

# Selecting climate simulations for impact studies based on multivariate patterns of climate change

Thomas Mendlik<sup>1</sup> · Andreas Gobiet<sup>1</sup>

Received: 20 August 2014 / Accepted: 13 December 2015 / Published online: 24 December 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** In climate change impact research it is crucial to carefully select the meteorological input for impact models. We present a method for model selection that enables the user to shrink the ensemble to a few representative members, conserving the model spread and accounting for model similarity. This is done in three steps: First, using principal component analysis for a multitude of meteorological parameters, to find common patterns of climate change within the multi-model ensemble. Second, detecting model similarities with regard to these multivariate patterns using cluster analysis. And third, sampling models from each cluster, to generate a subset of representative simulations. We present an application based on the ENSEMBLES regional multi-model ensemble with the aim to provide input for a variety of climate impact studies. We find that the two most dominant patterns of climate change relate to temperature and humidity patterns. The ensemble can be reduced from 25 to 5 simulations while still maintaining its essential characteristics. Having such a representative subset of simulations reduces computational costs for climate impact modeling and enhances the quality of the ensemble at the same time, as it prevents double-counting of dependent simulations that would lead to biased statistics.

## 1 Introduction

Currently, an increasing societal demand to anticipate the future impacts and costs of climate change challenges the climate research community, since projections of climate change impacts should be based on a robust and reliable climatological input. To estimate

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s10584-015-1582-0) contains supplementary material, which is available to authorized users.

---

✉ Thomas Mendlik  
thomas.mendlik@uni-graz.at

<sup>1</sup> Wegener Center for Climate and Global Change, University of Graz, Brandhofgasse 5, 8010 Graz, Austria

the bandwidth of possible future impacts, a balanced and unbiased estimate of the entire distribution of possible future changes is required. This bandwidth is often estimated by driving impact models with few selected climate scenarios, the selection being largely subjective or determined by practical reasoning such as availability. This work presents a method for selecting the climate input for climate change impact assessments in a more objective way, which aims to avoid sampling biases, to properly account for uncertainties and to save computational resources.

The most detailed information on future climate is provided by General Circulation Models (GCMs), often refined with regional climate models (RCMs) and empirical-statistical post-processing methods (e.g. Maraun 2013; Themeßl et al. 2011). However, as with any future scenario, they are subject to considerable uncertainties (e.g. Tebaldi and Knutti 2007) originating from the chaotic behavior of the climate system and the unknown future evolution of greenhouse gas concentrations and other forcing agents of the climate system, as well as simplifications and errors in climate models. Those inherent uncertainties are often investigated using multi-model ensembles (MMEs), which challenges climate change impact assessments to base their investigations on multi-model climatological input.

Assuming an unbiased MME, the model selection should conserve the statistical properties of the original MME as far as possible. A further complication arises if MMEs cannot be regarded as unbiased, which is usually due to sampling and model interdependence issues, as discussed in the next section. A sensible selection of climate simulations as input for climate change impact studies is needed in any case, either to limit computational demand and/or to mitigate biases in the ensemble statistics. Currently, such selection is often done “by opportunity” based on the ease of access to climate simulations or by subjective criteria. Only few systematic methods for model selection have been published so far (Smith and Hulme 1998; Knutti et al. 2010a; McSweeney et al. 2012; Whetton et al. 2012; Evans et al. 2013; Cannon 2015). Our study adds to this field by describing a multivariate method to select representative RCM or GCM simulations from a larger MME for impact investigations based on cluster analysis. The basic aim of the method is to provide a flexible tool for model selection, that can easily be adapted to different applications in climate change impact research. The method mitigates sampling and interdependence-biases, and effectively reduces the ensemble size with a minimal loss of information.

## 2 Design of multi-model ensembles

According to Masson and Knutti (2011) the goal of ensemble design should be to maximize model diversity in order to capture model uncertainty properly while ensuring good model performance. The same criteria are valid for the selection of a sub-ensemble from a larger ensemble for applications in climate change impact research.

As MMEs usually do not systematically sample components of model uncertainty (e.g. parameterizations), and thus do not stem from an experimental design in a statistical sense, they cannot be expected to represent unbiased distributions of possible future climate states (Knutti et al. 2010b). For example, interdependence between models may induce biases, since interdependent models gain too much weight in ensemble statistics if they are counted as independent ones. Studies by Pennell and Reichler (2010) and Masson and Knutti (2011) showed that GCM ensembles feature considerable model dependence, leading to a smaller effective ensemble size than the number of models in the ensemble. Such dependencies can lead to biased estimations of both mean and width of the distribution.

This problem is even worse in GCM-RCM MMEs, where an additional layer of uncertainty is introduced by nesting regional models into global models. Frequently some GCMs are used to provide boundary conditions for several RCMs in the ensemble, while others are used only once. Similar boundary conditions generate substantial interdependencies between the results of RCMs, leading to an unbalanced ensemble. Some methods to mitigate this problem by statistically reconstructing RCM simulations in order to obtain an ensemble in which each GCM-RCM combination appears exactly once are described in the literature (e.g. Déqué et al. 2011; Heinrich et al. 2014).

