Reviewer 1

The manuscript presents a hydrological modeling study in a glacier-influenced catchment. The work explores the value of auxiliary datasets, namely water isotope composition, snow cover area, and glacier mass balance in model calibration in a GLUE framework. The model structure allows tracer simulations and comparison with spatially variable datasets. The works finds different datasets have more power in model calibration in different hydrological seasons: isotopes during baseflow, and snow and glacier related observations during the melt period.

I liked the systematic approach for including model validation datasets of very different origin to model evaluation scheme. The GLUE uncertainty analysis framework for the work is in my judgement valid. The overall approach the authors develop to explore parameter sensitivity to model validation objectives and stream source water contribution are in my opinion of interest to the community. I recommend the work to be published after addressing my comments below:

MAJOR COMMENTS

I'd like to see better presentation of the stable water isotope data. You have only isotope data of the streamflow validation, it remains unclear how representative the input precipitation data is of the catchment. Do any of the references cited for the model development have any comparison data for simulated precipitation, snow or groundwater isotope composition? Having even cursory validation of the simulated isotope composition in different model compartments (snow, glacial melt, groundwater mainly) in the would give more credibility that the streamflow isotopes are correctly simulated and informative for the right reasons. On that note, I'd like to see a figure of the stream isotope data and model simulation fit to stream isotopes.

Response:

We will add more descriptions about the isotope of various water bodies.

- Precipitation: Our previous evaluation of isoGSM (Nan et al., 2021) indicated that it can effectively capture the seasonal variation in precipitation $\delta^{18}O$, but exhibited a systematic overestimation bias in the study region and performed relatively poorly in accurately capturing the isotope signature of specific events (see the Figures 1 and 2). We adopted the corrected isoGSM product from Nan et al. (2022) as the input data, in which the bias of isoGSM was adjusted based on a linear regression with altitude. Note that the corrected isoGSM directly incorporates measured precipitation $\delta^{18}O$ data for locations and dates with observations, so comparing the corrected isoGSM with measured data is not meaningful. Instead, we only show the relationship between the original isoGSM and measurements to illustrate the capability of the original isoGSM in simulating precipitation $\delta^{18}O$.

- Glacier melt: The $\delta^{18}O$ of glacier meltwater was calculated using the offset-parameter method, in which the glacier-melt $\delta^{18}O$ was assumed to be temporally constant and 5 ‰ lower than the weighted average of local precipitation $\delta^{18}O$. The value of the offset parameter (5 ‰) was estimated from the data collected by Boral and Sen (2020).
- Groundwater: We were not able to collect groundwater samples for isotope validation. Nonetheless, the characteristics of groundwater $\delta^{18}O$ are clear: it exhibits much lower temporal variation than precipitation or streamflow, due to the long water travel time.
- Snowmelt: We were also not able to collect snowmelt water samples for isotope validation. So the parameters constrained by isotope didn't lead to significant difference in snow simulation and the estimation of snowmelt runoff. The snow simulation is mainly constrained by the snow cover area data.
- Streamflow: Yes. We will add a figure to show the simulation and observation of stream water $\delta^{18}O$.

