FINAL MANUSCRIPT
CLIMATE RESEARCH
Vol. 78: 179–187, 2019

A new method to identify robust climate analogues

Carsten Walther , Matthias Lüdeke, Ramana Gudipudi

Potsdam Institute for Climate Impact Research, PO Box 601203, 14412 Potsdam, Germany

ABSTRACT: Climate analogues are a comprehensive approach for learning how to deal with the expected climatic future from current examples. Existing literature on climate analogues struggles with 2 methodological challenges: how to deal with the unavoidable uncertainty of climate projections and how to define reasonable lower limits of similarity for climate analogues. Here, we suggest a new method to identify robust climate analogues (RCAs) based on a clustering approach in climate space where each spatial grid element is represented by 3 points: its current climate, and a lower and upper bound of the climate projections. If the upper and lower bound of the projections for such a grid element share the same cluster and, additionally, this cluster contains current climate points, then the grid elements related to the latter are defined as RCAs. This definition divides the map of the investigated region into areas with RCAs and uncharted areas where, under the current uncertainty range of climate projections, such an attribution is not justified. An exemplary application of the algorithm for Europe shows that RCAs can be identified for 37% of the land area and that the new method allows selection of socioeconomically reasonable RCAs from climatologically equivalent (given the current uncertainty) grid elements.

KEY WORDS: Climate analogue · Clustering algorithm · Adaptation · Uncertainty · Regional climate model

1. INTRODUCTION

The development of appropriate strategies and tech niques for adaptation to projected climate change is an important (e.g. Pielke et al. 2007, Lobell et al. 2008, Reckien et al. 2015) but challenging task. It comprises comprehensive climate impact assessments, the development of catalogues of related adap tation options, and the choice of the most efficient options from these catalogues. Despite large and successful efforts towards this objective (e.g. Nasso poulos et al. 2012, Krysa nova & Hattermann 2017, Rosenzweig et al. 2017), there is no guarantee that all relevant climate-sensitive sectors and all possible adaptation options have been considered. This has motivated the development of complementary and more holistic perspectives. A promising approach in this context uses the spatial variability of the current climate to study how socio-ecological systems (SESs; e.g. urban agglomerations or rural areas) are affected by and deal with different climatic situations (e.g. arid or tropical climates). Given the projected climate for a specific location 'A' and following the space-for-time idea (e.g. Pickett 1989, Rastetter 1996), one may ask where the current climate equals this projection. Such locations are then called climate analogues (CAs) of location A. In the best case — such CAs exist and at least one of them also resembles the expected (future) socio-ecological situation of location A — successful strategies and techniques currently applied in the analogue area can be considered as useful adaptation options for location A. This approach has been followed in several studies in recent decades (e.g. Parry & Carter 1989, Dehn 1999, Dawson et al. 2009, Pugh et al. 2016) and requires an adequate choice of climate variables (depending on the considered SES), the definition of a distance measure in the climate space spanned by these variables and, based on this, an explicit definition of CAs.

Here, we present a new method for the identification of CAs which addresses 2 major, interlinked methodological problems: how to deal with the unavoidable uncertainty in climate projections and how to determine the degree of similarity justifying the notion of CAs. In earlier work, the dependence of the location of CAs on the chosen climate projection was recognized (e.g. Kopf et al. 2008) and partly used to discuss consequences of projection uncertainty (Hallegatte et al. 2007). However, these studies did not address the existence and location of robust climate analogues (RCAs), which are valid despite this uncertainty. More recent literature attempted to address this methodological issue either qualitatively (Arnbjerg-Nielsen et al. 2015) or quantitatively (Hibino et al. 2015). Arnbjerg-Nielsen et al. (2015) projected future rainfall for Denmark by taking measured historical intensity–duration curves (IDFs) from future Denmark's CAs. They used only IDFs from regions which appeared to be good analogues (measured by their ranking) under different climate projections. A more explicit approach to include projection uncertainty was recently presented by Hibino et al. (2015), which included the local uncertainty of climate projections in the parametrization of their 'similarity score'. This measure compares the standard deviation of different projections for grid element A with the (Euclidian) distance in climate space between this grid element and the potential CA for all considered climate variables. Each climate variable contributes to the similarity score only if the distance is smaller than the standard deviation, i.e. it is possible that the projection of grid element A is equal to the present climate of grid element B. This is a promising perspective on projection uncertainty and therefore we will follow the basic idea behind this approach in our CA definition.