A somewhat different approach to obtain a balanced ensemble is suggested by Whetton et al. (2012). They recommend shrinking the ensemble to a set of representative simulations that capture certain characteristics of the whole sample. This subset should then be used as a consistent forcing for various impact models. Methods to conduct such a selection are discussed in the next section.

### 3 Review of existing selection methods

For climate impact modelers dealing with climate simulations, certain objective criteria need to be fulfilled in order to make a smooth study possible. Such criteria include model performance in the past, spread of climate change signals and independence.

One of the first published approaches tackling GCM model selection with formal criteria stems from Smith and Hulme (1998). They propose several criteria such as vintage (considering the latest generation of climate simulations only), resolution (the higher the resolution, the better), validity (model performance in the past), and representativeness (picking simulations from the high and low end of the range of climate change signals of temperature and precipitation to obtain a representative sub-sample). This method has been adopted by the IPCC guidelines for climate scenarios (IPCC-TGICA 2007). Such a selection of GCMs has been applied by e.g. Murdock and Spittlehouse (2011) focusing on the region of British Columbia by analyzing the models based on the spread of change in temperature and precipitation. A discussion of sub-selecting climate simulations for hydrology studies has been published by Salathe et al. (2007). They propose sampling driving climate simulations by considering the projected model spread for hydrology-relevant parameters (temperature and precipitation change) to find groups of similar simulations and selecting representative climate models. A generalization to a multivariate setup has recently been presented by Cannon (2015). His proposed method maximizes model diversity by selecting the most extreme simulations. All those studies have a non-probability sampling scheme in common: instead of assigning probabilities to the simulations and sampling them randomly, the selection is based on qualitative characteristics which are relevant for the researcher (Mays and Pope 1995). The aim is to maximize diversity of these characteristics.

The Good Practice Guidance Paper on Assessing and Combining Multi Model Climate Projections (Knutti et al. 2010a) gives some more recent recommendations for model selection, also addressing the issue of model dependence. Knutti et al. (2010a) argue that agreement between models may arise due to the fact that models use similar simplifications and may feature similar errors. This means that models do not represent independent information and should be down-weighted in order to avoid biases in the statistical analysis of the ensemble, which are induced by double-counting similar models (Pirtle et al. 2010). Model selection can be regarded as a binary 0–1 weighting that should address these issues. Several impact studies address this problem of double-counting (e.g. Finger et al. 2012). Evans

et al. (2013) presents a selection method taking into account model performance and independence in climate change signals. This method selects models that are most independent from the rest of the entire ensemble.

In the literature, models are often selected based only on their performance in the past, without regarding spread in the climate change signals, with the aim to use only the “best” models. However, correlations between past performance and future climate change signals are known to be very weak (Knutti et al. 2010b), which means that there is no clear indication that the best performing models in the past are most realistic with regard to climate change signal. In addition, the ranking of models with regard to performance in the past is highly dependent on the definition of the performance measure (e.g. Jury et al. 2015), which leads to a very subjective ranking. Therefore it seems reasonable that model performance in the past should rather be used to detect and remove severely unrealistic models that cannot be trusted in their future projections for some clearly argued reasons, but not to select a few “best performing” models, since there is no indication that they are more realistic in their future projections than other reasonably performing models.

This leads to a further model selection criterion, namely the conservation of statistical properties of the climate change signals: the sub-sample of the selected simulations should properly represent uncertainties. Recently, methods have been published based on this idea, partly combined with some pre-selection based on model performance (e.g. McSweeney et al. 2012; Bishop and Abramowitz 2013).

## 4 Method

The following model selection method samples one representative climate simulation out of groups of models with similar characteristics, to obtain a sub-set of independent simulations that cover the multi-model ensemble spread. These groups are found using clustering techniques. The model spread and the similarity measure can be defined using an arbitrary number of climate parameters and indicators. In addition, multiple spatial regions and seasons of interest can be freely defined. Therefore, this method is not limited to the commonly used temperature and precipitation changes of a single region, but is rather a multivariate extension.

Having such a complex set-up, it is necessary to decrease the dimensionality of the climate parameters to eliminate collinearities and to reduce random noise. This is done by using Principal Component Analysis (PCA) to identify *patterns of climate change* as step (1) (Jolliffe 2002). Step (2) finds model similarities with a hierarchical clustering algorithm (Huth et al. 2008) and finally, step (3) involves sampling the simulations from each cluster detected. We assume that unrealistic simulations have been sorted out in advance of the study. The next sections explain these steps in more detail.