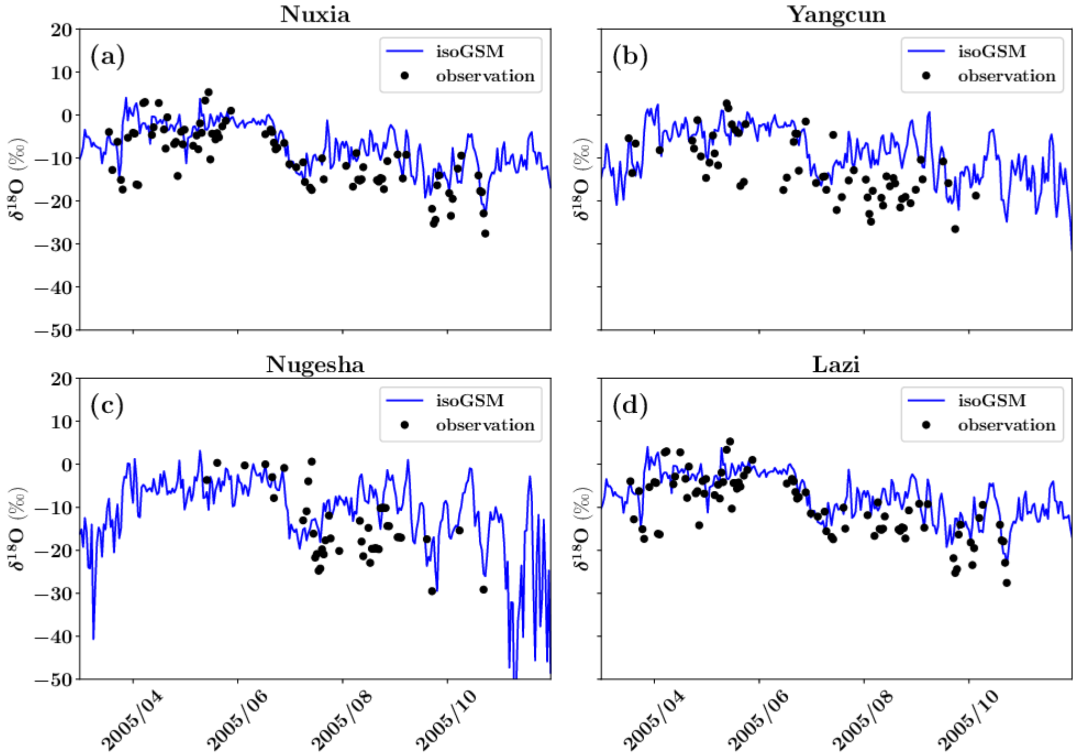

Fig. 1: Temporal variations in the precipitation $\delta^{18}O$ derived from observation and isoGSM data at four stations.

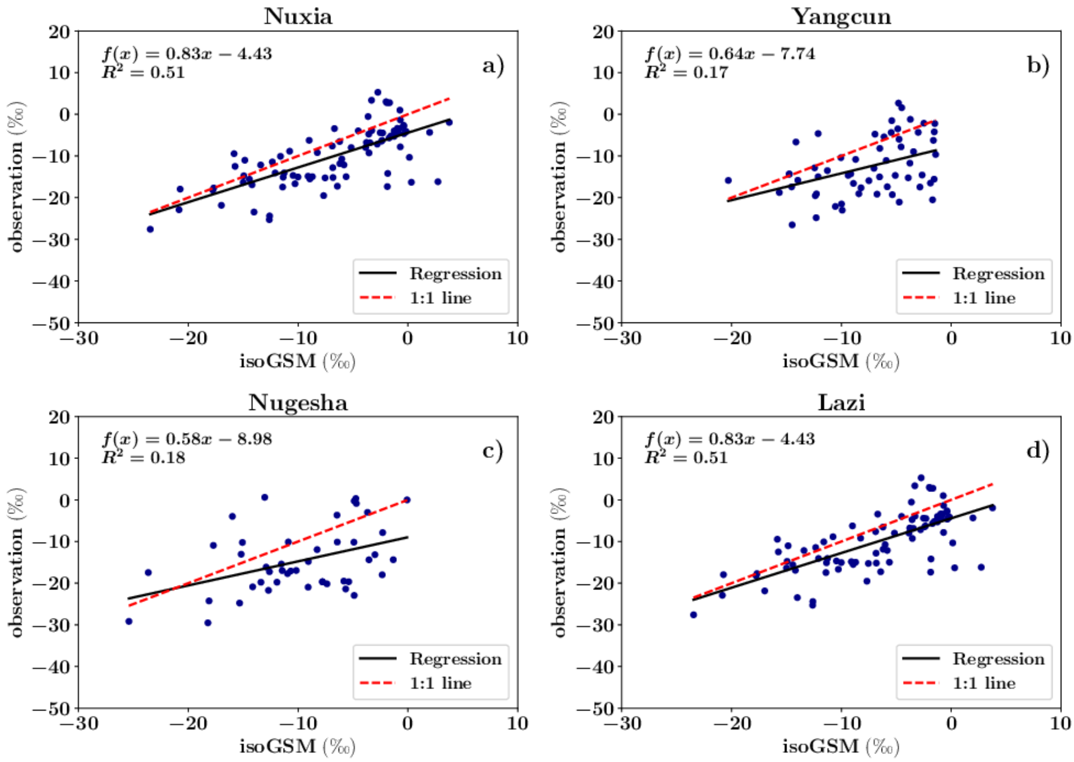

Fig. 2: The scatter diagrams between the isoGSM and measured isotope data at four stations.