Regarding the second methodological problem of necessary minimum similarity for CAs, to our knowledge, the existing literature identifies a CA by looking for a grid cell which is most similar in climate space and is thus essentially a relative measure. Formally, these approaches result in maps which show the degree of similarity of all grid elements relative to a fixed grid element A for which the CA is sought. The peak of the similarity landscape (or the minimum in the dissimilarity landscape) denotes the location of the CA. Sometimes a more or less subjective threshold is introduced to define a minimum similarity in climate space for a CA, which potentially results in locations without a CA (e.g. Hallegatte et al. 2007). We consider that looking at single maps of the aforementioned kind is only a first step, and we advocate analyzing the whole climate space, including the present climate and the different climate projections, to systematically identify reasonable requirements for CAs.

To solve the 2 above-mentioned problems, we suggest characterizing projection uncertainty for each grid element by a lower and upper bound for each climate variable. Each grid element thus generates 3 points in climate space: the present, the lower and the upper bound of projected climates. Then we perform an appropriate cluster analysis in climate space and define grid element B as a RCA of grid element A if and only if the present climate of B shares the same cluster with the lower and upper bound of climate projections for A. Thereby, the deviation in the projections is typically about the same magnitude as their distances to the current climate point, which takes up the basic idea on projection uncertainty consideration by Hibino et al. (2015). Furthermore, the subjective distance threshold definition is now substituted by the cluster partition which is determined by the data structure. Usually data structure excludes several cluster numbers, but some choice remains regarding this parameter. This choice is guided by the tradeoff between the need for sufficient differentiation, necessary for interpretation (pro large cluster number), and the extent of areas without RCAs (pro small cluster number). For a more comprehensive explanation and the interpretation of non-RCA point constellations, see Section 2. The proposed cluster method has some similarities with classical climate classifications (e.g. Geiger 1961). However, the major difference is that it generates a classification of the combined current and projected climate space — therefore a direct comparison of the resulting maps is not possible.

In this paper, we illustrate the suggested method by analyzing CAs in Europe with regard to monthly mean values of precipitation and temperature. These climate variables and the study region were mainly chosen to allow for a comparison of our new method's results with existing CA studies. Regarding the lower and upper bound of the climate projections, we used the model ensemble results of EURO-CORDEX for the period 2071–2100 under RCP8.5 (e.g. Jacob et al. 2014). For the present climate, we took the European Climate Assessment Data (ECAD) for the period 1971– 2000 (for more details see Section 2).

In the next section, we introduce the method in more detail and describe the data used for the illustrative example. In the following sections, we present the RCAs for Europe, discuss their peculiarities, and compare them with the results of Hallegatte et al. (2007) which used the same climate variables.

2. METHODS

2.1. Definition of CAs

Our approach considers a gridded area described by N grid elements, $i = 1, \ldots, N$. We assume climate is characterized by V appropriate normalized climate variables (e.g. monthly precipitation, total annual precipitation, monthly temperature), $j = 1, \ldots, V$. Each grid element i is described by the climate variables for its current climate *(C)* as well as for a lower *(L)* and upper bound of the climate projections *(U)* spanning the projection uncertainty. So each grid element i constitutes 3 data points in the V-dimensional climate space: $C_i = \{x^c_{i,1}, \ldots, x^c_{i,j}, \ldots, x^c_{i,V}\}$, $L_i = \{x^l_{i,1}, \ldots, x^l_{i,j}, \ldots, x^l_{i,V}\}$ and $U_i = \{x^u_{i,1}, \ldots, x^u_{i,j}, \ldots, x^u_{i,V}\}$ where x denotes the value of the climate variable j in grid element i of the current, lower and upper bound of the projected climates, $d \in \{c,l,u\}$. We then perform a cluster analysis on these $3 \times N$ objects in climate space. For further details on the choice of clustering method used in this analysis see Supplement 1 as well as Jans sen et al. (2012).

To illustrate the interpretation of the resulting clusters in terms of robust analogues, we show in Fig. 1 a stylized example for only 5 grid elements and 2 climate variables. Cluster I includes the lower and upper bound of the projections for grid element 1, as well as the current climate of grid elements 2 and 3. Therefore, in this example, grid elements 2 and 3 are RCAs to grid element 1. As the current climate of grid elements 1, 4, and 5 belongs to other clusters, these are not CAs for grid element 1. Since they do not share the lower and upper bound of the projections of a grid element within their respective clusters, they are not robust CAs for any grid element.

The blue lines in Fig. 1 illustrate how our approach relates to Hibino et al. (2015). In cluster I, the difference between L_1 and U_1 is in the same order of magnitude as the distance between the average climate projection and the current climates of grid elements 2 and 3. In contrast, C_4 and C_5 show a significantly larger distance than the uncertainty in climate projection for grid element 1 as they belong to other clusters. This takes up the basic idea of the recent suggestion by Hibino et al. (2015) to include projection uncertainty into CA definition where only distances that are smaller than or equal to the variance of projections contribute to a 'similarity index'. Another aspect of our method relates to the question of what degree of similarity is necessary to justify the notion of a CA. Here, cluster membership is the decisive property, i.e. instead of somewhat subjective thresholds (e.g. Hallegatte et al. 2007) the global data structure in climate space plays a large role in determining which climates can be defined as similar. This notion also results naturally in the identification of climates which disappear (cluster III in Fig. 1) and grid elements without (robust) CAs 'novel climates' after Williams et al. 2007 (represented by cluster IV in Fig. 1).

