Response to Editor

Comment:

After considering the comments of the three reviewers and your replies, a major revision of your manuscript is needed. I think the reviewers raised some important concerns about the calibration and the validation approaches, and their descriptions. Specifically, please consider showing in the manuscript or in the supplementary material simulation results for both a calibration and a validation period, as suggested by reviewer 2. Furthermore, please address all methodological limitations highlighted by the three reviewers (e.g., uncertainty due to the lack of local data, choice of a simulated linear temperature increase).

Response:

Many thanks for the efforts of editor and reviewers on providing valuable comments on our manuscript. We have revised the manuscript thoroughly according to the comments. In the previous response during discussion, we explained that the temporal validation was not conducted due to the high consistency of model performance during calibration and validation periods based on our previous studies. Now we have additionally collected the discharge data at outlet station during 1991~2000, and extended the simulation period to 1991~2015. We still calibrated the model based on data in 2001~2015, and validated the discharge simulation during 1991~2000, to bring little change to the calibrated parameters and the sensitivity analysis results. Also we have addressed all the limitations highlighted by the three reviewers. Please find the response to each comment as followed. The responses to each comment are colored in blue, and the revisions we have made to manuscript are colored in red.

Response to Reviewer #1

Comment 1:

This article calibrates a sophisticated hydrological model with diverse datasets in a large basin in the Tibetan plateau and then uses the model to understand sensitivity of hydrological fluxes to possible changes in temperature and precipitation. The entire calibration exercise is very interesting, as it is done with varied dataset (streamflow, snow coverage, glacier mass balance, and stable water isotopes). By calibrating across these datasets, it is assumed that the model parameters closely mimic the underlying hydrologic processes. The article then does climate perturbation studies and predicts that temperature increase changes internal water flux partitioning within the catchment with limited changes in absolute amount of streamflow, whereas precipitation changes has a significant impact on streamflow amounts (and not on flux partitioning). The article also highlights certain interesting threshold processes occurring in this region, where a small increase in temperature may lead to decrease in streamflow whereas a larger increase in temperature increases streamflow due to enhanced glacial melt. Additionally, there are lots of other small interesting results within the article which might be very relevant to researchers working in this region.

The flow of the article is well drafted and it explains modeling and results reasonably well. I have a few suggestions to further improve the article and a few questions regarding the model application.

Response:

Thank you very much for your appreciation. We have revised the manuscript thoroughly according to your comments.

Comment 2:

A key advantage of using a hydrologic model is that it allows estimating uncertainty within the modeling result. However, the article does a very poor job in describing uncertainties in model outputs. Except in Figure 5(a, d), uncertainty in model outputs have not been described in text or in figures. For e.g. in Table 4, change in CP days (<2 days in the most extreme scenario) have been inferred which might be deceiving unless the uncertainty bounds introduced by natural climate variability is taken into account. It is quite possible that the day of peak runoff varying by 2 days can be explained by natural variability. To ensure the robustness of these results, I suggest adding uncertainty to most plots and do statistical significance tests. I have pointed this out findings in my more detailed comments below. My general suggestion will be to introduce uncertainty bounds in every figure (wherever feasible) and include statistical significance of any future results (e.g. runoff decrease of 20%under a given scenario at what significance level?). This will make the discussions more robust.

Response:

Thank you very much for your suggestion. The model uncertainty is indeed an important issue for runoff projection and sensitivity analysis. A common method to estimate the uncertainty in runoff projection studies is to adopt the output of multiple general circulation models (GCM) to drive the model and provide the uncertainty range of results produced by different GCM product (e.g., Lutz et al., 2014; Cui et al., 2023). However, in this study, the model was driven by only one dataset (i.e., perturbing the historical temperature/precipitation by a certain degree) in each scenario, thus the uncertainty range cannot be produced in a similar way. Actually, the bounds in Figure 5a-d do not refer to the uncertainty, but the range of interannual variability, to show the conditions in relatively wet/dry years. Based on the above considerations, we adopted the standard deviation (STD) of annual variables (runoff, contribution of runoff components, CP/CR) to represent the uncertainties introduced by natural climate variability. We have added the STD of annual variables in Figure 4~7 and Table 4. In addition, we have conducted the t-Test of paired two samples for these variables in the baseline and perturbed temperature/precipitation scenarios. We have clarified this (L285~L287) and described the statistical significance in the revised manuscript (e.g., L369~L370, L386~L388, L425~L428, L503~L512).

