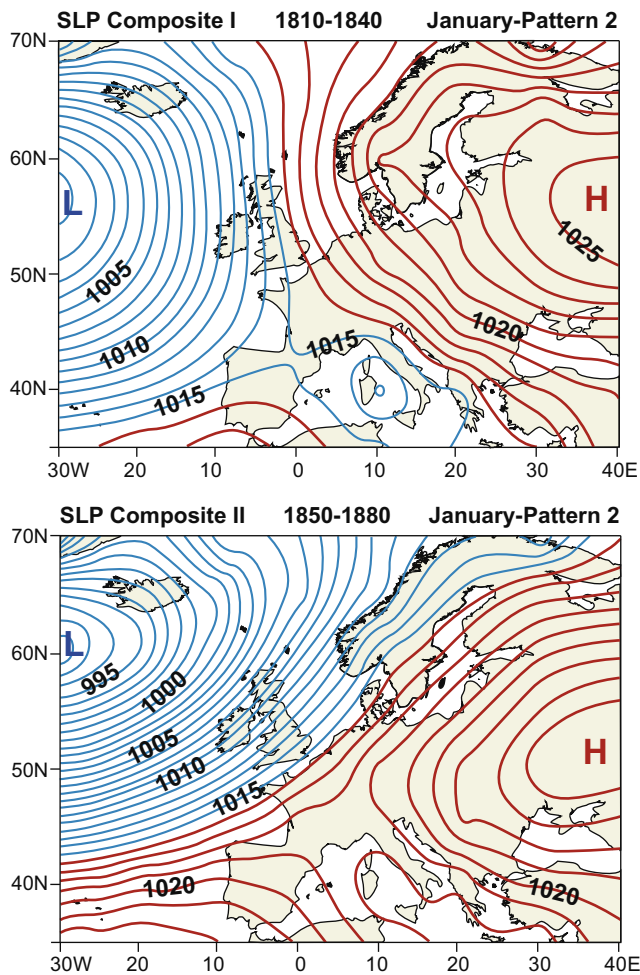
Classifications may be used to study temporal variations of the atmospheric circulation, most easily in terms of frequency changes of (individual or grouped) circulation types (e.g. Kysely and Huth, 2006; Philipp et al., 2007). Additionally, investigations on changes in persistence or mean residence times of circulation types have been done based on classifications thereby identifying the well-known increase in westerly-type persistence during the last decades in winter (Werner et al., 2000; Kysely and Domonkos, 2006). But also changes within particular circulation types have been addressed as for example in Fig. 3, showing SLP composites of the same January pattern (Russian High and mid-Atlantic Low) for different 30 year periods in the first and the second half of the 19th century (Jacobeit et al., 2001). An important within-type change from easterly to westerly components above Central Europe can be seen which has taken place around the mid-19th century according to sub-type frequency time series in Jacobeit et al. (2003) characterizing the transition from Little Ice Age to modern conditions in climate. Furthermore, analyses of circulation dynamics by applying particular classifications may also be extended to future climate change as for example done by Huth (1997, 2000) comparing frequencies and flow configurations of mid-tropospheric circulation types for GCM control and scenario runs.

Of course such results do not depend on the existence and application of classifications (see the comparison to other kinds of analyses in Section 4), but they reflect the scientific potential of classifications for studies on circulation dynamics.

#### 3.2. Circulation–climate relationships

This is a vast area in climate research, and only some particular examples in connection with circulation classifications can be discussed in this paper. An important approach is the study of trends in surface climate elements conditioned by circulation types. By this way Huth (2001) has shown for the Czech Republic that recent climate trends in summer have been unrelated to circulation changes whereas during winter a part of the observed warming could be explained by trends in the mid-tropospheric circulation.





**Fig. 3.** SLP composites (hPa) for different 30 year periods of a basic January circulation pattern characterized by a Russian High and an East Atlantic Low (modified after [Jacobeit et al., 2001](#)).

Similar seasonal differences arise from estimations of Central European temperature anomalies by composite and regression models based on seasonal frequencies of cluster circulation types ([Philipp et al., 2007](#)): explained variances increase to nearly 60% for winter whereas in the other seasons only values well below 40% are achieved. However, frequency changes of circulation types are only one factor influencing surface climate, another one are changes within circulation types ([Barry and Perry, 1973](#); [Brinkmann, 1999, 2000](#)), affecting dynamical (e.g. vorticity, intensity) as well as climatic (e.g. temperature, precipitation) characteristics of these types ([Beck et al., 2001](#); [Jacobeit et al., 2003](#)). Thus, in a study based on the same circulation types as in [Philipp et al. \(2007\)](#), [Jones and Lister \(2009\)](#) reveal that warming of some of these types has been a major factor for general climate warming trends compared to frequency increases of circulation types favouring positive temperature deviations in regional climate. The importance of within-type variations for Central European temperature and precipitation variability has already been shown by [Beck et al. \(2007\)](#) based on the above-mentioned objectified Großwettertype classification (Section 2.2.) and on a formula by [Barry and Perry \(1973\)](#) separating the frequency- and within-type-related variations in surface climate: except of spring time the larger part is always due to within-type variations reaching up to around 80% for the summer season.

Special focus in circulation–climate relationships is on dynamics of extremes. [Yiou and Nogaj \(2004\)](#) use so-called weather re-

gimes (classified states of the mid-tropospheric circulation) for assessing which one of them is the most important for the occurrence of temperature and precipitation extremes in different regions of the North-Atlantic European area, concluding that the Scandinavian blocking regime has a dominant impact on extremes in much of southern and high-latitude Northern Europe. [Jacobeit et al. \(2009\)](#) use the SLP cluster classification of [Philipp et al. \(2007\)](#) for identifying those circulation types being conducive to the occurrence of Central European precipitation and temperature extremes. [Fig. 4](#) gives an example referring to negative temperature extremes (below the 2nd percentile) during summer (JJA) 1850–2003:

