1. I had the pleasure to read your paper and in the following I provide some comments and suggestions aimed at improving your paper in a potential revision. I think that your paper has potential to advance the application and hydrological assessments using tracer-aided models in large-scale, mountainous catchments with an important glacier and snow contribution to streamflow. You clearly outline the challenges associated with working in such extreme environments and the difficulty to account for spatio-temporal variability of model inputs and outputs used for model evaluation. In this context, the use of isotope-enabled GCMs as input data for a hydrological model is a key result to me with potential for other applications elsewhere. Having said that, I have some doubts and suggestions you could consider incorporating into a revised paper

Response 1: Many thanks for your comments and suggestions. We will revise the manuscript thoroughly according to your comments.

2. The model experiments are not clearly presented, which makes it in parts difficult to follow up on the results and conclusions.

Response 2: Thanks for your comments. We will try to improve the presentation of experiments and make them clearer and easier to follow in the revised manuscript.

3. The model THREW-T, setup and data is poorly described raising questions about how the model treats isotope transport within model compartments. For example, I kept wondering throughout if and how the model treats glacier and snow isotope processes. To me, the glacier meltwater contribution (water and isotope) to streamflow is a simulated process within the model and part of the model evaluation. Your statements about a fixed glacier meltwater isotope signature to force the model (experiment 1) is not clear. Further, are there no model parameters associated to isotope transport and mixing? How did you arrive at the single (optimum?) benchmark parameter set for model comparison? The model calibration and how uncertainty was treated is also not clearly explained and no posterior parameter ranges after calibration are presented.

Response 3: Thanks for your questions. The work presented in this manuscript is on the basis of the model and isoGSM evaluation which was done in Nan et al. 2021a and 2021b. Consequently, the description of model was simplified in this paper to avoid too much repetition. Nonetheless, we will add some necessary descriptions of model in the revised manuscript to make it clear how isotope processes were simulated. Responses to specific questions are as followed:

(1) Simulation of snow/glacier melt processes and their isotope compositions

The THREW model simulated the accumulation and melting processes of snowpack, so the isotope compositions of snowpack and snowmelt were determined similarly with other water storages. Different with snow processes, the model simulated the glacier in a simplified manner. The model did not simulate the evolution of glacier in a centrary and even longer timeframe, but only calculated the glacier meltwater according to the degree-day factor, and the glacier meltwater was assumed to contribute to streamflow directly through surface runoff pathway. Consequently, the isotope composition of glacier was not simulated, and was assumed to be temporally constant.

(2) Parameters associated to isotope:

There is no parameter associated to isotope transport and mixing. We assumed the isotope mixed completely in each hydrological simulation unit within each step, which was reasonable to some extent because the structure of model is distributed. The water content of each unit and the fluxes among them has been calculated in the hydrological model, so the mass and concentration of isotope can be simulated.

(3) The benchmark parameter set:

Actually, the single benchmark parameter in Table 5 was only used to produce a series of stream isotope data, which is regarded as 'measurement' data for model calibration in experiment 3. This is because in some scenarios in experiment 3 (like RT_YTR_2year), the assumed stream isotope data availability was beyond the actual measurement dataset. Consequently, we solely picked out a parameter set from the behavioral parameter sets of triple-objective calibration according to several aspects of model performances (Table 5), and this parameter set was not necessarily an optimal set.

(4) Model calibration and parameters:

We will add more details about the calibration process, and use some figures to show the calibrated parameters.

4. The snow cover seems a crucial information for model calibration, but how was it derived and then simulated is not explained. Consider further sub-dividing the effect of snow cover on the model calibration compared to only streamflow and streamflow plus stream isotope calibration.

Response 4:

Thanks for your question and suggestion. The snow cover area (SCA) in this study was extracted from a multi-source fusion dataset produced by Chen et al. (2018). The extracted SCA data was regarded as measurement data, and was used for model calibration. In THREW model, the processes of snowfall, snow accumulation and snowmelt were simulated to estimate the variation of snow water equivalent (SWE). The SCA was then calculated according to SWE by a snow cover depletion curve. We will provide descriptions about this in the revised manuscript.

This study focused on the value of isotope data on improving model performance, and the influence of isotope data situation, thus isotope was calibrated based on the calibration variant towards the full-element dataset (e.g., streamflow plus SCA). Nonetheless, the suggestion you proposed is also useful to illustrate the value of isotope. When calibrating model towards streamflow plus stream isotope, we can explore the model performance on snow simulation even though SCA was not calibrated. However, this is not in good agreement with the purpose of this study, so we will do analysis in future works.