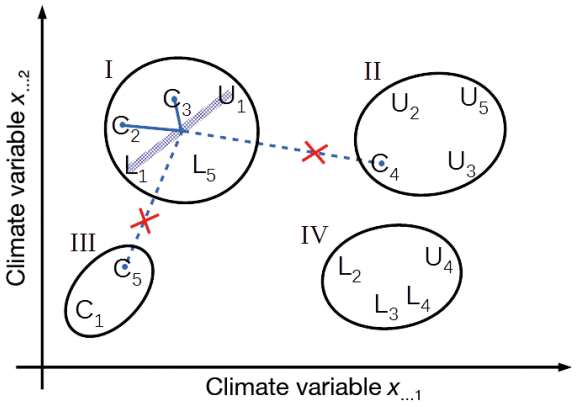

Fig. 1. Stylized illustration of the proposed approach for 5 grid elements $i = 1, \ldots, 5$ and 2 climate variables. C_i: current climate of grid element i; L_i: lower bound of the climate projections of grid element i; U_i: upper bound of the climate projections of grid element i. The 5 × 3 points in the 2-dimensional climate space form 4 clusters. Cluster I: grid elements 2 and 3 are robust climate analogues (CAs) for grid element 1; Cluster II: grid element 4 is not a robust CA to grid elements 2, 3, and 5 (it is only similar to their upper bound of the projections); Cluster III: grid elements 1 and 5 are not CAs for any grid element (disappearing climate of Williams et al. 2007); Cluster IV: grid element 4 has no present analogue (novel climate of Williams et al. 2007). For grid elements 2, 3, and 5, upper and lower bound of the projections belong to different clusters, i.e. they have no robust CAs. Blue lines: differences to and communalities with the approach of Hibino et al. (2015) (see Section 2.1)

Finally, we discuss a specific variant of RCAs as defined above. Assume cluster I would also include C_1, i.e. the projection uncertainty for grid element 1 does not allow one to distinguish between current and future climate. This formally fulfills our definition of a RCA (grid element 1 would be its own RCA) but is not interpretable regarding the transfer of adaptation options, and we therefore exclude these grid elements from our final RCA results and show them separately (see Fig. S3 in Supplement 5).

2.2. Determining lower and upper climate projections

Climate projections entail different sources of uncertainty: forcing, model response, internal variability, and the downscaling process (Tebaldi & Knutti

2007, Kopf et al. 2008, Hawkins & Sutton 2011, Taylor et al. 2012, Knutti & Sedlá<ek 2013). We ad dress this by analyzing the outputs of a climate model ensemble (e.g. from EURO-CORDEX) for a given forcing: for each grid element and climate variable, we consider the distribution of the climate variable over all climate models and calculate the value for the lower and upper bound of the projections according to fixed percentiles, in our case the 15^{th} and 85^{th} percentile, to get rid of outliers. Then, the 15^{th} (85^{th}) percentiles of the model distributions of all climate variables constitute the lower (upper) bound of the climate projections for a grid element. This characterization of the uncertainty range implies a simplification. The hyper cuboid in climate space, spanned by all combinations of upper and lower values of the climate variables for one grid element (2^{V} vertices), is approximated by one of its main diagonals. As the applied cluster algorithm tends to identify spherical clusters (see Supplement 1), this is an acceptable approximation which should not affect the identification of robust analogues significantly. Furthermore, it should be noted that these upper and lower vertices describe the uncertainty ranges spanned by all available models and are not necessarily physically consistent projections themselves.

While developing the presented algorithm we also considered the possibility to first apply the clustering to each single model projection and to identify afterwards communalities in the large number of resulting maps. One major problem of this procedure is that each single result will have a different optimum cluster number which makes the comparison almost impossible. This led us to the decision to first integrate the projections by using the upper/lower bound or envelope concept described above.

2.3. Applied data

Many existing CA studies are based on monthly mean temperature, monthly mean and sometimes yearly total precipitation (e.g. Hallegatte et al. 2007). As this study focusses on a new method, we limited our analysis to these climate variables to allow for comparison of the results. We selected Europe, a region with a broad range of climatic, cultural, and socioeconomic regions, as an appropriate study area. Since a large fraction (75%) of the European population resides in urban areas, mainly cities were used to present and compare our results.