Comment 3:

One key limitation in the model calibration is that the calibration period uses the entire period of collected hydrometeorological data, and there is no proper validation period. In a way, the article replaces time with space, where calibration is done over time and validation is done scantly over space. Please provide a justification for this calibration scheme, and why hasn't the model been calibrated and validated in a conventional way?

Response:

Thank you very much for pointing out this limitation. The main reason for such calibration scheme was the large basin area. In our previous studies in this region (Nan et al., 2021, 2022), we found that the model performance during validation period was highly correlated with that of calibration period, which was partly due to the strong linearity of precipitation-discharge relation in the large basin. But the model performance at internal stations had a large uncertainty when the discharge at outlet station was simulated well. Consequently, we used the discharge data of internal stations to validate the model. Nonetheless, we have also collected the streamflow data at the outlet station during 1991~2000, and extended the simulation period, to conduct the temporal validation on streamflow simulation. We have clarified the calibration scheme (L240~L248) and added the result of validation period in the revised manuscript (L307~L309).

Comment 4:

The hydrological model consistently underestimates peak streamflow across sites (only exception is 2008 period of Figure 3f). This suggests that the model has a consistent bias which hasn't been adequately discussed in the article. Please provide more context to this, why this bias exists and what should be done to remove this bias.

Response:

Thank you very much for pointing out this shortage of streamflow simulation. The major difficulty of hydrological simulation in this region is the uncertainty in precipitation input data. The Tibetan Plateau is a typical data scarce region, especially in high elevation regions, making the precipitation product poorly validated by station precipitation data (Li et al., 2021). Previous studies indicated high precipitation in some high elevation regions (e.g., Jiang et al., 2022), which was however not integrated in the gridded precipitation dataset. Results also showed that most precipitation products performed poorly on accurately capturing the amount of specific precipitation events (Xu et al., 2017), which had important influence on the simulations of peak flow. Consequently, the underestimations of peak streamflow could be attributed to the underestimated precipitation, especially for the source region with high elevation. Further correction on the precipitation product based on more station data could be helpful to remove this bias. Nonetheless, our results showed that the annual runoff and seasonality were reproduced relatively well for most stations, and our study did not focus on the sensitivity of flood processes. Thus we believe that the sensitivity analysis results are mostly meaningful. We have added description and discussion about the bias in the revised manuscript (L305~L307, L669~L672).

Comment 5:

L201-203: How different is glacier meltwater isotopes from snowmelt isotopes, and how have these two been differentiated within the tracer module of the model?

Response:

The isotope composition of glacier meltwater is very steady, because the glacier is formed by the precipitation within past hundreds of years and updated slowly. On the contrary, the snow is updated quickly in each year as indicated by the intra-annual variation of snow cover area. In the tracer module, the isotope composition of glacier meltwater was assumed as a constant value, which was lower than the average precipitation isotope to reflect the fractionation effect during melting process. The isotope composition of snowmelt was updated based on the water and isotope mass balance of the snowpack, similar with other water storages. We have illustrated it in the revised manuscript (L202~L207).

Comment 6:

In the mathematical equations, a lot of variables are depicted by small words (like PET). I will suggest using a single alphabet for a variable, in-line with proper mathematical convention. Please address this in all the equations.

Response:

Many thanks for your suggestion. We have changed the small words in Eq. 7 to single alphabets. However, the evaluation metrics such as NSE were not revised.

Comment 7:

Figure 4. Looking at difference between Y-axis of Fig 4a and 4d, its clear that annual runoff is not significantly impacted by increasing temperature. Is the conclusion about reduction in annual runoff in different temperature scenarios statistically significant?

Response:

Many thanks for your comments. We have conducted the t-Test of paired two samples for annual runoff and the reduction in different temperature scenarios is statistically significant, although the decreasing rate is small. We have described statistical significance of runoff change in the revised manuscript (L369~L370, L386~L388).

Comment 8:

Figure 5. One interesting takeaway from the figure is that precipitation increase has a disproportionate impact on summer month runoff whereas temperature increase has higher impact on winter runoff. This is a very interesting result as it suggests winter baseflow is more influenced by temperature changes. This should be highlighted in the text.

Response:

Many thanks for your high evaluation on this result. We have highlighted it in the revised manuscript (L414~L418).

Comment 9:

L366-367: concentration ratio and concentration period have not been defined in the text. Please define them

Response:

Many thanks for your suggestion. We have added the definitions of equations of CR and CP (L290~L297, Eq. 8~10).