The fractions for snowmelt surface runoff and glacier surface runoff seem low to me. Can you provide comparison with fractions found in other montanous snow and glacier influenced sites? Quite often end-member mixing analysis fraction estimations are done for three end members: snow, rain and glacial melt. In your model analysis groundwater is explicitly considered as a component, but isotopically it is essentially composed of rain, snow and glacial melt. This in my opinion creates a bit of confusion, and makes the glacial and snow melt seem less important for the regions water resources. I don't think there is an error in your analysis, but would be good to clarify the concepts further, to make your results more relatable to other literature.

Response:

- There are two common ways to define runoff components. One is based on water source, describing where the water originates; under this definition, the three end-members are rainfall, snowmelt, and glacier melt. The other is based on the runoff-generation pathway, describing how water produces runoff; here, the two end-members are surface runoff and subsurface runoff. In both cases, the sum of component contributions equals one. Because groundwater is also important, including all components would require reporting two separate sets of results (e.g., rainfall 80%, snowmelt 10%, glacier melt 10%; surface runoff 40%, subsurface runoff 60%), which can be confusing. Therefore, in this study we combined these definitions and defined four runoff components.

- The contribution of runoff component is highly dependent on the calculation definition and the dataset used for model validation. In our result, the low contribution of snow and glacier runoff is partly due to our component definitions and the correspondingly high share of subsurface runoff. The contributions of snowmelt and glacier meltwater in the study area are highly uncertain, ranging from less than 5 % to over 30 %. Nonetheless, studies using snow and glacier data to validate the model—thereby enhancing reliability—have all estimated relatively low contributions. For example, Chen et al. (2017) estimated snowmelt and glacier-melt contributions as 10.6 % and 9.9 % (ratios defined as SM/Q and GM/Q), whereas Zhang et al. (2025) reported corresponding values of 6.0 % and 6.2 % (ratios defined as SM/(RF + SM + GM) and GM/(RF + SM + GM)). When calculated in the same manner, our results closely match these estimates.

MINOR COMMENTS

L12: I perceive GW-SW interactions as specific water exchange processes between surface and subsurface water. As you don't really delve deeper into GW-SW interactions in your simulations, I'd propose that you stick with talking only about baseflow, not GW-SW interactions (which baseflow generation if of course a manifestation of)

Thank you for this clarification. We agree with your observation and will avoid using the term "groundwater–surface water interactions" in the revised manuscript, as our analysis does not explicitly address these processes. Instead, we will refer to subsurface flow, which is more consistent with the scope of our simulations and aligns with Reviewer 2's comments

L54-L66: seems like the research questions are to some extent repeated. Suggest to review and rewrite more concisely.

Thank you for the suggestion. We acknowledge the redundancy and will revise the paragraph to make the research questions more concise and focused in the revised manuscript.

L110: Do you think snow sublimation would be a significant flux in your region, possibly influencing the snow storage and isotope composition of the snowpack consequently snow melt?

Some studies indicated that due to the wet condition in the YTR basin, the sublimation losses are relatively small, accounting for only 2-3% of the annual snowfall (Lutz et al., 2016, Khanal rt al., 2021). Consequently, our model doesn't consider the influence of snow sublimation, as some other modeling studies in this basin (e.g., Chen et al., 2017, Sun at al., 2024).

L213: not clear how the simulations comprising the pareto front (red markers in

are selected. seems like the number of the included simulations is fairly low, around 15.
Thank you for your observation. The red markers correspond to simulations on the Pareto front, identified in the bi-objective space shown in each panel. The relatively small number of red points (~15) reflects the fact that only a limited subset of simulations are non-dominated with respect to both objectives. We emphasize that the Pareto front is computed over the entire simulation ensemble, independently of any behavioral classification. The red points should therefore not be interpreted as behavioral simulations, but rather as Pareto-optimal solutions based solely on the two performance metrics shown. This behavior is consistent with what was reported by Di Marco et al. (2021), who found that the number of Pareto-optimal simulations was substantially lower than the number of behavioral ones, highlighting how multi-objective trade-offs can lead to highly selective optimal subsets. We will revise the manuscript and the figure caption to make this distinction clearer.

