16th International Conference on Environmental Science and Technology
Rhodes, Greece, 4 to 7 September 2019

How to identify a representative subset of hydro-climatic simulations for impact modelling studies?

Pechlivanidis I.G.[1,*], Gupta H.[2], Bosshard T.[1]

[1] Swedish Meteorological and Hydrological Institute (SMHI), Norrköping, Sweden
[2] Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA

*corresponding author:e-mail: ilias.pechlivanidis@smhi.se

Abstract
Uncertainties in hydro-climatic projections are (in part) related to different components of the modeling chain. Although a combination of numerous projections (ensemble) would be needed to characterize the overall uncertainty, in practice a small set of scenario combinations are constructed to provide users with a subset that is manageable for decision-making. The approach is based on a framework, rooted in the information theoretic Maximum Information Minimum Redundancy (MIMR) concept, for identifying a representative subset from an available large ensemble of hydro-climatic projections. We analyze an ensemble of 16 precipitation and temperature projections for Sweden, and use these as inputs to the HBV hydrological model to simulate river discharge until the mid of the 21st century. Representative subsets are judged in terms of different statistical characteristics for precipitation, temperature and discharge and the sensitivity of the identified subset is assessed for different seasons and future periods. Results indicate that a 20-35% subset of the available set of projections can represent a large fraction (more than 80%) of the ensemble range of hydro-climatic changes. We find that the identified representative subsets are sensitive to the regional hydro-climatic characteristics and the choice of variables, seasons and future periods.
Keywords: Representative projections, information theory, climate change impacts,

1. Introduction

Uncertainties (epistemic and aleatoric) in climate models, e.g. structural hypotheses, boundary conditions, and chaotic behavior of the climate system, can influence the future climatic projections and, further, their sectorial impact []. Production of a large multi-model ensemble of climate simulations suitable for impact modeling studies is an attempt to enhance our knowledge about the associated uncertainties in future projections [*Pechlivanidis et al.*, 2017]. Nevertheless, due to the limited availability of resources, practical designs of sectoral impact assessment studies are typically restricted to using small numbers of climate projections; and communication of uncertainty between modelers and end-users can be challenging.

Identifying representative subsets from a high-dimensional sample space is not a straightforward task. Ideally, one would like the selected subset to capture as much of the full range in simulated future changes as possible, and failure to incorporate relevant scenarios in the analysis could lead to inadequate assessments of uncertainty in the indicators of interest.

Here, we use the MIMR concept [Li et al., 2012] to analyze the representativeness of selected subsets with respect to both the input and the output variables of hydrological impact models and pose the following scientific questions: 1) How representative is the chosen subset? 2) How much do identified subsets vary between hydro-climatic variables, seasons and future periods?

2. Study Area and Methods

2.1. Study area

The climatic patterns over Sweden can be clustered into four regions following analyses of long-term records and scenario modeling (Fig. 1a). The river basins in these regions also show similarities in morphology, but also represent the catchments of the marine basins.

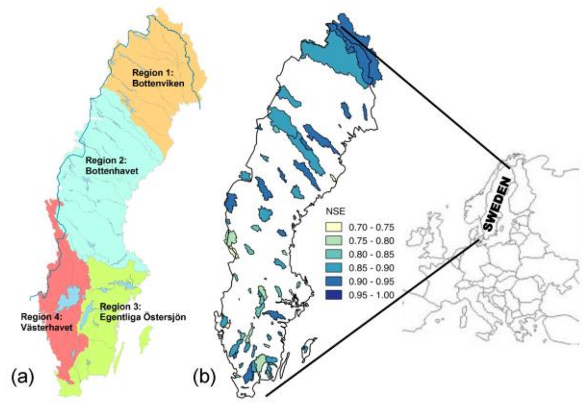

Figure 1. (a) The climate regions in Sweden, and (b) the HBV model performance (using NSE) at 69 test-basins.

The ensembles of climate projections consists of 16 available hydrological climate projections for Sweden [see *Pechlivanidis et al., 2018*]. They include 3 different emission scenarios, 5 GCMs, 6 RCMs, and two spatial resolutions. In addition, there are cases in which the model combination (modeling chain) is the same but projections are generated at different spatial resolutions. The PTHBV 4x4 km dataset for observed precipitation and temperature was used as a reference (1961-1990) for the bias-adjustment of the climate projections.

The HBV model is applied over Sweden with a spatial resolution of some 1000 basins (average 450 km^2). The model produces daily values for hydrological variables over historical periods and in forecasting. The model is reasonably well calibrated with a Nash-Sutcliffe Efficiency (NSE) > 0.70 at all 69 test-basins (Fig. 1b).