Comment 10:

L369-372: Was CP decrease of 2 days statistically significant?

Response:

Thank you very much for your question. We have conducted the t-Test of paired two samples for CR and CP. We found that the change of CR is significant at significance level of 0.01 in all scenarios, but the change of CP is insignificant in some scenarios. In specific, the decrease in CP for the warming of 1℃ is insignificant at the level of 0.01 ($p=0.014$), but significant in higher warming scenarios. The changes of CP in two reducing precipitation scenarios are significant, but is insignificant in two scenarios with increasing precipitation ($p=0.02$ and 0.12 for 110% P and 120% P, respectively). We have added these results in the revised manuscript (L425~L428).

Comment 11:

L380-383: if subsurface runoff is 70%, rainfall runoff 20%, glacial melt runoff 10% and snowmelt runoff 5%, then where is the remaining 5% water flux? I suggest using uncertainty bounds (or standard deviation values) with these fluxes to solve these issues

Response:

Many thanks for your comments. This is an inaccurate description of the results. The contributions of subsurface runoff, rainfall runoff, glacier melt runoff and snowmelt runoff in annual runoff are 67.8%, 21.6%, 7.4% and 3.2%, respectively. We have described them more accurately in the revised manuscript (L434-L438), and added the standard deviations of the proportions (Figure 6, Supplementary Table 2 and 4).

Comment 12:

L404-408: I don't understand the discussion around proportion of three other components. Also these values look very very small to make meaningful conclusions. Please include error bands around these values

Response:

I am sorry to make you confused about this. Here we are just reporting the changing rate of the components' contribution with the changing precipitation. We have added the standard deviation of these annual values in the revised manuscript (Supplementary Table 1~4).

Comment 13:

L453-457: This is a very interesting result, it highlights the most dynamic regions within the basin, which can keep shifting between energy vs water limited stages. Have past studies identified such regions? If yes, please mention them in discussions. Also, please highlight this part in the abstract and conclusions part.

Response:

Many thanks for your high evaluation on this result. As far as we know, there are no past studies identifying such regions. The past studies in this region focused on the hydrological responses either at the large basin scale (e.g., the whole YTR basin) or at the sub-catchment scale (e.g., the Lhasa River basin), but did not explore the spatial pattern of catchment scale hydrological characteristics within the whole basin, thus cannot identify the particularity of such regions. We have highlighted it in the abstract (L39~L40), main text (L628~L629) and conclusions (L718~L720).

Comment 14:

L478-479: Please show the other figures with insignificant correlations in the supplementary material

Response:

Many thanks for your suggestion. We have shown them in supplementary materials (Figure S3).

Comment 15:

L488-489: How can glacier coverage be higher in warmer regions? Shouldn't the presence of glaciers be more in colder places?

Response:

Many thanks for your correction. We made a mistake here and have corrected it.

Comment 16:

Figure 10: This is a useful figure but I think change in runoff is very low in T+5 scenario (maybe bulk of points lie between -10% to 10%). What is the uncertainty range of simulation of runoff?

Response:

Many thanks for your comment. The runoff change in response to warming temperature is indeed very low in most regions, but the change could be rather significant in regions with large glacier area ratio or large precipitation. We have calculated the statistical significance of runoff change in each REW. We find that the runoff change in response to perturbed precipitation is significant in all REWs, but things are different for warming temperature scenarios. The runoff change in response to increasing temperature is insignificant (at significance level of 0.01) in about 26% and 15% area of the basin for warming of 1°C and 5°C, respectively. The statistical significance in response to warming temperature is related to the runoff change magnitude and the drainage area. We have described these results in the revised manuscript (L503~L512), and provided the related figures in the supplementary materials (Figure S1 and S2).

Comment 17:

Figure 10: I suggest reproducing this sort of a figure for increasing precipitation scenario. In the text, it was mentioned that in certain areas streamflow increased by >10% if precipitation increased by 10% and in certain areas streamflow increased by <10% if precipitation increased by 10%. In a way, different regions are showing different streamflow elasticity. I suggest

producing a figure of streamflow elasticity in increasing precipitation scenarios and highlight which areas are closest to 1, as they would likely show non-monotonic behavior i.e. shift from <1 to >1 in different precipitation. Does glacier area ratio play a pivotal role there as well?

