Weather and Climate Extremes 9 (2015) 17–24

Contents lists available at ScienceDirect

Weather and Climate Extremes

journal homepage: www.elsevier.com/locate/wace

Spatial clustering of summer temperature maxima from the CNRM-CM5 climate model ensembles & E-OBS over Europe

Margot Bador [a,*], Philippe Naveau [b], Eric Gilleland [c], Mercè Castellà [d], Tatiana Arivelo [e]

[a] Climate Modelling and Global Change Team, CERFACS/CNRS, Toulouse, France
[b] Laboratoire des Sciences du Climat et de l'Environnement, CNRS-CEA-UVSQ, Gif-sur-Yvette, France
[c] Research Applications Laboratory, National Center for Atmospheric Research, Boulder, United States of America
[d] Centre for Climate Change, Department of Geography, University Rovira i Virgili, Tortosa, Spain
[e] African Climate Policy Centre, United Nations Economic Commission for Africa, Addis Ababa, Ethiopia

ARTICLE INFO

Article history:
Received 19 December 2014
Received in revised form
20 May 2015
Accepted 22 May 2015
Available online 27 May 2015

Keywords:
Spatial clustering
Climate extreme
Data ensemble
Multivariate extreme value theory

ABSTRACT

Reducing the dimensionality of the complex spatio-temporal variables associated with climate modeling, especially ensembles of climate models, is a challenging and important objective. For studies of detection and attribution, it is especially important to maintain information related to the extreme values of the atmospheric processes. Typical methods for data reduction involve summarizing climate model output information through means and variances, which does not preserve any information about the extremes. In order to help solve this challenge, a dependence summary measure appropriate for extreme values must be inferred. Here, we adapt one such measure from a recent study to a larger domain with a different variable and gridded data from observations and climate model ensembles, i.e. E-OBS observations and the CNRM-CM5 model. The handling of such ensembles of data is proposed, as well as a comparison of the spatial clusterings between two different ensembles, here a present-day and a future ensemble of climate simulations. This method yields valid information concerning extremes, while greatly reducing the data set.

1. Introduction

One aim of statistical climatology is to reduce complex spatio-temporal variations of atmospheric and oceanic variables to a small number of statistical quantities, often means and covariances, that summarize the present state and future evolution of the climate system. These summaries involve the analysis of spatio-temporal averages and variances of climate variables.

Empirical Orthogonal Functions (EOF) represent the canonical example of such statistical techniques taking their roots in correlation structures. EOF analysis has been instrumental in identifying modes of climate variability such as the North Atlantic Oscillation and North Pacific Pattern (e.g., see review by Hurrell et al., 2003). A key element to justify this reliance on averages and covariances is their strong mathematical link with the Gaussian probability density function (pdf), which is entirely characterized by these two mathematical summaries. In addition, the central limit theorem states that averages over large blocks in space or time can be well approximated by the Gaussian pdf (e.g., see page 35 of von Storch and Zwiers, 2002).

From a risk perspective, extreme weather events that strongly depart, not only from the mean, but also lay outside of the usual range of climate variability, are of the highest interest because of their strong potential of having a devastating impact on society. In such cases, the idea of summarizing the distributional features of extreme events via means and correlations is entirely inappropriate. Further, statistical analysis must be addressed within a very different probability framework. Here, we rely on the multivariate extreme value theory (EVT, e.g., see the books of Resnick, 2007; de Haan and Ferreira, 2006; Beirlant et al., 2004; Coles, 2001).

Just as EOF analysis seeks to reduce the dimensionality of climate variability by identifying spatial domains with highly correlated climate variability, we aim to reduce the dimensionality by clustering the time series of maxima taken over a block size of interest (season, year, decades, etc.) spatially.

We need here to measure the spatial dependence among block maxima, e.g. largest summer temperatures at different locations. Cooley et al. (2006) proposed a convenient distance adapted from the variogram distance used in geostatistics (e.g., Wackernagel, 2003). Vannitsem and Naveau(2007) applied this distance to precipitation measured in Belgium. It essentially compares the ordering of extreme events between two time series of maxima.

* Corresponding author.
E-mail address: margot.bador@cerfacs.fr (M. Bador).

http://dx.doi.org/10.1016/j.wace.2015.05.003

Consequently, this rank-based approach bypasses the complex step of fitting a parametric marginal law at each location.

For clustering, Bernard et al. (2013) recently proposed a rank-based algorithm and applied it to weekly maxima of hourly precipitation in autumn from 1993 to 2011 over France. In this paper, we make use of their clustering algorithm, where the new contributions derive from the following unique challenges.

Here, we analyze summer maxima of daily temperatures over Europe, which involves a much larger domain, as well as a variable with very different characteristics. Another important departure from Bernard et al. (2013) is that we work with gridded numerical model output, an historical (1950–2005) ensemble and a future (2006–2100) ensemble of the CNRM-CM5 climate model, as well as E-OBS gridded temperature observations. Working with climate model output will shed some light on potential future changes of spatial clustering of temperature maxima, a topic rarely covered in the Detection and Attribution literature.

The use of a different data set also raises important methodological questions. Compared to precipitation recorded from 1993 to 2011, the hypothesis of stationarity used by Bernard et al. (2013) is on shaky ground for temperature maxima over the period 1950–2005. The trend effect is removed from the observed and simulated temperatures before computing our distance between time series of maxima, thus explicitly focusing on spatially analyzing temperature maxima patterns that result from internal variability of the climate system. This approach provides complements to the large number of studies that have focused on changes in absolute temperature extremes over this time period in climate model ensembles (e.g., Kharin et al., 2007; IPCC, 2012; Sillmann et al., 2013).

A second methodological point is to determine how to compare two maps of clusters, for example the one obtained from present-day runs with the one derived from future runs. In this paper, we will offer a simple algorithmic solution to this nontrivial problem, see the end of Section 3.3.