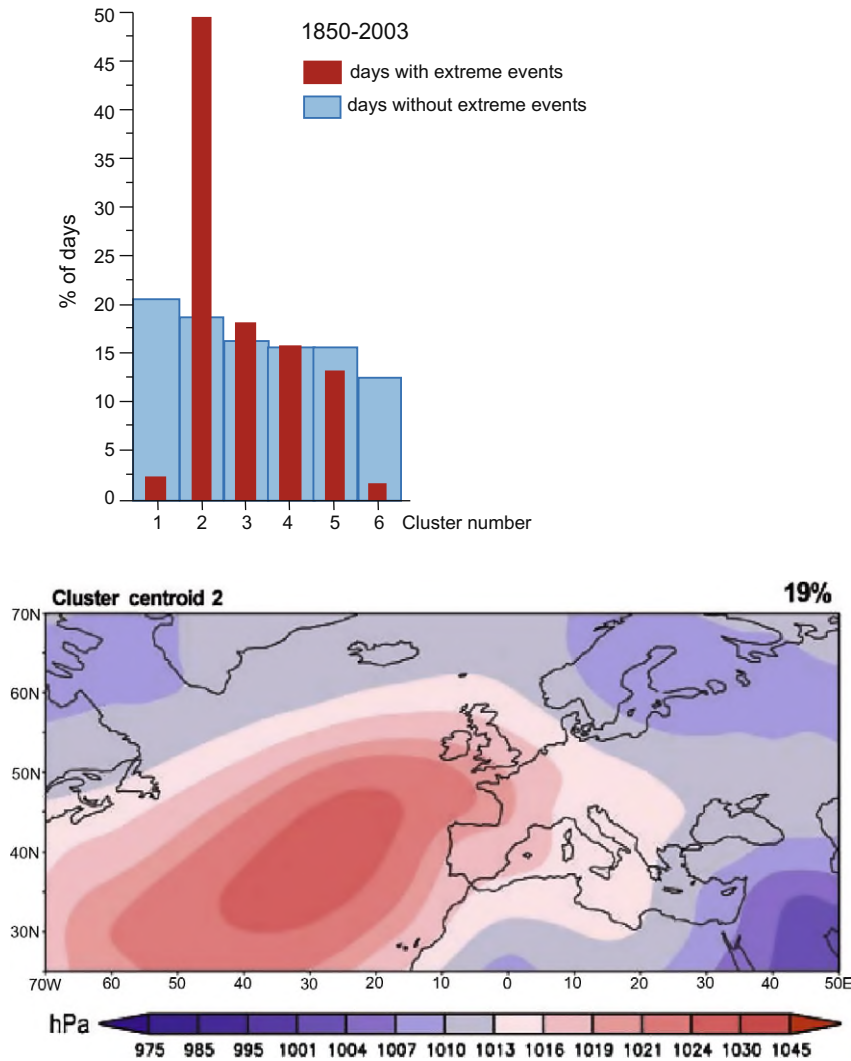
In terms of the ratio between the percentages of extreme days and of non-extreme days (number of extreme (non-extreme) days for a particular circulation type related to the overall occurrence of extreme (non-extreme) days), it is just one circulation type (see [Fig. 4](#)) out of six for the summer season for which this ratio is significantly (0.01 level) greater than 1 (tested by Monte-Carlo resampling of circulation type time series, i.e. for each circulation type this ratio is re-calculated for 1000 random circulation type time series and the actual ratio from [Fig. 4](#) is tested for being greater than the 99th percentile of all these random ratios). This type includes cold front penetrations from northerly directions and is remarkably linked to strong rain-out effects around Greenland which may be reconstructed on long-term time scales from stable isotope variability in ice-cores ([Rimbu and Lohmann, 2009](#)). A further aspect concerning links between circulation and extremes has been addressed by [Kysely \(2007\)](#) using the Hess–Brezowsky classification: comparing temperature anomalies from the mean annual cycle on successive days of heat and cold waves reveals that the observed intensification of positive temperature extremes is linked to a higher persistence of circulation types being conducive to the occurrence of these extremes.

Circulation–climate relationships are furthermore used for assessing future climate change by statistical downscaling techniques, and in this context circulation classifications have also been applied (e.g. [Goodess and Palutikof, 1998](#)). This may include that empirical transfer functions between the large-scale circulation and regional or local climate are established in dependence of classified circulation types (conditional downscaling) thereby improving explained variances of local weather elements ([Enke and Spekat, 1997](#)). Further developments of this approach ([Enke et al., 2005](#)) have resulted in a statistical regionalisation model (WETTREG) providing highly resolved climate projections for Germany in the 21st century ([Spekat et al., 2007](#)).

Finally it should be mentioned that weather or synoptic types are also used to study impacts on other quantities like for example particulate concentrations ([Davis and Gay, 1993](#); [Buchanan et al., 2002](#)) or viticultural potentials ([Jones and Davis, 2000](#)).

#### 4. Circulation types versus dynamical modes

For studying circulation dynamics and circulation–climate relationships scientists are not only using various classifications, but also – and very extensively – different kinds of EOF or PCA techniques providing basic circulation patterns or modes of variability or even circulation regimes. In this context we will not continue the discussion about these different terms (see [Stephenson et al., 2004](#)) and not focus on particular levels of frequency or persistence which might justify to speak about regimes, but we aim at pointing out fundamental differences in methods and corresponding fields of application for classification techniques on the one hand and EOF or PCA approaches on the other hand (for the latter see for example the excellent book publication from [Jolliffe, 2002](#)).



**Fig. 4.** Upper part: percentages of days with extremely low Central European temperatures (below the 2nd percentile) and without such extremes for the six summer (JJA) circulation types of a SLP cluster analysis, related to the overall occurrences in the 1850–2003 period. Lower part: SLP (hPa) centroid pattern for Cluster 2 according to the SANDRA classification for summer (JJA) 1850–2003 (modified after Philipp et al., 2007).

At first we have to consider particular modes of analysis (Richman, 1986) being relevant in this context. Since mostly data with spatial and temporal resolution are processed, S- and T-mode are the dominant ones. As pointed out by Compagnucci and Richman (2008), the former provides teleconnection patterns, the latter atmospheric circulation patterns, due to the basic difference that principal components represent (with respect to a particular quantity like SLP or geopotential heights) characteristic time series for spatial centres of variation in the first case (examples in Jacobeit et al., 1998 or Kysely and Huth, 2006), but spatial distribution patterns being dominant at particular time points in the second case (e.g. Huth, 1993; Jacobeit, 1993). For a comparison with circulation type classifications T-mode analysis is therefore the appropriate candidate. In order to have a clear-cut difference in technical terms against circulation types resulting from classification approaches, we will further speak about dynamical modes concerning T-mode PCA, and this must not be confused with so-called modes of variability as a result from S-mode analyses.