Response:

Thank you very much for your valuable suggestion. We reproduced this figure for precipitation scenarios and found similar results (as shown in the figure below). There are some REWs shifting from ΔQ/ΔP <1 to >1, and these REWs forms the boundary of REWs that ΔQ/ΔP <1 and >1. However, there is little difference from the results for temperature scenarios. There are only three non-monotonic REWs, providing less confidence to the boundary line. As a result, there are some REWs lying lower than the boundary line but the streamflow changing rate is lower than that of precipitation. We have described this result in the revised manuscript (L584~L593).

Comment 18:

L542-543: Can you point to the figure which highlights this?

Response:

Many thanks for your suggestion. This could be inferred from Figure 10. We have pointed it in the revised manuscript (L619).

Comment 19:

L553-554: Decrease of 0.1%-3% is likely very small and not statistically significant. Can you clarify?

Response:

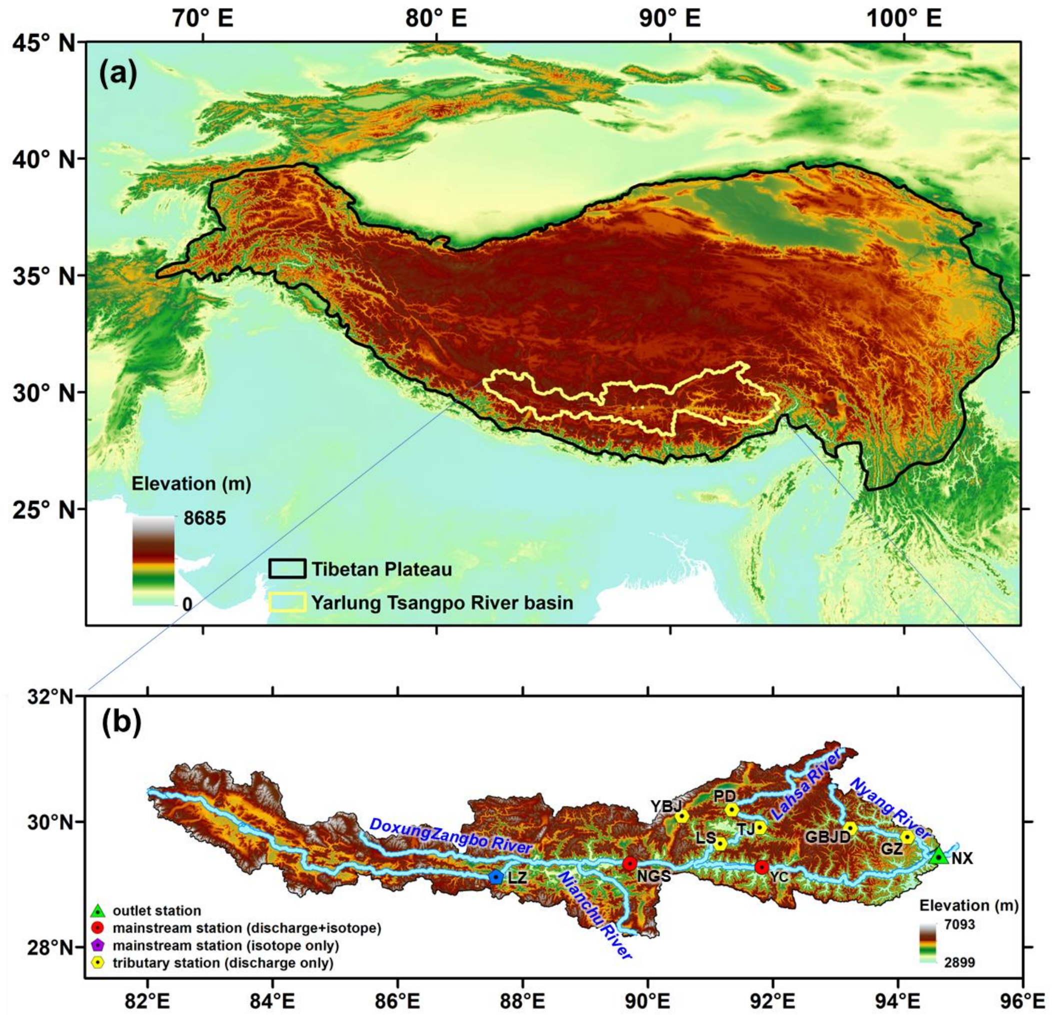
Many thanks for your question. This is the result from Cui et al. (2023) but the statistical significance was not discussed in the paper. We collect the original data from the Supplementary Information of that paper. The projected runoff change for the warming of 1.5°C in the YTR is -2.14(±4.30)%. Although the average change rate is even smaller than the standard deviation, the decreasing trend is statistically significant ($p=0.001$).

