

Dear Editor,

We thank you for handling our manuscript and we appreciate the effort that Referees have put into their assessment. Please enclosed you can find the revised version of the manuscript titled “Reducing Hydrological Uncertainty in Large Mountainous Basins: **The Role of Isotope, Snow Cover, and Glacier Dynamics in Capturing Streamflow Seasonality**”, reference number **EGUSPHERE-2025-664**.

After having carefully read their comments, we believe we fully addressed each point as reported in the attached rebuttal document. Please also find enclosed a pdf document which details in track changes mode all the revisions we included into the revised manuscript.

We do believe that the revised manuscript improved significantly and meets the quality standards of the HESS journal.

Diego Avesani
on behalf of the authors

Address
Department of Civil, Environmental and
Mechanical Engineering, University of Trento,
via Mesiano 77, 38123 Trento, Italy
email: diego.avesani@unitn.it

Reply to Editor and Reviewers

We thank the Editor and the Referees for the valuable comments. Below we reply point to point and describe the modifications introduced in the revised version of the manuscript. Our replies are evidenced in green.

Reply to Editor

Dear authors,

As you have seen, two reviewers have provided excellent, constructive and very detailed comments on your manuscript. They both, overall, appreciate your analysis and think that it can be a very valuable contribution to literature. However, they both also flag a number of critical issues that need to be resolved. I largely agree with that assessment.

From my perspective, the two most relevant points arising are the following:

(1) the choice to limit the distinction of water pathways to only two components is rather simplistic and may lead to misinterpretation of the results. Both reviewers have provided alternative approaches that can provide a bit more process detail and that may eventually strengthen the overall findings of your analysis. It will thus be a good idea to heed the reviewers advice and explore different options to test whether more information about the hydrological functioning can be obtained from defining more endmembers.

(2) although you have replied to the reviewer concern about the data availability, the data policy of HESS is unambiguous: "If the data are not publicly accessible at the time of final publication, the data statement should describe where and when they will appear, and provide information on how readers can obtain the data until then. Nevertheless, authors should make such embargoed data available to reviewers during the review process in order to foster reproducibility. The Copernicus review system allows to define such assets as 'access limited to reviewers' and reviewers must then sign that they will use such data only for the purpose of reviewing without making copies, sharing, or reusing. In rare cases where the data cannot be deposited publicly (e.g., because of commercial constraints), a detailed explanation of why this is the case is required. " (https://www.hydrology-and-earth-system-sciences.net/policies/data_policy.html). I thus politely request you to follow this policy and to add the required information in the revised version of the manuscript.

Once you have addressed and incorporated these and all other reviewer comments, I am looking forward to receiving a revised version of your manuscript.

Best regards,
Markus Hrachowitz

Reply

We thank the Editor for his assessment and for the opportunity to submit a revised version of the manuscript. We took in great considerations all Referees' comments and in the revised manuscript we introduced the following modifications:

- Regarding the data availability issue, we fully acknowledge the data policy of HESS and are committed to complying with it. As the Yarlung Tsangpo River is a transboundary river, streamflow measurement data are classified as nationally confidential by Chinese authorities and cannot be publicly released. This constraint is also emphasized in a recent perspective article (Lin et al., 2024), which highlights the particular sensitivity of water data in transboundary river basins and regions affected by geopolitical tensions. Consequently, we had to obscure the y-axis in figures that display observed streamflow. Nevertheless, in line with the journal's policy, we have made the simulated streamflow data openly available via Zenodo, and have provided the corresponding link in the revised Data Availability section. We believe that sharing the simulation outputs – together with a clear statement in the manuscript explaining the data limitations – offers transparency and supports reproducibility to the extent permitted by national regulations.

Reply to Review 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 work 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:

Reply

We thank the Referee for the overall positive assessment of our study and the encouraging comment.

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.

Reply

We thank the reviewer for raising this point. We have added more detailed descriptions of the isotope characteristics for various water bodies, see lines 103-110 of the revised manuscript. Specifically:

- Precipitation: We now clarify that our previous evaluation of isoGSM (Nan et al., 2021) showed it can reasonably capture the seasonal variation in

precipitation $\delta^{18}\text{O}$, though it tends to overestimate values and struggles with event-scale variability (see Figures S1 and S2 of supplementary material). To address this, we used a corrected isoGSM product developed in Nan et al. (2022), which adjusts the original isoGSM values through a regression-based bias correction with altitude. Importantly, this corrected product assimilates observed $\delta^{18}\text{O}$ data when available, so it is not appropriate to compare it directly against observations; we therefore present comparisons only between the original isoGSM and the observed data, to illustrate the model's raw performance.

