The effect of domain size on the relationship between circulation type classifications and surface climate

Christoph Beck,* Andreas Philipp and Florian Streicher

Institute of Geography, University of Augsburg, Augsburg, Germany

ABSTRACT: The ability of circulation type classifications (CTCs) to resolve surface climatic and environmental variables is essential with respect to their application in synoptic climatological applications. This 'synoptic skill' depends on several factors including inherent properties of classification methods but as well varying boundary conditions. In this contribution the relevance of the size of the spatial domain for which CTCs are derived is investigated. To this end varying automated CTCs are applied to daily gridded sea level pressure data for 1950–2010 and in each case eight spatial domains of varying size centred around 44 locations spread over the greater north Atlantic European region. For the resulting more than 7000 CTCs the synoptic skill for daily temperature and precipitation taken from the E-OBS v4.0 data set has been estimated using varying evaluation metrics. Resulting values of evaluation metrics aggregated according to varying domain sizes reveal a distinct influence of the size of the domain on the synoptic skill of CTCs. In general highest skill appears to be achieved for domain sizes with a horizontal dimension of roughly 1300–1800 km (in west–east direction) thus covering most frequent size ranges of synoptic scale systems. However, optimal domain sizes tend to be smaller for precipitation (compared to temperature) in summer (compared to winter) and in more continental regions (compared to more oceanic regions). Distinct deviations from the overall finding of relatively small optimal domains emerge for temperatures above/below certain thresholds for which in certain locations and seasons continental scale domains yield highest synoptic skill. Finally the comparison of varying CTCs concerning the effect of the domain size for synoptic skill shows marked differences between methods and moreover clearly elucidates that differences in synoptic skill that can be attributed to varying domain sizes reach comparable magnitude than those related to differing methods.

KEY WORDS synoptic skill; synoptic climatology; atmospheric circulation; circulation types; classification; spatial domain; North Atlantic; Europe

1. Introduction

Classifications of atmospheric circulation types are an important and often used tool for the synoptic climatological analysis of large-scale circulation dynamics and its influence on local and regional surface climatic and environmental variables. Numerous different methodological approaches are used for circulation classification (Huth *et al.*, 2008; Philipp *et al.*, 2010) and resulting circulation types are related to varying climatic or environmental variables and phenomena, i.e., air temperature (Beck *et al.*, 2007; Kysely, 2007), precipitation (Goodess and Jones, 2002; Beck *et al.*, 2007), drought (Vicente-Serrano and Lopez-Moreno, 2006; Fleig *et al.*, 2010), storminess (Leckebusch *et al.*, 2008), snowfall (Esteban *et al.*, 2005), snow avalanches (Garcia *et al.*, 2009), wild fires (Kassomenos, 2010), lightning (Pineda *et al.*, 2010), visibility (Dayan and Levy, 2005), natural (Makra *et al.*, 2006) and anthropogenic (Buchanan *et al.*, 2002; Demuzere *et al.*, 2011) aerosols and ozone (Demuzere *et al.*, 2009; Tang *et al.*, 2009). In the context of long-term climate and environmental variability, circulation type classifications (CTCs) may be applied in two different ways. First, for the classification of observed (or reanalysed) or reconstructed atmospheric circulation data and the investigation of the linkage between large-scale circulation – reflected in frequency and within-type changes of circulation types – and local climatic or environmental variables (Beck *et al.*, 2007). Second for the downscaling of local climatic or environmental variables from – most reliably reproduced – large-scale circulation data from GCMs (Enke and Spekat, 1997).

Irrespective of the specific synoptic climatological application (including downscaling) for which CTCs are used it can be stated in general that the utility and reliability of all approaches depend on the ability of the used circulation classifications to resolve the variations of the respective target variables. This essential capability of CTCs to clearly separate between different magnitudes of a target variable has been qualitatively characterized and quantitatively assessed by several authors. However, different authors use partly different terms for the same feature; i.e., 'synoptic-climatological applicability' (Huth, 2010), 'discriminative power' (Beck and Philipp,

* Correspondence to: C. Beck, Institute of Geography, University of Augsburg, Universitätsstr. 10, D-86135 Augsburg, Germany. E-mail: christoph.beck@geo.uni-augsburg.de

2010), 'resolution' or 'predictive capability' (Schiemann and Frei, 2010) or – referring to the cause of ambiguities in the relationships between classifications and target variables – 'within-type variability' or 'within-group variability' (Yarnal, 1993; Brinkmann, 1999; Beck *et al.*, 2007). In the end all the mentioned studies investigate the ability of CTCs to provide reliable approaches for quantifying the linkage between atmospheric circulation and environmental variables for synoptic climatological analyses. Therefore, the term 'synoptic skill' will be used in this context for the remainder of this paper.

Although studies on the estimation of the synoptic skill of CTCs and respective comparisons have been sporadically conducted in the past (Buishand and Brandsma, 1997; Philipp, 2009), respective systematic evaluation and comparison studies utilizing a comprehensive data base of CTCs have been performed not until the implementation of the EU funded COST733 Action 'Harmonisation and Applications of Weather Types Classifications for European Regions' (Huth *et al.*, 2010). Within the COST733 Action evaluation and comparison studies have been performed using the database of CTCs for Europe developed during the action (Philipp *et al.*, 2010) utilizing different evaluation metrics and focusing on varying target variables (Beck and Philipp, 2010; Casado *et al.*, 2010; Demuzere *et al.*, 2011; Fleig *et al.*, 2010; Huth, 2010; Schiemann and Frei, 2010; Tveito, 2010).

Main results from these analyses show distinct variations in synoptic skill of CTCs between different regions (e.g. better skill in more oceanic compared to more continental regions) and seasons (mostly better skill in winter than in summer), between different target variables (e.g. better skill for temperature than for precipitation) and between different classification methods. Whereas the first three points merely reflect spatiotemporal variations in circulation–climate relationships in northern hemisphere mid-latitudes the last point refers to the effect of different methodological approaches on synoptic skill. Results achieved so far provide no clear evidence for any basic or specific classification approach that in general appears to be particularly suitable (Beck and Philipp, 2010). Moreover, it turned out that beside the relevance of inherent properties of the methodological approach the synoptic skill of CTCs depends substantially on several basic settings – or boundary conditions – of the classification that are independent from the classification method. These settings are the number of classes, the type and number of variables which are used for classification, the length of the classified period (1 or more days) and the temporal subset (classifications applied to 3-month seasons or to the whole year).

This contribution focusses on the effect of the size of the spatial domain to which circulation classifications are applied. The relevance of this additional boundary condition of CTCs for synoptic skill already arose from respective initial analyses within the COST733 Action, but has so far been only casually addressed in the literature. Yarnal (1993) concluded that CTCs applied to domains of different size provide complementary relevant information related to variations in glacier mass balance in British Columbia. Saunders and Byrne (1999) showed that a classification based downscaling approach for generating precipitation in Alberta yields better results when applied to a larger domain (including large parts of the northern Pacific) than to a smaller – regional scale – domain. Analysing relationships between precipitation in south-west Western Australia and large-scale circulation Hope *et al.* (2006) on the other hand found that a CTC derived on the basis of a smaller domain (excluding the eastern parts of Australia) is more favourable. Wetterhall *et al.* (2007) evaluating different approaches for downscaling precipitation in Sweden detected seasonal differences concerning the ideal size of the domain. D'Onofrio *et al.* (2010) documented distinct effects of the domain size on the results of a weather type based downscaling of precipitation in Argentina, with best results achieved using a domain of roughly 13.5° by 13.5°. Yarnal (2006) discussed the effect of the domain size – among other factors – for the synoptic skill of circulation classifications and Brinkmann (2002) additionally stressed the importance of the location of the domain relative to the target region. However Enke *et al.* (2005) although revealing differences among western and eastern regions in Germany concerning the ideal grid domain for classification based downscaling of temperature and precipitation conclude that domain characteristics have only minor relevance. For varying statistical downscaling approaches (not classification based ones) Benestad (2001) and Spak *et al.* (2007) found that the domain size distinctly affects downscaled results for temperature in Norway and in the eastern United States, respectively, whereas Huth (2002) reached no comparable conclusion with respect to Central European temperature.

Given these sparse and partly contradictory findings concerning the relevance of the spatial characteristics of the domains used for synoptic climatological analyses and especially CTCs Sheridan and Lee (2010) explicitly stated the need for more systematic comparative analyses.

Against this background the main aim of the present contribution is to investigate in a systematic way how far the synoptic skill – concerning daily temperature and precipitation – of different automated CTCs depends on the size of the spatial domain underlying the classifications and to determine the respective seasonal and regional variations. To this end different automated CTCs are applied to domains of varying sizes distributed within a greater North Atlantic European region. The synoptic skill of the resulting classifications is evaluated using several objective measures and compared with respect to varying domain sizes.

The paper is organized as follows: In Section 2 the data sets used for determining and evaluating CTCs are introduced. The methodological approach for analysing circulation type–surface climate relationships in dependence of the domain size is explicated in Section 3. Relevant results are presented in Section 4 and briefly discussed in Section 5 where also the main conclusions are derived.

2. Data

Two different data sets have been used for the classification of circulation types and the evaluation of classifications with regard to surface climate variables respectively.

2.1. Reanalysis data for performing CTCs

All circulation classifications in this study have been generated using 2.5° by 2.5° gridded daily 12 UTC sea level pressure (SLP) data that are available for the period from 1948 to present from the NCEP/NCAR reanalysis 1 data archive (Kalnay *et al.*, 1996). For generating varying circulation classifications different spatial subsamples of the globally available SLP data have been selected within the greater North Atlantic European region (62.5°W to 82.5°E; 12.5°S to 87.5°N) for the period 1950 to 2010.

The relatively coarse spatial resolution on the one hand implicates that potentially relevant convective-scale features are not considered in the classifications. On the other hand the longer availability of the NCEP/NCAR reanalysis 1 data – compared to most other reanalysis data sets – increases sample sizes for estimating circulation type characteristics. The 2.5° by 2.5° resolution is furthermore in accordance with the coarse resolution of reconstructed atmospheric circulation data to which CTCs may be applied for the analysis of long-term climate variability (Philipp *et al.*, 2007).

2.2. Gridded observational data for the evaluation of CTCs

For quantifying European surface climate characteristics related to atmospheric circulation types daily mean temperatures (TM), daily minimum temperatures (TN), daily maximum temperatures (TX) and daily precipitation sums (RR) have been extracted from the 0.25° by 0.25° gridded regular grid version of the E-OBS v4.0 dataset (Haylock *et al.*, 2008). The temporal coverage of the E-OBS data from 1950 to 2010 implicates the period considered in this study. The reliability of the E-OBS gridded data set has been recently analysed and discussed in detail by Hofstra *et al.* (2009) who found evidence for numerous inhomogeneities in the gridded temperature and as well as precipitation data. Moreover, for several regions the E-OBS data show distinct biases of varying sign and absolute value compared to available high-resolution regional data sets.

However, these documented deficiencies of the E-OBS data are only of minor relevance for the analyses presented in this paper as neither estimations of long-term changes in surface climate nor any comparisons with other data sets are performed. The comparisons of different variants of circulation classifications are not expected to be seriously biased by the restrictions of the E-OBS data.

3. Methodology

The methodological approach underlying this study comprises (1) the definition of spatial domains of varying size for the classification of circulation types to which (2) different classification methods have been applied and (3) the evaluation of the resulting classifications with respect to their relevance for surface climate characteristics using several objective evaluation criteria.

3.1. Definition of spatial domains of varying size

A visualization of the location and the extent of the spatial domains defined for the analyses are provided in Figure 1. The spatial domains have been centred over totally 44 locations (each of them marked by a dot in Figure 1) spread over the region 20°W to 40°E and 35°N to 65°N and separated by steps of 10° in zonal and steps of 5° in meridional direction, respectively. Due to missing surface climate data in the respective regions no analyses have been performed for the five southernmost locations on 20°W. First, for each of the 44 selected locations a central inner domain has been defined. These core domains are the smallest domains to which varying circulation classifications have been applied. Furthermore, all circulation classifications – no matter to which spatial domain centred around the respective location they have been applied – are evaluated on the basis of surface climate data from these core (or target) domains. Starting from the smallest initial domains for each location seven additional spatial domains of increasing size have been defined.

