Hydrol. Earth Syst. Sci. Discuss., 10, C5181–C5184, 2013
www.hydrol-earth-syst-sci-discuss.net/10/C5181/2013/
© Author(s) 2013. This work is distributed under
the Creative Commons Attribute 3.0 License.

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Influence of downscaling methods in projecting climate change impact on hydrological extremes of upper Blue Nile basin" *by* M. T. Taye and P. Willems

Anonymous Referee #2

Received and published: 24 September 2013

This paper compares two methods of downscaling GCM output over the Blue Nile basin for simulations of future hydrology. The two methods are a quantile perturbation method (QPM) which adjusts GPM output based on what amounts to binned differences between GPM current climate and observations, and a weather generator, parameters of which are adjusted based on differences between GPM current and future climate.

It is a bit difficult to follow some aspects of the paper. Probably most important is how the spatial dimension of the problem is dealt with. I am left with the impression that the hydrology model VHM is forced with basin average quantities (in part due to the

statement on p. 7868 that "The comparison is done after calculating the basin-wide rainfall from both the observations and the control runs." This is a little hard to believe, and perhaps I don't have this right – the Blue Nile is such a large basin that spatial differences in forcings and response have to be important, and most modern era hydrology models would be spatially distributed at some level (i.e., if conceptually lumped, at least applied to subcatchments with different forcings, and more commonly run over a spatial grid of some kind). This is important, as both the GPM and weather generator discussions seem to be mapping directly to/from an observational time series, but these is only one such time series as nearly as I can tell. At one point in the paper, 11 stations are mentioned, but whether these have somehow been averaged to produce a single series isn't at all clear. In their "flow chart" (p. 7866), there is nothing to indicate looping over sub-basins, or stations. Furthermore, the stations are essentially points, whereas the GCM represents large grid cells, and the scale mismatch isn't addressed.

In the implementation of the weather generator, they state explicitly that it is implemented for each (of the presumably 11) stations historically. I interpret this as meaning that point data could now be synthesized for each of these 11 points for current climate, and as a next step, they adjust the weather generator parameters to reflect future climate. But, a key problem with weather generators is that to my knowledge, they don't account for coherence between the multiple points – so while the statistics should reflect spatial variations, the time step simulations do not (i.e., they don't reproduce storm characteristics, such as movement across a large river basin). This is a critical consideration for the Blue Nile, which again, is quite large, and has to be strongly affected not only by temporal, but by spatial variations in the inputs as well.

The authors conclude (probably correctly) that the GCMs don't do a very good job of reproducing observations for their current climate runs. They attribute this to the fact that the basin averages are computed from the GCM for its grid, and the points (stations) from the observations. There is no doubt some issue with the mismatch, but this isn't likely to be the major source of uncertainty (one could probably demonstrate this using, for instance, the reanalysis, which effectively is gridded model output, and station data someplace (like continental U.S.) where there are a lot of stations – I can state with pretty high likelihood that once you get beyond a fairly modest number of stations (and there are some papers that have done theoretical analyses of this – see papers in the 70s by Rafael Bras and Ignacio Rodriguez-Iturbe, among others) the average of the gridded fields and the observations will be pretty close. The problem is much more likely to be dynamical errors in the GCMs. That being the case, the authors may want to do some screening of GCM output (e.g. from the CMIP5 archive) to select a subset of GCMs that reproduce the variables of interest to them plausibly well.

The key result in the paper that QPM downscaling implies decreasing mean annual precipitation, whereas the weather generator implies increases, simply is not plausible. A minimal requirement of any downscaling method is that at least approximately, it needs to reflect the dominant signals from the GCMs. So if the global model says that precipitation should be increasing, and the downscaling method says decreases, this has to be an implausible artifact of the downscaling. As I tell my students when they bring me results that don't make sense, I don't necessarily know why, but it is wrong. At this point, in my view they should have stopped, rewound the entire process, and tried again.

One final comment – less important than those above: The temperature indexed PET (Hargreaves) is suspect – it's quite likely that it overestimates the sensitivity of PET to changing climate. For information on the pitfalls of using temperature indexed PET for projections outside the observed range (e.g., for climate change studies) see a 2011 J. Great Lakes Research paper by Lofgren et al; and also recent work on the subject by Chris Milly.

This is a paper that I think would best be sent around the track again. Perhaps ultimately, something useful can be produced, but it would be best that the current version not be in the archival literature.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 7857, 2013.