- Glacier melt: We now explain that glacier melt $\delta^{18}\text{O}$ was estimated using the offset-parameter method, assuming a constant value 5‰ lower than the altitude-weighted average of local precipitation $\delta^{18}\text{O}$. This offset was based on data from Boral and Sen (2020), and the value is supported by previous studies in similar environments.
- Groundwater: While groundwater samples were not available for isotope validation, we discuss that groundwater $\delta^{18}\text{O}$ typically shows low temporal variability compared to precipitation or streamflow, due to the long residence time. This characteristic has been included in our discussion of isotope contributions.
- Snowmelt: We were also unable to collect snowmelt $\delta^{18}\text{O}$ samples. Consequently, isotope likelihoods did not significantly constrain snow simulations or snowmelt runoff estimates. These remain primarily informed by snow cover area data, as described in the revised methods section.
- Streamflow: We have added a figure (now Figure~S5 in the Supplementary Material) that compares the observed and simulated $\delta^{18}\text{O}$ in streamflow, to demonstrate the model's capability in reproducing isotopic dynamics.

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.

Reply

We thank the reviewer for raising the important point regarding the definition of runoff components. In the revised manuscript, we have added a discussion to clarify the rationale behind our choice and how it compares with previous studies.

Specifically, in lines 146–150 of the revised version, we now explain that two common approaches exist to define runoff components: (1) by source (rainfall, snowmelt, glacier melt) and (2) by generation pathway (surface vs. subsurface runoff). Since groundwater is also important in the region, reporting results for both sets of definitions would require parallel accounting (e.g., rainfall 80%, snowmelt 10%, glacier melt 10%; surface runoff 40%, subsurface runoff 60%), which could cause confusion. Therefore, we adopted a hybrid approach and defined four components to provide a more informative and concise framework.

We also added that the estimated contribution of each component depends strongly on the definition adopted and the datasets used for model validation. The relatively low contribution of snowmelt and glacier melt in our results is partly due to the inclusion of a large share of subsurface runoff. However, our estimates are consistent with other studies that use snow and glacier data for model validation. For example, Chen et al. (2017) reported contributions of 10.6% and 9.9% for snowmelt and glacier melt, respectively, while Zhang et al. (2025) reported 6.0% and 6.2% using different normalization schemes. When adopting similar definitions, our estimates closely align with theirs, reinforcing the robustness of our results.

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)

Reply

Thank you for this helpful clarification. We agree that the term “groundwater–surface water interactions” may be misleading in the context of our study, as our model does not explicitly simulate these processes. In the revised manuscript, we have removed this term and now refer more appropriately to “subsurface flow,” which better reflects the structure of our model and is consistent with the scope of our analysis. This revision also aligns with the observation raised by Reviewer 2.

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

Reply

Thank you for the suggestion. We acknowledge the redundancy in the original paragraph and have revised it in the updated manuscript to make the research questions more concise and clearly focused.

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?

Reply

Thank you for raising this point. Some studies have indicated that, due to the relatively wet conditions in the YTR basin, sublimation losses are minor, typically accounting for only 2–3% of annual snowfall (Lutz et al., 2016; Khanal et al., 2021). Accordingly, and consistent with other modeling studies in the region (e.g., Chen et al., 2017; Sun et al., 2024), we did not include snow sublimation processes in our model. This assumption has now been clarified in the revised manuscript at lines 125-128.

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.

Reply

Thank you for your observation. The red markers in the figure represent simulations that lie on the Pareto front, identified in the bi-objective space illustrated in each panel. The relatively small number of these points (~15) is due to the fact that only a limited subset of simulations are non-dominated with respect to both objectives. We would like to clarify that the Pareto front is computed across the entire simulation ensemble and is not influenced by any behavioral classification criteria. Therefore, the red points should not be interpreted as behavioral simulations but rather as Pareto-optimal solutions based solely on the two plotted performance metrics. This is consistent with the findings of Di Marco et al. (2021), who also reported that the number of Pareto-optimal simulations is often substantially lower than that of behavioral ones, underscoring how multi-objective trade-offs can yield highly selective solution subsets. We have revised the figure caption and the corresponding text in the manuscript to clarify this distinction (lines 229-243 in the revised version).