Thereby in order to provide spatial domains of approximately comparable size (in terms of horizontal distances in kilometres) at all locations the definition of the longitudinal dimension of the domains (expressed as multiples of the resolution of the SLP data) varies with the latitude of the target locations. Accordingly the longitudinal extent of the core domains varies between 15° at 65°N and 7.5° at 35°N and that of the largest domain 8 varies between 85° at 65°N and 40° at 35°N (Figure 1). The approximate dimensions of the eight domain sizes (expressed in kilometres) are 700, 1300, 1800, 2200, 2500, 2900, 3300, 3700 in the zonal direction (at the central latitude of the respective domain) and 1100, 1600, 2200, 2800, 3300, 3900, 4400, 5000 in the meridional direction. The number of grid points varies between 665 for the largest variant of domain 8 and 20 for the smallest variant of domain 1. These domain definitions cover quite well the spectrum of domain sizes that – according to a comprehensive inventory of CTCs in Europe (Huth *et al.*, 2008) – are most commonly used for the application of CTCs. According to this 50% of the classifications have continental scale (more than 30° latitude by 40° longitude), 20% have subcontinental scale (20° latitude by 20° longitude)

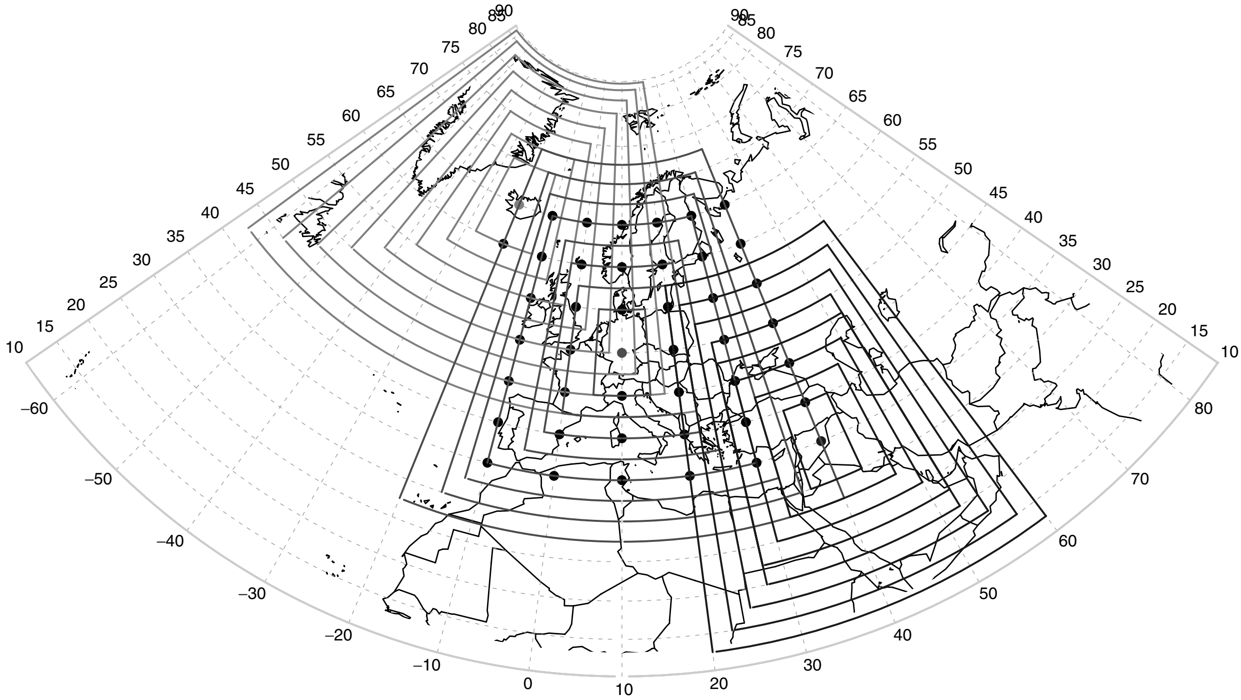

Figure 1. Location of the 44 grid points and respective spatial domains of varying size for the application and evaluation of circulation type classifications. As an example, for the locations 20°W/65°N, 10°E/50°N and 40°E/35°N the respective sequences of increasing domains – in each case ranging from the smallest domain 1 to the largest domain 8 – are indicated by the rectangles.

Table 1. Circulation type classifications used in this study.

Classification method	Abbreviation	Method group (according to Philipp *et al.* 2010)	Number of circulation types	Main reference
Grosswettertype classification	GWT	Threshold based method	10, 18, 26	Beck *et al.* (2007)
Classification of leading s-mode principal components	KRZ	PCA-based method	9, 18, 27	Kruizinga (1979)
T-mode principal component analysis (PCA)	PTT	PCA-based method	10, 18, 26	Huth (1993)
S-mode principal components extreme scores	PXE	PCA-based method	10, 18, 26	Esteban *et al.* (2005)
correlation based Lund classification	LND	method based on leader algorithm	10, 18, 26	Lund (1963)
k-Means cluster analysis	CKM	Optimization method	10, 18, 26	Enke & Spekat (1997)
k-Means cluster analysis	DKM	Optimization method	10, 18, 26	Philipp *et al.* (2010)

and another 20% cover one single country. Furthermore, our domain sizes correspond with those that were used within the framework of the COST733 Action 'Harmonisation and Applications of Weather Type Classifications for European regions' (Philipp *et al.*, 2010).

Exemplary for the locations 20°W/65°N, 10°E/50°N and 40°E/35°N the respective sequences of eight spatial domains of increasing size (in each case from the smallest domain 1 to the largest domain 8) are highlighted in Figure 1. In the end for each of the 44 central locations eight spatial domains of varying size – totally 352 domains – are available for the application of varying CTCs, which are evaluated taking into account surface climate data from the respective smallest spatial domain.

3.2. Application of different automated CTCs

To investigate the influence of the size of the spatial domain used for circulation classification on the linkage between CTCs and surface climate seven classification methods have been applied to daily gridded SLP data for each of the spatial domains described above. As indicated in Table 1 the selected classification methods are representative for different basic methodological approaches for the classification of atmospheric circulation types. In detail these basic groups of classification methods

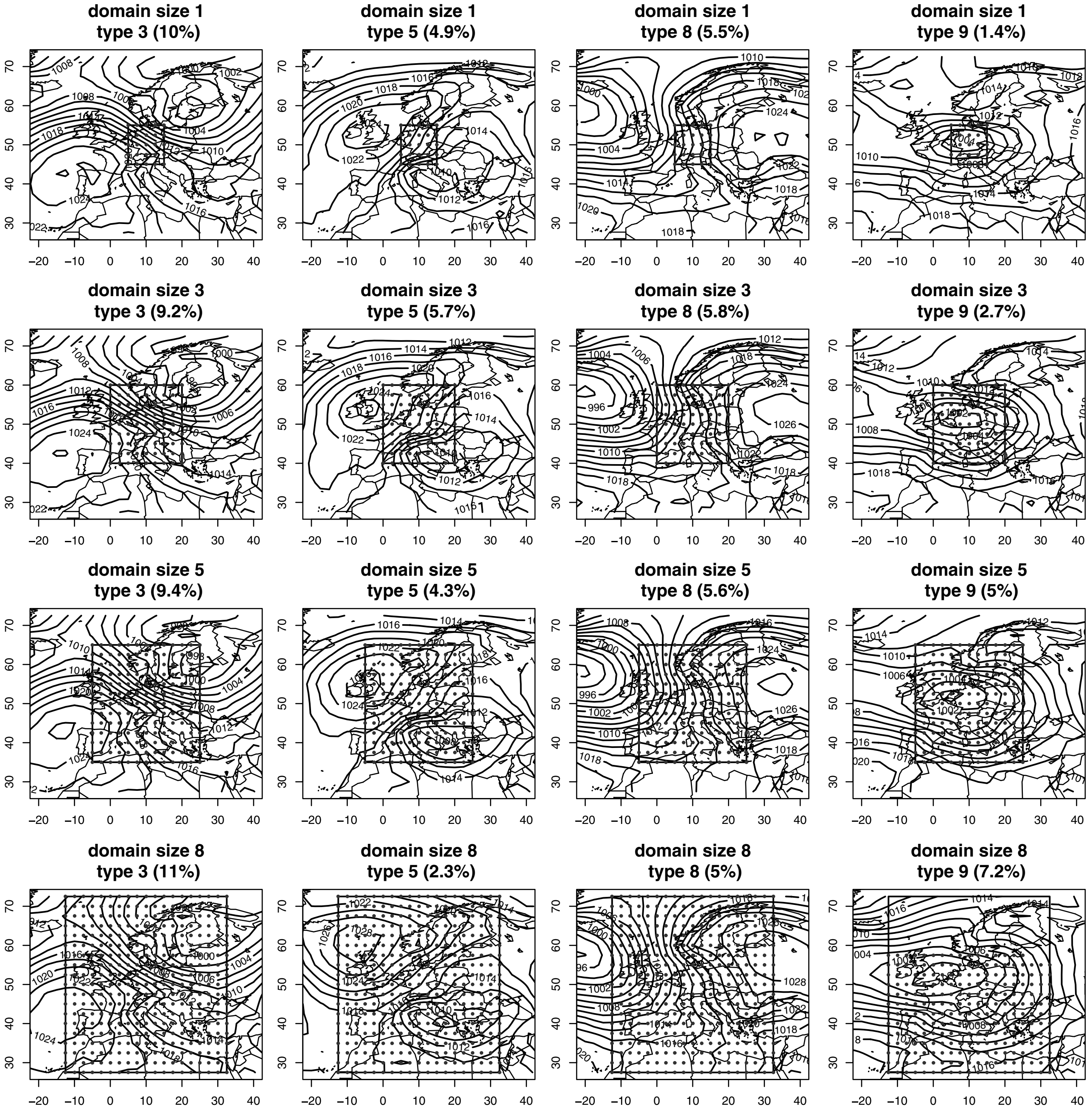

Figure 2. SLP (hPa) composites for winter (DJF) 1950–2010 for selected circulation types (types 3, 5, 8, 9; in columns from left to right) derived from the application of the GWT classification (comprising 10 types) to spatial domains of varying size (domain sizes 1, 3, 5, 8; from the top row to the bottom row). The size of the domain and the grid points used for each classification are indicated by grey rectangles and points, respectively. Relative type frequencies (in %) are given in brackets.

– according to Philipp *et al.* (2010) – are *Threshold-based methods* that use threshold values or explicitly defined decision rules applied to varying variables (e.g. wind direction, SLP) to delineate different circulation types. *PCA (principal component analysis) based methods* using the similarity between individual cases (atmospheric fields) and principal components (retrieved in t- or s-mode) for the categorization of circulation types. *Methods based on the leader algorithm* which first determine key patterns to which individual cases are iteratively assigned using varying similarity measures and finally *optimization methods* that arrange cases (daily fields) within a number of classes trying to achieve the minimization of the variance within classes. A more detailed description of concepts and characteristics of these basic method groups of circulation classification can be found in Philipp *et al.* (2010).