### 4.1 Common patterns of climate change: PCA

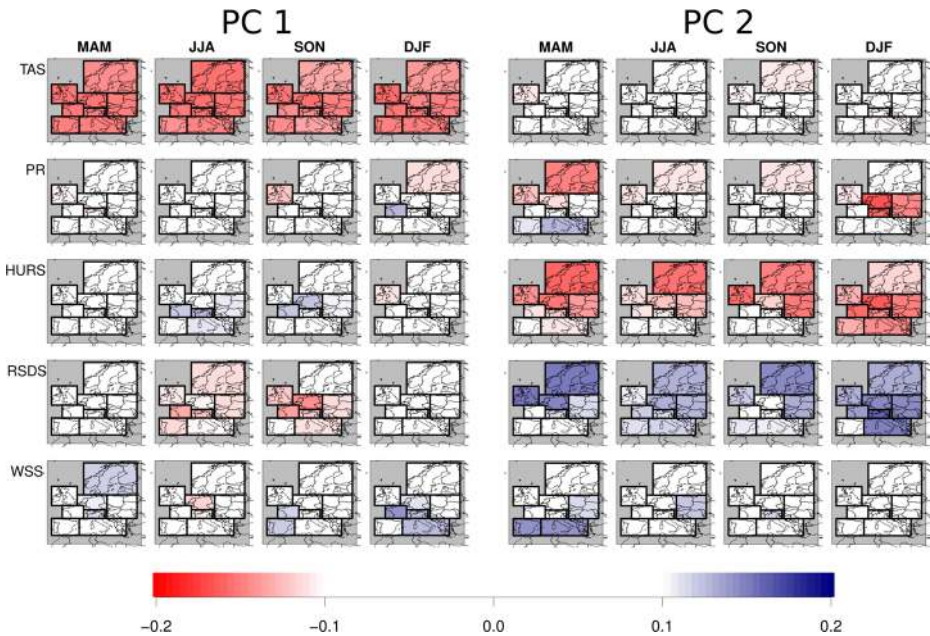
First, all area-averaged climate change signals of the meteorological parameters (like temperature and precipitation) have to be standardized, as different variables may have different units and variabilities. Then, with a PCA for each simulation, we transform the climate change signals to a linear combination of those variables. Those transformed meteorological variables are formally called principal components (PCs) and they form the *common*

patterns of climate change (see Fig. 1). The transformations, which are the coefficients of the linear combinations, are called *loadings* and they describe which meteorological variables are combined to form a particular pattern of climate change. The PCs of each simulation (i.e. patterns of climate change) are treated in the same way as meteorological variables, but they differ in being stochastically independent of each other, which is necessary for the subsequent cluster analysis.

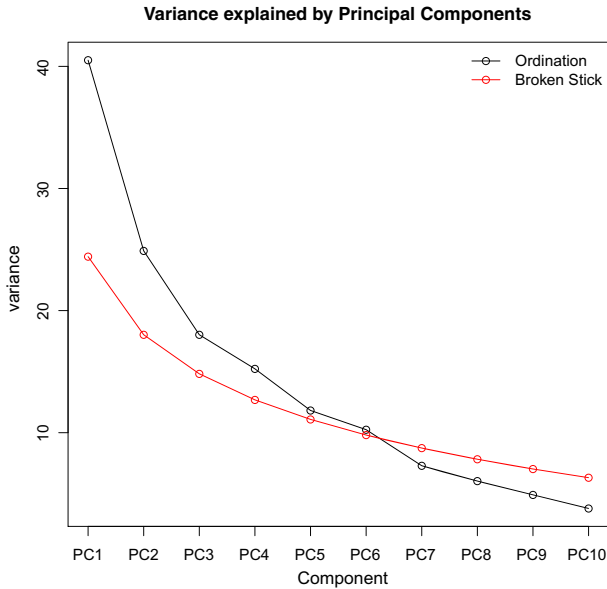
In the next step the most dominant patterns of climate change signals have to be detected in order to reduce noise and make the subsequent cluster analysis more robust. We use the broken-stick method (Jolliffe 2002), which compares the variances of individual PCs of the used dataset with the theoretical variances of a randomly generated dataset. If those random variances are equal or larger than the observed ones, the corresponding PC can be regarded as noise and should be excluded (Fig. 2).

### 4.2 Model similarity: cluster analysis

The aim here is to find groups of simulations based on their behavior with regard to the common climate change patterns obtained from the PCA. Those groups of simulations are found based on a hierarchical clustering algorithm which works like this: first, each simulation is assigned to its own cluster, and then the algorithm proceeds by iteratively joining the two closest clusters in each agglomeration step until one cluster remains. The measure of distance is based on Ward’s criterion, which finds new clusters with minimal variance. This



**Fig. 1** The coefficients of the linear combinations (loadings) yielding the first 2 dominant patterns of climate change (PCs) of the ENSEMBLES RCMs across Europe. *Blue boxes* indicate increase and *red boxes* decrease of the corresponding parameter