The observed climate of the recent past (current climate) was collected from the gridded observational E-OBS data provided by the European Climate Assessment & Data (ECA&D) (Haylock et al. 2008). This data set covers the whole of Europe with a comparably high spatial resolution of 0.22 degrees and daily temporal resolution. For the projections of climate, 10 regional climate model (RCM) outputs taken from the EURO-CORDEX project (Jacob et al. 2014) were considered in this study. These regional climate projections with a spatial resolution of 0.11 degrees represent downscaled projections from CMIP5 general circulation models (GCMs) (IPCC 2013). The RCM data is available only in a rotated grid and was remapped and aggregated to the regular E-OBS grid. The extent of the study area was 13° W to 41° E and 34° to 72° N — this correlates to 32 832 grid cells out of which 16 844 are land cells.

The E-OBS and the EURO-CORDEX RCM data were the basis for the calculation of the 12 monthly mean temperatures, 12 monthly precipitation sums, and the yearly total precipitation under climate change. To account for potential biases of the models, the difference of the temperature variables gained from the climate model outputs of the future period (2071–2100) and the historical time period (1971–2000) were added to the observed baseline data of E-OBS for the period 1971–2000 (delta method; e.g. Mote & Salathé 2010). For precipitation-related variables, the ratio and multiplication was used instead. To make the different temperature and precipitation variables comparable to each other, we mapped them linearly onto the interval [0, 1] (unity based normalization). This allows for the calculation of the Euclidian distance in climate space which is necessary for the cluster algorithm. This choice of the normalization method uses the observed range of a climate variable within the study area as a unit. It is applicable as there are no extreme outliers.

3. RESULTS

The choice of an adequate number of clusters strongly influences the subsequent RCAs. After de - termining the set of numbers of clusters adequate to the data structure (for details see Supplement 2 and Sietz et al. 2011), we selected 10 clusters for further analysis. Once the algorithm delivered the allocation of the data points to the 10 clusters, the data set was redistributed to 3 maps of the cluster membership in position space for the current (Fig. 2b) and the upper and lower bound of the projections. Then we identified the intersection of the 2 future maps, which denotes regions where all EURO-CORDEX projec-