L238: can you further explain where the prior parameter distributions in Fig.4 comes from. Is it the parameters with >0 NSE for streamflow?
We agree that the origin of the prior parameter distributions shown in Fig. 4 requires further clarification. The prior parameter distributions are derived from model parameter sets that resulted in a positive Nash-Sutcliffe Efficiency (NSE > 0) for streamflow. This filtering step ensures that only behaviorally plausible parameterizations—those capable of reproducing streamflow dynamics to a reasonable degree—are included in the prior. We will revise the manuscript to explicitly state this criterion in the main text, and we will update the caption of Fig. 4 accordingly for clarity.

L304: I don't fully understand why the sensitive LL parameter does not manifest in the snowmelt fraction.
From a water balance perspective, the contribution of snowmelt is primarily governed by the fraction of snowfall in total precipitation, which depends on the temperature threshold used for rainfall/snowfall partitioning. While the LL parameter mainly affects the spatial extent of snow cover — and thus the spatial and temporal distribution of snowmelt — its influence on the total amount of snowmelt remains limited. We will clarify this aspect more explicitly in the revised manuscript.

L307: the narrower ranges for isotope simulations are not evident visually compared to the Q simulations. Would any statistical test either looking for differences in central values or variability in the distributions be helpful in identifying the differences?
We will provide some quantitative results to illustrate this.

L310: incomplete sentence?
Thank you for pointing this out. We will correct the manuscript accordingly.

L341-343: not very clear how successful the snow cover extent simulations are in the first place. The NSE metric is not very intuitive for snow cover extent variable. If for example the extent in area does not quantify, if the snow cover is simulated in the correct location. Similarly as requested for the isotopes, can you provide the timeseries of observed vs simulated snow cover extent to identify and discuss some potential some biases.

Thank you for this comment. We agree that NSE, while commonly used, may not fully capture spatial aspects of snow cover dynamics. Our analysis focuses on the catchment-integrated snow-covered area (SCA), where NSE remains an informative metric for evaluating the agreement between observed and simulated temporal patterns of areal extent. To better illustrate model performance, we have included the time series of observed vs. simulated SCA in Figure 3, along with corresponding comparisons for glacier mass balance (GMB) and isotopic signatures in Figures 4 and 5. These figures allow the reader to visually assess the temporal dynamics and potential biases for each variable. As detailed in Sections 3.3 and 2.2, the figures also represent the posterior predictive uncertainty ranges for streamflow, SCA, GMB, and isotopic data. These will be provided as part of the Supplementary Material of the revised manuscript. We believe these additional results will help clarify how successful the simulations are in reproducing the observed seasonal and interannual variability.