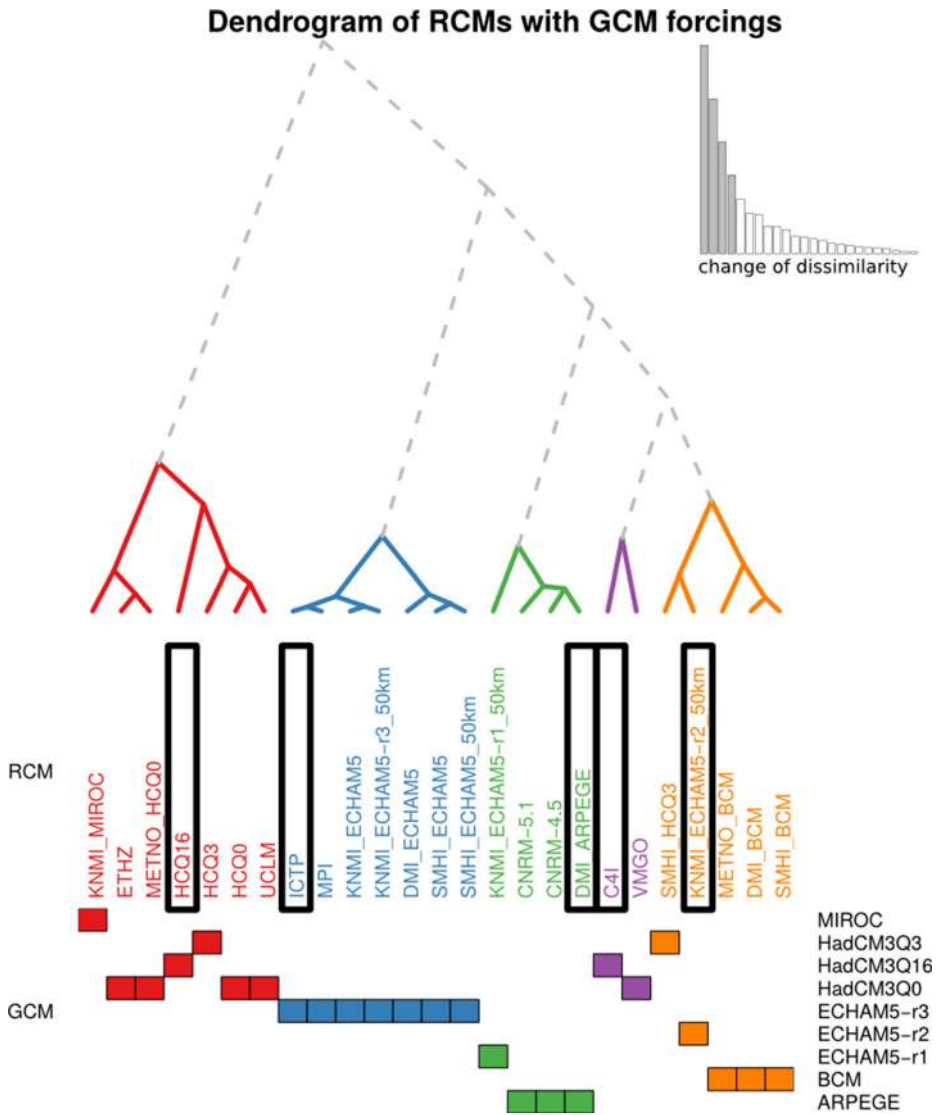
**Fig. 2** Scree plot depicting the variances explained by the individual PCs (patterns of climate change, *black line*). The *red line* represents the variances which would occur in a randomly generated dataset with same dimensionality. PCs with variances close to this red line can be regarded as noise and are excluded from the analysis

procedure tends to find compact and spherical groups of data. Hierarchical clustering results in a tree-like similarity structure, which is meaningful if we believe that some clusters might be more closely related to other clusters.

Having obtained a tree-like structure of the dependence of the simulations (Fig. 3), the open question of how many simulations to actually select from the ensemble still remains. There is no unique and best solution to this problem, but there are some criteria for how to obtain an optimal number of clusters. Our approach is to consider the distance criterion for each agglomeration step, which in the case of Ward's criterion is the variance increase of the newly merged two clusters. We cut the tree where this increase of variance does not change considerably (Fig. 3).

### 4.3 Model selection: sampling

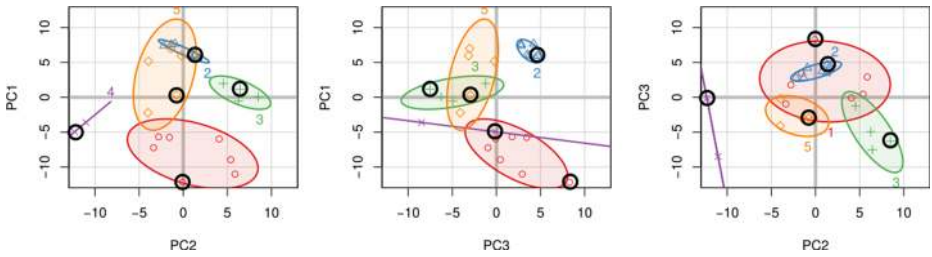
We extend the idea of non-probability sampling (see Section 3) by sampling one simulation out of each group of similar models obtained from the cluster analysis. This approach is also known as quota sampling, where one selects members out of each group using key information or characteristics relevant to the phenomenon being studied. This can be done by picking simulations from the scatter plot of the PCs (Fig. 4). Another way would be to look at the distribution of the PCs for each simulation individually, which can be displayed with a bar-chart denoting their location within the scatter plots (Fig. 5). One starts with an average simulation where all bars/PCs are close to 0 and then selects one simulation with distinct extreme characteristics from each cluster.



