Manuscript: Large-basin hydrological response to climate model outputs: uncertainty caused by the internal atmospheric variability

Major remarks

The authors present a robust analysis of a large ensemble of GCM-HM simulations to investigate the impact of internal variability on simulated river runoff. The study is interesting and worth publishing but a few things need to be addressed before.

[1] It needs to be pointed out clearly that the considered time scales are important for the validity of results of the study. To separate the study from other research working on longer, climatological time scales, time scales longer than one year should be notably excluded, i.e. the impact of internal variability diminishes compared to other uncertainty sources if, e.g., multi-year monthly or annual means are considered (e.g. Déqué, M., D. Rowell, D. Lüthi, F. Giorgi, J.H. Christensen, B. Rockel, D. Jacob, E. Kjellstrom, M. de Castro and B. van den Hurk (2007) An intercomparison of regional climate models for Europe: assessing uncertainties in model projections. Climatic Change 81, Supplement 1, 53-70)

[2] Studies such as Deque et al. (2007) or (Hagemann, S., H. Göttel, D. Jacob, P. Lorenz and E. Roeckner, 2009: Improved regional scale processes reflected in projected hydrological changes over large European catchments. Climate Dynamics 32 (6), doi: 10.1007/s00382-008-0403-9: 767-781) considering uncertainty introduced by internal variability at longer time scales should also be referred to in the introduction section.

[3] In the conclusions section it would be interesting to address the following question based on the results: What are the implications for seasonal to decadal predictions using GCMs?

[4] Technically I recommend a careful checking regarding the use/non-use of ‘a’ and ‘the’ in the manuscript. These seem to be missing at many places.

In summary, I suggest minor revisions to be conducted before the paper may be accepted for publication.

Minor Comments

In the following suggestions for editorial corrections are marked in *Italic*.

p. 2306 – line 25
… mean *value, which* indicates …

p. 2306 – line 26
It is written:
“…a considerable portion of the observed trend can be externally driven.”

As you only deal with simulations I would not recommend using the word “observed” in this context.

p. 2311 – line 25

In *Section 5, runoff* characteristics …

p. 2313 – line 14
… Geophysics*; Motovilov et al. 1999a) has* been…

p. 2313 – line 23-24
(SWAP*; Gusev and Nasonova 1998) has* been …

p. 2314 – line 10-11
It is written:
"Some key-parameters of the models are calibrated against streamflow measurements and …"

Some more information on the calibration and the respective parameters is desirable.

p. 2314 – line 25-27
It is written:
"In particular, ECHAM5 similar to majority of climate models (Flato et al., 2013; IPCC AR5) simulates colder climate in winter in high latitudes of the Northern Hemisphere …"

I doubt this statement. Hagemann et al. (2006, 2013) show a distinct warm bias of ECHAM5 (AMIP simulation, but also coupled to an ocean model) in the winter over the high northern latitudes land area (or the area covered by the six largest Arctic rivers).

References:
Hagemann. S., K. Arpe and E. Roeckner, 2006: Evaluation of the hydrological cycle in the ECHAM5 model. J. Climate, 19, 3810-3827
Hagemann, S., A. Loew, A. Andersson, 2013: Combined evaluation of MPI-ESM land surface water and energy fluxes. J. Adv. Model. Earth Syst., 5: 259-286, doi:10.1029/2012MS000173.

p. 2317 – line 10-11
It is written:
"One can see from this Figure that the applied post-processing allowed us to obtain rather similar fields of the above listed variables.."

The similarity between the model data and observations used for corrections is rather trivial as this can be expected from a bias correction approach. It would be of interest to show the uncorrected (original) fields in addition to see how large the correction actually is.

p. 2325 – line 12
It is written:
"… which is particularly noticeable for the winter season, when the SD-estimates are sometimes lower by hundreds percent in comparison with their observed variability."

Maybe it should be noted that discharges in winter are usually small for high latitude rivers so that even absolute small differences may yield large relative differences.

p. 2327 – line 7-8

It is written:
“Importantly, the role of the internal atmospheric variability is most visible for the time scales from years to first decades …”

This is only true if one does not consider multi-annual monthly or annual means. See major remark [1].

p. 2328 – line 2
… runoff *trend, were* estimated.

p. 2338 – Fig. 2
The top left panel is a duplicate of the top right panel. I assume, it should show temperature, not precipitation.

p. 2340 – Fig. 4
Instead of showing one curve per panel, the panels for the same river should be merged to allow an easier comparison between the two models.

p. 2342 – Fig. 6
Panels for the same river should be merged to allow an easier comparison between the two models.

p. 2343 – Fig. 7
Panels for the same river should be merged to allow an easier comparison between the two models. If this is not feasible, please use at least the same y-axis scaling for panels belonging to the same river.

p. 2344 – Fig. 8
Instead of showing one curve per panel, the panels for the same river should be merged to allow an easier comparison between the two models.

p. 2345 – Fig. 9
Panels for the same river should be merged to allow an easier comparison between the two models.

p. 2346 – Fig. 10
Panels for the same river should be merged to allow an easier comparison between the two models. If this is not feasible, please use at least the same y-axis scaling for panels belonging to the same river.