Comment 20:

Figure 1b: The yellow color of tributary station YBJ is blending with the yellow color of the underlying DEM. Figure 2b should be redrawn and the contrast between legends and background should be increased. Its currently very hard to read

Response:

Many thanks for your suggestion. We have revised the Figure 1 as below:

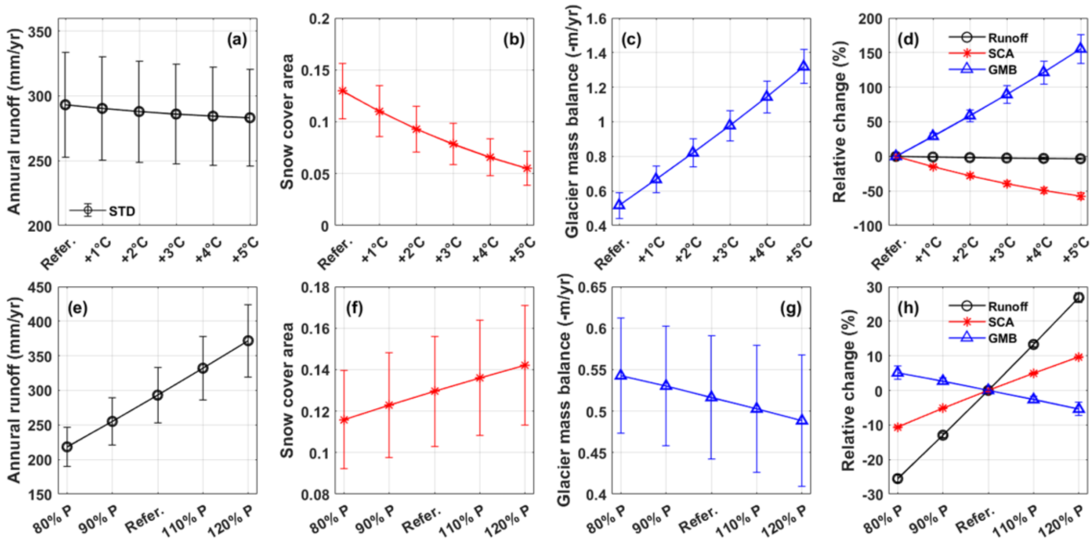


Comment 21:

Figure 4: Please describe subplots (d) and (h) in the figure description. Add error bands in all the subplots

Response:

Many thanks for your suggestion. We have changed the figure description to "The sensitivities of annual runoff, snow cover area, and glacier mass balance to the perturbed temperature (a-d) and precipitation (e-g). Subplots (d) and (h) are the relative changes of runoff, SCA and GMB compared to the reference scenario". Figure 4 has been redrawn as below.

