Hydrology and
Earth System
Sciences

Discussions

Interactive comment on "Impacts of climate change on extreme floods in Finland – studies using bias corrected Regional Climate Model data" *by* Noora Veijalainen et al.

Anonymous Referee #2

Received and published: 5 January 2018

The authors employ the DBS method (in two versions) to bias-adjust RCM projections before using them to drive a hydrological model in order to assess the climate change impacts on discharge in Finland. In the DBS method, both a single and a double gamma approach is used, and the results are compared with earlier results btained using Delta Change (DC). The impacts vary within Finland, largely depending on whether floods are generated by snow-melt or rainfall, and the differences between the two DBS versions are overall small.

The topic is interesting and relevant, the material/methods/results are accurate as far as I can judge and the presentation is overall fine. However, I share several of the

concerns raised by reviewer #1, in particular:

- Limited novelty and general significance. Assessing the impact of climate change on extreme floods in Finland is, as I see it, today an engineering excerise rather than scientific research (if "standard methods" are used, as here). And the added value of assessing the impact of "gamma-type" in the DBS method limited, see further next item.

- Doubtful choice of study basins. It is hardly surprising that only very little impact of DBS version on the extreme floods is found, considering the large size of the basins and the associated large-scale nature of the flood generating mechanisms. Smaller/faster basins are probably needed to detect differences related to bias-adjustment of daily P extremes.

- Unclear significance/relevance of the precipitation data used. Essentially no information is given or assessment made about neither the observations nor the RCM data used. About observations, we need to know how they were obtained (interpolated from gauges?), spatial resolution, etc. About RCM data, how does the ENSEMBLES projections relate to more recent, more high-resolution, RCP-based projections (EU-ROCORDEX)? There is nothing wrong with using the ENSEMBLES projections but I think some effort should be spent in putting them in an updated context considering that they are about a decade old.

Having said that, there are interesting parts that could perhaps be further explored, in particular related to the "adequacy issue":

- Rejecting individual projections based on their historical performance is getting more and more accepted, I think. The authors look at how well the hydrological results (extreme floods) based on bias-adjusted historical forcing agree with observations and then identify "adequate simulations". The implications for the final results are however rather briefly discussed and it is a bit hard to see the "big picture". A more systematic analysis covering all projections and basins would be interesting.

- It is further interesting (although perhaps not surprising) that when considering only "adequate simulations", the results approach those obtained by a DC approach. This highlights the trade-off between additional flexibility on one hand and additional complexity/uncertainty on the other hand, in DBS (or, more generally, quantile mapping type of bias adjustment) as compared with DC. This is definitely not the first case when the results of relatively elaborate and complex bias-adjustment applications end up very similar to much simpler DC calculations.

- The adequacy of the hydrological model is very important but only rather briefly discussed. Maybe inadequate models should have been disqualified from the start? Generally, the authors could have expored the "adequacy issue" further, assessed the impact on the results/conclusions in amore systematic way, maybe suggested some kind of general guidelines for how to use this concept.

Specific comments (page/line(s)):

- 6/20: "range 0-106 %", is this correct?

- Table 2: I think you need to include also the mean/median of the RCMs, not only the range.

- 8/8-9: "The average"..."was 22-95 %", difficult to interpret.

- Fig. 4: I have some problems distinguishing between the dark green and dark grey lines, more different colours would help.