The individual classification methods used within the analyses presented here can be briefly characterized as follows:

The threshold-based Grosswettertypes or prototype classification (GWT) arranges cases into types according

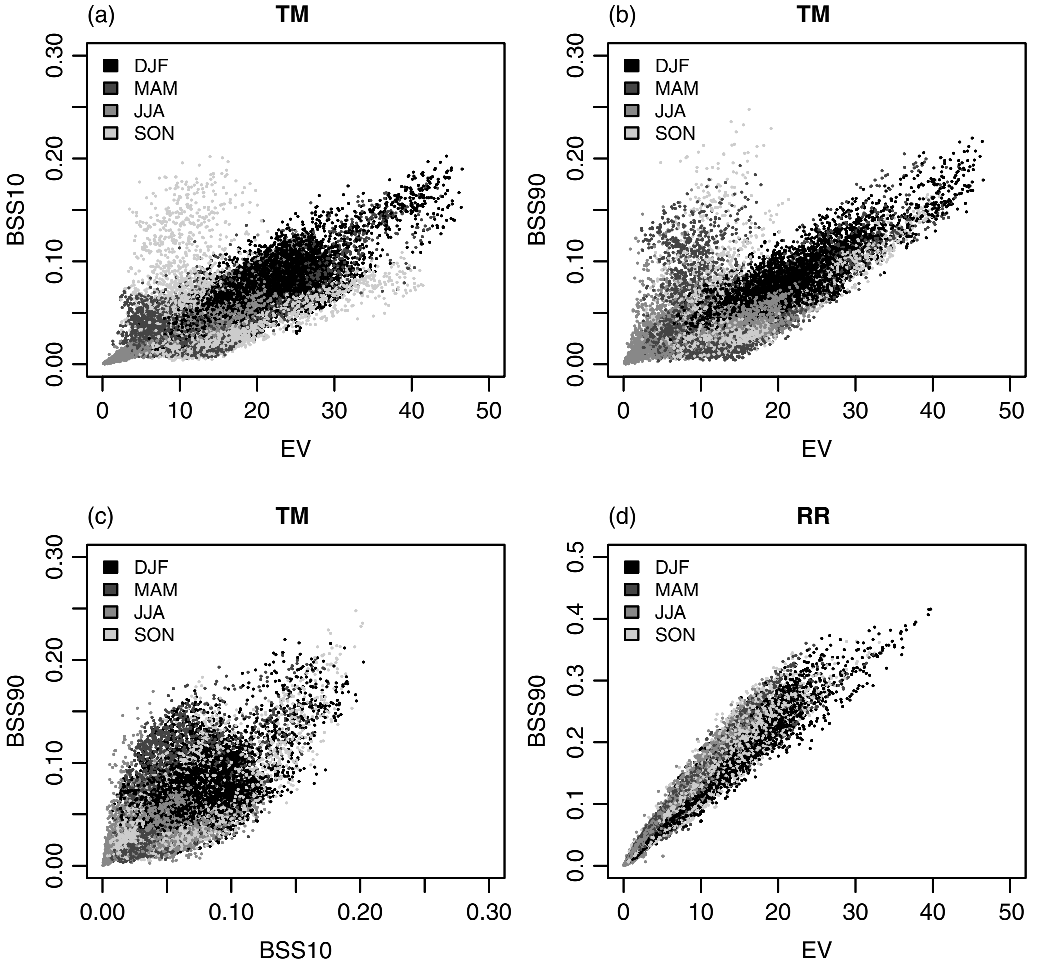

Figure 3. Scatterplots illustrating the relationships between evaluation metrics EV (%), BSS10 and BSS90 determined for daily mean temperature (TM) and daily precipitation sum (RR). Gray levels refer to seasonal subsets (winter – DJF, spring – MAM, summer – JJA and autumn – SON).

to varying degrees of zonality, meridionality and vorticity determined as spatial correlation coefficients between daily SLP fields and prototypical flow patterns (Beck *et al.*, 2007).

Three methods are from the group of classifications that are based on PCA. The Kruizinga empirical orthogonal function classification (KRZ or P27 scheme) utilizes the three leading s-mode principal components to assign cases to varying numbers of circulation types on the basis of characteristic combinations of the respective component scores (Kruizinga, 1979). S-mode principal component scores are also the basis for the PCA extreme scores (PXE) classification that defines initial circulation types considering cases (daily SLP fields) featuring distinctly high (extreme) scores for specific principal components and subsequently assigns remaining cases to these types according their respective minimum Euclidean distance. The classification by PCA in t-mode (*PTT*) assigns cases to classes according to the respective maximum loadings of obliquely rotated t-mode principal components (Huth, 1993).

The Lund classification (LND) represents the classical approach based on the leader algorithm utilizing the spatial Pearson correlation coefficients between SLP fields as similarity measure (Lund, 1963). Finally two classifications from the group of optimization methods have been applied that are based on the non-hierarchical k-means clustering algorithm (Hartigan, 1975). Both methods (CKM, DKM) first determine an initial starting partition based on the most dissimilar cases included in the data set (Enke and Spekat, 1997). Additionally *CKM* retains only those classes (clusters) that reach relative frequencies not less than 5% in order to derive only representative types. For more details of the individual classification methods see Philipp *et al.*, 2010 and references therein and references indicated in Table 1.

Each of the classification methods has been applied to all 352 spatial domains. Thereby each classification has been run for three different numbers of types (approximately 10, 18, 26 types representing low, medium and high numbers of types commonly used in classification schemes) resulting in a total number of 7392 classifications. Thus the influence of varying sizes of the spatial

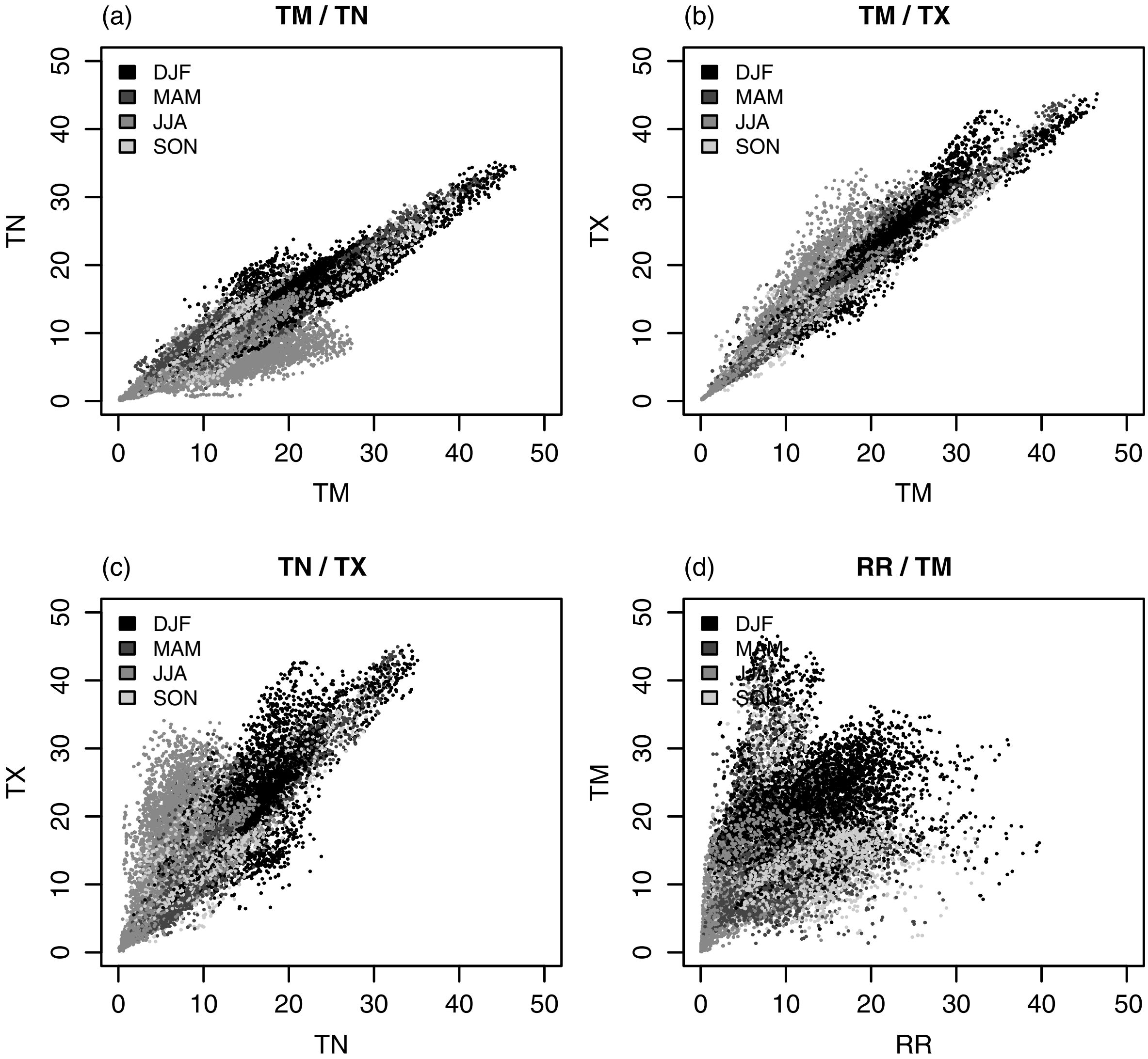

Figure 4. Scatterplots showing the relationships between evaluation results (explained variance EV in %) achieved for the surface climatic target variables daily mean temperature (TN), daily minimum temperature (TN), daily maximum temperature (TX) and daily precipitation sum (RR). Gray levels refer to seasonal subsets (winter – DJF, spring – MAM, summer – JJA and autumn – SON).

domains used for circulation classification can be analysed on a wide data basis of different CTCs and respective variants. This on the one hand reduces the risk that findings derived on the basis of individual classifications are overestimated and are misinterpreted as universally valid results. And on the other hand using this broad data base additionally offers the possibility to depict differences among varying classification approaches concerning the relevance of the domain size.

SLP composites for winter (DJF) for selected circulation types resulting from the application of the GWT classification to spatial domains of varying size centred on 50°N/10°E are shown in Figure 2. It is apparent that even classifications applied to the smallest domain size 1 provide physically meaningful circulation patterns, not only considering the classified domain, but as well with respect to the large-scale configurations beyond the incorporated domain. Although circulation types derived for varying domain sizes exhibit some distinct similarities in large-scale patterns, they feature specific characteristics related to the differing spatial scales. Finally differences between classifications for varying domain sizes are visible from varying occurrence frequencies.

3.3. Evaluating the relationship between CTCs and surface climate variables

The performance of classifications concerning their association to surface climate conditions (or the 'synoptic skill' of classifications) depending on domain size can be estimated using different quantitative measures.

From the varying approaches for the evaluation of the synoptic skill of classifications that were utilized within the framework of the COST733 Action 'Harmonisation and Applications of Weather Types Classifications for European Regions' (Beck and Philipp, 2010; Huth, 2010; Schiemann and Frei, 2010; Tveito, 2010) the following ones have been applied in this study.

Table 2. Pearson correlation coefficients between the three evaluation metrics EV, BSS10 and BSS90, determined separately for the surface climate variables TM, TN, TX and RR and for the four 3-month seasons winter (DJF), spring (MAM), summer (JJA) and autumn (SON).[a]

	TM				TN				TX				RR			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
EV/BSS10	**0.81**	0.74	***0.94***	0.14	**0.84**	0.68	***0.91***	0.08	0.76	0.74	***0.94***	0.33	***0.97***	***0.92***	***0.96***	***0.95***
EV/BSS90	**0.88**	0.51	0.60	0.52	**0.87**	0.41	0.35	0.26	**0.83**	0.50	0.69	0.51	***0.97***	***0.97***	***0.97***	***0.96***
BSS10/BSS90	0.55	0.68	0.51	0.78	0.57	0.69	0.34	**0.86**	0.40	0.62	0.54	**0.81**	**0.90**	**0.82**	**0.90**	**0.87**

[a]All correlation coefficients shown in the table are statistically significant (for $\alpha = 0.05$) according to a t-test. Correlation coefficients above 0.8 are in bold, correlation coefficients above 0.90 are in bold and italics.

The explained variation (EV) determines the relation between the variance among classes (circulation types) and the total variance of the variable under consideration in order to – like several other measures (Beck and Philipp, 2010) – quantify the discriminatory power of a classification:

$$\mathrm{EV} = \frac{\sum_{k=1}^{K} N_k \left(\overline{a}_k - \overline{a}\right)^2}{\sum_{i=1}^{N} \left(a_i - \overline{a}\right)^2}, \qquad (1)$$

N is the number of cases, K is the number of classes (circulation types), a_i is the value of the target variable for case i and a and a_k are the overall and type-specific mean values, respectively. The higher the values of EV, the better is the discriminatory power of a classification for the analysed variable.