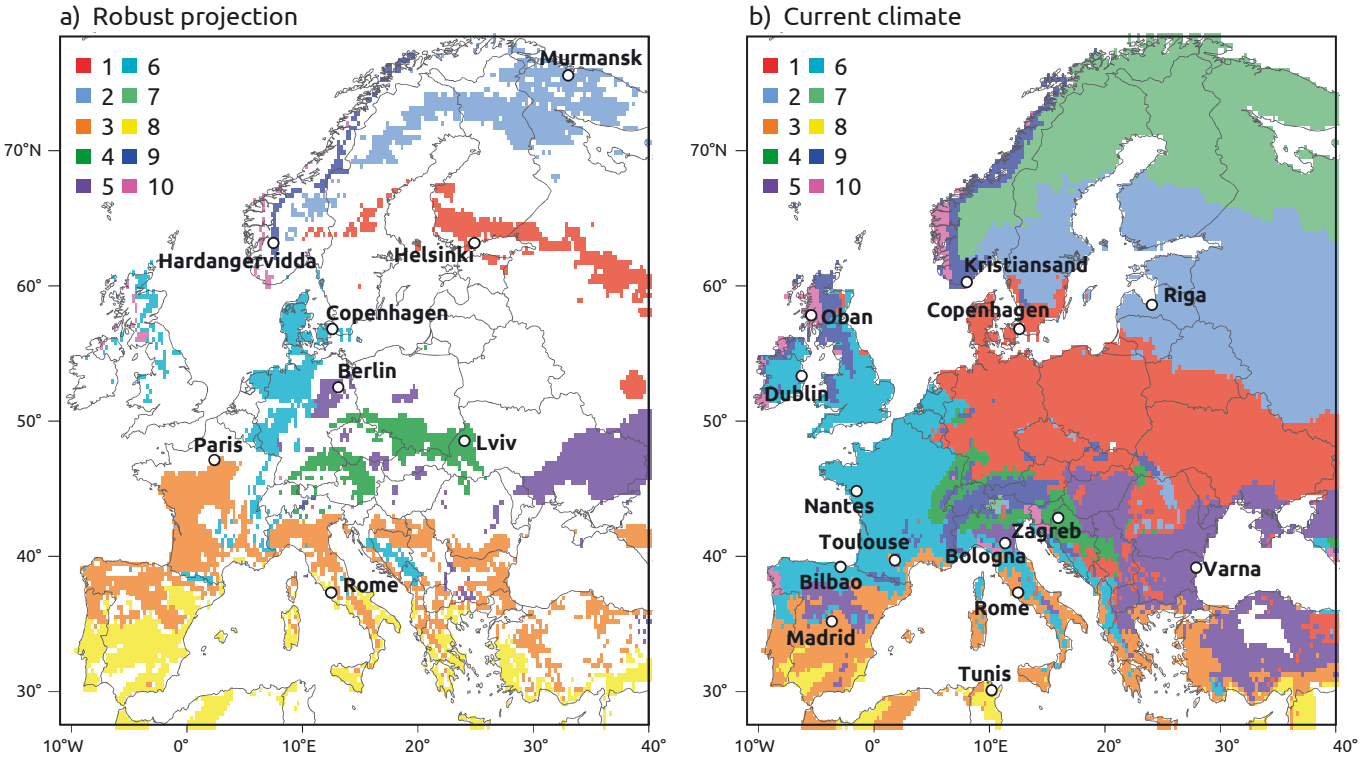

Fig. 2. (a) Ten climate analogue clusters, represented by their robust projections gained by overlapping lower and upper limits of the projections. For each cluster, one representative city is shown. (b) Current climate with a selection of possible robust climate analogue cities in each of the clusters. White grid cells: regions with deviating cluster allocations, overlapping or missing data, or oceans

tions belong to the same cluster (Fig. 2a). In our stylized illustration in climate space (Fig. 1), this intersection corresponds to (L_1, U_1) in cluster I and (L_4, U_4) in cluster IV. As the current climates of grid elements 2 and 3 (C_2 and C_3) are also members of cluster I they are RCAs for grid element 1. In terms of the maps in Fig. 2, this means that for a region A taken from the robust climate projections map (Fig. 2a), the corresponding cluster (color) of the current climate map (Fig. 2b) represents the CA region for the future of A.

Taking the example of Berlin, we find that it is allocated to the violet cluster (5) in the projection map (Fig. 2a). The points in climate space that belong to cluster 5 and describe the current climate (see Fig. 2b) are located in the region north of Madrid (Spain), in the river Po valley (Italy), and many grid cells in eastern Europe (e.g. Hungary, North Serbia, Macedonia, large parts of the coast of the Black Sea, and the higher altitudes of Turkey). With regard to the changes in climate, these grid cells can be seen as the RCAs to Berlin's climatological future. Depending on probable future socioeconomic or environmental features of Berlin, one can now select the most appropriate city within the RCA grid cells. In general this would be cities like Bologna or Toulouse rather than e.g. Varna in Bulgaria. But the de tailed criteria for this choice depend on the focus of the envisaged interpretation. This topic is of relevance in the context the further interpretation of CAs but is beyond the scope of this paper, which focusses only on the climatic aspects.

A look at the other CA clusters discloses more interesting examples of RCAs. We select one city or region from each of the clusters' exemplarily and present respective RCAs (Table 1). For the climatic properties of the clusters, see the feature plots (Kok et al. 2016) in Supplement 3.

Helsinki (Finland) finds its future climate in e.g. Copenhagen (Denmark), which is located in the red cluster (1) for the current climate. Copenhagen's future climate is represented by the cyan cluster (6), covering France, northern Spain, and parts of the Adriatic Sea region in the current climate. Appropriate options within the RCAs of Copenhagen might be Nantes (France) or Bilbao (Spain). The future climate of Paris is situated in the orange cluster (3), a climate pattern which is the second warmest and second driest in summer among the climate change patterns. The currently comparable regions defined by this cluster are the southern parts of the Mediterranean.

Table 1. Overview of the 10 climate analogue clusters, color, example cities for City/region of interest and optional robust climate analogue (RCA), as well as the variance of each cluster calculated by the mean within-cluster sum. NA: not available

Cluster no.	Color	City/Region of interest and example for each cluster	Optional RCA	Variance
1	Red	Helsinki (Finland)	Copenhagen (Denmark)	0.045
2	Light blue	Murmansk (Russia)	Riga (Latvia)	0.043
3	Orange	Paris (France)	Madrid (Spain), Rome (Italy)	0.101
4	Green	Lviv (Ukraine)	Zagreb (Croatia)	0.063
5	Violet	Berlin (Germany)	Bologna (Italy), Toulouse (France)	0.059
6	Cyan	Copenhagen (Denmark)	Nantes (France), Bilbao (Spain)	0.081
7	Turquoise	NA	NA	0.060
8	Yellow	Rome (Italy)	Tunis (Tunesia)	0.083
9	Blue	Hardangervidda (Norway)	Kristiansand (Norway)	0.155
10	Magenta	Strip parallel West Coast (Norway)	Oban (Scotland)	0.342

Plausible RCAs could therefore be Madrid (Spain) or Rome (Italy).

