Reply to reviewer #2

The manuscript addresses the different sources of precipitation and the effect on dis- charge, and also did the attribution analysis of discharge bias to forcing variables and model structure by using the Budyko assumption. The analysis of resources of precipitation and discharge is very important for the water cycle of the cold regions, and the attribution analysis contributes to the understanding and the simulation of water cycle. The research in this paper is very interesting and important. The content and the quality of the paper fit into the HESS standards. However, the innovation is not outstanding and some of the results is not well explained. Specific comments are as follows:

1) The logical structure, the abstract as well as the conclusions should be adjusted to highlight the innovation of the manuscript.

Reply: Thanks for the suggestions. With considering the following comments together, we realized that the paper can be improved if we explain the bias sources and the proposed methods in a better way and in an earlier place. Therefore, we plan to revise the paper in the following ways.

a. We revised the abstract (attached in the end of this document) by clearly separating it into four parts (i.e. brief background, the proposed method, the general results and a short conclusion). The innovation of the manuscript is therefore easily to be found and emphasized in the abstract.

b. The most significant change in the structure is that we plan to move the subsection 4.3 (the introduction of biases and their classification as forcing or model structure) and part of subsection 4.4.1 (introduction of the Bukydo analytical tool for bias analysis) forward to the end of the methodology part, which seems more logical. Then we plan to add more explanations on the strategy we take to attribute the bias sources of different variables.

c. The conclusion part will be revised accordingly. We propose to shorten the conclusion but emphasize the framework we introduced as we did in the revised abstract.

2) More details on the methodology to identify the bias from the model structure and forcing inputs, as well as the evaluation index should be provided.

Reply: To better organize the paper structure (Reply to question 1), we propose to move the current subsection 4.3 and 4.4.1, where different biases and their classifications as forcing bias or model bias and the analytical approach, to the methodology part. So that the biases will be introduced in methodological part which is more logical and easier to find for readers. In the end of methodological part (section 3), the introduction of the evaluation approach as well as the calculation index will be provided.

3) For bias from the model structure, in section 4.1.2 the bias of the model structure is identified as the discharge difference between WFDEI-CCG and WFDEI-CCG-SF. While in section 4.4.1, the change caused by ET with unchanged P and PET is identified as the bias affected by

the model structure. They are inconsonant, and what is the difference between these two?

Reply: We attribute the biases in the model simulation to either the forcing or the model structure. The forcing bias includes biases in precipitation and other atmospheric variables (e.g. radiation, wind etc.). Because the potential evapotranspiration (PET) is mainly determined by forcing variables, the bias in PET belongs to the forcing bias. Any other biases are involved in the bias from model structure.

The snow and soil freezing scheme (Sect 3.2.2, 'SF' as an abbreviation) is improved in the model used for WFDEI-CCG-SF simulation compared to the old version used for WFDEI-CCG simulation. The differences are induced by model structures because the forcing for these two simulations are the same. The new snow and soil freezing scheme determines the snow processes and the water movement in the soil, which will further affect the energy balance and evaporable water. Hence, the final response to the new scheme is the changes in ET and the discharge is changing as a result.

In the assumption 3 of subsection 4.4.1, the P and PET which are related to forcing conditions are considered right. The difference of estimated discharge from the observations is the bias of ET. The bias is caused due to discrepancies of ET processes in the models rather than the forcing. Therefore, the ET bias is also belonging to category of 'model structure'. It is consonant with the previous bias (SF module), as the SF module also changes the ET estimation at last. Bias in any processes that change the ET will be involved as the bias from model structure.

With the revisions finished for Comment 1 and 2, the biases and their sources will be presented in a clearer way in the revised manuscript at the end of the methodology part, which facilitates the understanding of the differences between biases.

4) For the results analysis, more in-deep reason should be put forward instead of just describing the phenomenon. For example, why there exits deviation between WFD-CRU and WFDEI-CRU after 1990, as shown in Page 12? Why the discharge correlation decreases for the upper Aksu, shown in Page 13.

Reply: a. The precipitation difference between WFD-CRU and WFDEI-CRU (Figure 3b) is only determined by the CRU versions used for correcting the precipitation. As introduced in 3.1.2 Near-surface atmospheric conditions, the precipitation in WFD was corrected using CRU TS 2.10 and that in WFDEI was corrected using CRU TS 3.1(Weedon *et al*., 2014). The two CRU versions differ in the time period, in the stations used and the methods employed (Harris *et al*., 2014). The version difference induces the final difference shown in Figure 3b and the difference enlarges after 1990.

b. The correlation between the simulated discharge and observation at Aksu decreases in experiment WFDEI-CCG to that of WFDEI-CCG-SF (Figure 6b). The decrease in the correlation can be understood by checking the annual cycle of discharge between the two experiments for Aksu (Figure 5-e) as "An early discharge peak exists in May, while not enough runoff is generated in the summer period (Figure 5-e)" (Line 8-9, Page 11). The difference between the two experiments only attributes to the usage of new developed snow and soil freezing module (introduced in subsection 3.2.3). Therefore, we consider the change is due to better description of the new SF module because the snow-melting in spring is better captured in the WFDEI-CCG-SF simulation (March-May, Figure 5-e). Although the correlation decreases, it does not mean the model/simulation deteriorates because correlation only evaluates the similarity of temporal variation but ignores the fact that the discharge amount has been better estimated (Figure 5-e and Figure 6-a).