As in Bernard et al. (2013), a byproduct of our analysis will be a reduction of the dimensionality in time (by taking the maxima over a block size of three months), and in space (the clustering reduces a grid of a few hundred points into a much smaller, well-chosen set of grid points that represent the center of each cluster).

2. Data and methods

2.1. Summer maxima of daily temperature maxima

In this paper, we focus on JJA (June–July–August) maxima of daily maximum temperatures over Europe. We work with two types of gridded data: E-OBS temperatures from 1950 to 2013 (0.5° horizontal resolution; Haylock et al., 2008) and climate model runs from the CNRM-CM5 model (1.4° horizontal resolution; Voldoire and et al., 2012). This climate model was used in the Coupled Model Intercomparison Project Phase 5 (CMIP5) and has 10 historical simulations over the period 1850–2005, 5 simulations over the future period (2006–2100) under the Representative Concentration Pathways (RCP) 8.5 (for a total radiative forcing pathway leading to 8.5 W/m^2 in 2100). In this paper, we refer to the E-OBS observation dataset as EOBS and to the historical and future ensembles as HIST (1950–2005) and RCP8.5 (2006–2100), respectively. To compare spatial clusterings between observations and model outputs, EOBS temperatures are interpolated on the model grid using the *remapcon* command of the Climate Data Operators software (http://www.mpimet.mpg.de/cdo), a conservative regridding method.

Simulated temperatures have first been linearly de-trended from the non-physical long-term changes (model drift) found in climate models (Gupta et al., 2013). This drift is estimated from the control simulation, a long simulation (850 years) in which the natural (solar activity and volcanic aerosols) and anthropogenic (greenhouse gases and aerosols) forcings remain constant at their pre-industrial 1850 values.

In a second step, we de-trend observed and simulated temperatures from the long-term warming trend, following a two-step procedure. First, we remove the multi-year climatological average from daily temperature maxima within every dataset. Then, from these temperature residuals, the historical and future trends are estimated by the HIST and RCP8.5 ensemble average (respectively), for every grid point and each calendar day. These 10- and 5-member means (respectively) are filtered by a 91-day running average. The future trend is removed from the simulated temperatures of the RCP8.5 ensemble, whereas the historical trend is removed from both observed temperatures of EOBS and simulated temperatures of the HIST ensemble.

Removing the same trend from observations and simulations adds the assumption that the model correctly captures the mean response to anthropogenic greenhouse gases over Europe. However, our algorithm always works with the ranks of individual time series, not their absolute values (see Eq. (1)) and this reduces the complex issue of trend removal by subtracting a linear or other fit. Such methods are extremely sensitive to the method of fitting.

Removing the warming trend is not a compulsory step of our algorithm. Still, interpreting clusters of raw temperatures is more difficult than from anomalies, as the trend itself could play a role in the clustering. Such an issue is especially prominent at the end of the 21st century and under the RCP8.5 scenario, which projects a strong warming over Europe. Here, we prefer to focus on detrended temperatures, a reflection solely of internal variability.

Before explaining the clustering algorithm used here, it is important to comment on the recurrent confusion concerning the interpretation of results based on block maxima, here summer extreme temperatures. Seasonal maxima at two different locations do not need to occur on the same day. In other words, the block size, the season here, removes all information about the timing within a block. Hence, any distance measuring the proximity between two series of maxima over 55 years provides “climatological” rather than “weather” information. In particular, it is misleading to speak about the temporal synchronicity of extreme “events” or “episodes” here. Having two close by stations with respect to our distance simply means that their maxima behave similarly, in a distributional sense at the yearly scale, see Fig. 3. Of course, it may be possible (although rarely) that the hottest days for some year happen simultaneously at two locations, but this temporal feature is never taken into account with our clustering analysis.

2.2. Clustering algorithm

Following the notations of Bernard et al. (2013), we denote by i and j two grid point locations, and their two associated time series of seasonal maxima by $M_i^{(t)}$ and $M_j^{(t)}$ where t represents a given year. In order to spatially cluster temperature maxima, we need to somehow calculate, not the geographical distance, but a type of dissimilarity measure between two time series of maxima. Here, we use a rank-based distance defined by

$$\hat{d}_{ij} = \frac{1}{2T}\sum_{t=1}^{T} |R_i^{(t)} - R_j^{(t)}| \tag{1}$$

where $R_i^{(t)}$ corresponds to the rank of the tth year within the time series of maxima recorded at location i. We opt for this distance for two reasons. First, being based on ranks, it is possible to compare

different objects in the sense that there is no need to "standardize" or "transform" the temperature maxima (no need to fit a Generalised Extreme Value distribution). Second, Eq. (1) can be interpreted as a "degree of dependence" within the mathematical framework of bivariate EVT (e.g., see Resnick, 2007; de Haan and Ferreira, 2006; Beirlant et al., 2004; Coles, 2001; Fougères, 2004). To see this, one has to remember the basic probability definitions of independence and complete dependence between two equiprobable events, say A and B with $\mathbb{P}(A) = \mathbb{P}(B)$. In particular, we can write

$$\mathbb{P}(A \text{ and } B) = \mathbb{P}(A)^{\theta}. \tag{2}$$

where the cases $\theta = 2$ and $\theta = 1$ define the independence and complete dependence between A and B, respectively.