A variant of the Brier skill score (BSS; Wilks, 2006) has been introduced by Schiemann and Frei (2010) to quantify the ability of CTCs to represent exceedances of certain thresholds of surface climate variables. The BSS takes values between 0 and 1, with higher values indicating better synoptic skill and is defined as

$$\mathrm{BSS} = \frac{\frac{1}{N}\sum_{k=1}^{K} N_k \left(f_k - \overline{o}\right)^2}{\overline{o}\left(1 - \overline{o}\right)}. \qquad (2)$$

Here N is the number of cases, K is the number of classes (circulation types). f_k is the observed or conditional frequency of events (e.g. exceedances or shortfalls of a certain threshold) in class k and o are the unconditional frequency of events. In this contribution the BSS has been determined utilizing empirical quantiles [90% (BSS90) for precipitation data, 90% and 10% (BSS10) for temperature data] – estimated over the whole period – as thresholds for the definition of events, thereby explicitly considering extremes in the determination of each classification's synoptic skill.

One adverse characteristic of the evaluation metrics is their sensitivity to the number of classes (Beck and Philipp, 2010; Huth, 2010; Schiemann and Frei, 2010). However, as fixed numbers of circulation types are used this does not affect the analyses presented here. Each evaluation metric has been determined for each grid point of the respective target domain. To account for seasonal variations in the circulation–climate relationships all evaluation criteria have been determined separately for the four 3-month seasons winter (DJF – December, January, February), spring (MAM – March, April, May), summer (JJA – June, July, August) and autumn (SON – September, October, November), thereby considering data from the whole available period 1950 to 2010.

4. Results

It is at first investigated to what extent the varying measures for the synoptic skill of circulation classifications yield comparable results. Secondly it is determined in how far estimates of the synoptic skill differ among surface climatic target variables and finally the dependence of synoptic skill on varying domain sizes is revealed.

4.1. Comparison of evaluation metrics

The pairwise relationships between EV, BSS10 and BSS90 for the two surface climate variables TM and RR are illustrated by the scatterplots in Figure 3. Additionally Pearson correlation coefficients between the three different evaluation metrics estimated for all surface climate variables (TM, TN, TX, RR) are given – separately for the four seasons – in Table 2.

For RR, a distinct similarity between the two skill metrics EV and BSS90 during all seasons is visible from Figure 3(d) and from the respective seasonal correlation coefficients – which all reach values above 0.95 – in Table 2.

With respect to the temperature variables TM, TN and TX it can be stated that all three variables feature essentially the same characteristics concerning the relations between evaluation metrics. From the respective scatterplots given for TM in Figure 3(a)–(c) and the corresponding correlation coefficients from Table 2 the most pronounced attributes can be deduced as high correlations between EV and BSS10 in summer, winter and (to a lesser extent) in spring. Between EV and BSS90 the most distinct linkage shows up for winter while a remarkable dependency between BSS10 and BSS90 exists primarily in autumn.

In summary – whereas with regard to precipitation one evaluation metric (either EV or BSS90) seems to be

Table 3. Pearson correlation coefficients between evaluation metrics for the surface climate variables TM, TN, TX and RR determined separately for the three evaluation metrics EV, BSS10 and BSS90 and for the four 3-month seasons winter (DJF), spring (MAM), summer (JJA) and autumn (SON).[a]

	EV				BSS10				BSS90			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
TM/TN	***0.92***	***0.93***	0.72	***0.94***	***0.92***	**0.90**	0.79	***0.98***	***0.93***	***0.94***	0.72	***0.93***
TM/TX	***0.96***	***0.97***	***0.95***	***0.94***	***0.94***	***0.94***	***0.94***	***0.96***	***0.91***	***0.97***	***0.92***	***0.97***
TN/TX	**0.83**	**0.85**	0.64	**0.82**	**0.82**	0.74	0.70	**0.91**	0.78	**0.87**	0.52	**0.86**
RR/TM	0.34	0.27	0.75	0.28	0.25	0.13	0.72	−0.19	0.39	−0.14	0.30	−0.09
RR/TN	0.42	0.22	0.49	0.32	0.35	−0.14	0.60	−0.24	0.38	−0.20	−0.04	−0.18
RR/TX	0.35	0.42	0.80	0.34	0.17	0.29	0.71	−0.11	0.52	−0.07	0.51	0.01

[a]All correlation coefficients shown in the table are statistically significant (for $\alpha = 0.05$) according to a t-test. Correlation coefficients above 0.8 are in bold, correlation coefficients above 0.90 are in bold and italics.

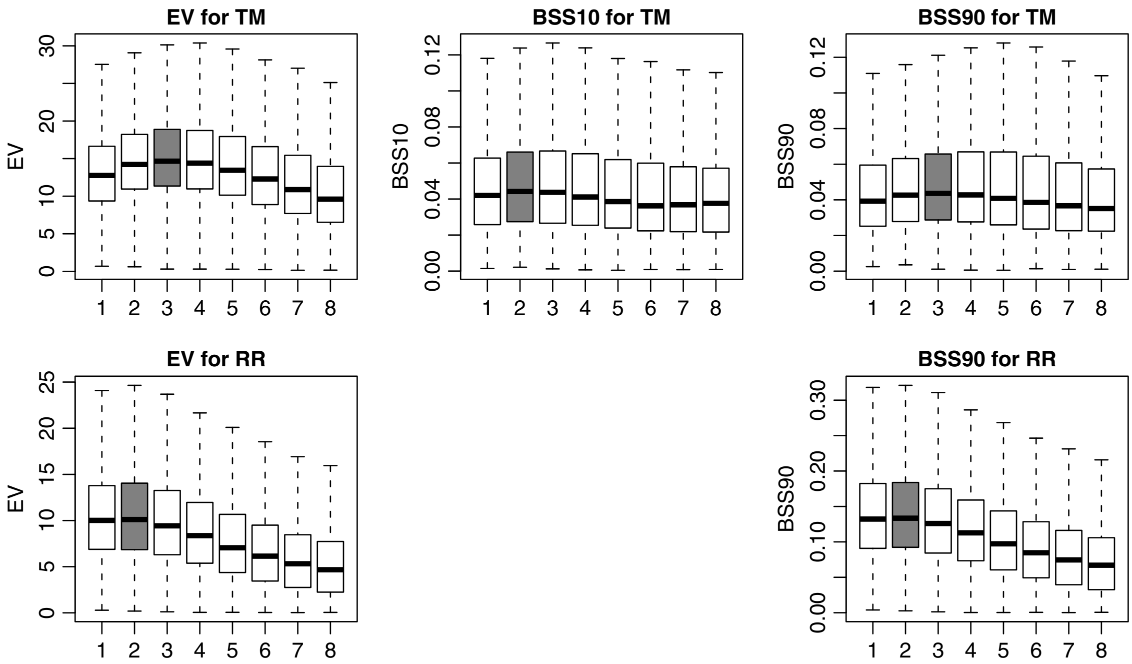

Figure 5. Boxplots of evaluation metrics for circulation type classifications grouped according to the size of the spatial domain used for classification, with domain sizes increasing from 1 to 8 along the x axis. The evaluation metrics explained variation (EV in %) and Brier skill score (BSS10, BSS90) have been estimated over all seasons, all target locations and all classification methods for daily mean temperature (TM) and daily precipitation sum (RR). Gray shaded boxes refer to the respective domain size featuring the highest median value. Upper/lower whiskers indicate the 1.5 interquartile range from the upper/lower quartile.

sufficient to quantify main characteristics of the synoptic skill of CTCs – for temperature variables obviously different metrics are necessary in order to adequately capture the synoptic skill of CTCs focusing on different characteristics (mean and extremes) of the frequency distribution.

4.2. Comparison of synoptic skill derived for different target variables

From Figure 4 and Table 3 a close relationship between synoptic skill estimated for TM and the other two temperature variables TN and TX is apparent for all evaluation metrics and during all seasons with the only exception of TM and TN in summer exhibiting distinctly lower correlations. The relationship between TN and TX on the other hand appears generally less pronounced with the least agreement in summer. Therefore, the skill for TM may be utilized as a proxy for the temperature related skill of a CTC in general. Turning to the relations between RR and temperature variables Figure 4 and Table 3 reveal remarkable relations primarily on the basis of EV and during summer. For all other comparisons including RR distinctly lower correlations compared to those derived among temperature variables are apparent.

4.3. Dependence of synoptic skill on domain size

Taking into account the findings from the preceding paragraphs the presentation of results will focus on TM and RR as target variables and – at least for RR – on EV as evaluation metric.

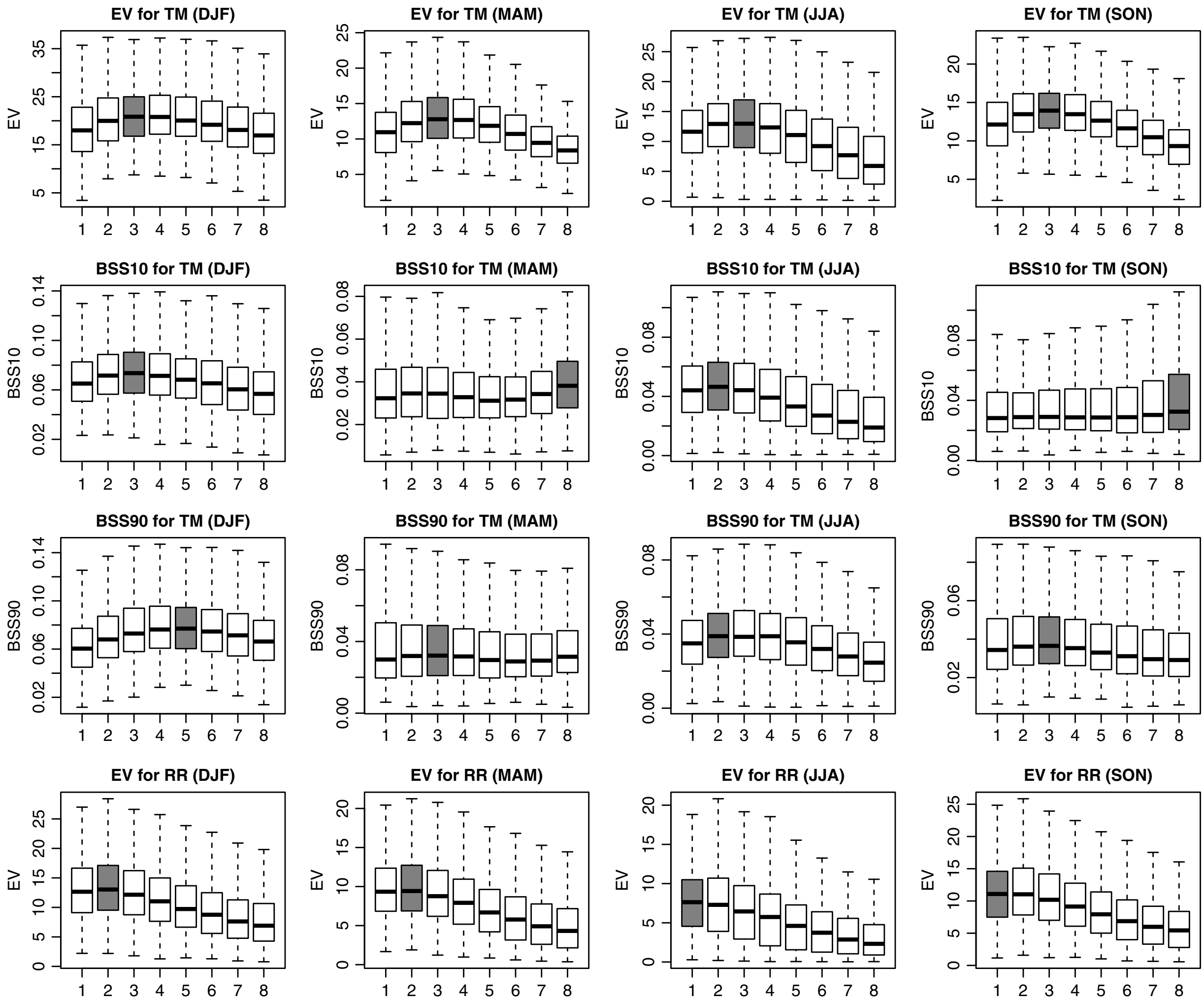

Figure 6. Boxplots of evaluation metrics for circulation type classifications grouped according to the size of the spatial domain used for classification, with domain sizes increasing from 1 to 8 along the x axis. The evaluation metrics explained variation (EV in %) and Brier skill score (BSS10, BSS90) have been estimated separately for the four 3-month seasons winter (DJF), spring (MAM), summer (JJA) and autumn (SON) over all target locations and all classification methods for daily mean temperature (TM) and daily precipitation sum (RR). Gray shaded boxes refer to the respective domain size featuring the highest median value. Upper/lower whiskers indicate the 1.5 interquartile range from the upper/lower quartile.

