Environmental Modelling & Software 78 (2016) 191–201

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Selecting regional climate scenarios for impact modelling studies

Renate A.I. Wilcke [a, *], Lars Bärring [a, b]

[a] Rossby Centre, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
[b] Centre for Environment and Climate Research, Lund University, Sweden

A R T I C L E I N F O

Article history:
Received 21 January 2015
Received in revised form
5 January 2016
Accepted 7 January 2016
Available online xxx

Keywords:
Climate models
Optimised subset
Ensembles
Selection
Impact studies
Climate change impact
Simulation

A B S T R A C T

In climate change research ensembles of climate simulations are produced in an attempt to cover the uncertainty in future projections. Many climate change impact studies face difficulties using the full number of simulations available, and therefore often only subsets are used. Until now such subsets were chosen based on their representation of temperature change or by accessibility of the simulations. By using more specific information about the needs of the impact study as guidance for the clustering of simulations, the subset fits the purpose of climate change impact research more appropriately. Here, the sensitivity of such a procedure is explored, particularly with regard to the use of different climate variables, seasons, and regions in Europe. While temperature dominates the clustering, the resulting selection is influenced by all variables, leading to the conclusion that different subsets fit different impact studies best.

1. Introduction

In climate change research one uses simulations of past and future climate to explore the possible changes in future climate. Different global climate models (GCMs) have been developed by a number of research groups. To stimulate development of these models and reap maximum benefits of these efforts the World Climate Research Programme (WCRP) has established the framework Coupled Modelling Intercomparison Programme which now is in its fifth phase, CMIP5 (Taylor et al., 2012). Regional climate models (RCMs) are then used to downscale the global information onto regional scales which helps in assessing climate change information on the scale relevant to the impact of a changing climate. The modelling and downscaling is coordinated in frameworks such as ENSEMBLES (van der Linden and Mitchell, 2009) and CORDEX (Giorgi et al., 2009; Jones et al., 2011) which produce ensembles of GCM and RCM combinations. The motivation behind the use of multiple models in climate change research is to cover different sources of uncertainties, for more details see e.g. Hawkins and Sutton (2009, 2012), and Deser et al. (2012). Due to limited computing resources those matrices with GCM-RCM combinations are not complete, thus only part of the known uncertainty is covered.

To further explore how a changing climate is affecting us and the environment, impact models use climate model output data for the simulation of future climate change impacts, such as crop yields, or in hydrological models simulating the run-off in local areas. Even though it is advised to take all available climate model data into account (e.g. Knutti et al., 2010; Tebaldi and Knutti, 2007; Palmer et al., 2004), often it is not feasible in research projects. The problem which projects face is that the ensembles of GCM-RCM simulations are too big to be handled by many impact modellers.

Until now the GCM-RCM ensembles have often been reduced by hand-picking climate simulations depending on the partners involved in the project. Recently, more thoughtful choices were backed up by considering the different climate change signals of temperature, and sometimes of precipitation too (e.g., Fig. 2.1 in Wilcke et al., 2012; Mendlik et al., 2015; Gobiet et al., 2012). Also Murdock and Spittlehouse (2011) use the spread in the change of temperature and precipitation in a study for British Columbia. In few studies in the field of climate research, e.g. Logan et al. (2011), cluster analysis has been applied to select a subset from an ensemble of climate simulations. Only recently Cannon (2015) presented a sophisticated method applying the Katsavounidis-Kuo-Zang (Katsavounidis et al., 1994) algorithm on a large ensemble of global climate simulations.

* Corresponding author.
E-mail addresses: renate.wilcke@smhi.se (R.A.I. Wilcke), lars.barring@smhi.se (L. Bärring).

http://dx.doi.org/10.1016/j.envsoft.2016.01.002

By reducing the model ensemble one also reduces the information about the uncertainty in the projections and the ensembles. Note, by reducing the ensemble one does not reduce the uncertainty as such, only the information available about the uncertainty. Thus the task is to maintain maximum information within the limitations dictated by climate impact studies. In projects like CLIP-C and IMPACT2C (Vautard et al., 2014) a clear demand is expressed for a more systematic approach to select subsets of ensembles of climate simulations tailored to the needs within climate change impact studies.

In this paper we present a selection procedure founded on the same basic clustering method as introduced by Mendlik and Gobiet (2015) and explore how it performs for various combinations of variables and climate indices. Specifically, our approach differs in the details of selecting how many principal components to retain, method for choosing the number of clusters and finally how to select one member from each cluster. Based on experiences from current projects, we designed an experiment matrix which focus the evaluation on varying combinations of study regions, climate variables, climate indices, and seasons to test different impact study situations. The study regions included here are examples from Northern Europe. However, the method is to generally select members out of an ensemble, independent of the region.

The paper is structured as follows. Starting with the data description in Section 2, which is followed by the explanation of the ensemble reduction method in Section 3. The experimental set-up is described in Section 4 and the results presented in Section 5. In the summary we draw some conclusions in Section 6.