In the case of a bivariate max-stable random vector with unit-Fréchet margins, i.e. $A = \{M_i \le u\}$, $B = \{M_j \le u\}$ with $\mathbb{P}(A) = \mathbb{P}(B) = \exp(-1/u)$ for $u > 0$, it is possible to write (2), not only for the two cases of independence and complete dependence but for any max-stable structure. More precisely, we can always write for bivariate max-stable vectors with unit-Fréchet margins that

$$\mathbb{P}(M_i \le u;\, M_j \le u) = \left[\mathbb{P}(M_i \le u)\, \mathbb{P}(M_j \le u)\right]^{\theta_{ij}/2}.$$

The so-called "extremal coefficient" θ_{ij} (e.g., Schlather, 2002; Schlather and Tawn, 2003; Naveau et al., 2009) summarizes the degree of dependence between the time series M_i and M_j.

To make the link with (1), Cooley et al. (2006) showed that the extremal coefficient contains the same information as our rank-based distance. In particular, one can be deduced from the other via

$$\hat{d}_{ij} = \frac{1}{2}\frac{\hat{\theta}_{ij} - 1}{\hat{\theta}_{ij} + 1}. \tag{3}$$

Knowing how to measure the proximity of two grid points with respect to their maxima, we can now apply a clustering algorithm. A classical candidate is the k-means algorithm that creates cluster centers by averaging points within a cluster. But averaging breaks the concept of max-stability, the mean of two maxima is not a maximum anymore and the interpretation with the extremal coefficient cannot be used within a k-means algorithm. As already noted by Bernard et al. (2013), the Partitioning Around Medoids (PAM) algorithm proposed by Kaufman and Rousseeuw(1990) has the advantage of preserving maxima in the sense that each cluster centers remains a time series of maxima, not an average of time series.

The PAM algorithm divides a dataset of N objects into K clusters. Three pre-processing steps are needed before implementing PAM. First, the distance matrix defined by (1) needs to be computed. Second, the number of clusters K must be chosen and third, the algorithm requires an initial set of K medoids, which are randomly selected.

Then, the PAM algorithm can be run as follows:

(A) Assign each grid point to the nearest medoid with respect to the distance (1).
(B) For each cluster, find the new medoid for which the total intra-cluster distance based on d_{ij} is minimized.
(C) Repeat steps (A) and (B) until the clusters converge and remain unchanged through one iteration

To choose a relevant number K of clusters and to assess if a weather station is well classified, Rousseeuw (1986) developed the so-called "silhouette coefficient" that compares cluster tightness (small d_{ik} within the cluster k) with cluster dissociation (see $\delta_{i,-k}$ defined below). After running the PAM algorithm with a given K, each location i is associated with a medoid k. The silhouette coefficient for the weather station i is defined as follows:

$$s_i(K) = 1 - \left(d_{ik}/\delta_{i,-k}\right), \tag{4}$$

where d_{ik} represents the average intra-cluster distance between station i and all other stations associated with medoid k. The real $\delta_{i,-k}$ corresponds to the smallest of the $k-1$ average distance between station i and all other stations associated with a medoid different from k. For the PAM algorithm procedure, $s_i(K)$ necessarily belongs to the interval $[-1, 1]$. If $s_i(K) \approx 1$, it means that the intra-cluster distance is much smaller than the inter-cluster distances. Consequently, the maximum M_i can be considered as well classified. In contrast, if s_i is near zero, the clustering is viewed as non-informative, meaning that M_i could have been in a different cluster with the same relevancy. To summarize the quality of a partitioning into K clusters, one can use the average silhouette coefficient

$$\bar{s}(K) = \frac{1}{N}\sum_{i=1}^{N} s_i(K), \tag{5}$$

or other statistics from the set $\{s_1(K), \ldots, s_N(K)\}$. Such summaries will be used in our application. To implement our approach, a package for the open-source statistical R software is available on the homepage of the second author.

3. Clustering of summer temperature maxima

3.1. Clustering EOBS

The clustering algorithm applied to the de-trended EOBS dataset of JJA seasonal maximum of daily maximum temperature is provided in Fig. 1 for the particular number of clusters $K=15$. As expected, the insignificant points are located mostly at the boundaries between two clusters, which is a positive sign concerning the robustness of our clustering algorithm. Medoids are not necessary located in the center of a cluster.

Concerning the climate interpretation, the output with $K=15$ clusters show coherent structures and highlights regional climates, e.g. over the Iberia peninsula and the United Kingdom. These clear spatial coherent structures must have a physical cause, likely due to synoptic-scale weather patterns. However, no geographical information like latitudinal and longitudinal coordinates is used in the clustering algorithm, which is only driven by the ranks of time series of JJA seasonal maxima of daily maximum temperatures.

Each cluster strength can be analyzed via its cluster-averaged silhouette coefficient. Highest values reveal strong clusters and lowest values weak ones. We have noticed that when the number of clusters increases by one unit, the weakest cluster (or one of the weakest) often splits into two clusters, whereas the strongest clusters are the most stable when varying K in the input of the clustering. This result leads to the delicate question of how to choose the appropriate number of clusters. Given K, the algorithm will always divide the domain into K subdomains, even if the optimal number of clusters is below K. In addition, one of our goals is to reduce dimensionality. Hence, it makes sense to choose the smallest K possible whenever adding an extra cluster does not bring any significant information.

To illustrate, Fig. 2 displays how the average silhouette coefficient $\bar{s}(K)$ defined by Eq. (5) varies as a function of K. From the observations (top panel), one can clearly see that $\bar{s}(15)$ is the highest value, and that using $K \ge 16$ does not improve the spatial clustering over Europe. In addition, Fig. 1 confirms that the clustering with $K=15$ is spatially and climatologically coherent.