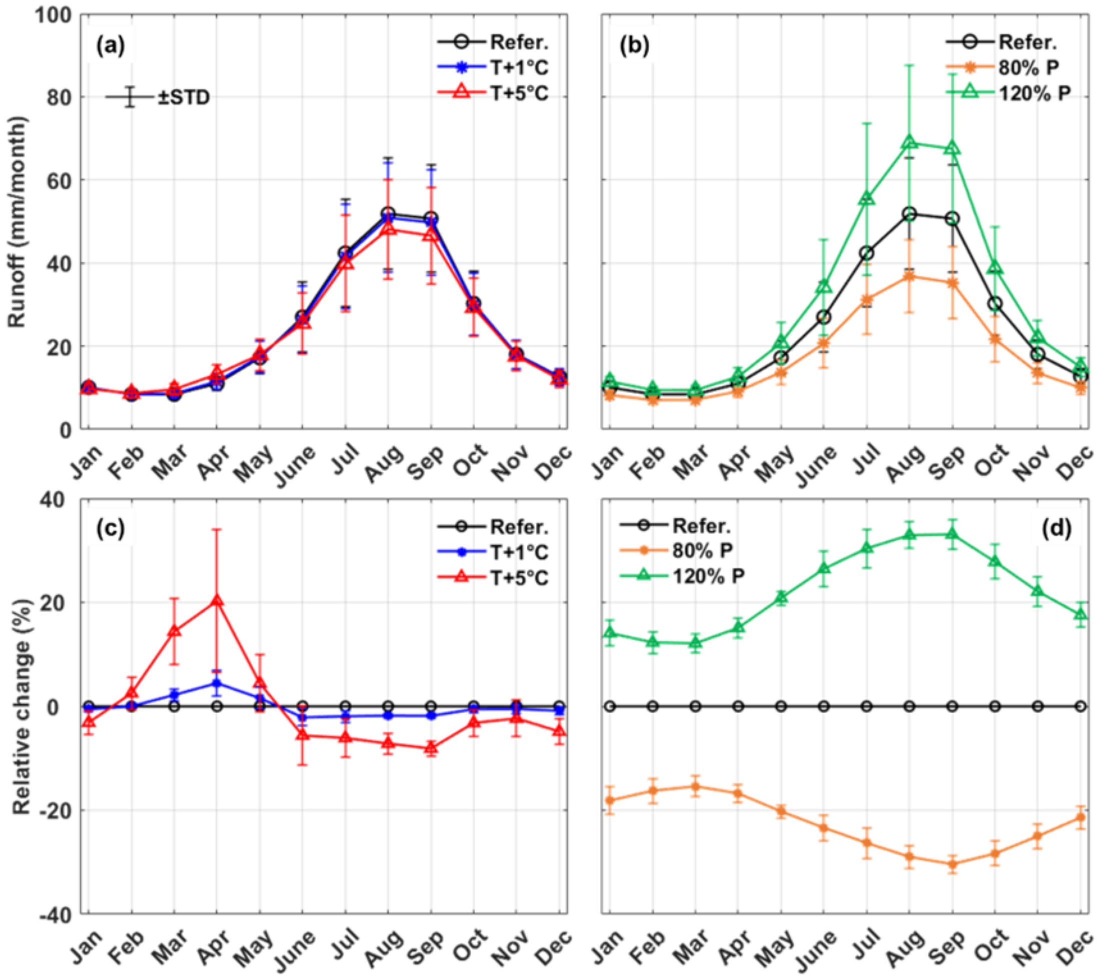


Comment 22:

Figure 5a. X-label 350 matches with "0" of subplot (b) making 350 look like 3500. Please resolve this. Also add error bands to Figure 5e and 5f

Response:

Many thanks for your suggestion. We have redrawn this Figure as below. Because another reviewer thought that the subplot (a) and (b) provided similar information as subplot (c) and (d), we have deleted the original subplot (a) and (b).

