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 #1

Received and published: 4 January 2018


General statement: The manuscript presents projected changes in the 100-year precipitation and the 100-year discharge for a set of 67 catchments in Finland, including detailed results for four catchments and a general overview of the results for the full set of catchments. The projected changes are based on precipitation and temperature series from 5 RCM runs from the ENSEMBLES project. These series are bias cor-
rected using both single gamma and double gamma distribution-based correction before further analysis, and the results are then compared. The results are also compared with hydrological projections derived using change factors applied directly to observed timeseries. The authors conclude that the differences between the projected floods estimated with precipitation adjusted using a single vs. a double gamma distribution-based correction are small, and in addition, the results based on bias-corrected RCM are generally similar to those estimated using change factors.

The assessment of likely changes in flooding under a future climate using a systematic approach, such as described in this work, is clearly of national importance for Finland. Comparison of the performance of different RCMs in producing floods of the correct magnitude during the control period is also a noteworthy component of this procedure. However, I fail to see that this manuscript, in its present state, presents methods or results which can be considered to be both new and of international relevance. It seems that the authors are aware of the need to highlight important methodological issues and have chosen to focus on a comparison of bias adjustment using single vs. double gamma functions. Simultaneously, the four case study catchments they present are all relatively large (> 2160 km²), two of them have a high lake percentage (18-20%) and at least one of them is significantly affected by regulation. In these cases, it is very unlikely that two methods which differ only in the correction of the highest quantiles (i.e. > 95%) of DAILY precipitation will give systematically different results which can be distinguished in the flood quantiles. The authors briefly acknowledge this in their discussion in that they suggest that it is the 5-15 day precipitation sums that are important for driving floods (outside of the snowmelt season) in many Finnish catchments. However, this raises the question as to why these two bias correction methods, which only differ in their adjustment of the highest quantiles, are being compared in these catchments (i.e. catchments with relatively long response times). It may well be that some of the 67 catchments in the national data set are small enough to have shorter response times such that the differences between the two methods could be discerned. No information on the catchment characteristics of the large data set is, however, pro-
vided. The authors also seem to be unaware of several recent studies which have compared bias correction and other adjustment methods (e.g. delta change and, more generally, perturbation methods) for extreme precipitation (e.g. Sunyer, et al., 2015, HESS; see also Willems and Vrac, 2011, J. Hydrol.) and their effect on projections of extreme discharge (e.g. Hundecha, et al., 2016, J. Hydrol., Osuch, et al., SERRA, 2016).

In summary, I suggest that the authors need to reformulate their objectives and focus before this study can be deemed suitable for publication in HESS and of interest to an international audience. A number of specific suggestions in this direction are given below:

Specific comments:

1. The ‘delta change’ method is presented as a conventional method used in standard practice and that the use of bias adjusted daily data from RCMs have ‘recently become more common’ (pg. 2 l. 17). This statement may well have been true 10 years ago, but climate impacts research is indeed a fast-moving research area. There are innumerable examples of the use of daily bias-adjusted RCM data for hydrological impacts research available today in the literature. I suggest that this section of the manuscript could be improved significantly by including a more comprehensive review of current work considering the effect of bias correction on extreme precipitation and projected changes in flooding. Some of the relevant work is already mentioned in the discussion (e.g. Gudmundsson, et al., 2012; Beldring, et al., 2007) but should come earlier in the manuscript. Other sources are given above in the general comments, others can be found in Madsen, et al., J. Hydrol., 2014, or indeed by a more comprehensive literature search than is evident in this manuscript.

2. The preservation of trends is an important point to consider when using bias correction, and this is also mentioned here. It has, however, been suggested by Hempel, et al. (Earth Sys. Dynamics, 2013) that a ‘trend-preserving bias correction’ in which
residuals are corrected after removing the trend, generally gives better results. Has this been considered or tested? I realise that a full implementation of this approach would require that the entire work be redone, but the implications of this could at least be tested for precipitation.

3. The spatial resolution of the ENSEMBLES RCM grids is not mentioned at all in this manuscript, although it is clearly relevant for the ability of an RCM to reproduce spatial and temporal patterns of extreme precipitation. The most recent publications presenting changes in hydrology under a future climate are based on EUROCORDEX RCM runs, including runs with grids having a higher spatial resolution than ENSEMBLES runs.

4. The focus of the extreme value analysis here is the assessment of the 100-year flood as it is of national importance for climate change adaptation. Given the uncertainties introduced by fitting an extreme value function and the short annual maximum series available for each time slice, it would be more reasonable to first assess changes in the mean annual maximum precipitation and mean annual maximum flood and/or changes in average over threshold values, before presenting results for longer return periods.

5. The use of a Gumbel distribution for modelling extreme discharge values is justified by the application of the likelihood ratio test. Was this test applied to all of the data series for each catchment (i.e. including both control and future periods modelled from the corrected RCM data) or has this only been assessed based on observed time series? The use of a Gumbel distribution as a basis for assessing changes also raises the question as to whether or not changes in the tails of the distribution (i.e. the most extreme values) can be fully detected by such an approach. This is a particularly important point here in that changes in extreme precipitation are assessed based on a comparison of GEV fits (i.e. with a shape parameter) and for discharge with a Gumbel distribution (i.e. without a shape parameter).

Other minor technical comments
1. The use of English is generally very good, although there are numerous small errors, particularly related to the use of indefinite and definite articles.

2. The title of the article is too general and does not highlight an important methodological or other contribution of this work.

3. As mentioned above, additional information regarding the characteristics of the 67 catchments representing the national dataset should be provided. This can be given as, for example, mean, minimum and maximum values of variables such as catchment area, lake percentage, degree of regulation (if possible). Some information as to the seasonality of maximum discharge (or ideally the major process driving flooding) would aid significantly in interpreting the spatial distributions shown in figures 5 and 6.

4. Page 11, l. 2-3. ‘This separation made for every month leaves the extreme 5% scant and makes it sensitive to few observed occurrences of heavy precipitation, which are relatively rare in the Finnish climate’. The last part of this sentence doesn’t entirely make sense, in that it is the upper 5% of the total distribution for each month which is being considered and fitted to a gamma function. For dry months without a large number of rainy days, it could well be the case that there are relatively few values over the 95% threshold, but there are then also proportionally fewer days under the 95% threshold. So, this would not reflect ‘few observed occurrences of heavy precipitation’, but rather, ‘few occurrences of days with precipitation’. It is for this reason that some researchers have chosen to use 3-month blocks for bias correcting precipitation, so that more extreme values are available.

5. The source of the precipitation and temperature data used for bias correction should be discussed, including how areal values are derived for catchment precipitation. Has bias correction been performed relative to station data, or to grid points in a gridded dataset, or to areal values derived from station or gridded data?

6. In Table 1, the column ‘Emission Scenario’ is superfluous as all are run under the A1B scenario. This could be mentioned in the Table description.
7. The discussion section could benefit from a better structure. It is relatively long, and no overview is given highlighting what topics are covered.

8. A very minor point. . . but nevertheless important when one is reviewing a paper: The paragraph structure needs to be delineated with, for example, an indentation of the first line, otherwise it can be a bit ambiguous as to when a new paragraph begins.