2. Data

2.1. Climate simulations

The ensemble of climate simulations used in this study consists of 11 GCM-RCM combinations from the EURO-CORDEX initiative (Jacob et al., 2013) with a grid spacing of 0.44 × 0.44° (approx. 50 km × 50 km). The simulations have been produced assuming concentration pathway RCP8.5 (van Vuuren et al., 2011; Stocker et al., 2013) and are listed in Table 1. For this study 30 years of data from historical (1971–2000) simulation runs were used as reference. The future climate is represented with three 30 year periods from the scenario simulation runs: 2021–2050, 2051–2080, 2069–2089.

These are the simulations available on the Earth System Grid Federation data network[1] in November 2014. Thus, it is a real world situation with an imperfect and imbalanced RCM-GCM matrix resulting in a limited ensemble of opportunity. This fact does not influence the integrity of this study, moreover it can be taken as motivation for selecting simulations different from each other.

2.2. Variables and indices

The data used here are daily values of six model outputs over Northern Europe: 2 m mean temperature (*tas*), minimum temperature (*tasmin*), maximum temperature (*tasmax*), surface precipitation amount (*pr*), mean relative humidity at 2 m (*hurs*), and mean wind speed at 10 m (*wss*). From these model output variables we calculated seasonal averages (Table 2) for each grid cell.

Additionally, for each grid-cell we calculate five annual climate indices that are derived from the climate model output. For temperatures, these are the *beetle-degree-days* index (*BDD*) and *exceeding threshold* index (*ET*) which are related to climate impact research on spruce bark beetles accompanying this study (Jönsson and Bärring, 2011). The *BDD* are the degree days marking spruce bark beetle maturity and *ET* is a threshold for the beginning of the second life cycle of those beetles (c.f. Jönsson and Bärring, 2011, for technical definitions).

The change in cold days can be described, e.g. by the *frost days* (*FD*) index which counts the number of days with *tasmin* below 0° C (Frich et al., 2002). From precipitation the *wet day frequency* (*RR1*, days with *pr* >1 mm/d) was derived, and *Beaufort days* (*FG6Bft*) are the number of days with wind speeds above 6 Bft (10.8 m/s) (see also ECA&D indices of extremes[2]).

This study focuses on the future climate, therefore the climate change signals (*ccs*) of the model output and indices are calculated and used as information for the clustering (subsection 2.1).

Covering regional and seasonal differences in climate change and in model performance, the climate change signals were integrated over 6 sub-regions in Northern Europe (Fig. 1) and four seasons. Table 2 shows the quantities used here, which gives $n = 174$ variables (6 climate variables × 4 seasons × 6 regions + 5 climate indices × 1 season × 6 regions = 144 + 30 = 174). This results in a matrix A spanned by m simulations and n variables

$$A = \{a_{ij}\} \tag{1}$$

where a variable a_{ij} is defined (Equationfootnote:footnote 2) as the *ccs* of a seasonal averaged climate model output or climate index averaged (ϕ_j) over a region (x).

Table 1
GCM-RCM combinations from EURO-CORDEX RCP8.5 on 0.44° grid and their abbreviations used in this study. (Kotlarski et al., 2014; CLIVAR Exchanges, 2011).

GCM	RCM	Abbreviation
CanESM2	SMHI-RCA4	CanESM2-RCA4
CERFACS CNRM CM5	SMHI-RCA4	CERFACS-RCA4
IPSL CM5A MR	SMHI-RCA4	IPSL-RCA4
MIROC5	SMHI-RCA4	MIROC5-RCA4
HadGEM2-ES	SMHI-RCA4	HadGEM2-RCA4
M-MPI-ESM-LR	SMHI-RCA4	MPI-RCA4
NorESM1-M	SMHI-RCA4	NorESM1-RCA4
GFDL-GFDL ESM2M	SMHI-RCA4	GFDL-RCA4
EC-EARTH	SMHI-RCA4	EC-RCA4
EC-EARTH	DMI HIRHAM5	EC-HIRHAM5
EC-EARTH	KNMI RACMO22E	EC-RACMO22E

Table 2
Ingredients for variables a_{ij} in this study.

ϕ		Seasons		Regions	
Mean temperature	*tas*	winter	DJF	region 1	R1
Min temperature	*tasmin*	spring	MAM	region 2	R2
Max temperature	*tasmax*	summer	JJA	region 3	R3
Precipitation	*pr*	autumn	SON	region 4	R4
Rel. humidity	*hurs*	annual		region 5	R5
Wind speed	*wss*			region 6	R6
Beetle degree day	*BDD*				
Exceeding threshold	*ET*				
Frost days	*FD*				
Wet day frequency	*RR1*				
Beaufort day	*FG6Bft*				

[1] http://esg-dn1.nsc.liu.se/esgf-web-fe/.

[2] http://eca.knmi.nl/indicesextremes/indicesdictionary.php.