5. The water source contributions to streamflow (Table 5) don't seem to add up to 1? Also, many of the boxplots do not show a difference compared to the benchmark. Maybe consider a different visualization or quantification of differences and/or similarities might help to support the conclusions.

Response 5:

Thanks for your question and suggestion. The contribution of runoff component (CRC) was quantified based on two definitions as clarified in L217-223 and reviewed by He et al. (2021). The first definition is the water sources in the total water input, e.g., where the water come from. The runoff component is divided into rainfall, snowmelt and glacier melt in this definition, and the contributions of these three components add up to 1 (Table 4 and 5). The second definition is the runoff generation pathway. Considering the structure of THREW model, the runoff is divided into two components: surface runoff and subsurface runoff (baseflow). Considering the contributions of these two components add up to 1, only the contribution of baseflow was shown in Table 5 and Figures 5, 7 and 10.

We will adjust the y-axis scale of some figures to make different CRC among scenarios more visible. We are concerned about both mean value and the uncertainty range of CRC, so maybe boxplot is a proper way to present the results. Actually, the difference in CRC is rather significant in experiments 1 and 2. The difference in experiment 3 is relatively small, so we use mean absolute error or standard deviation to quantify the difference (Figure 12 and 14).

6. The bias-correction of the iso-GCM is described in equations 1-3, but the results not shown. How does the result perform against an inferred or measured isotope-elevation gradient and is the lapse the dominant driver as opposed to spatial variability over such a large catchment area? The results of model experiment 2 (Figure 8) should be evaluated against streamflow isotopes and not streamflow.

Response 6:

Thanks for your comments. We will present some features of the merged precipitation isotope data in the revised manuscript (the process is more a multi-source merging than bias-correction, and we will change the term in the revised manuscript). We have evaluated the performance of isoGSM in the study area in a previous work (Nan et al., 2021b). We will add more descriptions about the bias characteristic of isoGSM, but would not provide detailed information about this, to avoid duplication with the previous paper.

IsoGSM well captured seasonality of precipitation $\delta^{18}O$, but it performed relatively bad in two aspects. The first is that it cannot capture the exact $\delta^{18}O$ value of a specific precipitation event or a short period. Consequently, for the dates with observation precipitation isotope data, the observation data was used to denote the temporal fluctuation, and isoGSM was used to quantify the spatial variability. For the dates without observation data, the isoGSM was used to quantify both spatial and temporal variability.

The second shortcoming of isoGSM is that it overestimated the precipitation $\delta^{18}O$ in the study area. We inferred that the bias changed with elevation, and the changing rate (parameter a in equation 2) was estimated according to the bias at precipitation sampling stations. The major difference among scenarios in experiment 2 is the changing rate of bias, rather than the measured isotope-elevation gradient. We will correct this in the revised paper.

We will add figures to show isotope simulation in experiment 2, but we think evaluation against streamflow is important as well. The primary aim of isotope simulation and calibration is to aid hydrological simulation, rather than reproduce the variation of stream isotope itself. Although precipitation isotope data did not influence streamflow simulation, the calibration process making simulated stream $\delta^{18}O$ fit with observed values would influence the parameter, and consequently influence streamflow simulation.

7. The paper should also be thoroughly edited for language, as I detected many odd wordings and grammatical errors. I attached an annotated pdf with comments and suggestions for your

information.

Response 7: Many thanks for your revisions. We will revise the manuscript thoroughly according to your comments and suggestions.

Chen, X., Long, D., Liang, S., He, L., Zeng, C., Hao, X., and Hong, Y.: Developing a composite daily snow cover extent record over the Tibetan Plateau from 1981 to 2016 using multisource data, Remote Sen. Environ., 215, 284–299, https://doi.org/10.1016/j.rse.2018.06.021, 2018.

He, Z., Duethmann, D., and Tian, F.: A meta-analysis based review of quantifying the contributions of runoff components to streamflow in glacierized basins, Journal of Hydrology, 603, 126890, 10.1016/j.jhydrol.2021.126890, 2021.

Nan, Y., He, Z., Tian, F., Wei, Z., and Tian, L.: Can we use precipitation isotope outputs of isotopic general circulation models to improve hydrological modeling in large mountainous catchments on the Tibetan Plateau?, Hydrology and Earth System Sciences, 25, 6151-6172, 10.5194/hess-25-6151-2021, 2021b.

Nan, Y., Tian, L., He, Z., Tian, F., and Shao, L.: The value of water isotope data on improving process understanding in a glacierized catchment on the Tibetan Plateau, Hydrology and Earth System Sciences, 25, 3653-3673, 10.5194/hess-25-3653-2021, 2021a.