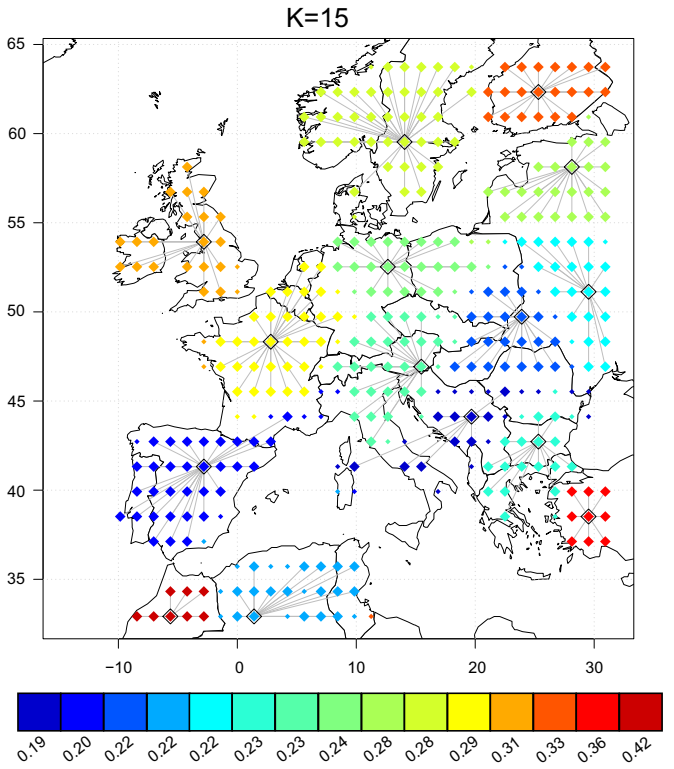

Fig. 1. Clustering of EOBS JJA maximum temperatures (de-trended) with $K=15$. Colors refer to the cluster strength from strong (reddish colors) to weak (blueish colors). Diamond shaped locations represent medoids (i.e. cluster centers) and points not linked by a grey line to a medoid have a silhouette coefficient below the minimum significance level of 0.1. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

To further develop this analysis, we have also rerun our clustering algorithm while only keeping significant points. The silhouette coefficients statistics from this second clustering are all shifted toward higher values, and the shape of the evolution is rather unchanged.

The bottom panel of Fig. 2 also displays how the average silhouette coefficient $\bar{s}(K)$ evolves within the HIST ensemble. A common feature between the observed and simulated evolutions is the first increase of $\bar{s}(K)$ towards its highest value, followed either by relatively stable or even decreasing values. Contrary to EOBS, the HIST ensemble highlights $K=11$ as the most appropriate number of clusters to choose as an input to the clustering algorithm.

In summary, $K=15$ appears to be a reasonable choice for EOBS, whereas $K=11$ seems more appropriate for the model. One possible reason for this difference may come from the original spatial resolution of EOBS, whose temperatures were gridded on a half-degree regular grid from a data set of stations. Local processes, such as soil moisture content variations for example, can have significant impacts in summer on daily maximum temperatures over Europe. Such local information could be included within the observations but not within the simulated temperatures because of a coarser spatial resolution (1.4-degree horizontal resolution) and also because of the limitations of the CNRM-CM5 land-surface scheme. In addition, differences between the observed and simulated atmospheric circulation could also explain these differences on the average silhouette coefficient.

To go one step further in understanding the spatial features of Fig. 1, the temporal evolution of significant members of clusters #7 (around Serbia), #8 (around Austria) and #15 (Turkey), is displayed in panels (b), (c), and (d) of Fig. 3, respectively. It is striking that although the amplitude can vary, the temporal synchronicity (ups and downs) of each series is very similar within each cluster. This

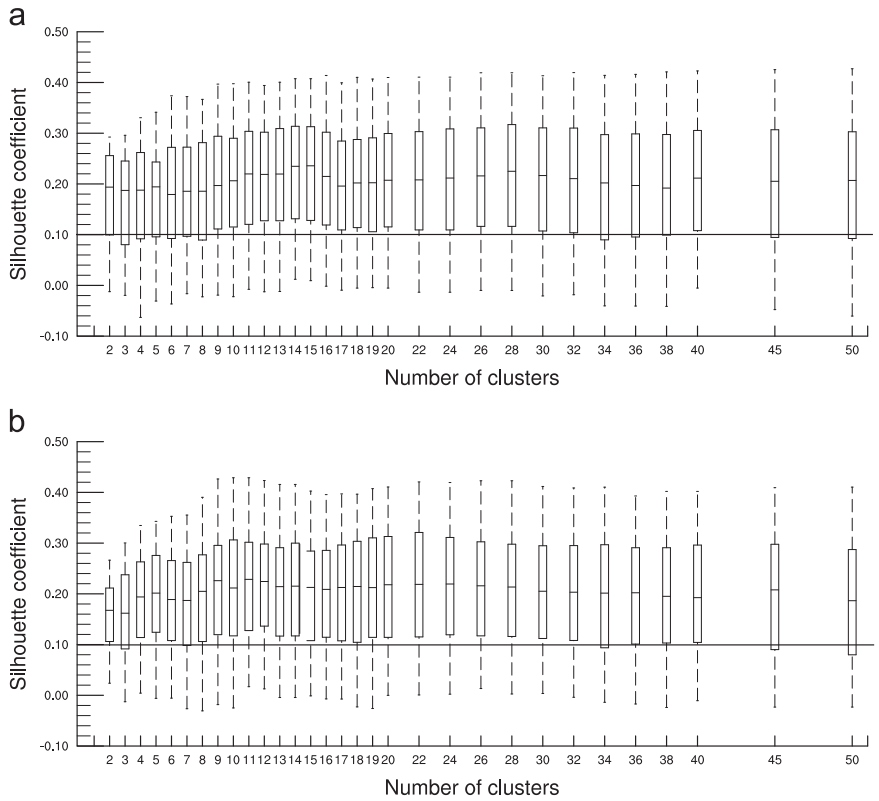

Fig. 2. Silhouette coefficient statistics (median, 10th and 90th percentiles) of $\bar{s}(K)$ defined by Eq. (5) as a function of the number of clusters from $K=2$ to $K=50$. A larger value corresponds to a more coherent clustering. The horizontal line indicates to the significance level of 0.1 and the dataset corresponds to EOBS JJA maximum temperatures (de-trended) for panel (a), and HIST JJA maximum temperatures (de-trended) for panel (b).