2.2. Methods

MIMR aims to select sets that: 1) maximize the overall information content (joint information), 2) maximize the overall information transition ability (transinformation), and 3) minimize the redundant information (total correlation). The multi-objective optimization problem is to select an optimal subset, which can convey effective information as much as possible, while retaining redundant information (if any) as little as possible.

For this analysis, we aggregated all data to the four climatic regions (Region 1-4), and the entire domain (All). At that level of spatial aggregation, we derived climate change signals of: 1) annual scales, 2) mean and extreme (10 and 100th percentile) statistical values, and 3) several different variables (precipitation, temperature and discharge). The climate change signals were estimated as differences between the corresponding statistics of the future 30-year scenario periods 2036-2065 (early century) and 2056-2085 (mid-century), and the reference period 1981-2010.

3. Results and Conclusions

We investigate the cumulative distribution of the joint information where projections are introduced into the subset in a step-wise manner (Fig. 2). This is done for all variables and their statistical properties, future periods and regions of the domain. The slope of the distributions differs between the variables, due to the different impact (relative change) that climate change can have in the four regions and overall. Precipitation and temperature cumulative distributions are characterized by sharp rises, in comparison to the smoother increase in total information associated with discharge.

If, for instance, 80% of the total information from the large ensemble is sufficient to represent the changes due to climate, our results indicate that one could select as few as 4 (or even 3) climatic projections and 5 (or 4) hydrological projections. This is a key finding, since any additional projections can only increase the total available information by about 20%.

Although the general pattern is that the ensemble spread gets wider with increasing lead-time, here many of the models project a similar pattern of climatic changes for the mid-century period. This is probably because we are still within the time horizon when internal variability (i.e. decadal variability) is dominant, and hence only a smaller subset of projections is needed when compared with the corresponding subset for the early-century. The homogeneity of mid-century climate change and its impact over the entire country results into a more consistent (less spread) distribution of total information between the regions.

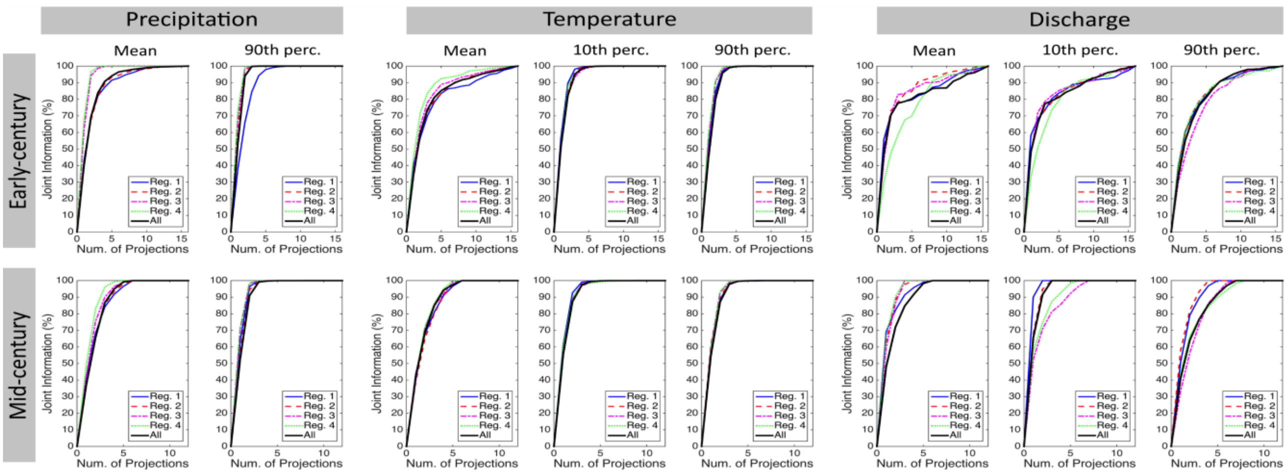

Figure 2. Cumulative joint information with increasing number of projections for different statistics of the annual hydro-climatic variables and for two future periods (early- and mid-century).

References

Li, C., Singh, V.P., Mishra, A.K., 2012. Entropy theory-based criterion for hydrometric network evaluation and design: Maximum information minimum redundancy. Water Resour. Res. 48, WR011251. doi:10.1029/2011WR011251

Pechlivanidis, I.G., Arheimer, B., Donnelly, C., Hundecha, Y., Huang, S., Aich, V., Samaniego, L., Eisner, S., Shi, P., 2017. Analysis of hydrological extremes at different hydro-climatic regimes under present and future conditions. Clim. Change 141, 467–481. doi:10.1007/s10584-016-1723-0

Pechlivanidis, I. G., H. Gupta, and T. Bosshard (2018), An Information Theory Approach to Identifying a Representative Subset of Hydro-Climatic Simulations for Impact Modeling Studies, *Water Resour. Res.*, *54*, 1–14, doi:10.1029/2017WR022035.