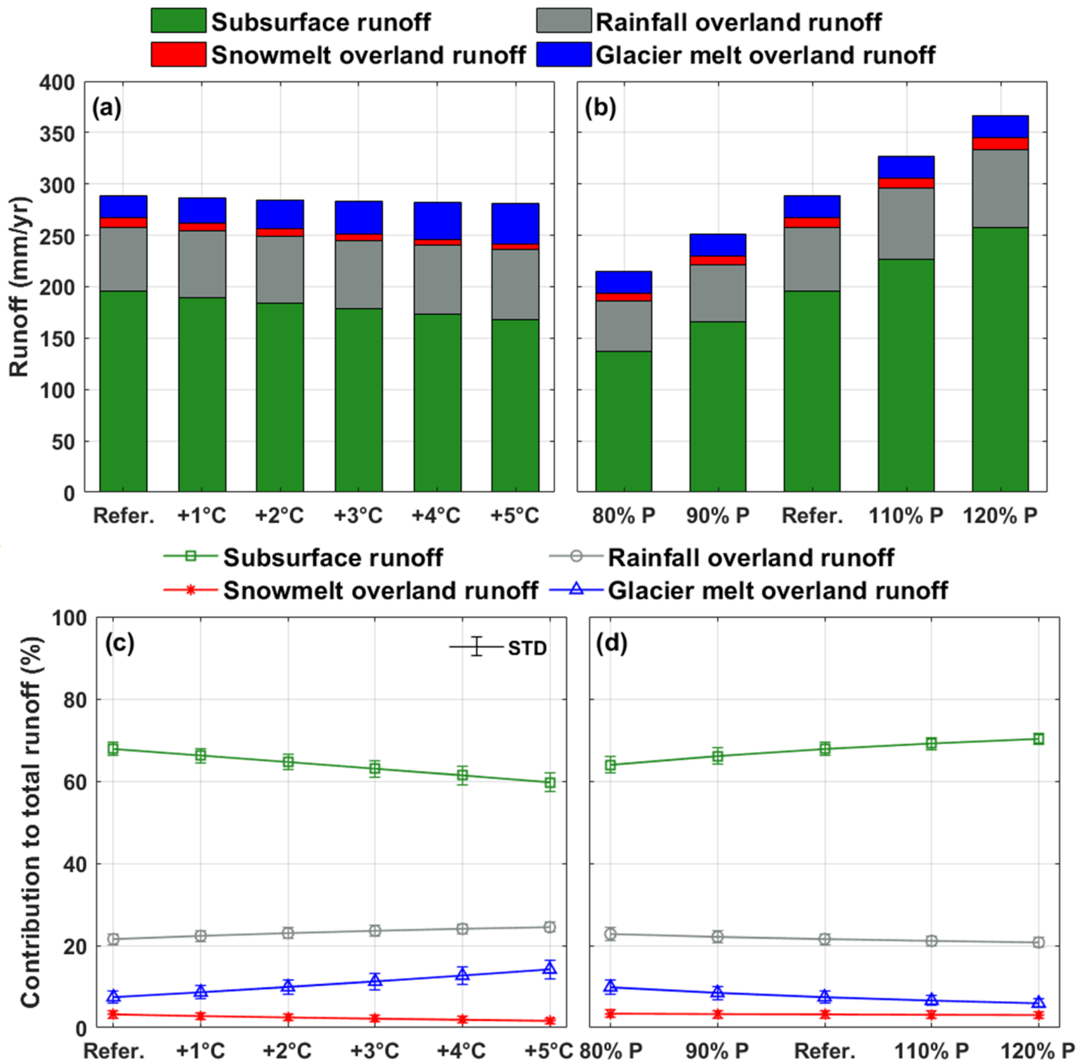


Comment 23:

Figure 6c,d: Add uncertainty bands

Response:

Many thanks for your suggestion. We have redrawn this figure as below. The interannual variations of the contributions of runoff components are small in all the scenarios, as indicated by the small STDs.