Dynamical modes are generated by some kind of eigenvector analysis allowing to derive a set of equations which links these modes to the original variables (e.g. individual pressure fields) in the following way:

$$V_i = \sum_{j=1}^m l_{ij} * DM_j + R_i$$

$i = 1, \dots, n$ ,  $m < n$ ,  $V_i$  is the original variable,  $DM_j$  is the dynamical mode,  $l_{ij}$  is the loadings of  $DM_j$ ,  $R_i$  is the residual term,  $n$  is the number of original variables,  $m$  is the number of dynamical modes. Thus, we get for each variable – in our context for a pressure field at each time point  $i$  – a set of indices ( $l_{i1}, \dots, l_{im}$ ) describing the multivariate relation to the set of  $m$  dynamical modes. In contrast to that, the result of a classification is just one figure  $k$  for each variable (e.g. for each pressure field) with  $k$  being the current class number. Obviously, there is less information than in the former case reflecting a certain degree of simplification. Fuzzy rules including multiple memberships have already been mentioned as an intermediate state between these opposite cases (Section 2.2). However, this difference is not even decisive, since a set of indices for each individual pressure field could also be generated for a classification into disjunctive circulation types, for example by correlation or regression analysis subsequent to the classification. Furthermore, results from T-mode analysis are also used in the sense of classifications, for example when studying within-mode variability (e.g. Jacobeit et al., 2003) since this may require an unambiguous assignment of

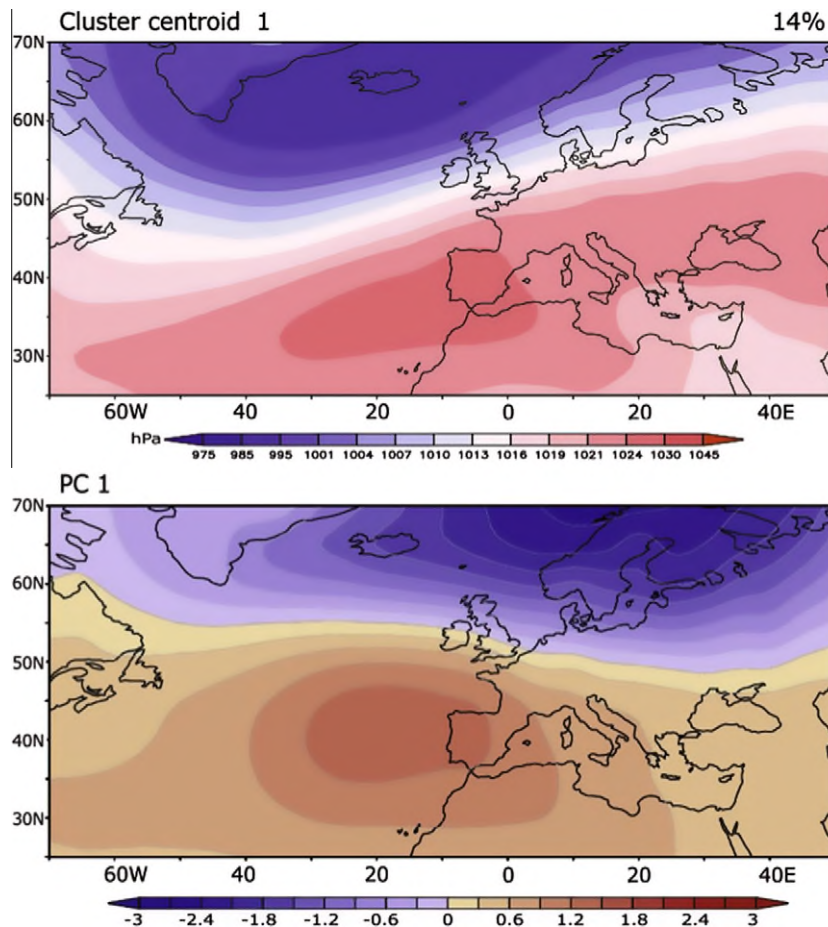
individual pressure fields to one of the dynamical modes (achieved by the maximum loading as already mentioned for the PCA-based classification approach, see Section 2.2).

In view of that we may ask for particular advantages of mode analysis compared to strict classifications. Three aspects should be considered in this context:

- (i) For complex data sets with no 'natural' separation into well-defined sub-sets of elements – as for example daily pressure fields with a large number of grid points for a long time period – the determination of an appropriate number of classes (in this case circulation types) may prove to be ambiguous (e.g. Philipp et al., 2007). Dynamical modes, however, which do not define a disjunctive separation of the data set but represent basic states with varying weight in the data set, may be extracted according to particular criteria being adjusted to data characteristics and scientific aims. Thus, Philipp et al. (2007) use a varimax-rotated T-mode PCA to determine (by a set of particular criteria) the number of orthogonal modes that fit into the daily mean SLP data set from 1850 to 2003 and take this number as the prescribed number of circulation types being classified by a simulated annealing clustering technique.
- (ii) In contrast to classification, mode analysis is not constrained to assign each object to one category, and this might be more appropriate for complex data sets without 'natural' separation – provided that such an assignment is not necessary

or convenient (as for example in studies referring to daily events). Thus, even the SANDRA classification including a highly effective optimization (see Section 2.2) is not able to increase the explained cluster variance for daily winter circulation types beyond 48.3% (Philipp et al., 2007) due to the large within-cluster differences arising with the complete assignment of individual pressure fields to circulation types. On the other hand, dynamical modes from a T-mode analysis of the same data – without using them for classification – are not intended to account for the whole variance in the data set, they do so by a certain percentage (88% in this case, referring to the original data and not to a classified structure as above) with the remaining part being negligible noise.