In the projections for the southern parts of the Mediterranean, we find the climate of the yellow cluster (8), with the driest summers and all-year hottest temperatures including the city of Rome (Italy). In the current climate, this cluster comprises regions in Algeria, Tunisia, northern Syria, and the region north-east of Huelva in Spain. So Tunis might be the most appropriate RCA for Rome. The turquoise cluster (7) is situated in the far north (see Fig. 2b) and shows a distinct colder current climate than all other clusters. Due to increasing temperatures, climate patterns typically move further north or to higher altitude, which leaves little space for this climate pattern. Accordingly, it is identified as a 'disappearing climate' within Europe. No grid elements that are robust regarding projection uncertainty and that switched cluster membership in the course of climate change remain by the end of the century. As no turquoise pixels occur in Fig. 2a, there is no location for which the turquoise pixels on the right hand side serve as a RCA. However, in Fig. S3, some 'no change' pixels occur in the North of Sweden around Abisko, depicting that these pixels stay in the same cluster during climate change— thus, in climatological terms, it is not exactly a disappearing climate.

The light blue cluster (2) is the second coldest region according to summer temperatures but not yet under risk of being categorized as a disappearing climate. Large regions of northern Scandinavia will exhibit the former climate of the northeastern Baltic Sea countries. The Russian city of Murmansk can find its RCA in e.g. the capital of Latvia, Riga, which is also a port city with a Soviet history and similar population size.

The blue (9) and the magenta (10) cluster are strongly influenced by the sea and represent the wet climate of the western shorelines. The blue cluster shows robust projections in the mountainous regions of Norway, starting around Hardangervidda and extending northwards. Villages in this sparsely populated region find their current analogues either closer to or at the Norwegian shoreline (e.g. Kristansand) or e.g. in the Scottish Highlands. Pixels with robust projections for the magenta cluster which change cluster membership under climate change include only a narrow strip west of the blue cluster in Norway, also sparsely populated. Villages in this region find their RCAs at the Norwegian, Scottish, and Irish West Coast.

An example for the green cluster (4) is the city of Lviv in Ukraine. RCAs for that town can be found in the lower Alps and northern parts of former Yugoslavia, e.g. Zagreb in Croatia.

Keeping in mind the above RCAs, one may ask if the degree of similarity of the mentioned city pairs is comparable. As the size of a cluster in climate space is a measure for this similarity, we show the cluster variance in Table 1, which is measured by the sum of the squared distances between the objects (grid cells) and their corresponding cluster centers. Summarizing these results, we obtain the 'sharpest' RCAs for Murmansk, Helsinki, Lviv, and Berlin. Intermediate similarity is observed for Copenhagen, Rome, and Paris, whereas the identification of RCAs using cluster 9 and 10 have to be taken with care. On the other hand, no better robust analogues can be found here, given the current climate model uncertainty.

For some cities it was possible to do a comparison between RCAs and the study results of Hallegatte et al. (2007), which used 2 different climate projections for Europe; namely, the HadRM3H model (stronger climate response) and the ARPEGE model (weaker). With regard to the climate indices considered, on average the values of these models lie within the range

of the EURO-CORDEX models used in our study. Therefore, with respect to the applied data, our study is comparable to that of Hallegatte et al. (2007). Further, Hallegatte et al. (2007) used a minimum distance approach for each of the models including externally determined and variable specific thresholds. For Berlin, this study presented 2 options for a CA: the region of Rome (Italy) for a weaker climate response (ARPEGE) and the region of Salamanca (Spain) for a stronger response (HadRM3H). In our study, the violet cluster includes Salamanca, but as regards Rome only the mountainous regions nearby are part of the climate cluster. For Copenhagen, the CAs are identical with the region of Paris and Albania. For Helsinki, Hallegatte et al. (2007) could not find an analogue fulfilling both the temperature and precipitation distance threshold. The regions of Bordeaux (France) and Cordoba (Spain) are CAs for Paris according to Hallegatte et al. (2007). This corresponds to our results for Cordoba but not for Bordeaux. For Rome, Hallegatte et al. (2007) did not find an analogue for both models while in our case Tunis was identified. The results of this comparison are summarized in Supplement 4 and will be discussed in the next section.