**Fig. 3** Cluster dendrogram for the first 3 PCs showing 5 clusters. The boxes on the bottom show the driving GCM of the corresponding RCM. The barplot on the top right shows the dissimilarity measure for each agglomeration step when merging the two closest clusters. The largest 4 dissimilarities, leading to this 5 cluster partition, are marked *gray*

### 5 Case study: model selection for the European domain

We apply the methods explained above within an example case study to select regional climate models for climate impact studies. The study is motivated by the EU-FP7 project IMPACT2C (e.g. Vautard et al. 2014) to seek driving data for multiple diverse impact

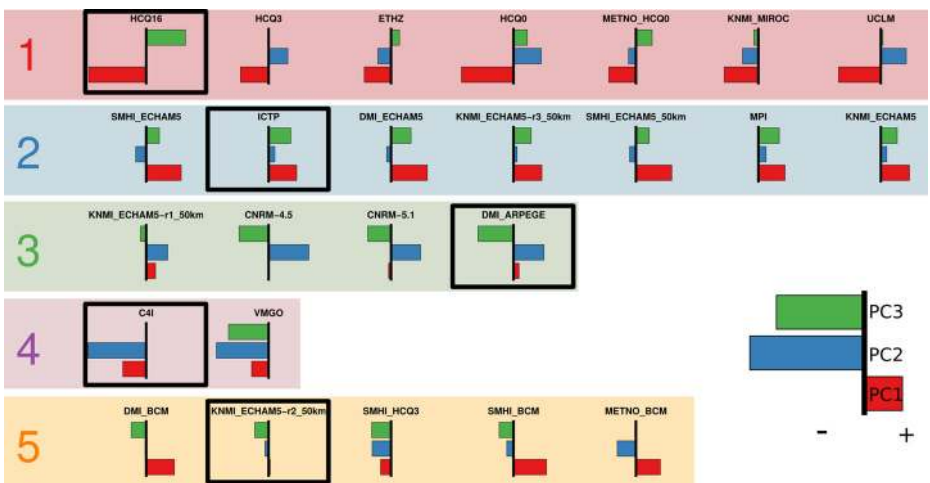


**Fig. 4** Climate change signals of the ENSEMBLES RCMs within the principal component space, showing the 75 % variance ellipsoid within each cluster. The selected models are highlighted

models spread across the whole European continent. The large number of different impact modeling groups drastically increases the number of meteorological variables that need to be considered. The analysis for this example, including the R code and the data, can be found in the [Electronic supplementary material](#) online.

**5.1 Data**

The MME used here is the multi-model dataset from the ENSEMBLES project. The 27 ENSEMBLES regional climate model simulations are driven by 6 different GCMs which are all forced by the SRES A1B emission scenario (Nakicenovic et al. 2000). The *ECHAM5* GCM appears three times using different initializations and the *HadCM3* GCM also shows up with three different parameterization schemes (*Q0*, *Q3* and *Q16*). One RCM (*KNMI-RACMO*) has been forced by all three *ECHAM5* realizations, but with a coarser horizontal resolution of 50 km. One other simulation has been driven with this resolution as well, while the remaining 23 simulations have a 25 km resolution. As one simulation (*GKSS-CCLM4.8-IPSL*) lacks the variables *HURS* and *RSDS*, it has been omitted in this study. In addition one simulation (*OURANOS-CRCM-CGCM3*) shows very noticeable biases and has been excluded as well, leaving  $n = 25$  regional climate simulations for the analysis. The baseline



**Fig. 5** Climate change signals for each RCM simulation in the principal component space. The simulations are split according to their clustering. The suggested selections are highlighted



period for the climate change signal (CCS) is the 30 years average of 1971 to 2000. The future scenario period to determine the climate change signal is chosen to be 2021 to 2050.

The most important meteorological drivers for climate change impacts in the European study have been defined on the basis of a user survey among project partners, and experience from previous projects. In total,  $p_{par} = 5$  parameters have been selected: mean air temperature (*TAS*), precipitation amount (*PR*), relative humidity (*HURS*), global radiation (*RSDS*) and wind speed (*WSS*). The climate change signals of these variables are standardized and analyzed in subregions of Europe (as in Christensen and Christensen 2007) by aggregating spatially over  $p_{spat} = 8$  domains for the  $p_{seas} = 4$  seasons summer (JJA), winter (DJF), spring (MAM) and autumn (SON). All subregions are equally weighted, however a weighting scheme according to each area would be possible as well. In total, this gives 160 different parameters. We obtained the climate change signals using the R package *wux* (Mendlik et al. 2015).

## 5.2 Common patterns of climate change

We use PCA to reduce the dimensionality  $p = p_{par} \times p_{seas} \times p_{spat} = 5 \cdot 4 \cdot 8 = 160$  of the  $n = 25$  climate models. The broken-stick method detects 3–4 robust PCs (Fig. 2), excluding the remaining PCs as being random noise. We decided to reduce the dimensionality to  $p_{red} = 3$  PCs.

The most dominant climate change pattern (PC1) is largely linked to the temperature change along all four seasons for all European subregions. PC2 shows a negative relationship between relative humidity (HURS) and global radiation (RSDS). This means that simulations projecting a higher change in HURS than others tend to project a lower change in RSDS. This anti-correlation seems to hold for the entire European region in winter (DJF) and for the northern and eastern parts of Europe in the remaining seasons, especially in spring (MAM) and autumn (SON). A positive correlation of humidity and precipitation can be detected for the Scandinavian region over the whole year and in winter for mid- and eastern Europe and the Alpine region. PC3 shows a humidity-precipitation pattern for the southern regions for MAM (not shown).