- (iii) A major difference between classification and mode analysis has to be considered with respect to the resulting circulation patterns. Fig. 5 reproduces the first ones from both analyses applied to the daily mean SLP grids from winter (DJF) 1850–2003. They look very similar both representing a westerly flow pattern which characterizes the positive phase of the NAO. However, there is a principal difference in origin: the centroid pattern from the cluster classification results from averaging all the different members of the corresponding cluster, whereas the T-mode scores are not averaged from individual objects, they represent a generic circulation pattern, some kind of a prototype with varying degrees of similarity to the original pressure fields (expressed by the



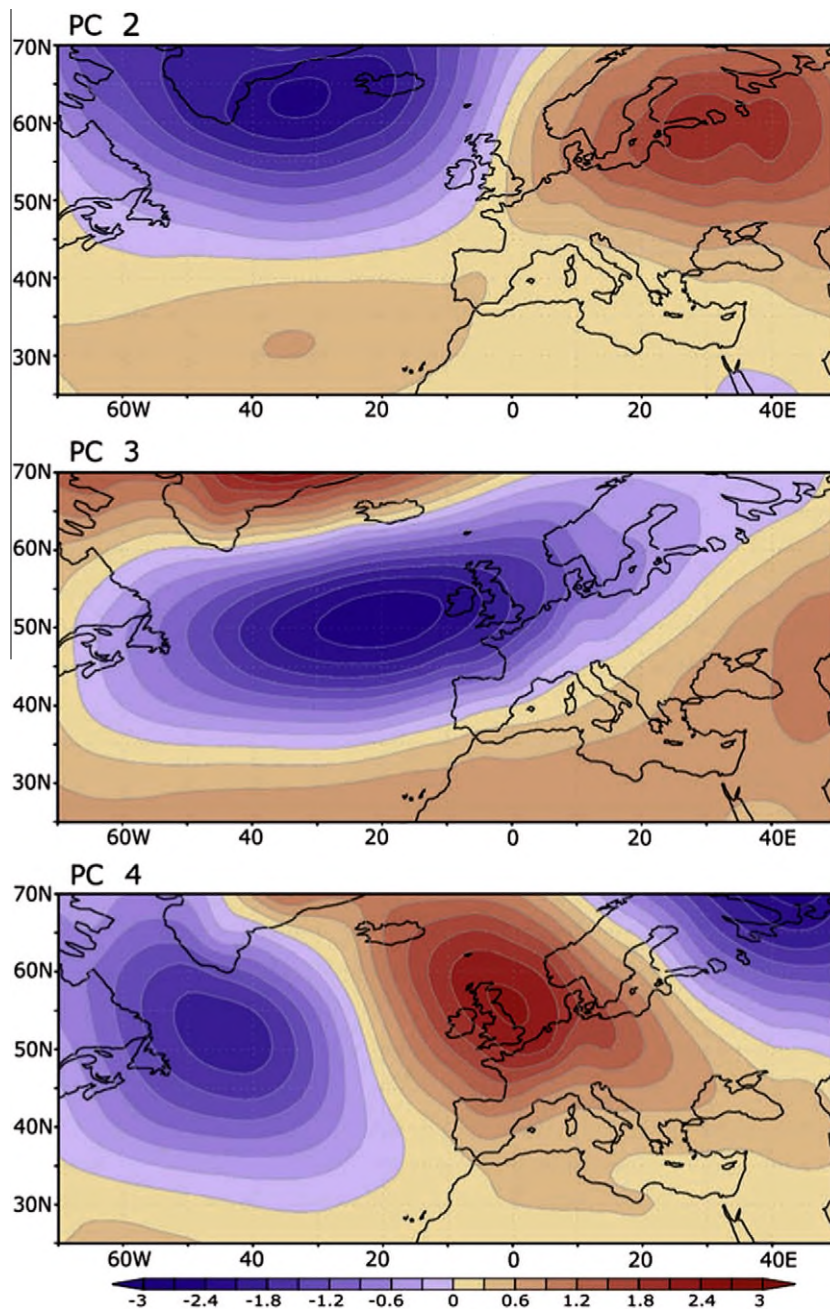
**Fig. 5.** Leading North-Atlantic European circulation patterns for winter (DJF) 1850–2003 derived from different analyses. Upper part: SLP (hPa) centroid pattern for Cluster 1 according to the SANDRA classification (modified after Philipp et al., 2007). Lower part: normalized scores of the first principal component (27.1% explained variance) resulting from a varimax-rotated T-mode PCA.



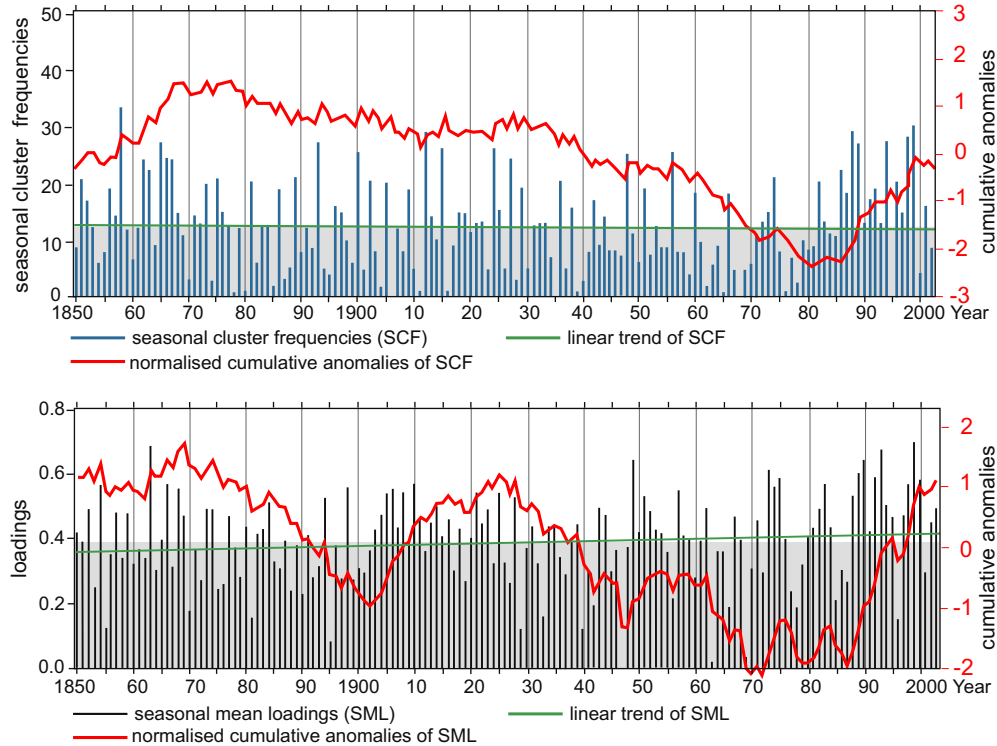
T-mode loadings). Thus, these prototypes should not be replaced by composites with a lot of objects entering which differ from the prototypes (and this is done by classification!). Instead we get the ensemble of individual pressure fields as a result of varying superimpositions of prototypes, or in other terms – following the idea that ‘dynamical modes’ are reflected by these prototypes – individual objects are composed of a combination of dynamical modes with varying weights (as expressed in the above-mentioned equation). This means that more than one dynamical mode contributes to the actual flow configuration with the dynamical modes themselves remaining undisturbed from classification. Furthermore, it is justified to speak of prototype patterns reflecting these modes, since they are no *artefacts*

but well-known from observation (see for example Fig. 6) with highly correlating representatives among the original pressure fields (e.g. greater than 0.9 for the PC patterns of Figs. 5 and 6). If a description of dynamical modes should still be preferred in terms of original data, these representatives being most similar to the prototypes should be used, but no composites including more and more dissimilar patterns (corresponding to a classification approach).