4. DISCUSSION AND CONCLUSIONS

The exemplary comparison between analogues found by Hallegatte et al. (2007) and by the method proposed in this paper illustrates relevant differences and commonalities. For Helsinki, the presented method identifies Copenhagen as a RCA while Halle gatte et al. (2007) did not find a CA because the least distances in climate space transgress the rather subjectively determined thresholds for both climate projections. In our method, the cluster result based on data structure and reasonably selected cluster number relates Helsinki to cluster 2 (the red cluster in Fig. 2) which has a relatively small extent in climate space (second smallest variance, see Table 1). This example shows the role of the difference in the determination of minimum distances. Contrasting to this, the determination of RCAs of Copenhagen shows agreement for both methods, i.e. the CAs for both climate projections lie within the RCA region identified by our algorithm. But in the method of Hallegatte et al. (2007) the climatic equivalence under uncertainty of the 2 cities is not recognized and one has to follow both cases for all further interpretations and conclusions. Furthermore, under the existing climate projection, we can choose e.g. Bilbao or Nantes which are certainly closer to the probable socioeconomic future of Copenhagen. The case of Rome shows that due to new climatic conditions the pointwise (e.g. Hallegatte et al. 2007) method failed to produce an analogue. This might be related to the spatial boundaries of the study area: a spatial extension further south could open up analogues also for regions like Rome in case of the existence of an adequate temperature–precipitation relation. Our algorithm found an RCA within the spatial boundaries (Tunis) but one has to consider that this was within the third-largest cluster in climate space (see Table 1) and therefore less significant than e.g. Murmansk-Riga case (half the cluster size). In the case of Berlin and Paris the comparison with the weaker climate signal analogue shows disagreement, which means that the respective analogues do not lie within the RCAs. The reason must be that for these 2 cities the weaker climate response model used by Hallegatte et al. (2007) is outside the (L, U) range defined by the EURO-CORDEX ensemble (see Fig. 1). Therefore this disagreement is caused rather by differences in the input data than the applied method.

In earlier approaches, the degree of similarity was determined by the global minimum in distance space together with rather subjective thresholds for maximum distances which are still acceptable (Hallegatte et al. 2007). These procedures suffer from the difficult justification of these thresholds and the point character of the analogue, which usually do not lead to an urban region with adequate socioeconomic properties. The approach presented here refers to the data structure of the combined climate data set. However, the choice of the adequate number of clusters is often not unique (2, 3, 5, and 10 in our example) but influences the subsequent RCAs regarding the variance within the clusters and, related to that, the size of the regions without a RCA ('uncharted regions' in Fig. 2a). Choosing a smaller number of clusters (e.g. 5) will lead to smaller 'uncharted areas' but larger variance within the CA clusters. The first means that robust statements can be made for a larger fraction of European cities while the latter means that these statements are weaker. Since CAs can only generate a useful guide to adaptation options with an adequate level of differentiation, we chose 10 clusters in the analysis presented here. Be - sides this, the detection of disappearing climates due to northward-moving climate patterns (cluster 7) is in line with earlier studies (e.g. Williams et al. 2007).

Climate projections will always be loaded with uncertainty to a certain extent (Meehl et al. 2007, Hawkins & Sutton 2011). But due to ongoing changes in climate, adaptation efforts are necessary even under this unfavorable data situation. It is therefore important to find and interpret CAs while simultane-

ously addressing the uncertainties, which is possible with our approach. In the case that our algorithm finds no analogue we can still identify the reason: if the upper and lower bound of a pixel's climate projections belong to different clusters, projection uncertainty is too large. If both bounds lie within one cluster which has no current climate members, the city's future climate currently does not exist (within the considered study region).

A further general advantage of the method that has to be acknowledged is that it produces a full and comprehensible picture. While other methodologies always lack the ability to deal with multiple regions of interest (Hallegatte et al. 2007, Williams et al. 2007, Arnbjerg-Nielsen et al. 2015, Hibino et al. 2015), we automatically deliver RCAs for larger areas. Essentially this means that we can display the whole RCA information in 2 maps (see Fig. 2). This enables another application of the method: the reverse usage of an analogue. As the method delivers a comprehensive picture of the study region, a city in the current climate (Fig. 2b) could identify a region in the robust projection map (Fig. 2a) as a potential demand region for a supplier of today's local climate-related applications. For example, a supplier of air humidifiers for public spaces could identify future sales markets. In other words, Helsinki is a RCA of Murmansk but Murmansk is a potential sales market of adaptation methods for Helsinki. Here, of course, the time scale of climate change has to be considered.

It is necessary to extend the set of climate variables when applying this method more specifically to urban areas. In particular, extreme weather event indices have to be included due to the disruptive impact of such weather events on various critical infrastructure and public health. Furthermore, the multi-grid-cell character of many urban areas must be considered. To better meet urban challenges, nearer time horizons and, if possible, higher spatial resolution data should be considered. A follow-up study should implement other RCPs as this study only focusses on RCP8.5. This reflects the stakeholder's perspective which does not distinguish between different kinds of uncertainty and will deliver CAs which are robust towards both model and scenario uncertainty. These different concentration pathways can easily be integrated into the treatment of uncertainty by the method proposed here.

Acknowledgements. This work was funded in part by the Senate of Berlin (AFOK) and the City of Potsdam ('Adaptation to climate change in the State Capital of Potsdam'). We thank Boris Prahl for helpful comments and Alison Schlums for proofreading the manuscript.