## 5.3 Model similarity

Based on the first  $p_{red} = 3$  PCs we performed a hierarchical cluster analysis as described in Section 4.2.

The tree-like dependency structure is visualized by a dendrogram in Fig. 3. The height of the branches depicts the measure of dissimilarity between simulations and clusters with regard to the common patterns of climate change. To detect the optimal number of clusters, we observe the dissimilarity measure at each step when incrementally merging the closest pair of clusters into a single one. Before the dissimilarity significantly increases, we stop merging to obtain a set of dissimilar clusters each obtaining similar simulations. The dissimilarity measure at each increment is shown as a bar plot in Fig. 3.

We show partitions with 5 clusters to visualize the range of reasonable clustering. Notably, simulations driven by the lateral boundary conditions of the GCMs *ECHAM5-r3*, *BCM* and *ARPEGE* show very strong GCM-specific clustering, meaning that those RCMs driven by the same GCM behave rather similarly with regard to the common patterns of climate change. Further, *ECHAM5-r3* and *BCM*-driven simulations tend to be more similar than *ARPEGE* models. Interestingly, the 50 km versions of the RCM *KNMI-RACMO* driven by *ECHAM5-r1* and by *ECHAM5-r2* behave rather differently than the *ECHAM5-r3*-

driven version and are spread among different clusters. Also forced with *ECHAM5-r3*, the simulations *KNMI-RACMO* and *SMHI-RCA* both have identical set-ups using both a 25 km and a 50 km resolution. In each case, the similarity is very high. On the other hand, RCMs driven by the three *HadCM3* GCMs show quite some heterogeneity: On the one hand they are split among two clusters of different sizes. On the other hand, RCMs driven by different *HadCM3* GCMs can be found in either cluster. Also, the *MIROC*-driven *KNMI-RACMO* does not form a cluster of its own but behaves similarly to *HadCM3*-driven RCMs.

## 5.4 Model selection

Our model selection approach identifies 5 groups of similar simulations. As shown in Fig. 3, these groups also show dependencies, some more than others, but by far not as strong as between the individual simulations. By selecting one simulation from each cluster, we definitely reduced the model dependency and obtained a more independent ensemble. We decided to select an average climate simulation and 4 extreme simulations to span the uncertainty range.

Figure 4 depicts the climate change signals of the regional climate simulations in the principal component space of the first three PCs. Simulations close to 0 can be interpreted as having an “average pattern” of the climate change induced by the corresponding principal component. The sign and order of magnitude in the scatter plot (Fig. 4) corresponds to the pattern described in the corresponding PC from the loadings plot in Fig. 1. For PC1 (warming pattern), simulations within cluster 1 show the highest changes, whereas clusters 2 and 5 tend to have cooler projections and cluster 3 is average. For PC2 (humidity pattern) cluster 4 and cluster 3 show distinct behavior. PC3 (precipitation and humidity pattern) mostly distinguishes between clusters 2, 3 and 5.

The individual locations within the scatter plots can be visualized with barplots (Fig. 5), which makes sampling easier. We started by choosing one average simulation closest to 0 along all PC axes. Then, out of each cluster, we took one extreme representative to obtain maximum diversity of our sub-sample. One possible model selection could be the following: *KNMI-ECHAM5-r2-50 km* (average behavior), *C4I-HadCM3Q16* (low PC2), *DMI-ARPEGE* (high PC2, low PC3), *ICTP-ECHAM5-r3* (high PC1, high PC3) and *HCQ16-HadCM3Q16* (low PC1, high PC3). This selection is marked in Figs. 3, 4 and 5.

Here, some driving GCMs appear two times, such as *ECHAM5* and *HadCM3Q16*, as the corresponding RCMs project very different patterns of climate change. However, other constellations capturing the extreme characteristics are also possible.

## 6 Summary and discussion

In order to provide sound meteorological input for climate change impact studies, it is important to address the uncertainty induced by climate simulations. We present a simple tool to aid users in selection of appropriate climate simulations (either GCMs or RCMs) as input for their studies. The aim of the proposed method is to sample the climate model uncertainty, find model similarities, and sub-select models that are as independent as possible while conserving the spread of the full ensemble.

Our method generalizes the pragmatic approach of finding the model spread of climate change signals of, say, temperature against precipitation (IPCC-TGICA 2007). It allows for simultaneous analysis of an arbitrary number of meteorological parameters over several spatial regions of interest, and brings forth dominant patterns of climate change. Model

similarities are detected based entirely on those patterns of projected change. This stands in contrast to most other studies, which find similarities in the 20th century historical runs (e.g. Abramowitz and Gupta 2008; Pennell and Reichler 2010; Bishop and Abramowitz 2013). An interesting aspect for further research would be the question of how model similarities in historical runs and dependencies in future projections relate over time.

