Replies to Referee #1

Does the weighting of climate simulations result in a more reasonable quantification of hydrological impacts?

Hui-Min Wang, Jie Chen, Chong-Yu Xu, Hua Chen, Shenglian Guo, Ping Xie, Xiangquan Li

We sincerely appreciate the referee's comments on this manuscript. These comments are all helpful to improve this manuscript. We have carefully studied and responded to all comments point-by-point as follows. For clarity, all comments are given in *italics* and responses are given in plain text. The manuscript will be modified correspondingly.

> *This is a well written paper that studies the added value of weighting GCMs within an ensemble as a function of hydrological performance rather than as a function of climatological performance as usually done. The paper discusses some interesting aspects (e.g. the difference in outcome if weighting according to precipitation or temperature under different hydrological regimes) and comes to the conclusion that if raw GCM data is to be used, ensembles should be weighted based on streamflow rather than temperature or precipitation. In exchange, there is not much added value with streamflow-based weighting if the underlying GCMs are duly bias corrected. This outcome is not entirely surprising (see detailed comments) but I think it is nevertheless interesting for the readers of HESS and thus worth publishing.*

We appreciate that the referee is in favor of the content of this research. Detailed comments have been replied as follows and will be addressed in the revision.

> *Detailed comments*
>
> *In this paper, the GCM weighing is tested for large catchments (» 10'000 km^2) that are simulated with a lumped model (GR4J) at a daily time step. With such a lumped model, it can a priori be assumed that the most important aspect of climate inputs for hydrological model performance and for future simulations is the actual precipitation and temperature bias. In fact, there is a whole body of hydrological literature on the importance of correct area-average precipitation estimates, which should perhaps be linked to this study. A starting point is the work of Lebel et al. 1987. Since the model is lumped, spatial differences between meteorological inputs derived from GCMs cannot show up in the simulation results otherwise than affecting the catchment scale average values (i.e. the bias). Differences between GCM outputs in temporal variability do most likely not show up because they are dampened by the model. The authors argue that the response of a catchment to climate input is nonlinear. This holds in general but if such a simple model is used, no surprising outcomes can be expected (not much difference between climate-based weighting and hydrological weighting in absence*

of major meteorological biases). This is a limitation of the study: major differences between climate-based weighting and hydrological weighting can a priori not be expected in the bias corrected set-up with such a simple model. This has to be discussed in sufficient detail in the paper and highlighted also in the perspectives. Finally: I am not an expert on bias correction methods. Accordingly I can only assume that this part of the work is state-of-the-art.

Thanks for the comment. We agree with the referee that it is a limitation that only large watersheds and one lumped hydrological model were considered in this study. When using a lumped hydrological model, the nonlinear relationship between the climate variables and the impact variable (streamflow) may not be sufficiently revealed. Spatial differences between different climate simulations only affect the basin-averaged inputs to the hydrological model but not directly affect the process of runoff generation and streamflow routing. Temporal variations of climate simulations may be partially reduced by the lumped hydrological model. With the help of other more sophisticated hydrological models (such as the distributed model, SWAT), the differences between climate-based weights and streamflow-based weights may become more obvious. For the experiment of raw GCM-simulated streamflows, the weights based on streamflow show better performances than those based on climate variables. This may be related to large differences between climate simulations. But in the experiment of bias-corrected GCM-simulated streamflows, no much differences in the performances between unequal and equal weighting may be because only a simple hydrological model is used. In other words, even though the performance of the bias correction method differs in climate model simulations, the remaining biases among corrected climate simulations may not be well presented in streamflow simulations. All analyses above will be presented in the revision of the manuscript.

In addition, the daily bias correction (DBC) method used in this study has been applied in many recent studies (e.g., Chen et al., 2017; Li et al., 2019). It can be considered as a superior bias correction method in terms of correcting the bias of precipitation frequency and the bias in the distributions of precipitation amounts and temperature.

References

Chen, J., Brissette, F. P., Lucas-Picher, P., and Caya, D.: Impacts of weighting climate models for hydro-meteorological climate change studies, Journal of Hydrology, 549, 534-546, https://doi.org/10.1016/j.jhydrol.2017.04.025, 2017.

Li, L., Shen, M., Hou, Y., Xu, C.-Y., Lutz, A. F., Chen, J., Jain, S. K., Li, J., and Chen, H.: Twenty-first-century glacio-hydrological changes in the Himalayan headwater Beas River basin, Hydrology and Earth System Sciences, 23, 1483-1503, https://doi.org/10.5194/hess-23-1483-2019, 2019.