Hydrol. Earth Syst. Sci. Discuss., 11, C2304–C2308, 2014
www.hydrol-earth-syst-sci-discuss.net/11/C2304/2014/

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Inter-comparison of statistical downscaling methods for projection of extreme precipitation in Europe" *by* M. A. Sunyer et al.

Anonymous Referee #1

Received and published: 10 July 2014

The authors apply different bias-correction/downscaling (BC/DS) methods to RCM-simulated daily precipitation time series in 11 European catchments. They evaluate how the methods differ both with respect to the agreement with observations in the control period and with respect to future changes, in both cases with focus on extreme values of duration between 1 day and 1 month. Overall limited differences are found with weak dependence on e.g. location and duration.

Precipitation BC/DS is generally a key activity in hydrological climate change impact studies and evaluation/comparison of methods is an impoartant acticity. The experiment is very comprehensive, spanning a wide range of climate projections, methods

and catchments. The outcome must be worth sharing in the scientific community, but I think substantial revision of this manuscript is needed before publication, mainly for the reasons discussed in the following.

- The overall conclusion is essentially that there is no best method but we must use many, and on an as large an ensemble of RCM-projections as possible. In reality, however, it is extremely rare to have such resources but the impact study must be limited to one BC/DS method applied to a small RCM-ensemble (or even just one projection, it is not unusual). Thus I think that a key objective of this kind of study must be to provide advice and recommendations for real-world, limited-resource impact studies. If precipitation extremes is the key focus, what should we do? With all these results and all these (prominent) authors it must be possible to provide more useful knowledge, conclusions and adivice than what is currently the case, only "highlight the need for considering an ensemble" is not of enough help. For example different methods are to various degree prone to different problems like (1) cannot be applied/create unrealistic extreme values under some circumstances, (2) increases the bias in precipitation extremes under some circumstances, (3) modifies the climate change signal with respect to changes in extreme precipitation under some circumstances, (4, 5. . .). Further different methods are more or less prone to deviate from the rest in other aspects. A systematic review of this kind of key issues would provide very useful information.

- As I understand it the authors have calibrated the methods on the full set of reference data available and then applied to the same set of data (as well as the future-period data). However, I think this type of study needs to also include some kind of cross-validation analysis for historical periods, i.e. calibrate on one period and verify for another. This may not be possible for all catchments but in many there are some 50 years of data so you could split equally, divide 30/20 or something else. I think this kind of analysis is crucial for assessing the uncertainty when applying the methods to future periods.

- I find the presentation of results hard to follow. Results from different periods focus-

ing on different aspects are mixed under the headings “all catchments” and “selected catchments”. I think the paper would benefit from being structured more like the bullets 1-4 on l 13-25, p 6181.

- Related, I think the paper would benefit substantially from more distinct objectives. Now they are rather vaguely and incompletely formulated like “comparing BC/DS methods” or “assess the changes in extreme precipitation”. It would be better to formulate some distcinct hypotheses to investigate. Focus on apparent current knowledge gaps that this study can help filling.

- The Results section is largely a textual description of the tables and figures with little interpretation or explanation or speculation of the resons behind the findings. Sometimes it is speculated, but in those cases they could often have gone back to the results and for verification. See further specific comments below.

Specific comments:

- 20, 6169: It sounds puzzling that bias correction improves the agreement with observations only “in most cases”, clarify that is concerns the extremes.

- 13-23, 6171: You need to indicate also what relevant knowledge that was found in these studies.

- 24, 6171: Is not comparing BC/DS methods rather the main focus?

- 3-6, 6173: Was the gridded data not based on observations?, if so how were they derived?

- 2.1,2.2: What was the volume resolution in the observations?, 0.1 mm?, was the same cutoff used for the RCM data for consistency?

- 5, 6174: Rummukainen is misspelled and not in referece list, check carefully before submitting, should not be my task.

- 10-12, 6175: I think harmonisation would have been much better, why not used?

- 22-23, 6176: What if there are more wet days in OBS than in RCM?

- 19, 6177 – 5, 6178: As acknowledged it seems questionable to train on RCM-ERA and then apply on RCM-GCM without any correction. Can you provide some information on the additional error introduced?

- 9-12, 6179: The intervals are very small (and different from BCQM). It is well known that there may be an enormous variation in the very highest quantiles (above 99 or so), I think it needs to be demonstrated that using 0.0005 works well, instead of smoothing out the fluctuations a bit with larger intervals.

- 15, 6185: What is a “threshold return level”?

- 20-27, 6185: Assessing to which degree BC/CS modifies the CC signal is important and needs more attention. More analysis and intrepretation is needed. What is the change increased after BC/DS? What is the reason for the regional changes? Dig in the results.

- 7, 6188: “most likely” – check in the data

- 21, 6188: “might be similar” – check in the data

- 26-27, 6188: Why expect a larger impact in TR in summer?

- 4.2: I think it would be more natural to have 4.2 before 4.1, i.e. first an evaluation in reference period and then results from future period.

- 13, 6190: Not “extreme value index” but only “extreme values”, right?

- 5-15, 6191: Key paragraph which is far too compact. Only listing values is not very helpful; go much further, dig in the data, find out concretely why and under which circumstances the BC/DS methods decrease agreement with observations.

- Table 3: Add the methods’ abbreviations.

- Table 3: That seasonality is not taking into account in BCQM is an application issue,

not a disadcantage of the method.

- Table 3: CFQP: ACF can be checked for all methods.

- Fig 5: Add legend, refering to other figure is not sufficient.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 11, 6167, 2014.