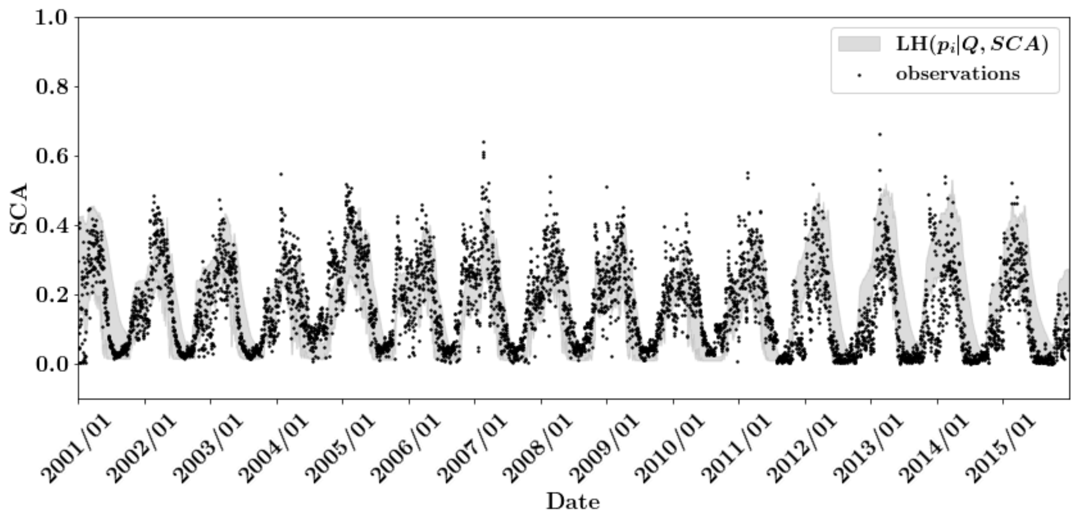

Fig. 3. Observed Snow Cover Area and posterior model ensemble LH(p_i | Q, SCA), based on joint conditioning with streamflow and snow cover area information.

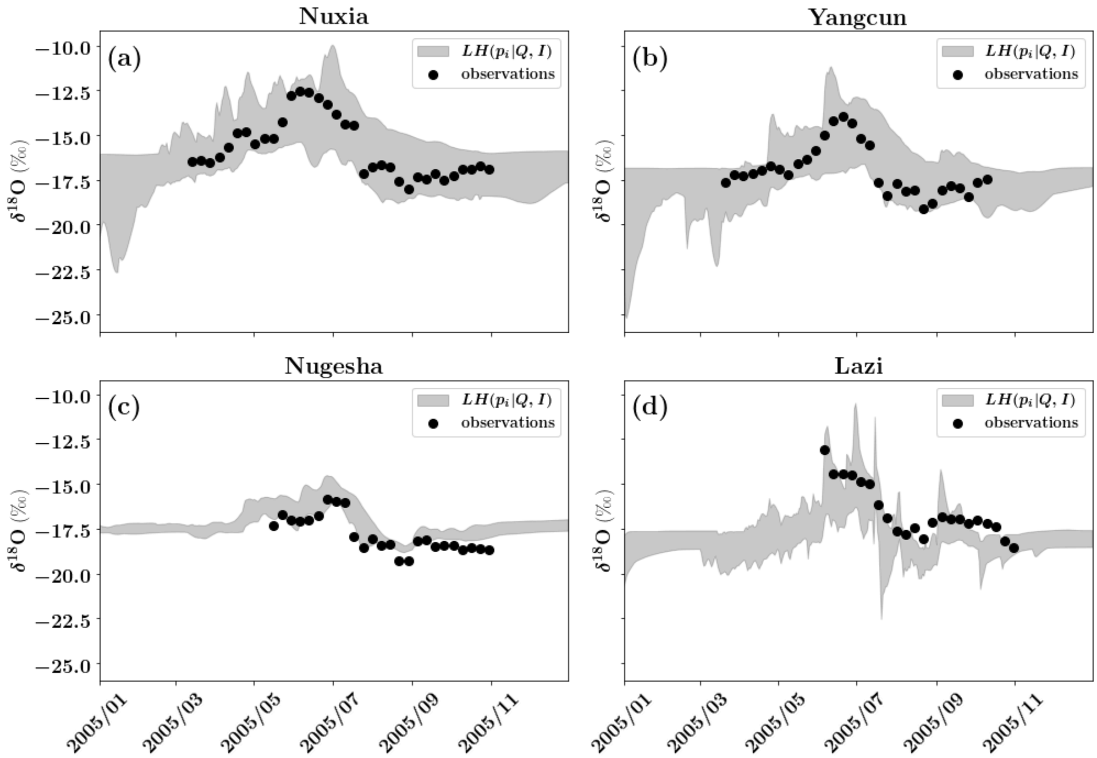

Fig. 4. Comparison between observed $\delta^{18}O$ and the model posterior ensemble $LH(p_i \mid Q, I)$, conditioned on streamflow and isotope information.