The sub-ensemble of selected simulations conserves the main climate change characteristics of the entire ensemble, but it might not share identical statistical properties like mean and standard deviation. This is a desirable property, as unbalanced ensemble designs often lead to biased estimates due to double-counting induced by model dependencies. For balanced and thus unbiased ensembles, like for reconstructed datasets (Heinrich et al. 2014), the statistical properties are conserved, as the sample size is decreased equally in each cluster.

We do not discuss model selection based on performance, as different models show different strengths and weaknesses depending on the metric (Knutti et al. 2010a). We tend to the pragmatic approach of excluding only simulations with severe and clearly demonstrated deficiencies and keeping as many simulations as possible as input for the model selection procedure.

It should be noted that the proposed method does not deliver one single and unique subset. Instead the user has to decide on how to select one simulation out of each cluster. This can be done with probabilistic (random) sampling or non-probabilistic sampling. We do not recommend any type of random sampling as it is vulnerable to random sampling error: the randomness of the selection can result in a subset which is not representative for the ensemble. The probability of such a misspecification increases with decreasing sample size. For such small sample sizes it is more advisable to take most extreme simulations so as to sample the entire model spread.

Cannon (2015) proposes a very interesting alternative model selection algorithm, addressing the same problems as are presented in this work (multivariate set-up and model dependency). However, in contrast to the method proposed in this work, in Cannon's method the selected models are uniquely identified. This surely makes model selection simpler, but there is no flexibility for the user to add some subjective selection criteria when sampling, like the inclusion of an extremely well-performing simulation.

We demonstrate the presented model selection procedure with the ENSEMBLES multi-model dataset (Section 5). Our results show that the first two most dominant patterns of change relate to temperature and humidity and that the dataset can be split into 5 groups of similar simulations. An important factor for model similarity in this setting is the GCM forcing of the RCM. Interestingly, some GCM forcings lead to very dense clusters (*ECHAM5-r3*), while others are very heterogeneous and may even be split among different groups of similarity. This is particularly the case for the different initial conditions of *ECHAM5* (*r1*, *r2* and *r3*), each of which induces a distinct behavior of the RCMs. On the other hand, some driving GCMs do not create their own clusters at all (e.g. *MIROC*). In our example application this leads to a selection where two GCMs appear twice, while others are omitted entirely. Selecting simulations from each GCM would not necessarily span the entire uncertainty range.

Note that our method is not restricted to the selection of RCMs, and can also be used to select suitable GCMs.

In summary, we present a flexible method to select models from an ensemble of simulations conserving the model spread and accounting for model similarity. This method reduces computational costs for climate impact modeling and enhances the quality of the ensemble at the same time, as it prevents double-counting of dependent simulations that would lead to biased statistics.

**Acknowledgments** We acknowledge the ENSEMBLES project, funded by the European Commission's 6th Framework Programme through contract GOCE-CT-2003-505539. We would like to thank the EU FP7 project IMPACT2C (grant agreement 282746) for funding this research as well as the FWF project (I 1295-N29) "Undersanding contrasts in high mountain hydrology in Asia" (UNCOMUN). Thanks to Erik van Meijgaard for providing us with additional simulations. Also many thanks go to Georg Heinrich, Renate Wilcke and Ernst Stadlober for fruitful discussions and support as well to Barbara Puchinger for proof-reading the manuscript. We also appreciate the detailed inputs from the anonymous reviewers, which greatly improved the quality of this work.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Abramowitz G, Gupta H (2008) Toward a model space and model independence metric. *Geophys Res Lett* 35. doi:[10.1029/2007GL032834](https://doi.org/10.1029/2007GL032834)
- Bishop CH, Abramowitz G (2013) Climate model dependence and the replicate Earth paradigm. *Clim Dyn* 42(3–4):885–900. doi:[10.1007/s00382-012-1610-y](https://doi.org/10.1007/s00382-012-1610-y)
- Cannon AJ (2015) Selecting GCM scenarios that span the range of changes in a multimodel ensemble: application to CMIP5 climate extremes indices. *J Clim* 28(3):1260–1267. doi:[10.1175/JCLI-D-14-00636.1](https://doi.org/10.1175/JCLI-D-14-00636.1)
- Christensen JH, Christensen OB (2007) A summary of the PRUDENCE model projections of changes in European climate by the end of this century. *Clim Chang* 81(1):7–30. doi:[10.1007/s10584-006-9210-7](https://doi.org/10.1007/s10584-006-9210-7)
- Déqué M, Somot S, Sanchez-Gomez E, Goodess CM, Jacob D, Lenderink G, Christensen OB (2011) The spread amongst ENSEMBLES regional scenarios: regional climate models, driving general circulation models and interannual variability. *Clim Dyn* 38(5–6):951–964. doi:[10.1007/s00382-011-1053-x](https://doi.org/10.1007/s00382-011-1053-x)
- Evans JP, Ji F, Lee C, Smith P, Argüeso D, Fita L (2013) A regional climate modelling projection ensemble experiment—NARCLIM. *Geosci Model Dev Discuss* 6(3):5117–5139. doi:[10.5194/gmdd-6-5117-2013](https://doi.org/10.5194/gmdd-6-5117-2013)
- Finger D, Heinrich G, Gobiet A, Bauder A (2012) Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century. *Water Resour Res* 48(2). doi:[10.1029/2011WR010733](https://doi.org/10.1029/2011WR010733)
- Heinrich G, Gobiet A, Mendlik T (2014) Extended regional climate model projections for Europe until the mid-twentyfirst century: combining ENSEMBLES and CMIP3. *Clim Dyn* 42(1–2):521–535
- Huth R, Beck C, Philipp A, Demuzere M, Ustrnul Z, Cahynová M, Kyselý J, Tveit OE (2008) Classifications of atmospheric circulation patterns: recent advances and applications. *Ann New York Acad Sci* 1146(1):105–152
- IPCC-TGICA A (2007) General guidelines on the use of scenario data for climate impact and adaptation assessment. version 2. Tech. rep.
- Jolliffe IT (2002) *Principal component analysis*, 2nd edn. Springer, New York
- Jury MW, Prein AF, Truhetz H, Gobiet A (2015) Evaluation of CMIP5 models in the context of dynamical downscaling over Europe. *J Clim* 28(14):5575–5582. doi:[10.1175/JCLI-D-14-00430.1](https://doi.org/10.1175/JCLI-D-14-00430.1)
- Knutti R, Abramowitz G, Collins M, Eyring V, Gleckler PJ, Hewitson B, Mearns L (2010a) Good practice guidance paper on assessing and combining multi model climate projections. In: Stocker T, Qin D, Plattner GK, Tignor M, Midgley P (eds) *IPCC Expert meeting on assessing and combining multi model climate projections*, p 1
- Knutti R, Furrer R, Tebaldi C, Cermak J, Meehl GA (2010b) Challenges in combining projections from multiple climate models. *J Clim* 23. doi:[10.1175/2009JCLI3361.1](https://doi.org/10.1175/2009JCLI3361.1)
- Maraun D (2013) Bias correction, quantile mapping, and downscaling: revisiting the inflation issue. *J Clim* 26(6):2137–2143
- Masson D, Knutti R (2011) Climate model genealogy. *Geophys Res Lett* 38(8). doi:[10.1029/2011GL046864](https://doi.org/10.1029/2011GL046864)
- Mays N, Pope C (1995) Rigour and qualitative research. *BMJ: Br Med J* 311(6997):109–112
- McSweeney CF, Jones RG, Booth BBB (2012) Selecting ensemble members to provide regional climate change information. *J Clim* 25(20):7100–7121. doi:[10.1175/jcli-d-11-00526.1](https://doi.org/10.1175/jcli-d-11-00526.1)
- Mendlik T, Heinrich G, Leuprecht A (2015) wux: Wegener Center Climate Uncertainty Explorer. R package version 2.1-0.