L238: can you further explain where the prior parameter distributions in Fig.4 comes from. Is it the parameters with >0 NSE for streamflow?

Reply

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 are included in the prior.

We have revised the main text to explicitly describe this criterion to improve clarity (lines 264–271 in the revised manuscript).

L304: I don't fully understand why the sensitive LL parameter does not manifest in the snowmelt fraction.

Reply

Thank you for this insightful comment. 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 have clarified this point in the revised manuscript (lines 274–277).

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?

Reply

We thank the reviewer for this insightful suggestion. We considered conducting statistical tests to assess differences in central tendency or dispersion between the posterior distributions. However, we found that emphasizing these differences visually was more effective in this context. To this end, we added an inset panel in Figure 5, which allows for a clearer comparison between the distributions. The quantification and interpretation of these differences are discussed in the revised Discussion section, where both the sharpness and the containment ratio are used as metrics to characterize the differences between the prior and posterior distributions.

L310: incomplete sentence?

Reply

Thank you for pointing this out. We corrected 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.

Reply

Thank you for this comment. We agree that NSE, while commonly used, may not fully capture spatial characteristics of snow cover dynamics. In our analysis, we focus on the catchment-integrated snow-covered area (SCA), for which NSE remains a useful metric to evaluate the agreement between observed and simulated temporal patterns of areal extent. To better illustrate model performance, we have included the time series of observed versus simulated SCA in Figure S3 of the Supplementary Material, along with analogous comparisons for glacier mass balance (GMB) and isotopic signatures in Figures S4 and S5. These visualizations allow the reader to assess the temporal evolution and potential systematic biases for each variable. As described in Section 3.4, and detailed in lines 375–381, of the revised manuscript, the figures also display the corresponding posterior predictive uncertainty ranges.

L248-249: Why does the KKA shows a noticeable convergence, but not KKD? They both are parameters that control the subsurface runoff outflow rates. Please clarify this point.

Reply

We thank the reviewer for this useful observation. We agree that the contrasting convergence behavior of parameters KKA and KKD warrants clarification. Although both parameters affect subsurface runoff outflow, their functional roles in the model differ. KKA is an exponential coefficient, and even minor variations in its value can produce strong nonlinear changes in the simulated outflow. This sensitivity makes KKA more responsive to the calibration constraints, resulting in a sharper posterior distribution. Conversely, KKD is a linear coefficient, whose impact on runoff is more gradual and can often be offset by compensatory effects from other parameters. This structural compensation reduces identifiability, leading to a flatter posterior and limited convergence. We have clarified this explanation in the revised manuscript (lines 282–287).

Data perspective:

Limited information is provided on the input data of this study. This could hamper the readers to interpret the results.

L79: The four river gauging are only given by names and no other information and data are available. It is recommended to provide details on the coordinates and elevations of the four river gauging stations in this mountainous basin, also their observed periods, frequency, and measurement method. Any observations

errors/failures in the winter low flow and high flow periods? These details are important to interpret the observed and simulated discharge.

Replay

We appreciate the reviewer's suggestion and have now included a table summarizing the basic information of the hydrological stations used in our study. However, we wish to note that, although these stations are part of China's national hydrological monitoring network, detailed metadata—such as measurement protocols or error characterizations under extreme flow conditions—are not publicly accessible. Our access to the discharge data was made possible through personal connections, which reflects broader challenges in water data availability across China. This limitation is consistent with the issues reported by Lin et al. (2023), who emphasized the restricted accessibility and usability of hydrological data in China and called for the development of a national water data infrastructure. These clarifications have been added to the revised manuscript (lines 95–97).

Reply to Review 2:

General comments:

This manuscript focuses on evaluating the value of snow cover area, glacier mass balance, and isotopes in reducing uncertainty and equifinality of hydrological modeling in a large mountainous basin in the Tibetan Plateau. The Bayesian approach and GLUE method are adopted to investigate the research questions. The research topic aligns with the journal scope and the research findings are potentially useful for the readers. I have a few concerns regarding the modeling procedure, the details of the input data, and the interpretation of the results before the paper being accepted for publication.