The necessary explanations to the two questions are planned to be added in the revised manuscript.

5) How do you calculate the biases range in Table 5 since there is no observations for each of the forcing variables? The manuscript declares that the observations in the nearby regions or the regions with similar climatic and regional characteristic are used. But only the underestimate or overestimate can be concluded compared to the observations nearby regions instead of an exact value.

Reply: As presented in the caption of Table 5, the bias implies the bias of the values in the current variables compared to the values they should be. The current value is the value estimated by model and the ideal value is obtained from the Budyko curve. For instance, with assumption 1 for the upper Yarkand, we consider P as the only factor biased. To meet the discharge observation, we got the 'corrected' P by Budyko approach as 453.4 mm/yr (Figure 8a, from point A to point B). Because the original P is 247.3 mm/yr, the bias is (247.3 – 453.4) / 453.4 * 100%= -43.2%. However, the value is obtained when we consider only P is incorrect, if there are other variables with bias, the possible bias of P could be smaller. So, the possible maximum range of P bias is -43.2% (Line 3-10, Page 16).

We do use the observations in nearby or similar regions, but these references are not used for obtaining the exact values in Table 5. They are used to analyze the possibility of the assumptions because we are not sure whether the assumptions exist (shown in Table 6). For instance, because the PET for Tibet Plateau ranges from 580-720 mm/yr (Chen, Liu and Axel, 2006), the PET (=1153.7 mm/yr) is therefore very likely overestimated for the Upper Hotan where the climate is not very different from the Tibet (Line 12-16, Page 17). On the other hand, it avoids over-correction. For example, if we consider only PET is biased in the Upper Yarkand, we have to correct it to 225 mm/yr to meet the observations, which is too low compared to the reference. The PET is thus not the only biased variable for this region (Line 16-18, Page 17). With the reference values, we can obtain the relative possibility of biases in variables P, ET and PET. Then, we can know the action we need to take first to improve our model simulation (Line 18-24, Page 17).

All the proposed changes will be traced in a new version of manuscript.

Reference:

Chen, S., Liu, Y. and Axel, T. (2006) 'Climatic change on the Tibetan Plateau: Potential evapotranspiration trends from 1961-2000', *Climatic Change*, 76(3–4), pp. 291–319. doi: 10.1007/s10584-006-9080-z.

Harris, I. *et al*. (2014) 'Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset', *International Journal of Climatology*, 34(3), pp. 623–642. doi: 10.1002/joc.3711.

Weedon, G. P. *et al*. (2014) 'The WFDEI meteorological forcing data set: WATCH Forcing data methodology applied to ERA-Interim reanalysis data', *Water Resources Research*, 50(9), pp. 7505–7514. doi: 10.1002/2014WR015638.

Abstract:

The bias in atmospheric variables as well as that in model computation are two major causes of failures in discharge estimation. Attributing the bias in discharge estimation becomes difficult if the forcing bias cannot be evaluated and excluded in advance in places lack of qualified meteorological observations. The problem is more complicated over the mountainous area where strong orographic effects exist and with severe heterogeneous geography (e.g. the Upper Tarim basin). In this study, we proposed an ORCHIDEE-Budyko framework which helps identify the bias range from the two sources (i.e. forcing and model structure) with a set of analytical approaches. A latest version of land surface model-ORCHIDEE was used to provide more reliable discharge simulations based on the most improved forcing inputs and model modules. The Budyko approach was then introduced to attribute the discharge bias with observations to two sources with prescribed assumptions. The possibility of these biases was discussed by referring to many other studies with similar climatic or land surface characteristics. Results show that as the forcing biases, the water inputs (rainfall, snowfall or glacier melt) are very likely underestimated for the Tarim headwater catchments (-43.2%~-21.0%). Meanwhile, the potential evapotranspiration is unrealistically high over the upper Yarkand and the upper Hotan (1240.4 mm/yr and 1153.7mm/yr respectively). Determined by the model structure, the bias in actual evapotranspiration is possible but not the only contributor to the discharge underestimation (overestimated up to 105.8\% for the upper Aksu). Based on a simple scaling approach, we estimated the water consumption by human intervention ranging from 213.50×10^{8} m^{3}/yr to 300.58×10^{8} m^{3}/yr up the Alar gauge station, which is another bias source in current version of ORCHIDEE. This study succeeded in retrospecting the bias from the discharge estimation to multiple bias sources of the atmospheric variables and the model structure, although the framework needs further argumentations about its robustness, it provides a unique method for evaluating the regional water cycle and its biases with our current knowledge of observational uncertainties.