Figure 5 provides an initial overview of the effect of the size of the spatial domain on synoptic skill. Integrating over all seasons, target locations and classification methods domains sizes 3 and 2 appear to be overall best suited for applying CTCs aiming at temperature (TM and – not shown – TN and TX) and precipitation (RR), respectively. The statistical significance of differences between varying domain sizes has been estimated by applying the nonparametric Kruskal–Wallis (K–W) test (Hollander and Wolfe, 1999) to samples categorized according to domain sizes 1 to 8. For all target variables and all evaluation metrics the null hypothesis of equality of the location parameter among all categories can be rejected (for $\alpha = 0.05$). An additional Mann–Whitney (M–W) test (Hollander and Wolfe, 1999) has been performed to derive the statistical significance of the differences in location between the optimal domain size and the combined remaining domains. Statistical significance (for $\alpha = 0.05$) is reached in all cases.

4.3.1. Seasonal variations

The general finding that the synoptic skill of CTCs for climatic target variables reaches maxima for domain sizes 3 and smaller – with optimal domain sizes tending to be smaller for precipitation than for temperature variables – is largely confirmed by the seasonal boxplots of EV (for TM and RR) and BSS10 and BSS90 (only for TM) conditioned by domain size (Figure 6). However, some interesting additional variations in the relationship between domain size and synoptic skill become apparent from Figure 6.

First, for both RR and – less distinct – TM the size of the spatial domain for which highest synoptic skill is achieved tends to be smaller in summer compared to winter. Concerning RR the smallest domain 1 reaches highest EV values in summer (and autumn) whereas superior synoptic skill is related to domain size 2 in winter and also in spring. For TM domain size 3 (2)

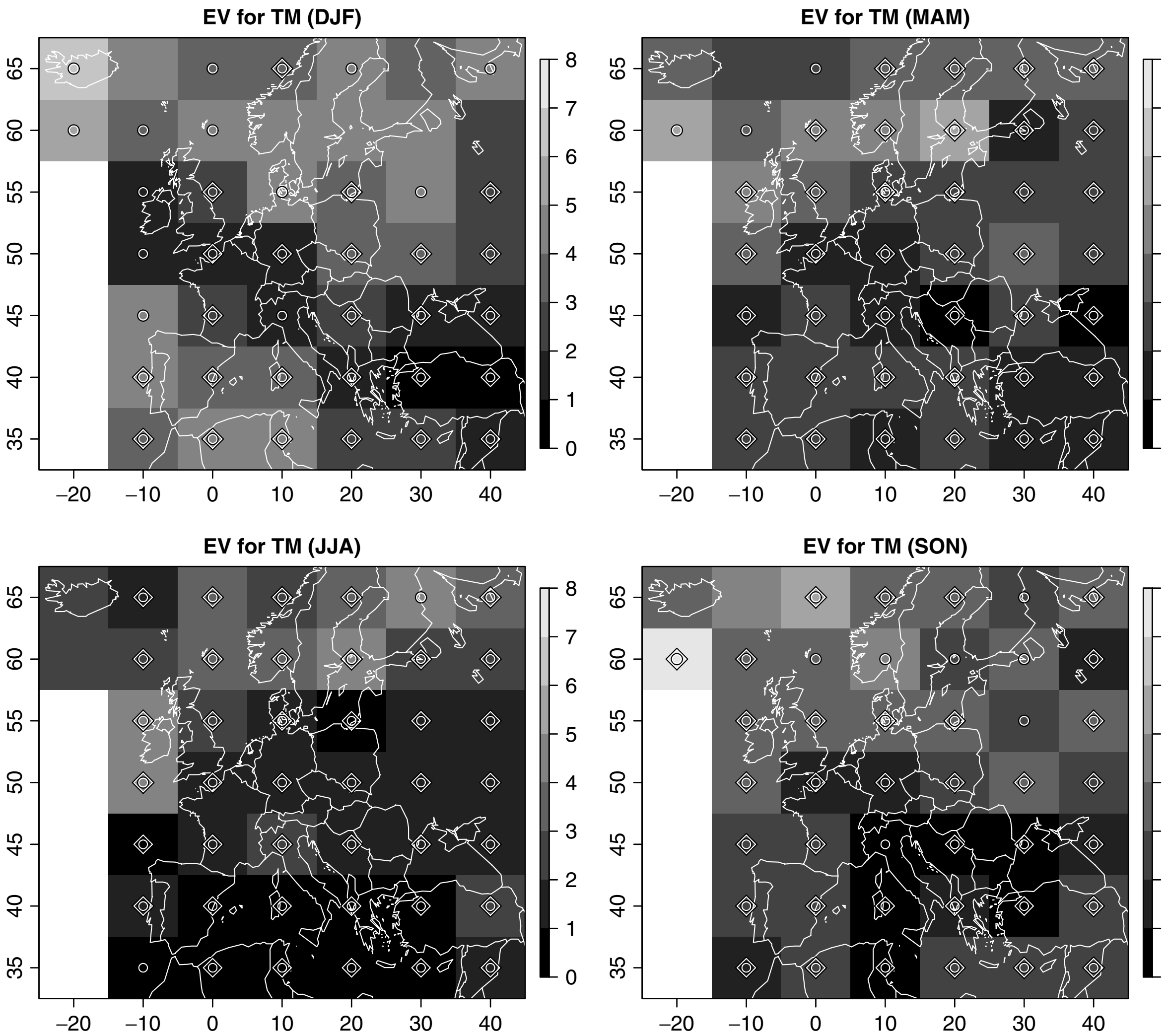

Figure 7. Spatial variation of domain sizes for which highest synoptic skill (EV) of circulation type classifications for the target variable TM is achieved in the four 3-month seasons winter (DJF), spring (MAM), summer (JJA) and autumn (SON). Optimal domain sizes – from smallest domain size 1 to largest domain size 8 – for each of the 44 target domains are indicated by gray shades. The optimal domain size for each location has been determined on the basis of the seasonal synoptic skill values averaged over all grid points within the respective domain. Circles at a location denote statistical significance (for $\alpha = 0.05$) of the influence of varying domain sizes on synoptic skill in the respective season (according to a Kruskal–Wallis test). Diamonds indicate additionally a statistically significant (for $\alpha = 0.05$) difference between location parameters estimated for the optimal domain size and estimated over all other domain sizes respectively (according to a two sample Mann–Whitney test).

is in most cases best in winter (summer). One distinct exception to this rule turns out for BSS90 in winter with larger domains of size 5 being related to best skill. During the transition seasons spring and autumn domain size 3 represents the optimal domain size for EV and BSS90. For BSS10 on the other hand best synoptic skill is achieved for the largest domain size 8. These rough major patterns of seasonal variations in optimal domain size are also reflected by respective boxplots for TN and TX (not shown) and boxplots of BSS90 for RR (not shown). K–W tests yield statistical significance ($\alpha = 0.05$) of differences in the location parameter between domain sizes in all cases. Additional M–W tests also reveal statistical significance (for $\alpha = 0.05$) of the differences in location between the optimal domain size and the combination of all other domains in all but one (BSS90 for TM in spring) case.

4.3.2. Spatial variations

Figures 7 to 10 indicate for varying target locations the domain sizes for which best synoptic skill is achieved for different climatic variables. To highlight respective variations among target domains rather than variations within target domains spatial averages of evaluation metrics have been used in this context.

The basic pattern of spatial variations in optimal domain size for TM is a decrease in domain size roughly from the north-west to the south-east (Figure 7). Distinct modifications of this spatial configuration show up especially in winter with relatively large optimal domain sizes in the south-west and in summer when relatively small optimal domain sizes are present in the north-west. For the majority of locations a K–W test applied to EV values categorized according to domain size yielded statistical significance (for $\alpha = 0.05$) of the

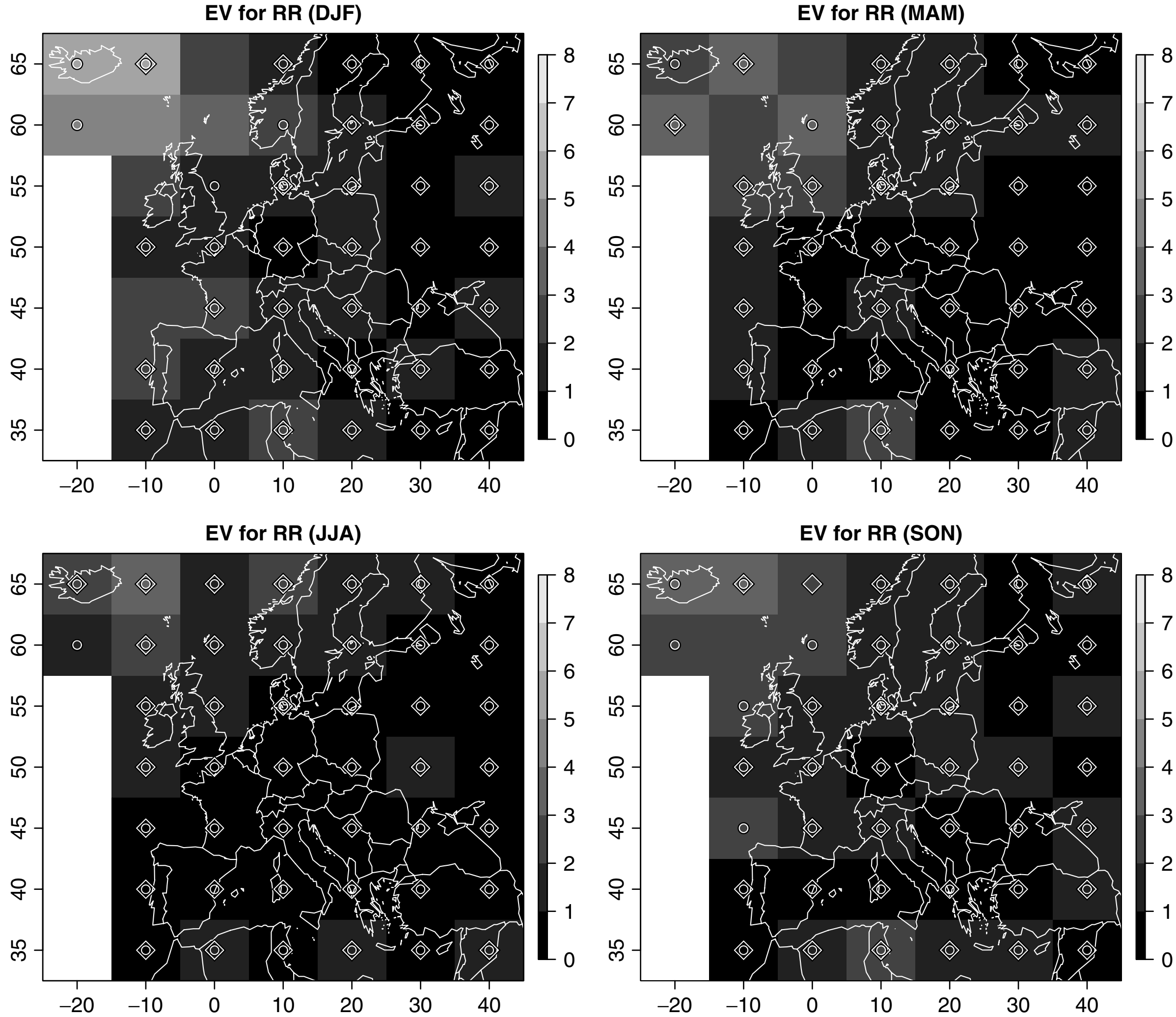

Figure 8. Spatial variation of domain sizes for which highest synoptic skill (EV) of circulation type classifications for the target variable RR is achieved in the four three-month seasons winter (DJF), spring (MAM), summer (JJA) and autumn (SON). Optimal domain sizes – from smallest domain size 1 to largest domain size 8 – for each of the 44 target domains are indicated by gray shades. The optimal domain size for each location has been determined on the basis of the seasonal synoptic skill values averaged over all grid points within the respective domain. Circles at a location denote statistical significance (for alpha = 0.05) of the influence of varying domain sizes on synoptic skill in the respective season (according to a Kruskal-Wallis test). Diamonds indicate additionally a statistically significant (for alpha = 0.05) difference between location parameters estimated for the optimal domain size and estimated over all other domain sizes respectively (according to a two sample Mann-Whitney test).