Considering these differences between classification and mode analysis, it is not surprising that the time coefficients of the two westerly patterns from Fig. 5 may differ quite substantially (Fig. 7). Seasonal cluster frequencies do not reveal a long-term trend and have only a few turning points in the cumulative anom-



**Fig. 6.** Normalized scores of principal components 2–4 (explained variances are 26.3%, 10.5%, and 8.9%) resulting from a varimax-rotated T-mode PCA of daily mean SLP for winter (DJF) 1850–2003.



**Fig. 7.** Seasonal time coefficients with normalized cumulative anomalies and linear trends for the leading North-Atlantic European circulation patterns from Fig. 5 (winter (DJF) 1850–2003). Upper part: Seasonal frequencies of Cluster 1 according to the SANDRA classification (modified after Philipp et al., 2007). Lower part: seasonally averaged loadings of the first principal component resulting from a varimax-rotated T-mode PCA. The upper level of the grey-shaded areas corresponds to the long-term mean value of the time coefficients.

alies. Seasonal mean T-mode loadings, however, include a long-term trend (though not significant in this case) and in particular several sub-periods with varying sign of the cumulative anomalies. These differences are mainly due to the fact that the SANDRA classification includes two further clusters with westerly patterns, one of them is reproduced in Fig. 8 showing in particular a significant (0.05 level) upward trend of its seasonal cluster frequencies.

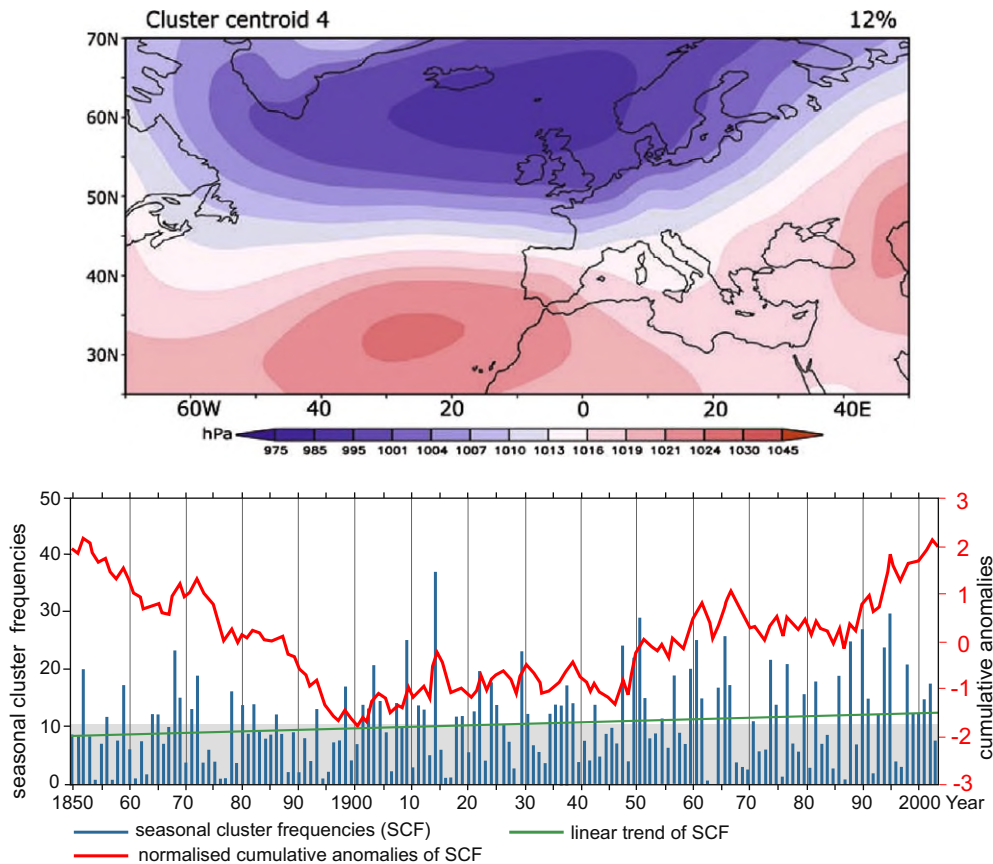
How to judge these differences? The cluster classification having no constraint of orthogonality is able to separate several westerly patterns with regional differences (e.g. concerning the detailed positions of major centres of action), and in fact this differentiation proves to be relevant for climatic extremes in Central Europe (Jacobeit et al., 2009): positive temperature extremes during winter (beyond the 98th percentile) are linked preferably to zonal circulation patterns as seen in Cluster 1 (Fig. 5) whereas heavy winter precipitation is distinctly associated with other westerly patterns (see one of them in Fig. 8) characterized by an eastward or southeastward shift of the subpolar centre of low pressure. On the other hand, the time coefficient of Cluster 1 (Figs. 5 and 7) being that with the highest NAO correlation does not reflect the real behaviour of the NAO mode since several components of this mode are dismembered into different clusters. Thus, the study of large-scale dynamics in the atmosphere seems to be more appropriate in terms of mode analysis (including modes of variability mentioned earlier), and this may cover a wide range of objectives: hemispheric (or continental-scale) circulation patterns and their low-frequency variations in observed data (e.g. Barnston and Livezey, 1987; Jacobeit, 1993), in reconstructed data (e.g. Luterbacher et al., 2002; Jacobeit et al., 2001, 2003) and in model data (e.g. Jacobeit, 1994, 1996; Huth, 1997), teleconnection studies (e.g. Philipp, 2003; Compagnucci and Richman, 2008), statistical downscaling from large-scale modes including seasonal forecasts (e.g. Jacobeit,

1996; Hertig and Jacobeit 2008a,b, submitted for publication), principal sequence pattern analyses (e.g. Compagnucci et al., 2001; Jacobeit et al., 2006) or investigations on coupled modes of variability between circulation and climate (e.g. by Canonical Correlation Analysis, see Dünkloh and Jacobeit, 2003; Della-Marta et al., 2007). Thus, an appropriate decision between classification and mode analysis depends on the particular focus of each study, and this will be summarized in the concluding section.