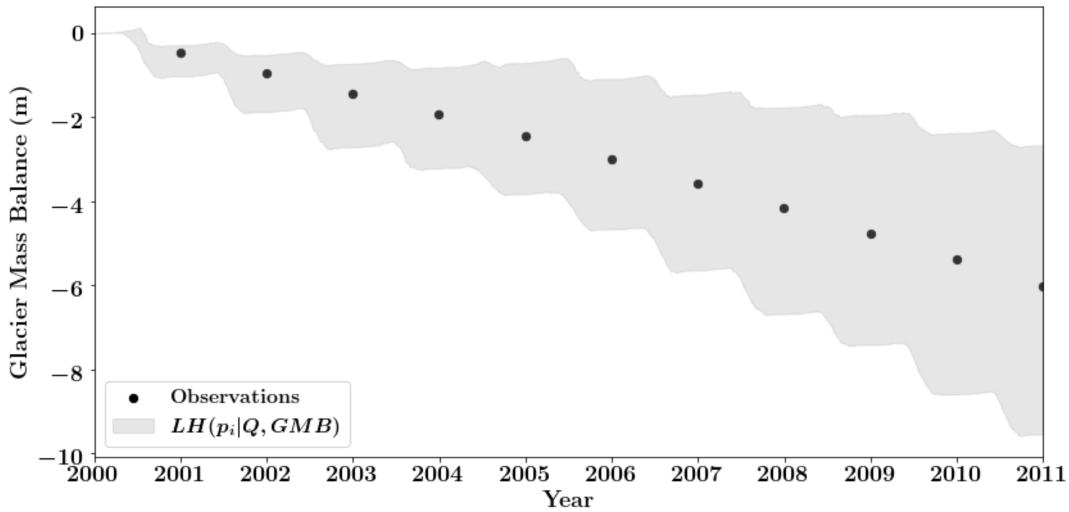

Fig. 5. Observed glacier mass balance and posterior model ensemble $LH(p_i \mid Q, GMB)$, based on joint conditioning with streamflow and mass balance information.

Bibliography:

Boral, S., & Sen, I. S. (2020). Tracing 'Third Pole' ice meltwater contribution to the Himalayan rivers using oxygen and hydrogen isotopes. Geochemical Perspectives Letters, 13, 48–53. https://doi.org/10.7185/geochemlet.2013

Chen, X., D. Long, Y. Hong, C. Zeng, and D. Yan (2017), Improved modeling of snow and glacier melting by a progressive two-stage calibration strategy with GRACE and multisource data: How snow and glacier meltwater contributes to the runoff of the Upper Brahmaputra River basin?, Water Resour. Res., 53, 2431–2466, doi:10.1002/2016WR019656

Di Marco, N., Avesani, D., Righetti, M., Zaramella, M., Majone, B., & Borga, M. (2021). Reducing hydrological modelling uncertainty by using MODIS snow cover data and a topography-based distribution function snowmelt model. Journal of Hydrology, 599, 126020. https://doi.org/10.1016/j.jhydrol.2021.126020

Lutz, A. F., Immerzeel, W. W., Kraaijenbrink, P. D. A., Shrestha, A. B., & Bierkens, M. F. P. (2016). Climate change impacts on the Upper Indus hydrology: Sources, shifts and extremes. PLOS ONE, 11(11), e0165630. https://doi.org/10.1371/journal.pone.0165630

Khanal, S., Lutz, A. F., Kraaijenbrink, P. D. A., van den Hurk, B., Yao, T., & Immerzeel, W. W. (2021). Variable 21st century climate change response for rivers in High Mountain Asia at seasonal to decadal time scales. Water Resources Research, 57, e2020WR029266. https://doi.org/10.1029/2020WR029266

Nan, Y., He, Z., Tian, F., Wei, Z., & Tian, L. (2021), 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(12), 6151–6172. https://doi.org/10.5194/hess-25-6151-2021

Nan, Y., He, Z., Tian, F., Wei, Z., & Tian, L. (2022), Assessing the influence of water sampling strategy on the performance of tracer-aided hydrological modeling in a mountainous basin on the Tibetan Plateau. Hydrology and Earth System Sciences, 26(15), 4147–4167. https://doi.org/10.5194/hess-26-4147-2022

Sun, H., Yao, T., Su, F., Yang, W., and Chen, D.: Spatiotemporal responses of runoff to climate change in the southern Tibetan Plateau, Hydrol. Earth Syst. Sci., 28, 4361–4381, https://doi.org/10.5194/hess-28-4361-2024, 2024.