influence of varying domain sizes on synoptic skill in terms of EV. Furthermore, subsequently performed two sample Mann–Whitney (M–W) tests revealed – in most cases – the statistical significance (for $\alpha = 0.05$) of differences between location parameters determined for the optimal domain size and location parameters derived on the basis of all other domain sizes, respectively. These spatial structures exemplified for TM apply as well to TN and TX (not shown).

For RR a somewhat similar spatial structure can be seen (Figure 8). As for TM this structure appears in all seasons, however less marked in summer. K–W tests and as well M–W tests yield statistical significant positive test results for almost all locations. Whereas for RR the spatial patterns of optimal domain sizes for EV are congruent to the respective patterns for BSS90 (not shown) this is not the case for TM. Optimal domain sizes for TM derived on the basis of BSS10 (Figure 9) and BSS90 (Figure 10) exhibit spatial variations that (particularly in spring and autumn) partly differ from those depicted for EV. For BSS10 in spring the above detected decrease in optimal domain size from the north to the south is maintained in the western part of the superordinate spatial domain, whereas in the eastern part the largest domain sizes 7 and 8 are optimal (at most locations supported by significance of respective K–W and M–W tests). Partly comparable spatial structures emerge for BSS10 in autumn and BSS90 in spring. However, with regard to BSS10 in autumn the dominance of the largest domain sizes is restricted mainly to the south-eastern parts and only for a minority of locations K-W and M–W test yield significance. For BSS90 in spring

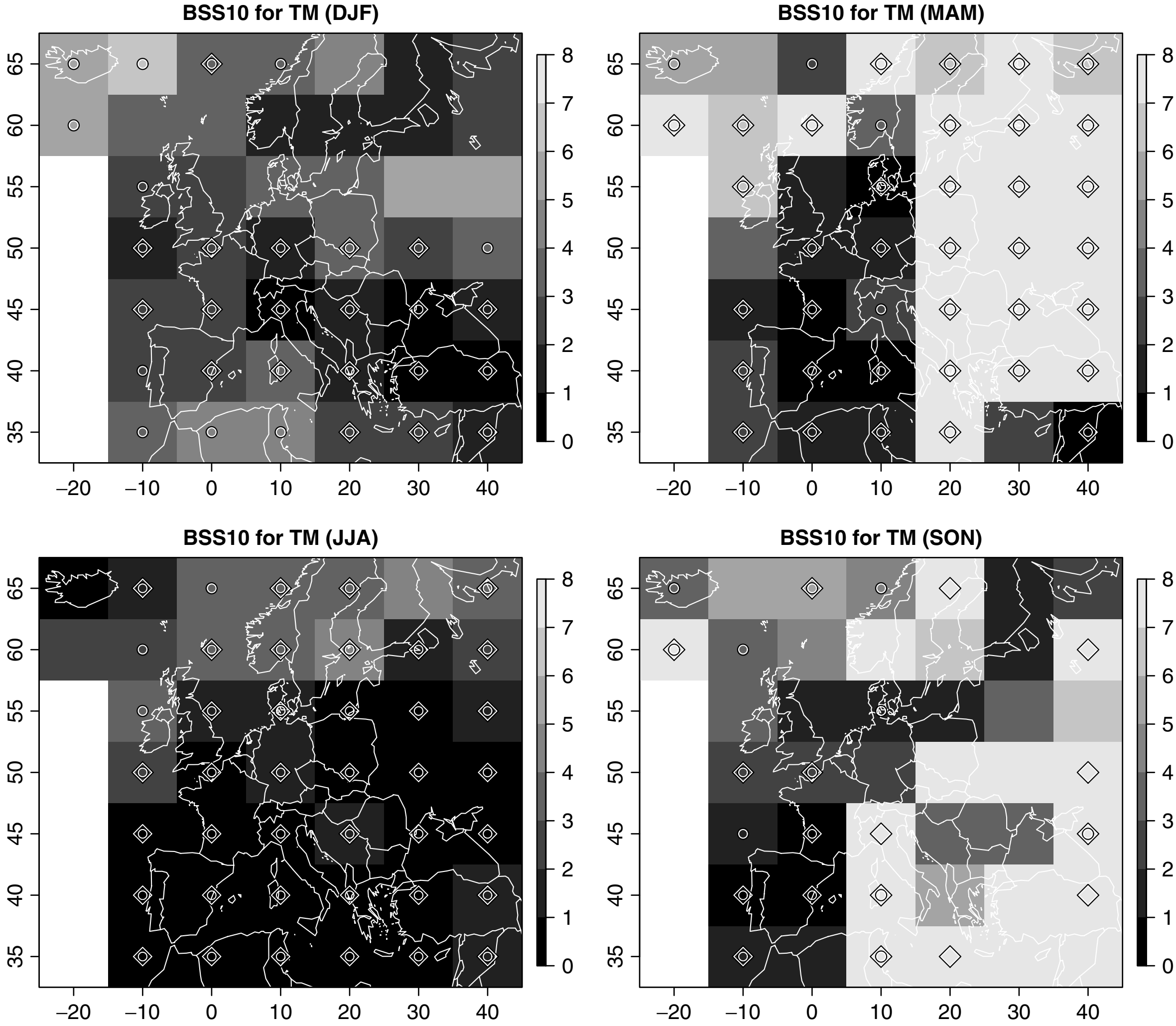

Figure 9. Spatial variation of domain sizes for which highest synoptic skill (BSS10) of circulation type classifications for the target variable TM is achieved in the four three-month seasons winter (DJF), spring (MAM), summer (JJA) and autumn (SON). Optimal domain sizes – from smallest domain size 1 to largest domain size 8 – for each of the 44 target domains are indicated by gray shades. The optimal domain size for each location has been determined on the basis of the seasonal synoptic skill values averaged over all grid points within the respective domain. Circles at a location denote statistical significance (for alpha = 0.05) of the influence of varying domain sizes on synoptic skill in the respective season (according to a Kruskal-Wallis test). Diamonds indicate additionally a statistically significant (for alpha = 0.05) difference between location parameters estimated for the optimal domain size and estimated over all other domain sizes respectively (according to a two sample Mann-Whitney test).

the largest domain size 8 is clearly favoured in the region east of 5°W and north of 45°N whereas smaller domain sizes (1 to 3) are best suited in the south/south-east (only partly supported by respective significant results of K–W and M–W tests). Finally, the congruence of TM, TN and TX stated above concerning spatial variations in optimal domain size for EV is also valid for BSS10 and BSS90.

4.3.3. *Variations among classifications*

In this section how far effects of domain size on synoptic skill vary among CTCs is investigated.

As an example Figure 11 shows synoptic skill – determined for the central European target domain centred around 10°E/50°N – in terms of EV (in %) for TM and RR in winter and summer in dependence on domain size separately for the seven CTCs listed in Table 1. Although there are partly marked differences in synoptic skill between the varying CTCs all classification methods feature similar characteristics with respect to the optimal domain size. For instance the optimal domain sizes are for TM in the range between 3 and 5 (winter) and 2 and 3 (summer) and for RR in the range between 1 and 3 (winter) and 1 and 2 (summer). Thus although not for all CTCs highest synoptic skill is achieved for the same domain size the different methods are at least in good agreement concerning the approximate spatial magnitude (‘small’, ‘medium’, ‘large’) of the optimal domain. This applies as well to other seasons and other locations (not shown). Concerning RR the general accordance among CTCs concerning the optimal domain size also exists for BSS90 (not shown). For BSS90 and BSS10 derived for temperature variables (TM, TN, TX) on the other hand such an agreement among CTCs does not exist. Figure 12

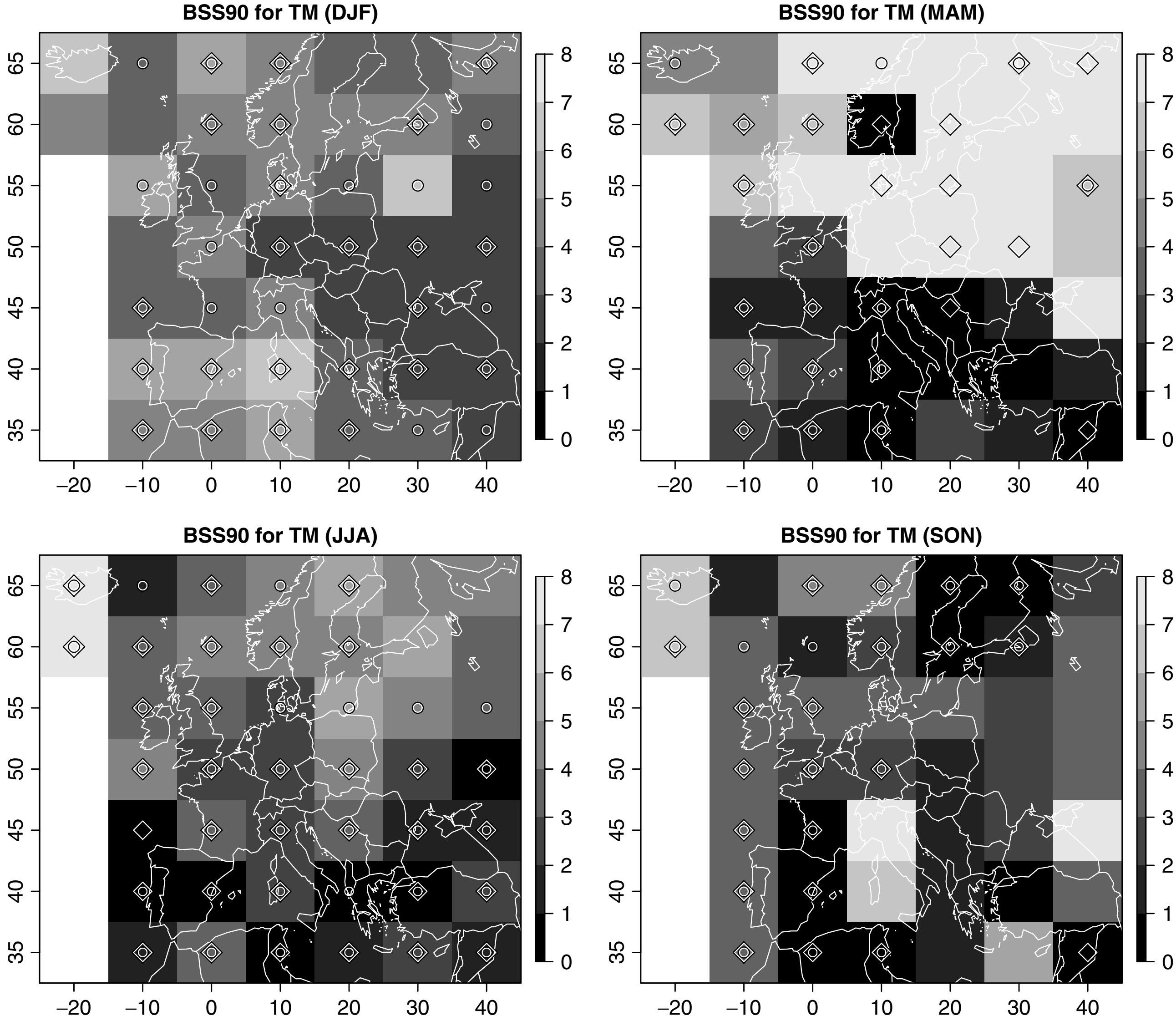

Figure 10. Spatial variation of domain sizes for which highest synoptic skill (BSS90) of circulation type classifications for the target variable TM is achieved in the four three-month seasons winter (DJF), spring (MAM), summer (JJA) and autumn (SON). Optimal domain sizes – from smallest domain size 1 to largest domain size 8 – for each of the 44 target domains are indicated by gray shades. The optimal domain size for each location has been determined on the basis of the seasonal synoptic skill values averaged over all grid points within the respective domain. Circles at a location denote statistical significance (for alpha = 0.05) of the influence of varying domain sizes on synoptic skill in the respective season (according to a Kruskal-Wallis test). Diamonds indicate additionally a statistically significant (for alpha = 0.05) difference between location parameters estimated for the optimal domain size and estimated over all other domain sizes respectively (according to a two sample Mann-Whitney test).

