Replies to comments by Yan Liu et al.

We would like to thank Yan Liu, Veit Blauhut, Amelie Herzog, Tunde Olarinoye and Ruth Stephan for their interest in our paper and for posting their comments on our manuscript.

Below, their comments are in italic font and our replies are in normal font.

Comments are from the discussion during a workshop by: Yan Liu, Veit Blauhut, Amelie Herzog, Tunde Olarinoye, Ruth Stephan

The study "Why does a conceptual hydrological model fail to predict discharge changes in response to climate change?" by Duethmann et al presents a very interesting topic, which tries to find important factors that influence the prediction capability of conceptual hydrological models, especially under climate change. In this study, the HBV model was used as one representative of conceptual hydrological models. Three aspects regarding precipitation input, model calibration period, and potential evapotransporation (LAI and NDVI were used to consider changes of vegetation dynamics and land cover) were investigated to discuss the causes why HBV model fails to predict discharge under changing climate. This study is in the scope of HESS and well written.

After reading and discussing this manuscript during a workshop, we thought that posting our comments might be helpful for improving the manuscript. We have following major and specific points:

Major points: 1) Title and abstract are a bit misleading because the results are not generalising for all hydrological models but using HBV as one representative. It would be better to explicitly state that the results are based on HBV model in the abstract. Using subtitle may also help clarifying this issue.

> We will revise the title, please also refer to the comment #1 by David Post. The new title reads 'Why does a conceptual hydrological model fail to correctly predict discharge changes in response to climate change?'. By referring to 'a conceptual hydrological model' and not 'conceptual hydrological models' we intend to indicate that we have tested one and not several or more models. We will make the abstract clearer and mention explicitly that the results are based on a HBV-type model (and catchments in Austria). To avoid abbreviations in the title we did not add 'a HBV-type model' there. Our results are based on a specific model and catchments in Austria, the problems we found may, however, also be relevant for other hydrological models and other regions (also see SC3 by Taehee Hwang).

2) The prior distribution of model parameters was assumed to be the beta-distribution. In such way, by giving shaping parameters _ and _ for the beta distribution, it seems that the optimal parameter ranges (high probability density part of the beta distribution) are known for the prior. That will affect the model calibration. To justify why using a beta distribution not a uniform distribution for the parameter prior distribution is needed in the method section.

> The a priori distributions for the model parameters were applied to be consistent with the study by Merz2011. It is assumed that we have more information on the likely parameter values than just the parameter range. We checked that including the penalty for deviating from the prior distributions does not have much influence on changes in model performance

over time (see Fig. 1 compared to Fig. 4 in the original paper). When the penalty for deviating from the prior distributions was omitted from the objective function, calibrating the model in S1 and applying it to 1978–2013 resulted in an average discharge trend of 118 ± 86 mm yr^{-1} per 35 yrs and thus virtually no effect compared to the original model.

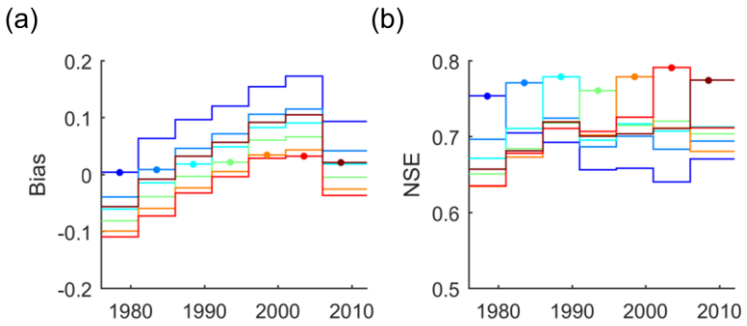

Figure 1 (a) Bias and (b) NSE for the different subperiods averaged over all study catchments when omitting the penalty for deviating from the prior distributions. Each line refers to models calibrated in one subperiod, showing bias and NSE during calibration (marked by the filled circle) and during evaluation in the other six subperiods.

3) Since the results were analyzed for the averages over 156 catchments, it would be better to see the probability density distribution of the bias (Qobs-Qsim) of all catchments for the prediction periods to support that the low predictability of the averages of all catchments is not due to several catchments that bring very big bias. Providing this information in the supplement will strongly support the results.

This is a good idea and we will add violin plots showing the distribution of the bias and NSE to the manuscript or the supplement.

Specific points:

1) A northern arrow is missing in Fig.1, the elevation legend is normally vertical. Fig. 2 is not very informative, maybe merge it with Fig. 1.

A northern arrow will be added. The elevation legend was set horizontal to better use the space. Fig. 2 might be moved to the supplement.

2) In Fig. 4, how was the bias calculated.

This will be added. The bias was calculated as $bias = \left(\sum_{t=1}^{n} Q_{sim,t} - \sum_{t=1}^{n} Q_{obs,t}\right) / \sum_{t=1}^{n} Q_{obs,t}$.

3) What does the unit "mm yr-1 per 35 yrs" mean? Is that the mean discharge (mm yr-1) over the 35 years?

The unit "mm yr-1 per 35 yrs" refers to trends, such as the trend in mean annual discharge over a period of 35 yrs.

4) In equation 8, definition of fbeta is missing. fp was not used.

Thanks, this will be corrected.

5) In Sect. 2.3.1, many model parameters were introduced, such as CR and Bmax, but these two parameters are not provided in Table 1.

Table 1 lists parameter ranges of the a priori distribution for the model parameters that were included in the model calibration. The parameters T_R, T_S, C_r and B_{max} were not included in the model calibration and set to constant values (described in P9, L7). They are therefore not listed in Table 1.

6) Table 2 contains almost all the details of hypotheses. But there is also quite long text in Sect. 2.4.2 that repeats the table. Table 2 is clear, try to reduce the duplicate text in Sect. 2.4.2.

We will have a look at this when we will revise the manuscript.

7) Hypothesis should be a result out of the introduction and be mentioned at the last paragraph in the introduction.

That is an alternative we have also thought about. However, the reason why we decided to introduce the detailed hypotheses after the model description and not in the introduction was that we assume it to be easier for a reader to follow the hypotheses and the corresponding modifications to the model after the model (including the input data) has been introduced.

8) In the discussion, very good literature review was done. But it should more highlight the findings of this study and relate and compare to literatures.

We will adjust the discussion to emphasize more the results of our study.

9) It is not clear that how the trend was calculated when using 25 years as the calibration period. Please clarify that.

As in the cases where the model is calibrated for 5 years periods, the parameters are applied to the entire study period and the trend over the entire study period is calculated.