Additionally, one thing I noticed here is that the time-series simulated and observed discharge does not have a y-axis (Fig.5), which is present on purpose due to data dissemination restrictions mentioned in the caption. However, this is not possible for readers to understand the model performance, and the magnitude of the simulated and observed discharge. A manuscript avoiding showing y-axis of time-series discharge plot in the results could potentially conflict with the basic principle of open science of HESS/Copernicus journals.

Specific comments:

Modeling perspective:

The subsurface is overly-simplified represented in the model. The subsurface flow generated from the model is composed of the subsurface lateral flow (“interflow”) in the unsaturated zone and the baseflow from groundwater to surface water in the saturated zone. These two subsurface flow components are simulated as a sum (L105 and Fig.1). It is thus not possible to conclude the role of groundwater in contributing to the streamflow and the groundwater- surface water interactions. The subsurface lateral flow can be high and not negligible in such large mountainous basin ($>2 \times 10^5 \text{ km}^2$). It is recommended to be cautious in interpreting and concluding the result regarding the baseflow. All mentioning of groundwater baseflow in the manuscript actually refer to the subsurface flow, i.e. the sum of both unsaturated and saturated zone, e.g. on L134, it is subsurface flow, but not baseflow. The presented modeling approach is not able to investigate groundwater alone.

Replay

We thank the reviewer for this detailed and constructive comment. In response, we revised the manuscript to clarify that in our modeling framework, subsurface flow comprises both lateral flow in the unsaturated zone (commonly referred to as interflow) and baseflow from the saturated zone. As these two components are treated as a single aggregated term, it is not possible to explicitly quantify

groundwater contributions or investigate groundwater–surface water interactions.

Accordingly, we have replaced the term *baseflow* with *subsurface flow* throughout the manuscript (e.g., line 134 and similar occurrences), to ensure terminological consistency. A clarification of this definition has also been added in the Methods section. Furthermore, we now explicitly acknowledge this structural simplification as a model limitation, and we caution against interpreting our results in terms of baseflow or groundwater dynamics. Lastly, in places where streamflow during dry periods is discussed, we replaced *baseflow* with *low flow*, to emphasize that we are referring to overall hydrograph behavior, rather than to baseflow in the strict hydrogeological sense.

Regarding the modeling, are the spatial zones delineated the same for both the surface and subsurface? (this could potentially fragment the aquifers located at the boundaries). Is the subsurface flow allowed to cross the delineated boundaries? The conceptualization of the subsurface processes in the model potentially limits the ability of the model for investigating the surface-subsurface interactions. The model limitation should be clearly discussed in Section 4.3

Replay

We thank the reviewer for raising this important point. In the model, only the runoff concentration process through the river network is allowed to cross the boundaries of the simulation units (i.e., Representative Elementary Watersheds, REWs), while runoff generation – both surface and subsurface – occurs entirely within each REW. The model accounts only for shallow groundwater, which is frequently recharged by infiltration, and does not simulate the deeper groundwater cycle. We acknowledge this as a structural limitation of the model, particularly in the context of the Tibetan Plateau, where previous studies have highlighted the existence of deep interbasin groundwater pathways. This limitation is now explicitly discussed in the revised manuscript (lines 480–485).

There are 4 discharge stations, but only the results at Nuxia Station are presented. The results for the other stations should be presented in the Supplementary Information. The authors should also clarify if the conclusions achieved at Nuxia Station are held the same as the other three stations.

Replay

We thank the reviewer for pointing this out. Although four national discharge stations exist in the basin, the data are not publicly accessible and can only be obtained upon request, subject to approval and specific conditions. We were able to acquire discharge records for only two additional stations, Yangcun and Nugesha, in addition to Nuxia. In the revised manuscript, we now include the corresponding results in the Supplementary Information. As shown in Figures

S6 and S7, and explicitly stated in the main text (lines 328–331), we clarify that the conclusions drawn at Nuxia are consistent with those observed at Yangcun and Nugesha, thereby strengthening the robustness of our findings.

Does glacier melt contribute to groundwater recharge? Or is it assumed that all glacier melt goes into streamflow? This assumption should be clear in the text as well.