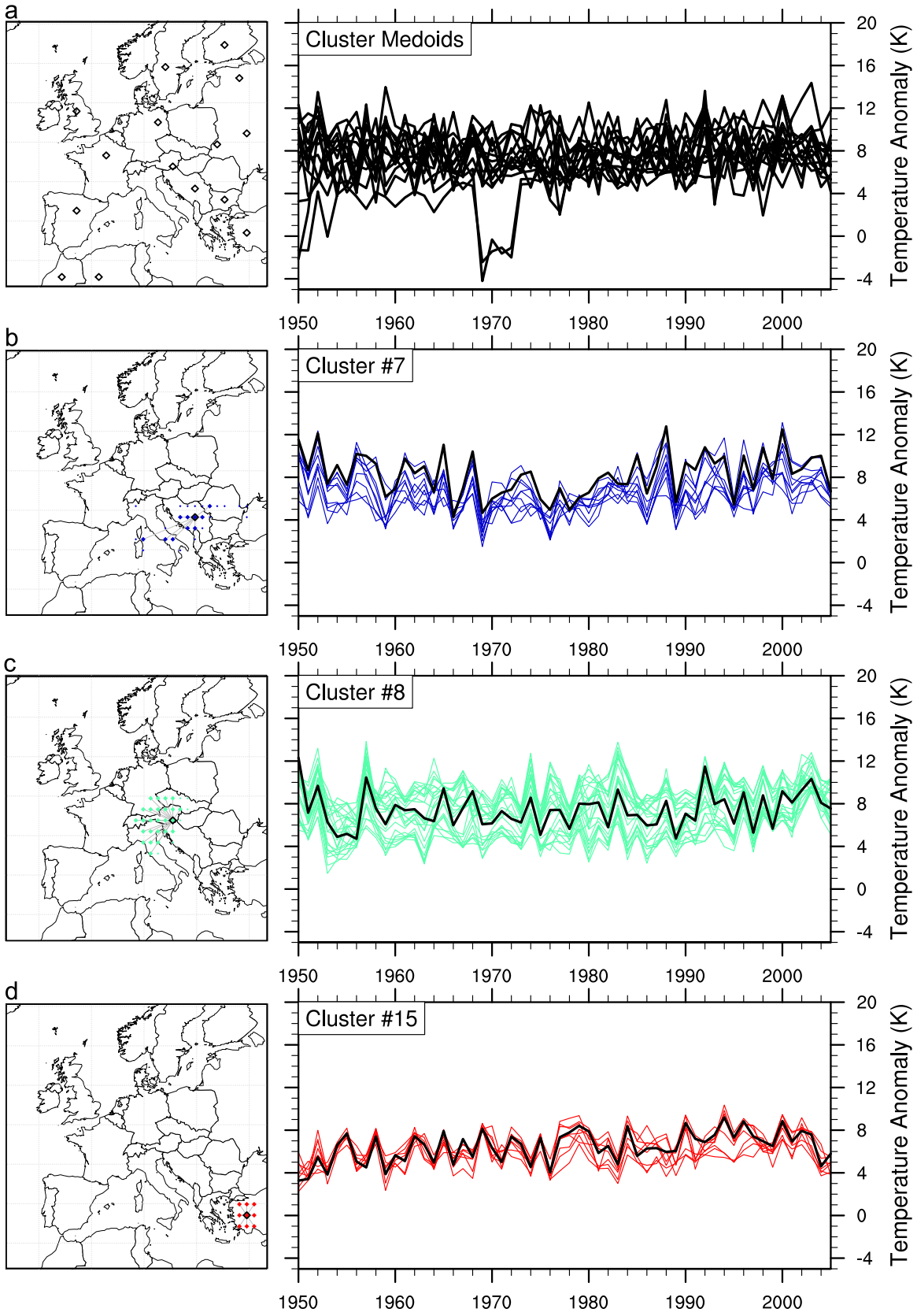

Fig. 3. Temporal depiction of the clustering of Fig. 1 with $K=15$. Panel (a): medoids time series. Panels (b), (c) and (d): temporal evolution of significant members (colored lines) of clusters #7 (around Serbia), #8 (around Austria) and #15(Turkey), respectively. The maps show the locations of the cluster medoids (diamonds) and the cluster members (colored points), where only significant members are considered ($s_i(K) > 0.1$). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

can be explained by our distance (1) based on ranks. Another clear feature in panels (b)–(d) is that cluster #15 is better self-organized than #7 and #8. This is reflected by the color code in Fig. 1 that is proportional to the averaged silhouette coefficient of the cluster.

Concerning the variability between clusters, panel (a) of Fig. 3 shows a strong variability among the medoids' temporal behavior. This result makes sense because we expect a similar conduct within a cluster but a large difference between clusters.

In summary, Fig. 3 emphasizes that medoids appear to be correctly selected because they accurately describe their own cluster and strongly depart from other medoids. In terms of spatial dimension reduction, it means that our 15 medoids could be used as representative time series for the whole European domain.

A medoid represents a member of a cluster that can provide similar up and downs (in time) than the elements of its cluster. In other words, a medoid should, in theory, indicate how the synchronicity of its largest values propagates within its cluster. However, the medoids do not summarize all characteristics of a cluster. Instead, rather like a EOF captures some elements (variances) of a data set, a medoid provides a particular feature (ranks dissimilarity) of the cluster, and regarding a given metric, the extremal coefficient in our case, a medoid is an optimal representative.