illustrates synoptic skill for TM grouped according to domain size for individual CTCs in terms of EV (in %), BSS10 and BSS90 in autumn determined for the eastern Mediterranean target domain centred around 30°E/40°N. For EV optimal domain sizes for the different CTCs are in the range between 1 and 4 whereas for BSS10 and BSS90 optimal sizes vary between 1 and 8. Concerning BSS10 four CTCs (GWT, KRZ, PTT, LND) show a marked increase in synoptic skill for TM with expanding domain size whereas the other three CTCs (PXE, CKM, DKM) reach maximum synoptic skill for the smallest domain sizes 1 and 2. For BSS90 a clear predominance of a specific domain size is clearly evident only for CKM, DKM and PXE with maximum synoptic skill related to the smallest domain size. These examples are only valid for one specific location and one season and can not be transferred to other locations and seasons. However differences of comparable magnitude between CTCs concerning optimal domain sizes with respect to BSS10 and BSS90 for temperature variables appear for other locations and seasons as well. Most pronounced in spring and autumn and in more continental regions. This seasonal and spatial distribution is partly reflected by the absence of statistical significance of K−W and M−W tests in Figure 9 and Figure 10 pointing to the fact that at these locations maximum values of synoptic skill can not clearly be attributed to a certain domain size.

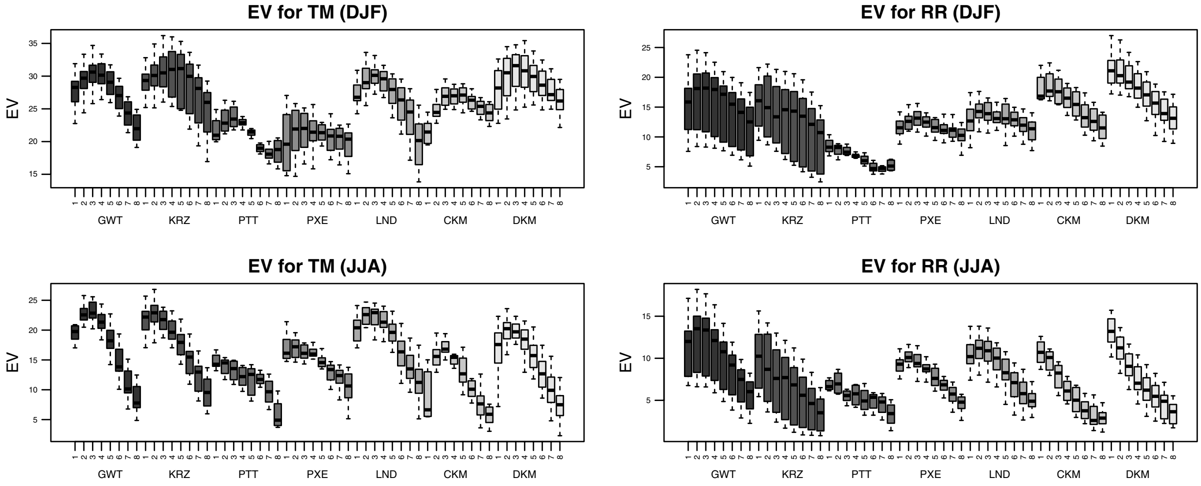

Figure 11. Boxplots of evaluation metric EV (in %) estimated for daily mean temperature (TM) and daily precipitation sums (RR) in winter (DJF) and summer (JJA), respectively, within the central European target domain centred around 10°E/50°N. EV values have been determined for domain sizes 1 to 8 separately for each of the circulation type classifications listed in Table 1. Upper/lower whiskers indicate the 1.5 interquartile range from the upper/lower quartile. Circulation type classifications and domain sizes are indicated along the x axis.

Besides the variations in optimal domain size among CTCs an additional important finding is that the differences in synoptic skill of CTCs that are due to varying domain sizes are comparable to those differences that are induced by varying classification methods. Looking for example at differences in EV (in %) for TM in summer between different classification methods applied to the same domain size (Figure 11) one can estimate a maximum difference of roughly 15%. The maximum difference between EV values for one classification method applied to domains of varying size is approximately also 15%.

5. Discussion and conclusions

The primary aim of this study was to determine the effect of the size of the spatial domain on the synoptic skill of CTCs.

A first major outcome of the analyses is that – integrated over all classifications, seasons, locations and climatic target variables – CTCs applied to domain sizes 2 and 3 (with horizontal dimensions of approximately 1300 and 1800 km in west–east direction) feature highest synoptic skill. According to Rudeva and Gulev (2007) the radii of northern hemispheric extratropical cyclones vary between 200 and 1400 km with modal values of the frequency distribution around 400–500 km and 500–600 km over land and ocean respectively. Thus domain sizes 2 and 3 cover the typical dimensions of these synoptic systems that are highly relevant for surface climate characteristics.

However, this overall finding needs to be further specified with respect to varying climatic target variables, seasons and locations.

First optimal domain sizes are systematically smaller for precipitation than for temperature pointing the fact that precipitation characteristics are more closely related to smaller scale circulation features especially during summer when optimal domain sizes for precipitation are further reduced. The decrease in optimal domain size in summer is also in accordance with the finding that the size of cyclones is reduced in summer compared to winter (Rudeva and Gulev, 2007). In numerous cases the smallest domain size 1 appears to be best suited to capture precipitation characteristics although this domain size (approximately 700 km in west–east direction) not necessarily covers the whole of synoptic systems but possibly only captures parts of them. This is in line with previous findings (Beck and Philipp, 2010) showing that best synoptic skill for precipitation in Central Europe is reached by CTCs that classify indices for vorticity and the direction of isobars determined only for the central parts of the domain. Thus it can be concluded that CTCs do not necessarily have to resolve whole synoptic systems in order to provide high synoptic skill but may as well focus on specific circulation features at a smaller spatial scale.

However, two questions arise from the application of CTCs to such small spatial scales as represented by domain size 1. (1) Can the synoptic skill of CTCs be further optimized through the use of circulation data with higher spatial resolution allowing to resolve relevant small scale processes? (2) Do CTCs applied to such small domains provide better synoptic skill than indices derived for single grid points or stations (whereby such local indices may enter into a synoptic classification (Kalkstein *et al.*, 1987) as well)? Both questions lead beyond the scope of this article but raise interesting perspectives for subsequent investigations.

A distinct deviation to the overall preponderance of relatively small optimal domain sizes has been detected concerning the synoptic skill for temperatures below (above) the 10% (90%) quantile particularly in spring and autumn (winter). For these cases domain sizes 5

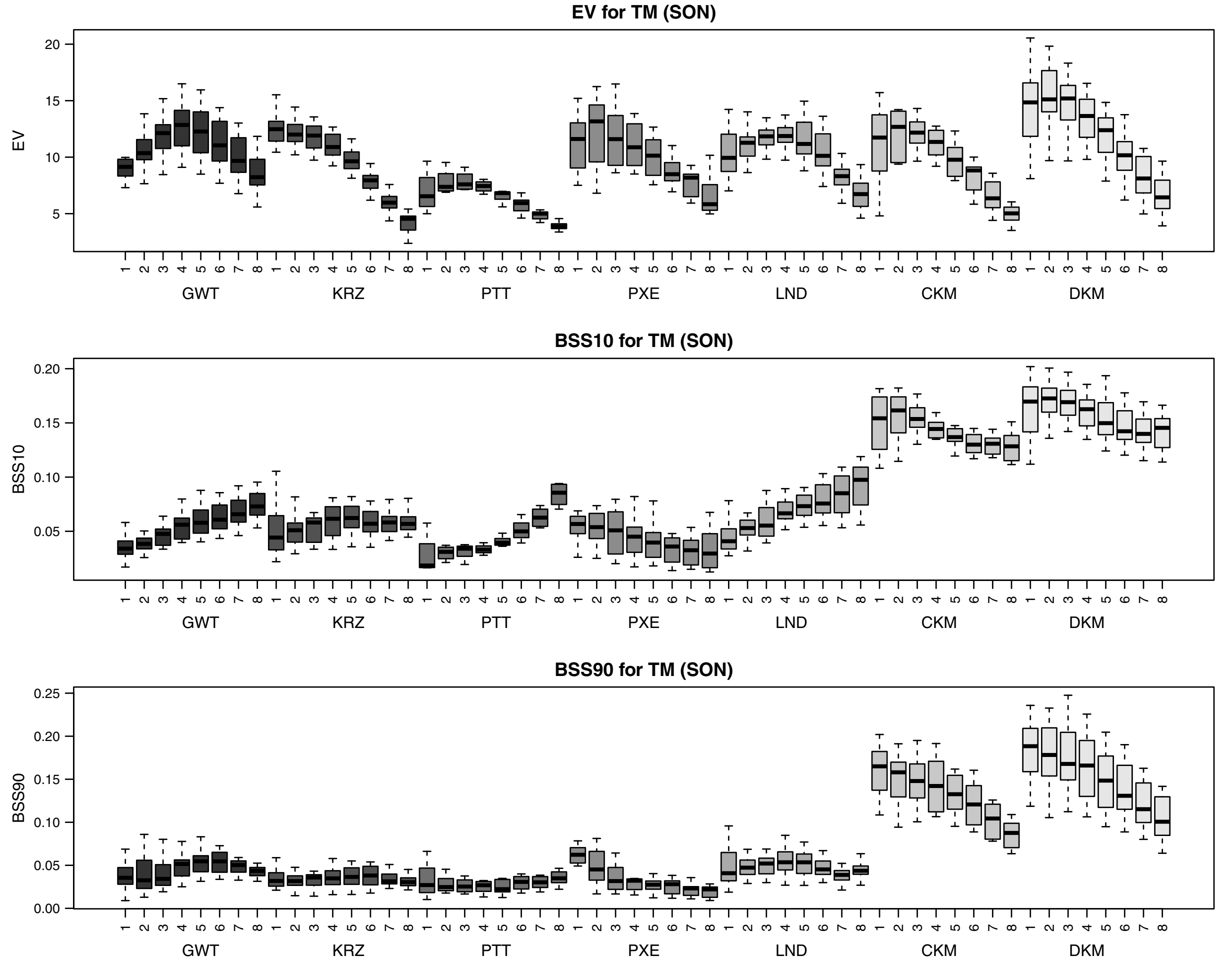

Figure 12. Boxplots of evaluation metrics EV (in %), BSS10 and BSS90 estimated for daily mean temperature (TM) in autumn (SON) within the eastern Mediterranean target domain centred around 30°E/40°N. EV, BSS10 and BSS90 values have been determined for domain sizes 1 to 8 separately for each of the circulation type classifications listed in Table 1. Upper/lower whiskers indicate the 1.5 interquartile range from the upper/lower quartile. Circulation type classifications and domain sizes are indicated along the x axis.

to 8 (approximately 2500 to 3700 km in west–east direction) appear to be best suited at numerous locations. This may be explained by the importance of large-scale atmospheric circulation structures for the advection of warm and cold air masses or for autochthonous conditions leading to local temperature anomalies of positive and negative sign respectively (e.g. an extended Russian High in spring leading to negative temperature anomalies particularly in the eastern part of the superordinate domain) and the fact that such continental scale features are better captured with larger domains.