Replay

We thank the reviewer for this comment. Yes, in our model setup, we assume that glacier melt contributes directly to streamflow via the surface runoff pathway due to the low permeability of glacier surfaces. This modeling assumption has now been explicitly clarified in the revised manuscript (see lines 134–137).

The simple degree-day-factor methods are used to solve snowmelt and glacier melt. Glacier mass balance is estimated with a simple volume-area scaling factor approach. The limitations of these adopted simple approaches for solving snow and glacier melts should be discussed in terms of modeling limitations.

Replay

We thank the reviewer for raising this important point. We agree that the adopted methods are simplified and that using a spatially uniform degree-day factor does not capture the heterogeneity of melt processes across the basin. In the revised manuscript, we now explicitly acknowledge this limitation and explain our rationale for the chosen approach (lines 131–132). Specifically, due to the large spatial extent of the study basin and the need for computational efficiency in the subsequent GLUE analysis, we adopted the degree-day factor method, which, despite its simplicity, is widely used and has proven effective for snow and glacier simulations, particularly at large spatial scales.

L213: how are the Pareto fronts defined? Please justify this threshold used to show the Pareto fronts in Figure 3 and the conclusions obtained from this result relating to this threshold.

Replay

We thank the reviewer for this important observation. The red points shown in Figure 3 correspond to Pareto-optimal solutions defined using the standard dominance criterion: no other solution performs better across all objectives and strictly better in at least one. The blue dashed lines are not part of the Pareto front itself but are included to indicate minimal performance thresholds (e.g., $NSE > 0$) that help distinguish solutions with some predictive skill from those that are clearly inadequate. Specifically, $NSE < 0$ indicates that a model performs worse than the mean of the observations, and $VE < 0$ suggests a glacier mass balance deviation worse than a null model.

We acknowledge that the rationale for choosing these thresholds was not clearly articulated in the original version. In the revised manuscript (lines 256–262), we have now added a dedicated paragraph to explain the purpose of these thresholds and clarify that they are used for interpretative purposes, to better visualize the trade-off space, without altering the actual definition of the Pareto front. We also confirm that the conclusions regarding trade-offs are robust to different threshold selections.

Section 2.1: What is the modelling period? Please detail the start and end dates of the meteorological data sets and the modelling period. Also add details on which years of DEM, land use data, soil data, snow cover, and glacier data are used in this modelling study.

Replay

We thank the reviewer for the request to clarify the usage of datasets in our modeling framework. We have now added a dedicated paragraph in the revised manuscript (line 98) that provides a more concise and structured description of the role of each dataset and its relevance within the modeling period (2001–2015).

L79-93: Are the gridded meteorological satellite data corrected with in-situ station data? How are the different resolutions of various types of gridded spatial data used in the hydrological model? Please provide details on this.

Replay

We thank the reviewer for the helpful comment. We clarify that no additional correction of satellite data using in-situ station observations was performed in our study. However, some of the datasets we use – such as precipitation and temperature from CMFD – already include such corrections as part of their original data processing (e.g., He et al., 2020). The only correction we performed ourselves was for the precipitation isotope data from isoGSM, which we adjusted based on station measurements; this is explained in detail in response to the following comment. These clarifications have been added to the revised manuscript (lines 103–110).

Regarding spatial resolution, we note that the model operates at the scale of representative elementary watersheds (REWs), with an average area of approximately 700 km², which is larger than the resolution of the gridded datasets used. All input data are aggregated to the REW level as areal averages prior to simulation. These clarifications have been added to the revised manuscript (lines 118-120).

The description of the streamflow sampling is very vague, which is simply stated as “Grab samples of stream water were collected in 2005 at four stations..”. Please provide details on how many samples and in which months the samples

were collected. Do the authors have the precipitation (rainfall, snow) isotopes in the same year (2005) or in a different year (2008)? Using streamflow and precipitation isotope data of different years in the same model can be inappropriate.

Replay

We thank the reviewer for this helpful comment. In the revised manuscript, we will provide a table summarizing the details of the precipitation and stream water samples, including the number of samples, sampling periods, and their isotopic characteristics. We confirm that precipitation samples were also collected in 2005, during the same period as the stream water samples. These precipitation isotope data were used to correct the gridded outputs of the isoGSM model, which serve as inputs to our hydrological model. A detailed description of this correction procedure has been added in the Data section of the revised manuscript (see lines 103–110).

