Review of Technical Note

'The impact of spatial scale in bias correction of climate model output for hydrologic impact studies'

by E.P. Maurer et al.

Dear Editor, dear Authors,

I have reviewed the aforementioned work. My conclusions and comments are as follows:

1. Scope

The article is within the scope of HESS.

2. Summary

The authors investigate the influence of the spatial resolution at which a bias correction of climate model output (daily precipitation and max and min temperature) is performed. The bias correction (quantile mapping) is set up in a calibration period (1969-1989) and applied in a validation period (1990-2011). Input data are daily observations interpolated on 1/8° grid and reanalysis data on 1.9° grid over several large river basins in the Western United States. The bias correction is performed on data mapped to resolutions ranging from 2.0° to 0.125°. Both the observed and bias corrected data are then further downscaled to 0.125° resolution and fed to hydrological models.

The observed and bias corrected data are compared in various ways: The cdf's for the three meteorological parameters are compared for a single grid cell, the river flow cdfs of daily streamflow are discussed for two gauges, and a test statistic (Mann-Whitney U test for agreement of observation- and reanalysis-based flow cdf median) of annual 3-day high flow and annual 7-day low flow for 185 gauges is discussed for the validation period.

The main findings are that i) there is no clear 'best bias correction resolution' for the meteorological variables. Instead, results vary with variable, location and quantile; ii) with respect to the hydrological output, differences in the quantiles of observation- and reanalysis-driven streamflow are considerable, with a (weak) agreement optimum for bias-correction at 0.5° resolution.

The authors conclude that bias-correcting at 0.5° may indicate an optimum between detail of the observations exploited and nonstationary effects between calibration and validation period minimized due to spatial averaging.

3. Overall ranking

The work is ranked 'Major revision'.

4. Evaluation

For the reasons detailed in Ehret et al. (2012), I still argue that bias correction of climate model projections for further application (e.g. in hydrological models) is highly questionable. However, as the study discussed here rather evaluates than applies a bias correction method, I set these reservations aside for now.

The study is generally short but thoroughly conducted (some minor objections are given below), what strikes me is the weak agreement of the cdf's for river flow (Figures 6 and 7) for the extreme high and low flows and for the agreement of the cdf's in general (Figures 8 and 9). For high flows, the difference between the reference and the bias-corrected reanalysis data is roughly that between a once-in-10 and a once-in-100 year flood (compare discharge from reference at P=0.99 and P from 2.0° bias corrected reanalysis at P = 0.9. Admittedly, from a 20-year data set, extrapolations to 100-

year floods are uncertain). For low flow cases this is even worse. As the data used here are reanalysis data, which, as the authors state (10899/6pp) should outperform predictions from free climate model runs, this makes the applicability of climate model output to address questions of hydrological extremes doubtful. Even if this is not the core topic of the manuscript, it should be discussed to put results into perspective.

Secondly, as the authors correctly state in the text (10895/14pp) quantile mapping bias correction affects both stationary and instationary parts of model-observation discrepancies, which means that it has the potential to worsen results in a validation period, which can partly be attributed to removing the physical coherence between the model output fields. If this procedure is generally accepted in the climate change community, I wonder why bias correction is not applied 'end-of-the-pipe', i.e. directly to the river flow, if this is the variable of interest. The advantage is that the meteorological fields remain coherent, and the undesirable step of disaggregating meteorological input to the scale of hydrological subbasins can be omitted (as the results are re-aggregated by the convolution of river flow anyways). And if application of bias correction to output of meteorological models is justified, why should it not be so for river flow? The only justification that comes to my mind to do the one (bias correction of climate model output) and leave the other (bias correction of hydrological model output) is to argue that the atmosphere-landsurface interface is one of weak interaction. It would be interesting to see whether bias correcting river flow yields better results in a validation period than correcting the meteorological drivers. I realize that this point is not directly related to the scope of this paper, but I wonder if the authors can comment on it.

Some specific points

- 10898/14: Please give some more information on the underlying observations (number of stations etc.)
- 10898/25pp: Please clarify whether the cdf's are determined individually for each grid cell or aggregated for larger regions or the entire domain.
- 10900/14pp: Please state in which period and with which forcing the hydrological model calibration was done.
- 10901/1pp and Figure 3: This does not tell us much about the bias in the extremes. Rather show differences for extreme rainfall and length of dry spells, which better corresponds to the hydrological statistics discussed later in the text.

Your sincerely,

Uwe Ehret

Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi and J. Liebert (2012): HESS Opinions "Should we apply bias correction to global and regional climate model data?". Hydrol. Earth Syst. Sci. 16 (9), 3391-3404, 10.5194/hess-16-3391-2012.