Besides large-scale spatial gradients in optimal domain size (e.g. decreasing domain size from north-west to south-east for RR) partly abrupt changes can be observed between neighbouring locations as well. On the one hand these differences in some cases may be related to corresponding changes in general land surface characteristics; e.g. the different sizes of synoptic systems over the ocean and over land (Rudeva and Gulev, 2007). On the other hand it has to be stated that most of the very distinct changes between neighbouring locations cannot be convincingly explained. This corresponds to other studies where local influences lead to low spatial autocorrelation concerning links between circulation and local climate (e.g. Philipp, 2009). Only for a minority of these cases the relevance of the discrepancies is reduced by the fact that the determination of the optimal domain size is not statistically significant according to K–W and M–W tests.

Finally the most important overall finding from our study is that the differences in synoptic skill between different domain sizes reach comparable magnitude than those between different methods applied to the same domain size. Thus we can conclude that with respect to the application of CTCs in synoptic climatological analyses – including classification based downscaling approaches – it is important not only to select the most appropriate classification method but it is just as important to choose the adequate spatial domain. However in the same way as it is not possible to provide universally valid recommendations concerning the most appropriate classification method (Beck and Philipp, 2010; Huth, 2010) no general rule concerning the optimal domain size that should be used for performing CTCs can be derived from the here presented results. Rather it is advisable to determine the most appropriate domain size – according to synoptic skill – for a specific application depending on the location, the season and the climatic or environmental target variable. With respect to recently developed advanced approaches for circulation type or weather type classification that utilize not only one but multiple parameter fields for classification such a screening for the optimal domain accordingly needs to be extended from the two-dimensional to the multidimensional space.

Acknowledgements

Main parts of the work presented in this paper are related to the COST733 Action 'Harmonisation and Applications of Weather Types Classifications for European Regions' that has been funded by the EU within the framework of COST (European Cooperation in Science and Technology). The authors thank two anonymous reviewers for their helpful comments on an earlier version of the manuscript.

References

Beck C, Philipp A. 2010. Evaluation and comparison of circulation type classifications for the European domain. *Phys. Chem. Earth* **35**: 374–387.

Beck C, Jacobeit J, Jones PD. 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.

Benestad RE. 2001. A comparison between two empirical downscaling strategies. *Int. J. Climatol.* **21**: 1645–1668.

Brinkmann WAR. 1999. Within-type variability of 700 hPa winter circulation patterns over the Lake Superior Basin. *Int. J. Climatol.* **19**: 41–58.

Brinkmann WAR. 2002. Local versus remote grid points in climate downscaling. *Int. J. Climatol.* **21**: 27–42.

Buchanan CM, Beverland IJ, Heal MR. 2002. The influence of weather-type and long-range transport on airborne particle concentrations in Edinburgh, UK. *Atmos. Environ.* **36**: 5343–5354.

Buishand TA, Brandsma T. 1997. Comparison of circulation classification schemes for predicting temperature and precipitation in the Netherlands. *Int. J. Climatol.* **17**: 875–889.

Casado MJ, Pastor MA, Doblas-Reyes FJ. 2010. Links between circulation types and precipitation over Spain. *Phys. Chem. Earth* **35**: 437–447.

D'Onofrio A, Boulanger JP, Segura EC. 2010. CHAC: a weather pattern classification system for regional climate downscaling of daily precipitation. *Clim. Change* **98**: 405–427.

Dayan U, Levy I. 2005. The influence of meteorological conditions and atmospheric circulation types on PM10 and visibility in Tel Aviv. *J. Appl. Meteorol.* **44**: 606–619.

Demuzere M, Trigo RM, Vila-Guerau de Arellano J, Van Lipzig NPM. 2009. The impact of weather and atmospheric circulation on O3 and PM10 levels at a rural mid-latitude site. *Atmos. Chem. Phys.* **9**: 2695–2714.

Demuzere M, Kassomenos P, Philipp A. 2011. The COST733 circulation type classification software: an example for surface ozone concentrations in Central Europe. *Theor. Appl. Climatol.* **105**: 143–166, doi: 10.1007/s00704-010-0378-4.

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, Schneider F, Deutschländer T. 2005. A novel scheme to derive optimized circulation pattern classifications for downscaling and forecast purposes. *Theor. Appl. Climatol.* **82**: 51–63.

Esteban P, Jones PD, Martin-Vide J, Mases M. 2005. Atmospheric circulation patterns related to heavy snowfall days in Andorra, Pyrenees. *Int. J. Climatol.* **25**: 319–329.

Fleig A, Tallaksen LM, Hisdal H, Stahl K, Hannah DM. 2010. Inter-comparison of weather and circulation type classifications for hydrological drought development. *Phys. Chem. Earth* **35**: 507–515.

Garcia C, Marti G, Oller P, Moner I, Gavalda J, Martinez P, Pena JC. 2009. Major avalanches occurrence at regional scale and related atmospheric circulation patterns in the Eastern Pyrenees. *Cold Reg. Sci. Technol.* **59**: 106–118.

Goodess CM, Jones PD. 2002. Links between circulation and changes in the characteristics of Iberian rainfall. *Int. J. Climatol.* **22**: 1593–1615.

Hartigan JA. 1975. *Clustering Algorithms. Wiley Series in Probability and Mathematical Statistics*. John Wiley & Sons: New York, pp. 351.

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M. 2008. A European daily high-resolution gridded dataset of surface temperature and precipitation. *J. Geophys. Res.* **113**: D20119, doi: 10.1029/2008JD10201.

Hofstra N, Haylock M, New M, Jones PD. 2009. Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature. *J. Geophys. Res.* **114**: D21101, doi: 10.1029/2009JD011799.

Hollander M, Wolfe DA. 1999. *Nonparametric Statistical Methods*, 2nd edn. Wiley: New York, 787 pp.

Hope PK, Drosdowsky W, Nicholls N. 2006. Shifts in the synoptic systems influencing southwest Western Australia. *Clim. Dyn.* **26**: 751–764.

Huth R. 1993. An example of using obliquely rotated principal components to detect circulation types over Europe. *Meteorol. Zeitschr.* **2**: 285–293.

Huth R. 2002. Statistical downscaling of daily temperature in Central Europe. *J. Clim.* **15**: 1731–1742.

Huth R. 2010. Synoptic-climatological applicability of circulation classifications from the COST733 collection: first results. *Phys. Chem. Earth* **35**: 388–394.

Huth R, Beck C, Philipp A, Demuzere M, Untstrul Z, Cahynova M, Kysely J, Tveito OE. 2008. Classifications of atmospheric circulation patterns: recent advances and applications. *Ann. N.Y. Acade. Sci.* **1146**: 105–152.

Huth R, Beck C, Tveito OE. 2010. Classifications of atmospheric circulation patterns – theory and applications – preface. *Phys. Chem. Earth* **35**: 307–308.

Kalkstein LS, Tan G, Skindlov A. 1987. An evaluation of three clustering procedures for use in synoptic climatological classification. *J. Appl. Meteorol.* **26**: 717–730.

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White J, Woolen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**: 437–470.

Kassomenos P. 2010. Synoptic circulation control on wild fire occurrence. *Phys. Chem. Earth* **35**: 544–552.

Kruizinga S. 1979. Objective classification of daily 500 mbar patterns. Preprints Sixth Conference on Probability and Statistics in Atmospheric Sciences, Banff, Alberta. American Meteorological Society, Boston, MA, 126–129.

Kysely J. 2007. Implications of enhanced persistence of atmospheric circulation for the occurrence and severity of temperature extremes. *Int. J. Climatol.* **27**: 689–695.

Leckebusch GC, Weimer A, Pinto JG, Reyers M, Speth P. 2008. Extreme wind storms over Europe in present and future climate: a cluster analysis approach. *Meteorol. Zeitschr.* **17**: 67–82.

Lund IA. 1963. Map-pattern classification by statistical methods. *J. Appl. Meteorol.* **2**: 56–65.

Makra L, Juhasz M, Mika J, Bartzokas A, Beczi R, Sümeghy Z. 2006. An objective classification system of air mass types for Szeged, Hungary, with special attention to plant pollen levels. *Int. J. Biometeorol.* **50**: 403–421.

Philipp A. 2009. Comparison of principal component and cluster analysis for classifying circulation pattern sequences for the European domain. *Theor. Appl. Climatol.* **96**: 31–41.

Philipp A, Della-Marta PM, Jacobeit J, Fereday DR, Jones PD, 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.

Philipp A, Bartholy J, Beck C, Erpicum M, Esteban P, Fettweis X, Huth R, James P, Jourdain S, Kreienkamp F, Krennert T, Lykoudis S, Michalides S, Pianko-Kluczynska K, Post P, Rassilla Alvarez D, Schiemann R, Spekat A, Tymvios FS. 2010. COST733CAT – a database of weather and circulation type classifications. *Phys. Chem. Earth* **35**: 360–373.

Pineda N, Esteban P, Trapero L, Soler X, Beck C. 2010. Circulation types related to lightning activity over Catalonia and the Principality of Andorra. *Phys. Chem. Earth* **35**: 469–476.

Rudeva I, Gulev SK. 2007. Climatology of cyclone size characteristics and their changes during the cyclone life cycle. *Mon. Weather Rev.* **135**: 2568–2587.

Saunders IR, Byrne JM. 1999. Using synoptic surface and geopotential height fields for generating grid-scale precipitation. *Int. J. Climatol.* **19**: 1165–1176.

Schiemann R, Frei C. 2010. How to quantify the resolution of surface climate by circulation types: An example for Alpine precipitation. *Phys. Chem. Earth* **35**: 403–410.

Sheridan SC, Lee CC. 2010. Synoptic climatology and the general circulation model. *Prog. Phys. Geogr.* **34**: 101–109.

Spak S, Holloway T, Lynn B, Goldberg R. 2007. A comparison of statistical and dynamical downscaling for surface temperature in North America. *J. Geophys. Res.* **112**: D08101, doi: 10.1029/2005JD006712.

Tang L, Cheng D, Karlsson PE. 2009. Synoptic circulation and its influence on spring and summer surface ozone concentrations in southern Sweden. *Boreal Environ. Res.* **14**: 889–902.

Tveito OE. 2010. An assessment of circulation type classifications for precipitation distribution in Norway. *Phys. Chem. Earth* **35**: 395–402.

Vicente-Serrano SM, Lopez-Moreno JI. 2006. The influence of atmospheric circulation at different spatial scales on winter drought variability through a semi-arid climatic gradient in northeast Spain. *Int. J. Climatol.* **26**: 1427–1453.

Wetterhall F, Halldin S, Xu CY. 2007. Seasonality properties of four statistical-downscaling methods in central Sweden. *Theor. Appl. Climatol.* **87**: 123–137.

Wilks DS. 2006. *Statistical Methods in the Atmospheric Sciences*, 2nd edn. Elsevier: Academic Press, 627 pp.

Yarnal B. 2006. The effect of weather map scale on the results of a synoptic climatology. *J. Climatol.* **4**: 481–493.

Yarnal B. 1993. *Synoptic Climatology in Environmental Analysis*. Belhaven Press: London.