How is the precipitation tracer estimated for rainfall and snow individually? This needs to be clarified in the manuscript.

Replay

As explained in the previous response, the precipitation isotope input used in our model was obtained from isoGSM outputs, which were corrected using observed precipitation isotope data. In the model, the isotopic compositions of rainfall and snowfall were assumed to be the same. However, the isotope composition of snowpack and snowmelt was dynamically simulated using mass balance equations for both water and isotopes, consistent with the treatment of other hydrological stores in the model.

Interpretation of the results:

L256-259: The SCA shows a higher influence on the posterior distribution of T0 than the GMB, which does not show the strongest influence as the authors interpreted. Could the authors please clarify why they see GMB as the strongest from this result figure (Fig.4j)?

Replay

Thank you for pointing this out. You are absolutely right: the snow-covered area (SCA) exerts a stronger influence on the posterior distribution of T0 compared to the glacier mass balance (GMB), as clearly shown in Figure 4j. We acknowledge the misinterpretation in our original statement.

L306-308: The isotope data have increased uncertainty of the simulated glacier melt runoff (Fig.6d), but they are helpful to constrain other surface runoff components (rainfall runoff, snowmelt). Please clarify this result.

Replay

We thank the reviewer for this observation. Glacier meltwater is assumed to generate runoff directly through the surface pathway in our model setup. Therefore, its contribution does not involve any partitioning between surface and subsurface components, the aspect for which isotopic data are most informative. As a result, the isotope likelihood does not help constrain the glacier melt contribution, leading to limited or no reduction in its associated uncertainty. This clarification has now been added to the revised manuscript (lines 459-467).

L272-277: Including the isotope data leads to a decreased containing ratio. This means a significant under capture of the extremely low and high streamflow. Why including isotope data has decreased the streamflow simulation performance? Please clarify this result.

Replay

We thank the reviewer for this insightful comment. The observed reduction in the containing ratio (CR) when including isotope data stems from the stronger constraints introduced by isotopic information. Isotopes provide orthogonal insights into flow partitioning and residence times, thereby narrowing the set of parameter combinations that are consistent with both streamflow and isotopic observations. This results in a sharper ensemble, where the spread of simulations is reduced and predictive confidence increases. However, this enhanced sharpness also increases the risk that observed streamflow values fall outside the uncertainty bounds, thus lowering the CR. Compared to other observational targets such as snow cover area (SCA) or glacier mass balance (GMB), which mainly constrain the seasonal water balance and storage dynamics, isotope data exert a stronger influence on internal hydrological processes. This leads to a more selective posterior and a reduced behavioral parameter space, highlighting the trade-off between sharpness and coverage (Beven et Binley, Gneiting et.al., 2007). While this underscores the diagnostic value of isotopic data in improving model consistency, it also suggests that further model improvements may be needed to achieve both sharpness and reliability. This explanation has been incorporated into the revised manuscript at lines 414-423.

Technical corrections:

- L8: It would be helpful to mention which type of hydrological model the THREW-T is in the abstract. i.e. fully-distributed, semi-distributed, or conceptual?
- L78: km2 should be straight upright, not italic. Please correct all formats of the units for similar cases.
- L80, L82: Please add years between which the mean annual precipitation and mean annual temperature are calculated.
- L100: distributed -> semi-distributed?

- L104: what is bare zone? Bare soil, bare rock?
- Figures 2 and 4: avoid using red and green colors together in the same figure to allow readers with colour vision deficiencies to correctly interpret your findings.
- Table 1: table caption should be on top of the table.
- L165: Please correct the formats of Equations 1-6 by following the journal guideline. e.g. the NSE should be straight upright, not italic. The text subscription should be straight upright as well.
- L210-211: NSE, VE should be straight upright, not italic. Please check the format of all such mentioning.
- L266-267, 281: Please remove the parentheses around the Section and Figure numbers, and correct all such mentioning in the manuscript.
- Figure 4 caption i) covered area -> snow covered area.

Replay

We thank the reviewer for these helpful technical corrections. We have carefully revised the manuscript and addressed all the suggested issues.