LITERATURE CITED

Arnbjerg-Nielsen K, Funder SG, Madsen H (2015) Identifying climate analogues for precipitation extremes for Denmark based on RCM simulations from the ENSEMBLES database. Water Sci Technol 71: 418–425

Dawson J, Scott D, Mcboyle G (2009) Climate change analogue analysis of ski tourism in the northeastern USA. Clim Res 39: 1–9

Dehn M (1999) Application of an analog downscaling technique to the assessment of future landslide activity — a case study in the Italian Alps. Clim Res 13: 103–113

Hallegatte S, Hourcade JC, Ambrosi P (2007) Using climate analogues for assessing climate change economic impacts in urban areas. Clim Change 82: 47–60

Hawkins E, Sutton R (2011) The potential to narrow uncertainty in projections of regional precipitation change. Clim Dyn 37: 407–418

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M (2008) A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. J Geophys Res D Atmospheres 113: D20119

Hibino K, Takayabu I, Nakaegawa T (2015) Objective estimate of future climate analogues projected by an ensemble AGCM experiment under the SRES A1B scenario. Clim Change 131: 677–689

IPCC (2013) Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Jacob D, Petersen J, Eggert B, Alias A and others (2014) EURO-CORDEX: new high-resolution climate change projections for European impact research. Reg Environ Change 14: 563–578

Janssen P, Walther C, Lüdeke M (2012) Cluster analysis to understand socio-ecological systems: a guideline. PIK Report No. 126. Potsdam Institute for Climate Impact Research, Potsdam

Knutti R, Sedlá<ek J (2013) Robustness and uncertainties in the new CMIP5 climate model projections. Nat Clim Chang 3: 369–373

Kok M, Lüdeke M, Lucas P, Sterzel T and others (2016) A new method for analysing socio-ecological patterns of vulnerability. Reg Environ Change 16: 229–243

Kopf S, Ha-Duong M, Hallegatte S (2008) Using maps of city analogues to display and interpret climate change scenarios and their uncertainty. Nat Hazards Earth Syst Sci 8: 905–918

Kottek M, Grieser J, Beck C, Rudolf B, Rubel F (2006) World Map of the Köppen-Geiger climate classification updated. Meteorologische Z 15: 259–263

Krysanova V, Hattermann F (2017) Intercomparison for climate change impacts in 12 large river basins; overview of methods and summary of results. Clim Change 141: 363–379

Lobell DB, Burke MB, Tebaldi C, Mastrandrea MD, Falcon WP, Naylor RL (2008) Prioritizing climate change adaptation needs for food security in 2030. Science 319: 607–610

Meehl GA, Stocker TF, Collins WD, Friedlingstein P and others (2007) Global climate projections. In: Solomon S, Qin D, Manning M, Chen Z and others (eds) Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, p 747–845

Mote PW, Salathé EP (2010) Future climate in the Pacific Northwest. Clim Change 102: 29–50

Nassopoulos H, Dumas P, Hallegatte S (2012) Adaptation to an uncertain climate change: cost benefit analysis and robust decision making for dam dimensioning. Clim Change 114: 497–508

Parry ML, Carter TR (1989) An assessment of the effects of climatic change on agriculture. Clim Change 15: 95–116

Pickett STA (1989) Space-for-time substitution as an alternative to long-term studies. In: Likens GE (ed) Long-term studies in ecology. Springer, New York, NY, p 110–135

Pielke R Jr, Prins G, Rayner S, Sarewitz D (2007) Lifting the taboo on adaptation. Nature 445: 597–598

Pugh TAM, Müller C, Elliott J, Deryng D and others (2016) Climate analogues suggest limited potential for intensification of production on current croplands under climate change. Nat Commun 7: 12608

Rastetter EB (1996) Validating models of ecosystem response to global change: How can we best assess models of long-term global change? Bioscience 46: 190–198

Reckien D, Flacke J, Olazabal M, Heidrich O (2015) The influence of drivers and barriers on urban adaptation and mitigation plans — an empirical analysis of European Cities. PLOS ONE 10: e0135597

Rosenzweig C, Arnell NW, Ebi KL, Lotze-Campen H and others (2017) Assessing inter-sectoral climate change risks: the role of ISIMIP. Environ Res Lett 12: 010301

Sietz D, Lüdeke MKB, Walther C (2011) Categorisation of typical vulnerability patterns in global drylands. Glob Environ Change 21: 431–440

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 93: 485–498

Tebaldi C, Knutti R (2007) The use of the multi-model ensemble in probabilistic climate projections. Phil Trans R Soc A 365: 2053–2075

Williams JW, Jackson ST, Kutzbach JE (2007) Projected distributions of novel and disappearing climates by 2100 AD. Proc Natl Acad Sci USA 104: 5738–5742