An application of this result could be found in the comparison of the time series of a medoid, say temperature maxima, with a complete different variable, say pressure fields. In this case, working with ranks provides a simple way to compare apples and oranges. The study of the medoid time series between two far away regions could also help seeking interconnections, through the identification of temporal synchronicities, even with large magnitude differences between these two regions.

3.2. *Clustering an ensemble*

To cluster a model ensemble, say with M members, two possibilities can be considered. One can apply the clustering algorithm to each member and then summarize the resulting M maps in one way or another. A second option is to merge all of the ensembles into one big matrix, each column representing the same grid point where each ensemble is stacked one after the other, and then apply the clustering algorithm at once.

Although the first option may appear simpler at first, it has two main drawbacks. Statistically, it is not easy to summarize M maps. Physically, the clustering from one member is a result of the particular trajectory of the climate undertaken in that member over the period of interest. Seeking the ensemble medoids from the clustering of each of the members should not be expected, because these different expressions of the internal variability cannot finally converge toward the clustering that best describes the climate (as simulated in the model, with a finite number of members). This behavior is particularly true over continental areas where internal variability has stronger impacts than over coastal areas. The second approach does not have these drawbacks. In terms of interpretation, the output simply depicts the spatial clustering over the period of interest, based on a sample of the different realizations (i.e. the members).

As we want to compare two ensembles (HIST and RCP8.5) in this paper, we need to develop a simple and fast criterion to assess how a cluster for ensemble A has changed in ensemble B, and vice-versa. Our idea is to study how medoids move around from one ensemble to the other. For example, suppose that the city of Trieste corresponds to a medoid location in the clustering of ensemble A. Then one can easily find the cluster that contains Trieste in ensemble B and its associated medoid, say Pisa for the sake of illustration. Now, it is possible to check if Pisa either belongs to the original cluster centered around Trieste in ensemble A or not.

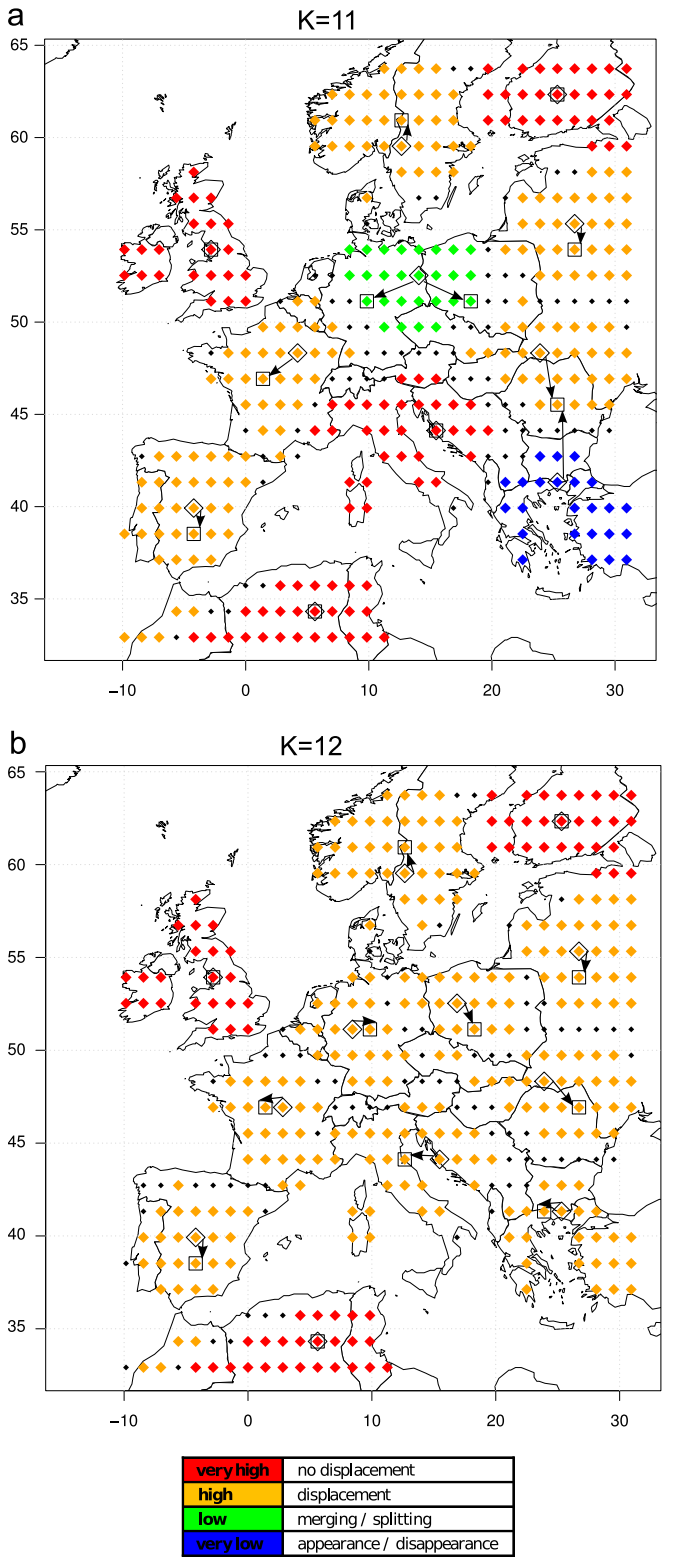

Fig. 4. Comparison of two clustering maps with $K=11$ obtained from HIST (1950–2005) and RCP8.5 (2006–2100) ensembles of the CNRM-CM5 model. Medoids from HIST ensemble are indicated by diamonds, while squares refer to medoids of the RCP8.5 clustering. The clustering map corresponds to the output from the HIST ensemble. Hot (cold) colors mean stability (change) of the spatial distribution of clusters over time. Black points refer to insignificant members ($s_i(K) < 0.1$). Arrows illustrate the displacement of the medoids between HIST and RCP8.5 clustering maps. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Basically, we need to look at how the medoids of ensemble A move to ensemble B and back to A (A → B → A). The opposite road (B → A → B) can also be explored and, at the end, we have four different cases:

very high medoids remain identical in their trajectories from A to B, and vice-versa.
high medoids change locations but remain in the same cluster in both directions.
low medoids change locations but remain in the same cluster in only one direction.
very low medoids change locations and clusters in any direction.