## 5. Conclusion

Classifications in climate research have seen significant progress during the last decades mainly due to the increasing application of multivariate statistical techniques and the inclusion of dynamical aspects related to the atmospheric circulation. Correspondingly, there has been a growing field of applications of these classifications in particular with respect to studies on circulation dynamics and circulation–climate relationships (see Huth et al., 2008). Therefore it should be emphasized which fields of application are the appropriate ones in contrast to cases for which we should better rely on techniques summarized as mode analyses.

Classifications as discussed in the present context imply a strict assignment of individual objects to disjunctive groups, and this may be a serious simplification leading to considerable dissimilarities within these groups, especially in case of complex data sets. However, there are many research questions for which such a strict assignment is necessary or at least favourable as for example any direct relation of the atmospheric circulation to events on the time scale of the temporal resolution (e.g. daily occurrences of weather characteristics, extreme events or particular phenomena). In this case some classification approach will be an appropriate tool to re-



**Fig. 8.** Cluster 4 according to the SANDRA classification of winter (DJF) daily mean SLP 1850–2003 (modified after Philipp et al., 2007). Upper part: SLP (hPa) centroid pattern. Lower part: seasonal cluster frequencies with normalized cumulative anomalies and linear trend. The upper level of the grey-shaded area corresponds to the long-term mean value of the cluster frequencies.

late the atmospheric circulation to these daily occurrences on a regional or local scale. Furthermore, the internal characterization of atmospheric circulation patterns in terms of dynamic parameters (e.g. relative vorticity or pressure gradient) and climatic variables (e.g. temperature or rainfall) likewise requires an unambiguous assignment of temporal units to classified states, and this is even true for dynamical modes and their within-mode variations which are often assessed on the basis of an implicit classification (e.g. by assigning each time unit to that mode with the highest loading on it, see Jacobeit et al., 2003).

On the other hand, if such requirements (as mentioned in the previous paragraph) do not exist, it is not necessary for investigations on climate and circulation dynamics to rely on classifications and to accept their shortcomings mentioned above. Instead we may benefit from multivariate representations of large-scale fields and their variations by different techniques of mode analysis. In particular the capability to derive generic circulation patterns (instead of composite patterns averaged from individual members of classified groups) and to represent individual pressure fields as superimpositions of such generic patterns with varying weights seems to be more appropriate for a complex reality than strictly assigning objects to disjunctive groups (as done by classifications). This should be considered especially for research on large-scale dynamics.

## References

Bahrenberg, G., Giese, E., 1975. Statistische Methoden und ihre Anwendung in der Geographie (Statistical Methods and their Application in Geography). Stuttgart, p. 308.

- Bardossy, A., Duckstein, L., Bogardi, I., 1995. Fuzzy-rule based classification of atmospheric circulation patterns. *Int. J. Climatol.* 15, 1087–1097.
- Bardossy, A., Stehlik, J., Caspary, H.-J., 2002. Automated objective classification of daily circulation patterns for precipitation and temperature downscaling based on optimized fuzzy rules. *Clim. Res.* 23, 11–22.
- Bardossy, A., 2010. Atmospheric circulation pattern classification for South-West Germany using hydrological variables. *Phys. Chem. Earth* 35, 498–506.
- Barnston, A.G., Livezey, R.E., 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Weather Rev.* 115, 1083–1126.
- Barry, R.G., Perry, A.H., 1973. *Synoptic Climatology*. Methuen, London.
- Beck, C., 2000. Zirkulationsdynamische Variabilität im Bereich Nordatlantik-Europa seit 1780 (Variability of circulation dynamics in the North-Atlantic European area since 1780). *Würzburger Geographische Arbeiten* 95, 350.
- Beck, C., Jacobeit, J., Philipp, A., 2001. Variability of North-Atlantic-European circulation patterns since 1780 and corresponding variations in central European climate. In: Brunet India, M., Lopez Bonillo, D. (Eds.), *Detecting and Modelling Regional Climate Change*. pp. 321–332.
- Beck, C., Jacobeit, J., Jones, P.D., 2007. Frequency and within-type variations of large scale circulation types and their effects on low-frequency climate variability in Central Europe since 1780. *Int. J. Climatol.* 27, 473–491.
- Bower, D., McGregor, G.R., Hannah, D.M., et al., 2007. Development of a spatial synoptic classification scheme for western Europe. *Int. J. Climatol.* 27, 2017–2040.
- Brinkmann, W.A.R., 1999. Within-type variability of 700 hPa winter circulation patterns over the Lake Superior Basin. *Int. J. Climatol.* 19, 41–58.
- Brinkmann, W.A.R., 2000. Modification of a correlation-based circulation pattern classification to reduce within-type variability of temperature and precipitation. *Int. J. Climatol.* 20, 839–852.
- Buchanan, C.M., Beverland, I.J., Heal, M.R., 2002. The influence of weather-type and long-range transport on airborne particle concentrations in Edinburgh, UK. *Atmos. Environ.* 36, 5343–5354.
- Compagnucci, R.H., Richman, M.B., 2008. Can principal component analysis provide atmospheric circulation or teleconnection patterns? *Int. J. Climatol.* 28, 703–726. doi:10.1002/joc.1574.
- Compagnucci, R.H., Araneo, D., Canziani, P.O., 2001. Principal sequence pattern analysis: a new approach to classifying the evolution of atmospheric systems. *Int. J. Climatol.* 21, 197–217.



