Response to Editor:

Thank you for your responses to the comments provided by the reviewers. I read the manuscript, as well as the comments of the three reviewers and your replies. Overall, your contribution about the evaluation of the effect of sampling strategies on the performance of a tracer-aided hydrological model is very valuable. However, as stated by reviewer 1 and reviewer 2, in the revised manuscript you need to improve the presentation of the model experiments, the model setup, the observed isotopic data, and the snow cover data (how snow cover data was obtained, and if and how it was used in the model calibration). Furthermore, I recommend you discuss more, as a limitation of your study, the fact that you have not considered the impact of a high temporal and spatial variation in the isotopic composition of snowmelt and ice melt (this is an important limitation, particularly, if you do not have any samples from these two water sources). As reported by the second reviewer, there are many studies observing how a high variability in the isotopic composition of snowmelt and ice melt can lead to a high uncertainty in the quantification of the contribution of these two water sources to stream runoff.

Response:

Many thanks for your effort of handling the manuscript and the comments. We have revised the manuscript thoroughly according to the comments of three reviewers. Meanwhile, we have discussed more about the temporal and spatial variation in the isotope composition of melt water (L679-690). Actually, the spatio-temporal variability of melt water isotope was considered in the model by simulating the isotope composition of snowpack storage, and estimating the glacier melt isotope according to the average local isotope composition of precipitation. However, it was difficult to valid whether they were characterized properly due to the limited isotope data for snow and glacier melt water. We could only infer that the simulation of melt water isotope was acceptable, by the fact that the model performs better on the simulation of discharge and stream isotope at both outlet and internal stations, compared to the result obtained by bi-objective calibration without calibrating isotope. More data of melt water isotope would be helpful to verify the isotope simulation and estimation of CRC.

Response to Reviewer 1:

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 have revised 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 have improved the introduction of experiments and make them clearer and easier to follow in the revised manuscript (Table 3, and the main text of section 2.4).

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 have added 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 century 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. We have added more details about above issues in L213-221 and L228-233 in the revised manuscript.

(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. We have clarified this in L222-228 in the revised manuscript.

(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. We have clarified this in L327-333, L350-352 and L397-399 in the revised manuscript.

(4) Model calibration and parameters:

We have added more details about the calibration process (L272-277 and L346-355), and added Figure 3 to show the uncertainty of 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 have added descriptions about the snow cover simulation and calibration in L144-146 and L213-218 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 L238-242 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, 10 and 13.

We have adjusted 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 have added some features of the merged precipitation isotope data in L449-458 and Figure 7 in the revised manuscript (the process is more a multi-source merging than bias-correction, and we have changed 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 have added some descriptions about the bias characteristic of isoGSM in L160-166, but did not provide much 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 have added Figure 8 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 have revised 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.

Response to Reviewer 2:

1. Data input. I did not see the description on the water isotopes original data. So what is the variability of precipitation isotopes? And how about snow-melt, glacier-melt and groundwater?

Response 1: Thanks for your questions. We have added the average value and standard deviation of $\delta^{18}O$ in precipitation and stream water in the Table 1. The isotope data of other water bodies was not collected in the field work, which was a limitation of this study.

2. The conclusion: 'Using a set of glacier meltwater δ18O that were 2‰~9‰ lower than the mean precipitation δ18O resulted in only small changes in the model performance and the quantifications of contributions of runoff components' was inconsistent to the existing findings. Most of the previous studies attach great importance to glacier melt and snow-melt water isotope change, because they think this will lead to great bias of hydrograph separation (See the following references). Please discuss more on this.

Response 2: Thanks for your comment. Most of the hydrograph separation works were based on the end-member mixing approach, which was applied in a short time scale, and was more dependent on the absolute isotope composition of each runoff component. However, this work applied the tracer-aided hydrological model in a longer time scale, where the temporal variability of isotope composition played a more important role than its absolute value, on the parameter calibration. Consequently, when the temporal variability of isotope composition of each water source was reproduced properly, the glacier melt $\delta^{18}O$ value in a reasonable range would have little influence on the model performance. We have clarified this in L586-596 in the revised manuscript.

3. The authors gave some suggestions on the sampling. For example, they concluded 'It is highly recommended to increase the number of stream water sampling sites rather than spending resource on extensive sampling of stream water at a sole site for multiple years'. But I think this is highly up to the research purpose. If one wants to see the seasonal variation of water source contribution related to climate change, the conclusion should be inverse. So I suggest the authors to draw the conclusion more seriously, or add some preconditions.

Response 3: Many thanks for your suggestion. We agree with you that the water sampling strategy is highly related to the research purpose. The sampling strategy proposed in this study is mainly aimed at capturing sufficient hydrological variability for establishing a tracer-aided hydrological model. We have made this clearer in the conclusion in the revised manuscript (L696 and L713).

Response to Reviewer 3:

1. I appreciate the description of the numerical experiment in Tab. 3. This is, certainly, a good way to summarize the faced setting. However, I think that the authors need to improve the model description in the section 2.4. In particular, it is necessary to integrate the experiment 1 section because is not clear how to justify the benchmark parameter and then, the calibration is not clear explained. I think that many information can be deduced from Nan 2021 but it is better for the reader to have an integration. Similar considerations can be achieved for the correction of precipitation (c.ca Line 601, iGCM).

Response 1: Thank you very much for your appreciation and suggestion. We have added more descriptions about the model (L213-237), calibration (L272-277 and L346-355) and iGCM correction (L160-177 and L449-458) in the revised manuscript.

2. Fig. 1 I suggest to remove the connection lines between the three figures. Then, I suggest to include the geographical info of Fig.1a and c inside the figure (similar to fig. 1b).

Response 2: Thanks for your suggestion. We think the connection lines between figures are necessary to present the relationship between the three figures, but we have reduced the width of the lines to make the figure more beautiful.

I suppose you mean to include the legend of Figure 1a and b inside figure similar to Figure 1c. We have adjusted the figure according to your suggestion in the revised manuscript.

3. Line 297-298 maybe you miss the word “precipitation” after the first “isotope”.

Response 3: Thanks for your comment. We have added the missing word in the revised manuscript.