These four medoid behaviors can be interpreted as no displacement, displacement, merging/splitting and appearance/disappearance, respectively, (see color code in Fig. 4).

3.3. Evolution of clusterings maps made from CNRM-CM5 outputs

To determine possible changes in internal variability of extreme summer temperatures between the historical and the future period (in the model world), we apply the clustering on HIST (1950–2005) and RCP8.5 (2006–2100) ensembles with $K=11$, and the top panel of Fig. 4 compares the two clusterings.

Out of the 11 medoids, 4 stay at the exact same location and 5 are slightly displaced. Overall, 9 clusters out of 11 do not present clear changes in the future ensemble, but two others do. Because one cluster disappears (blue) in the RCP8.5 clustering, one has to appear, and so is the green cluster split in two, as indicated by the two arrows.

The analysis of Fig. 4a with $K=11$ highlights two regions presenting possible changes in clustering of summer extreme temperatures, while the rest of Europe remains unchanged. To explore the robustness of these two regions, the analysis has also been performed with $K=12$ (Fig. 4b). The two regions that witness a change for $K=11$ disappear with $K=12$. The same conclusion also holds for $K=10$ (not shown).

Hence, we conclude that the spatial clustering of summer extreme temperatures within the CNRM-CM5 model world is robust over time. The changes detected for $K=11$ are an artifact of statistical instabilities rather than a strong signal. This conclusion is corroborated by the fact that the other nine regions remain invariant for $K=11$.

In addition, the changes between the two ensembles for $K=11$ are associated with clusters of low averaged silhouette coefficients in both ensembles (not shown). If we were detecting changes for the largest silhouette coefficients, then changes will be more significant, but it is not the case here, and for all these reasons, we cannot reject the hypothesis of temporal stationarity for the clusters spatial patterns.

Statistical instabilities have been found through the sensitivity to the sampling size. We tested the differences in the spatial clustering with the same RCP8.5 ensemble but with a half-reduced HIST ensemble size (for both ensembles to have 5 members). Results clearly show the dependence to the sampling size (not shown). Indeed, the HIST clustering maps present small changes, and so does the map of changes between the two ensembles. In our case, a larger RCP8.5 ensemble should probably lead to slightly different maps of changes between the two ensembles.

4. Discussion

Bernard et al. (2013) introduced a new framework for clustering events based on their extremes rather than solely on their means and variances, which is important for risk analysis concerning a changing climate. In this paper, we add to the method in order to account for a much larger domain, ensembles of data, and idiosyncrasies associated with a different variable. It is found that using different numbers of clusters can reveal associations at different scales of weather/climate phenomena, but for the most part, within reason, the results are robust to the choice of the number of clusters, which does need to be determined a priori. The most surprising, and positive, result is that, without using positional information, specific geographical structures nevertheless revealed themselves.

Reducing the dimensionality of a dataset while preserving information about the extremes is a difficult task because it falls within the realm of extreme value theory where handling multivariate/spatial data can be challenging. The distance measure used here, has fundamental links with multivariate extreme value theory via the extremal coefficient, but does not rely on estimating any parameters or dependence structures. In fact, the measure is very straightforward to compute, and it is appropriate for clustering extremes. Moreover, the framework is couched inside the usual, and well understood, clustering framework; specifically, in this case, the Partitioning Around Medoids (PAM) technique.

Another important contribution of this work is the handling of ensembles of data, which introduce unique challenges. On one hand, they can be handled by analyzing each member individually, and then summarizing the resulting ensemble of maps. Such a scheme is fraught with difficulties, despite its simple appearance. In particular, from a statistical point of view, it is difficult to summarize many maps, but also the results of each map is a direct consequence of the unique trajectories of each member, which do not translate well into a summary. On the other hand, the ensemble can be taken as a large multivariate variable whereby the clustering can be performed so that one unique map is determined. This latter approach does not suffer from the difficulties of the first, ostensibly more attractive method.

This study has been conducted using a single climate model, which does not show any significant changes in spatial clustering of summer temperature extremes for the future European climate. The present work also offers a blueprint to develop a methodology for future studies focused on extremes where it is desired to reduce spatio-temporal dimensions.

A multi-model analysis from the CMIP5 simulations could be interesting and helpful to understand the changes in interannual variability over Europe. These issues are actively studied by the climate community (e.g., Scherrer, 2005; Seneviratne et al., 2006; Parey et al., 2009; Fischer et al., 2012), but the spatial dependencies are less frequently examined.

Understanding how the internal variability is represented in every climate model is a crucial point to further investigate if changes are projected to occur in response to global warming. This could be done with the control simulations of the models, where the forcings remain constant over very long periods.

Models having a large number of members in their ensembles could be particularly useful for these issues. Indeed, the effect of the sampling has been highlighted in this study. Large ensembles would imply lower sampling effects, and a better description of the spatial correlation of extremes.

Acknowledgments

We acknowledge the World Climate Research Programme (WCRP) and the International Centre for Theoretical Physics (ICTP), which have supported this research through the WCRP-ICTP Summer School on Extremes (2014). This work contributes to the WCRP Grand Challenge on Extremes. Part of this work has been

supported by EDF and by the French National Research Agency (ANR) and its program«Investissementsd'avenir» under the Grant ANR-11-RSNR-0021, ANR-DADA, the LEFE-Multirisk, CHAVANA, AMERISKA and ExtremeScope projects, and also the National Science Foundation through Earth System Modeling (EaSM) Grant number AGS-1243030. The authors would like to thank the contributors of the statistical freely available software R http://www.r-project.org/. Analyses and graphics have also been done using the NCAR Command Language (Version 6.2.1, http://dx.doi.org/10.5065/D6WD3XH5).