- Creutzburg, N., 1950. Klima, Klimatypen und Klimakarten (Climate, climate types and climatic maps). *Petermanns Geogr. Mitt.* 94, 57–69.
- Davis, R.E., Gay, D.A., 1993. An assessment of air quality variations in the south-western USA using an upper air synoptic climatology. *Int. J. Climatol.* 13, 755–781.
- Della-Marta, P. et al., 2007. Summer heat waves over western Europe 1880–2003, their relationship to large-scale forcings and predictability. *Clim. Dynam.* 29, 251–275.
- Dittmann, E., Barth, S., Lang, J., Müller-Westermeier, G., 1995. Objektive Wetterlagenklassifikation (Objective weather type classification). *Ber. Dt. Wetterdienst* 197. Offenbach a. M.
- Düneloh, A., Jacobeit, J., 2003. Circulation dynamics of mediterranean precipitation variability 1948–1998. *Int. J. Climatol.* 23, 1843–1866.
- El-Kadi, A.K.A., Smithson, P.A., 1992. Atmospheric classifications and synoptic climatology. *Prog. Phys. Geogr.* 16, 432–455.
- Enke, W., Spekat, A., 1997. Downscaling climate model outputs into local and regional weather elements by classification and regression. *Clim. Res.* 8, 195–207.
- Enke, W.F., Schneider, F., Deutschländer, Th., 2005. A novel scheme to derive optimized circulation pattern classifications for downscaling and forecast purposes. *Theor. Appl. Climatol.* 82, 51–63.
- Gerstengarbe, F.-W., Werner, P.C., 2007. Der rezente Klimawandel (Recent climate change). In: Endlicher, W., Gerstengarbe, F.-W. (Eds.), *Der Klimawandel: Einblicke, Rückblicke und Ausblicke*, Potsdam, pp. 34–43.
- Gerstengarbe, F.-W., Werner, P.C., Fraedrich, K., 1999. Applying non-hierarchical cluster analysis algorithms to climate classification: some problems and their solution. *Theor. Appl. Climatol.* 64, 143–150.
- Goodess, C.M., Jones, P.D., 2002. Links between circulation and changes in the characteristics of Iberian rainfall. *Int. J. Climatol.* 22, 1593–1615.
- Goodess, C.M., Palutikof, J.P., 1998. Development of daily rainfall scenarios for southeast Spain using a circulation-type approach to downscaling. *Int. J. Climatol.* 18, 1051–1083.
- Hertig, E., 2004. Niederschlags und Temperaturabschätzungen für den Mittelmeerraum unter anthropogen verstärktem Treibhauseffekt (Precipitation and temperature assessments for the Mediterranean area under conditions of man-made increased greenhouse warming). – Ph.D. Thesis, University of Würzburg, pp. 274. <<http://opus.bibliothek.uni-wuerzburg.de/opus/volltexte/2004/874/>>.
- Hertig, E., Jacobeit, J., 2008a. Assessments of Mediterranean precipitation changes for the 21st century using statistical downscaling techniques. *Int. J. Climatol.* 28, 1025–1045.
- Hertig, E., Jacobeit, J., 2008b. Downscaling future climate change: temperature scenarios for the mediterranean area. *Global Planet. Change* 63, 127–131.
- Hertig, E., Jacobeit, J., submitted for publication. Predictability of mediterranean climate variables from oceanic variability.
- Hess, P., Brezowsky, H., 1952. Katalog der Großwetterlagen Europas. *Ber. Dt. Wetterd.* in der US-Zone 33. Bad Kissingen.
- Hewitson, B.C., Crane, R.G., 2002. Selforganizing maps: applications to synoptic climatology. *Clim. Res.* 22, 13–26.
- Huth, R., 1993. An example of using obliquely rotated principal components to detect circulation types over Europe. *Meteorol. Z.* 2, 285–293.
- Huth, R., 1996. An intercomparison of computer assisted circulation classification methods. *Int. J. Climatol.* 16, 893–922.
- Huth, R., 1997. Continental-scale circulation in the UKHI GCM. *J. Clim.* 10, 1545–1561.
- Huth, R., 2000. A circulation classification scheme applicable in GCM studies. *Theor. Appl. Climatol.* 67, 1–18.
- Huth, R., 2001. Disaggregating climatic trends by classification of circulation patterns. *Int. J. Climatol.* 21, 135–153.
- Huth, R., Beck, C., Philipp, A., Demuzere, M., Ustrnul, Z., Cahynova, M., Kysely, J., Tveit, O.E., 2008. Classifications of atmospheric circulation patterns: recent advances and applications. *Ann. NY Acad. Sci.* 1146, 105–152.
- Jacobeit, J., 1993. Regionale Unterschiede im atmosphärischen Zirkulationsgeschehen bei globalen Klimaveränderungen (Regional differences of the atmospheric circulation under conditions of global climate change). *Die Erde* 124, 63–77.
- Jacobeit, J., 1994. Atmosphärische Zirkulationsveränderungen bei anthropogen verstärktem Treibhauseffekt (Atmospheric circulation changes due to man-made enhanced greenhouse warming). *Würzburger Geographische Manuskripte* 34, p. 101.
- Jacobeit, J., 1996. Atmospheric circulation changes due to increased greenhouse warming and its impact on seasonal rainfall in the Mediterranean area. In: Nemesova, I. (Ed.), *Climate Variability and Climate Change Vulnerability and Adaptation*, Praha, pp. 71–80.
- Jacobeit, J., Beck, C., Philipp, A., 1998. Annual to decadal variability in climate in Europe – Objectives and results of the German contribution to the European climate research project ADVICE. *Würzburger Geographische Manuskripte* 43, 163.
- Jacobeit, J., Jones, P.D., Davies, T., Beck, C., 2001. Circulation changes in Europe since the 1780s. In: Jones, P.D., Ogilvie, A.E., Davies, T.D., Briffa, K.R. (Eds.), *History and Climate: Memories of the Future*. Kluwer Academic/Plenum Publishers, New York, Boston, Dordrecht, London, Moscow, pp. 79–100.
- Jacobeit, J., Wanner, H., Luterbacher, J., Beck, C., Philipp, A., Sturm, K., 2003. Atmospheric circulation variability in the North-Atlantic-European area since the mid-seventeenth century. *Clim. Dynam.* 20, 341–352.