- Murdock T, Spittlehouse D (2011) Selecting and using climate change scenarios for british columbia. Tech. rep., Pacific Climate Impacts Consortium. University of Victoria, Victoria
- Nakicenovic N, Alcamo J, Davis G (2000) IPCC special report on emissions scenarios. Cambridge University Press, Cambridge, United Kingdom and New York
- Pennell C, Reichler T (2010) On the effective number of climate models. *J Clim* 24(9):2358–2367. doi:[10.1175/2010JCLI3814.1](https://doi.org/10.1175/2010JCLI3814.1)
- Pirtle Z, Meyer R, Hamilton A (2010) What does it mean when climate models agree? A case for assessing independence among general circulation models. *Environ Sci* 13(5):351–361. doi:[10.1016/j.envsci.2010.04.004](https://doi.org/10.1016/j.envsci.2010.04.004)
- Salathe EP, Mote PW, Wiley MW (2007) Review of scenario selection and downscaling methods for the assessment of climate change impacts on hydrology in the United States pacific northwest. *Int J Climatol* 27(12):1611–1621. doi:[10.1002/joc.1540](https://doi.org/10.1002/joc.1540)
- Smith J, Hulme M (1998) Climate change scenarios. In: Feenstra J, Burton I, Smith J, Tol R (eds) UNEP handbook on methods for climate change impact assessment and adaptation studies. United Nations Environment Programme, Nairobi, Kenya and Institute for Environmental Studies, Amsterdam, pp 3–1–3–40
- Tebaldi C, Knutti R (2007) The use of the multi-model ensemble in probabilistic climate projections. *Philos Trans R Soc A Math Phys Eng Sci* 365(1857):2053–2075. doi:[10.1098/rsta.2007.2076](https://doi.org/10.1098/rsta.2007.2076)
- Themeßl M, Gobiet A, Leuprecht A (2011) Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *Int J Climatol* 31(10):1530–1544. doi:[10.1002/joc.2168](https://doi.org/10.1002/joc.2168)
- Vautard R, Gobiet A, Sobolowski S, Kjellström E, Stegehuis A, Watkiss P, Mendlik T, Landgren O, Nikulin G, Teichmann C, Jacob D (2014) The European climate under a 2 °C global warming. *Environ Res Lett* 9(3):034,006
- Whetton P, Hennessy K, Clarke J, McInnes K, Kent D (2012) Use of representative climate futures in impact and adaptation assessment. *Clim Change* 115:433–442. doi:[10.1007/s10584-012-0471-z](https://doi.org/10.1007/s10584-012-0471-z)