References

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of Extremes: Theory and Applications. John Wiley & Sons, New York.

Bernard, E., Naveau, P., Vrac, M., Mestre, O., 2013. Clustering of maxima: spatial dependencies among heavy rainfall in France. J. Clim. 26 (20), 7929–7937. http://dx.doi.org/10.1175/JCLI-D-12-00836.1.

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. 208 pp.

Cooley, D., Naveau, P., Poncet, P., 2006. Variograms for spatial max-stable random fields. In: Dependence in Probability and Statistics, Lecture Notes In Statistics, vol. 187. Springer, New York, pp. 373–390.

de Haan, L., Ferreira, A., 2006. Extreme Value Theory, An Introduction. Springer Series in Operations Research and Financial Engineering.

Fischer, E.M., Rajczak, J., Schär, C., 2012. Changes in European summer temperature variability revisited. Geophys. Res. Lett. 39 (19). http://dx.doi.org/10.1029/2012GL052730, URL ⟨http://doi.wiley.com/10.1029/2012GL052730⟩.

Fougères, A.-L., 2004 : Multivariate extremes. In: Extreme Values in Finance, Telecommunications, and the Environment, pp. 373–388.

Gupta, A.S., Jourdain, N.C., Brown, J.N., Monselesan, D., 2013. Climate drift in the CMIP5 models. J. Clim. 26 (21), 8597–8615. http://dx.doi.org/10.1175/JCLI-D-12-00521.1, URL ⟨http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00521.1⟩.

Haylock, M.R., Hofstra, N., Klein Tank, a.M.G., Klok, E.J., Jones, P.D., New, M., 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. J. Geophys. Res. 113 (D20), D20119. http://dx.doi.org/10.1029/2008JD010201, URL ⟨http://doi.wiley.com/10.1029/2008JD010201⟩.

Hurrell, J.W., Kushnir, Y., Ottersen, G. and Visbeck, Martin, eds. 2003. The North Atlantic Oscillation: Climate Significance and Environmental Impact. Geophysical Monograph Series, 134. American Geophysical Union, Washington, DC.

IPCC, 2012. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. Cambridge University Press, Cambridge, http://dx.doi.org/10.1017/CBO9781139177245, URL ⟨http://ebooks.cambridge.org/ref/id/CBO9781139177245⟩.

Kaufman, L., Rousseeuw, P., 1990. Finding Groups in Data: An Introduction to Cluster Analysis. Wiley, New-York.

Kharin, V.V., Zwiers, F.W., Zhang, X., Hegerl, G.C., 2007. Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. J. Clim. 20 (8), 1419–1444. http://dx.doi.org/10.1175/JCLI4066.1 ⟨http://journals.ametsoc.org/doi/abs/10.1175/JCLI4066.1⟩.

Naveau, P., Guillou, A., Cooley, D., Diebolt, J., 2009. Modeling pairwise dependence of maxima in space. Biometrika 96 (1), 1–17. http://dx.doi.org/10.1093/biomet/asp001.

Parey, S., Dacunha-Castelle, D., Hoang, T.T.H., 2009. Mean and variance evolutions of the hot and cold temperatures in Europe. Clim. Dyn. 34 (2–3), 345–359. http://dx.doi.org/10.1007/s00382-009-0557-0, URL ⟨http://link.springer.com/10.1007/s00382-009-0557-0⟩.

Resnick, S., 2007. Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer Series in Operations Research and Financial Engineering.

Rousseeuw, P., 1986. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20 (53–65).

Scherrer, S.C., 2005. European temperature distribution changes in observations and climate change scenarios. Geophys. Res. Lett. 32 (19), L19705. http://dx.doi.org/10.1029/2005GL024108, URL ⟨http://doi.wiley.com/10.1029/2005GL024108⟩.

Schlather, M., 2002. Models for stationary max-stable random fields. Extremes 5, 33–44.

Schlather, M., Tawn, J., 2003. A dependence measure for multivariate and spatial extreme values: Properties and inference. Biometrika 90 (1), 139–156.

Seneviratne, S.I., Lüthi, D., Litschi, M., Schär, C., 2006. Land-atmosphere coupling and climate change in Europe. Nature 443 (7108), 205–209. http://dx.doi.org/10.1038/nature05095, URL ⟨http://www.ncbi.nlm.nih.gov/pubmed/16971947⟩.

Sillmann, J., Kharin, V.V., Zwiers, F.W., Zhang, X., Bronaugh, D., 2013. Climate extremes indices in the cmip5 multimodel ensemble: Part 2. future climate projections. J. Geophys. Res. 118 (November (2012)), 2473–2493. http://dx.doi.org/10.1002/jgrd.50188.

Vannitsem, S., Naveau, P., 2007. Spatial dependences among precipitation maxima over Belgium. Nonlin. Process. Geophys. 14, 621–630.

Voldoire, a., et al., 2012. The CNRM-CM5.1 global climate model: description and basic evaluation. Clim. Dyn. 40 (9–10), 2091–2121. http://dx.doi.org/10.1007/s00382-011-1259-y, URL ⟨http://link.springer.com/10.1007/s00382-011-1259-y⟩.

von Storch, H., Zwiers, F.W., 2002. Statistical Analysis in Climate Research. Cambridge University Press, New York, NY, USA; Melbourne, VIC, Australia 484 pages.

Wackernagel, H., 2003. Multivariate Geostatistics. An Introduction with Applications, 3rd ed. Springer, Heidelberg.