- Jacobeit, J., Philipp, A., Nonnenmacher, M., 2006. Atmospheric circulation dynamics linked with prominent discharge events in Central Europe. *J. Hydrol. Sci.* 51, 946–965.
- Jacobeit, J., Rathmann, J., Philipp, A., Jones, P.D., 2009. Central European precipitation and temperature extremes in relation to large-scale atmospheric circulation types. *Met. Zeitschrift* 18, 397–410.
- James, P.M., 2007. An objective classification method for Hess and Brezowsky Grosswetterlagen over Europe. *Theor. Appl. Climatol.* 88, 17–42.
- Jolliffe, I.T., 2002. *Principal Component Analysis*, second ed. New York, p. 487.
- Jones, G.V., Davis, R.E., 2000. Using a synoptic climatological approach to understand climate viticulture relationships. *Int. J. Climatol.* 20, 813–837.
- Jones, P.D., Lister, D., 2009. The influence of the circulation on surface temperature and precipitation patterns over Europe. *Clim. Past Discuss.* 5, 535–555.
- Jones, P.D., Hulme, M., Briffa, K.R., 1993. A comparison of Lamb circulation types with an objective classification scheme. *Int. J. Climatol.* 13, 655–663.
- Kohonen, T., 2001. *Self-organizing Maps*. New York.
- Köppen, W., 1923. *Die Klimate der Erde (Climates of the Earth)*. Berlin/Leipzig.
- Kysely, J., 2007. Implications of enhanced persistence of atmospheric circulation for the occurrence and severity of temperature extremes. *Int. J. Climatol.* 27, 689–695.
- Kysely, J., Domonkos, P., 2006. Recent increase in persistence of atmospheric circulation over Europe: comparison with long-term variations since 1881. *Int. J. Climatol.* 26, 461–483.
- Kysely, J., Huth, R., 2006. Changes in atmospheric circulation over Europe detected by objective and subjective methods. *Theor. Appl. Climatol.* 85, 19–36.
- Lamb, H.H., 1972. British Isles weather types and a register of daily sequence of circulation patterns, 1861–1971. *Geophysical Memoirs* 116, HMSO, London, p. 85.
- Lauer, W., Rafiqpoor, M.D., 2002. *Die Klimate der Erde (Climates of the earth)*. Erdwissenschaftliche Forschung, vol. XL, Stuttgart.
- Lauer, W., Rafiqpoor, M.D., Frankenberg, P., 1996. *Die Klimate der Erde (Climates of the earth)*. *Erdkunde* 50, 275–300.
- Lauscher, F., 1972. 25 Jahre mit täglicher Klassifikation der Wetterlage in den Ostalpenländern (25 years with daily weather type classification in eastern alpine countries). *Wetter und Leben* 24, 185–189.
- Lund, I.A., 1963. Map-pattern classification by statistical methods. *J. Appl. Meteorol.* 2, 56–65.
- Luterbacher, J. et al., 2002. Reconstruction of Sea level pressure fields over the Eastern North Atlantic and Europe back to 1500. *Clim. Dynam.* 18, 545–561.
- Michelangeli, P.-A., Vautard, R., Legras, B., 1995. Weather regimes: recurrence and quasi stationarity. *J. Atmos. Sci.* 52, 1237–1256.
- Pédalaborde, P., 1957. *Le climat du Bassin Parisien. Essai d'une méthode rationnelle de climatologie physique*, Paris.
- Philipp, A., 2003. Zirkulationsdynamische Telekonnektivität des Sommerniederschlags im südhemisphärischen Afrika (Teleconnectivity of summer rainfall in southern hemispheric Africa with respect to circulation dynamics). Ph.D. Thesis, University of Würzburg, p. 400. <<http://opus.bibliothek.uni-wuerzburg.de/opus/volltexte/2004/810/>>.
- Philipp, A., 2008. Comparison of principal component and cluster analysis for classifying circulation pattern sequences for the European domain. *Theor. Appl. Climatol.* doi:10.1007/200704-008-0037-1.
- Philipp, A., Della-Marta, P.M., Jacobeit, J., Fereday, D., Jones, P., Moberg, A., Wanner, H., 2007. Long-term variability of daily North Atlantic-European pressure patterns since 1850 classified by simulated annealing clustering. *J. Clim.* 20, 4065–4095.
- Reusch, D.B., Alley, R.B., Hewitson, B.C., 2007. North Atlantic climate variability from a self organizing map perspective. *J. Geophys. Res.* 112, D02104.
- Richman, M.B., 1986. Rotation of principal components. *J. Climatol.* 6, 293–335.
- Rimbu, N., Lohmann, G., 2009. Daily atmospheric circulation patterns from the North Atlantic region as recorded in high-resolution stable isotope records from Greenland ice cores. *Geophys. Res. Abstr.* 11, EGU2009-6594-1.
- Schmutz, C., Wanner, H., 1998. Low frequency variability of atmospheric circulation over Europe between 1785 and 1994. *Erdkunde* 52, 81–94.
- Schüepp, M., 1959. Die Klassifikation der Witterungslagen (Classification of large-scale weather types). *Geofis. Pura Appl.* 44, 242–248.
- Spekat, A., Enke, W., Kreienkamp, F., 2007. Neuentwicklung von regional hoch aufgelösten Wetterlagen für Deutschland und Bereitstellung regionaler Klimaszenarios mit dem Regionalisierungsmodell WETTREG auf der Basis von globalen Klimasimulationen mit ECHAM5/MPI-OM T63L31 2010–2100 für die SRES-Szenarios B1, A1B und A2. Publ. d. UBA, Dessau, p. 149.
- Stephenson, D.B., Hannachi, A., Allen, M., 2004. On the existence of multiple climate regimes. *Quart. J. Roy. Meteor. Soc.* 130, 583–605.
- Thornthwaite, C.W., 1933. The climates of the earth. *Geogr. Rev.* 23, 433–440.
- Troll, C., 1958. Climatic seasons and climatic classification. *Orient. Geogr.* 2, 141–165.
- Werner, P.C., Gerstengarbe, F.-W., Fraedrich, K., Oesterle, H., 2000. Climate change in the North Atlantic-European Sector. *Int. J. Climatol.* 20, 463–471.
- Yarnal, B., 1993. *Synoptic Climatology in Environmental Analysis*. Belhaven Press, London, UK, p. 195.
- Yarnal, B., Comrie, A.C., Frakes, B., et al., 2001. Developments and prospects in synoptic climatology. *Int. J. Climatol.* 21, 1923–1950.
- Yiou, P., Nogaj, M., 2004. Extreme climatic events and weather regimes over the North Atlantic: when and where? *Geophys. Res. Lett.* 31, L07202.