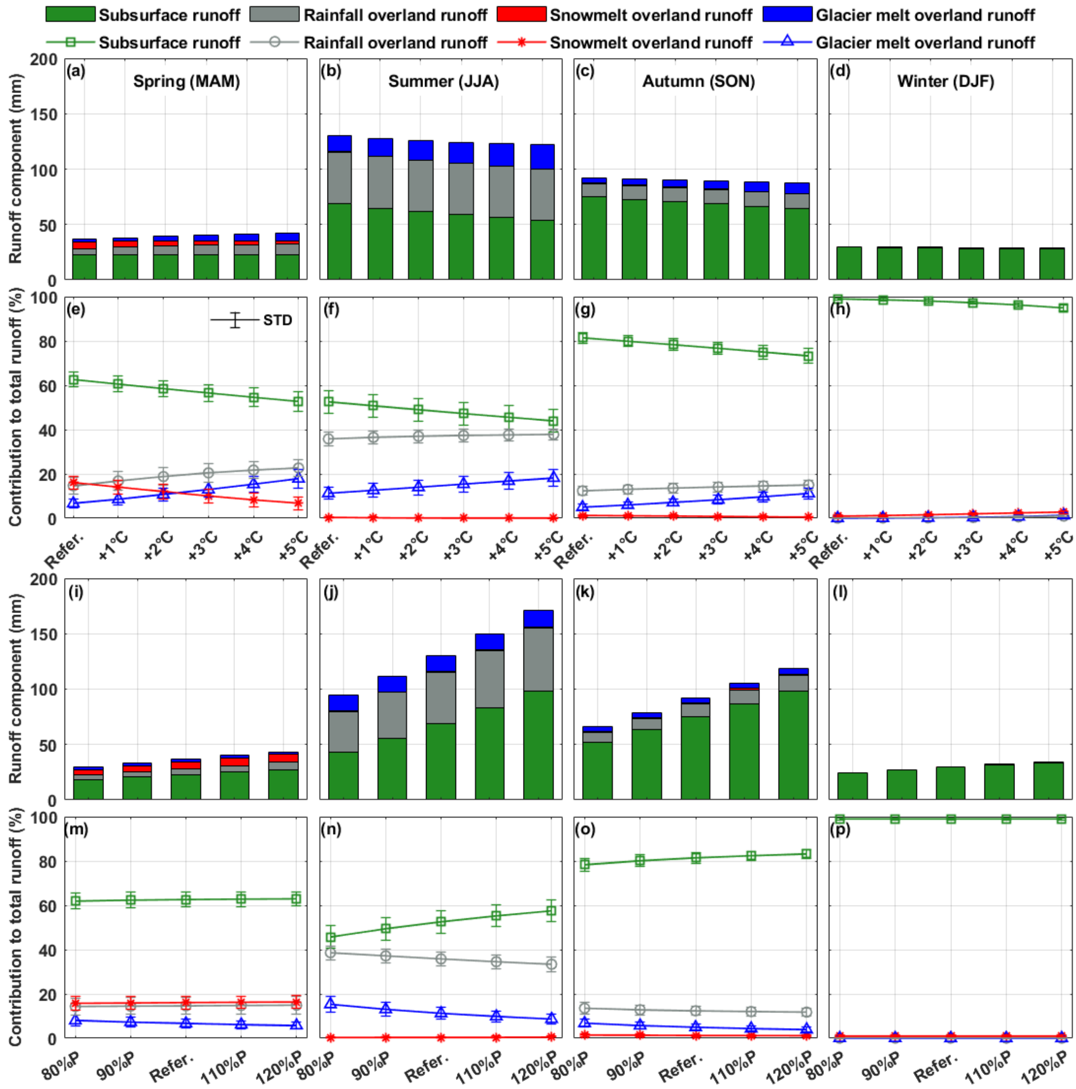


Comment 24:

Figure 7e-h, m-p: Add uncertainty bands

Response:

Many thanks for your suggestion. We have redrawn this figure as below.



Comment 25:

L16: replace exist with “existed”

Response:

Many thanks for your correction. We have changed the phrase accordingly.

Comment 26:

L23: can use “multiple datasets” instead of “Datasets of multiple objectives”. It makes reading easier

Response:

Many thanks for your suggestion. We have changed the phrase accordingly.

Comment 27:

L38-40: I cannot understand what is being said in this line. Please rephrase

Response:

Many thanks for your comments. We have rephrased this sentence as: Some regions had a non-monotonic runoff change rate in response to climate perturbation, which represented the most dynamic regions within the basin, as they kept shifting between energy and water limited stages. The GAR and mean annual precipitation (MAP) of the non-monotonic regions had a linear relation, and formed the boundary of regions with different runoff trends in the GAR-MAP plot.

Comment 28:

L81: replace “processes” with “models”

Response:

Many thanks for your correction. We have changed the phrase accordingly.

Comment 29:

L97: replace “likened” with “compared”

Response:

Many thanks for your correction. We have changed the phrase accordingly.

Comment 30:

L121: remove “and”. It can be better framed as “Snow, glacier, isotope data and observation ..”

Response:

Many thanks for your suggestion. We have reframed this sentence accordingly.

Comment 31:

L140: Missing unit “km2” next to 2x10^5

Response:

Many thanks for your correction. We have added the unit in the revised manuscript.

Comment 32:

L145-146: Why is the acronym of Yangjia “TJ”? It is not at all intuitive

Response:

Many thanks for your suggestion. We made a mistake and it should be Tangjia. We have corrected it in the revised manuscript.

Comment 33:

Table 2: Instead of using column name as “Sample number”, use “Number of samples”. Sample number gives the impression that it’s the laboratory sample number of a collected water sample

Response:

Many thanks for your suggestion. We have changed the term in the revised manuscript.

Comment 34:

Eq.6: Please move the equation to L255 as it is a continuation of that sentence

Response:

Many thanks for your suggestion. We have moved equation 6 (the current equation 7) forward.

Comment 35:

Table 4: Add error bands (or standard deviation values) in this table

Response:

Many thanks for your suggestion. We have calculated the STD values and revised the table as below.

		CR		CP (days)	
		Average	STD	Average	STD
Reference scenario		0.432	0.044	244.4	7.09
T scenario	+1°C	0.425	0.044	244.1	7.12
	+2°C	0.419	0.045	243.8	7.18
	+3°C	0.413	0.045	243.3	7.26
	+4°C	0.408	0.046	242.8	7.36
	+5°C	0.402	0.046	242.3	7.45
P scenario	80%	0.398	0.039	242.2	6.86
	90%	0.415	0.042	243.6	7.01
	110%	0.449	0.045	244.7	7.13
	120%	0.465	0.045	244.7	7.14

References

Cui, T., Li, Y., Yang, L., Nan, Y., Li, K., Tudaji, M., Hu, H., Long, D., Shahid, M., Mubeen, A., He, Z., Yong, B., Lu, H., Li, C., Ni, G., Hu, C., and Tian, F.: Non-monotonic changes in Asian Water Towers' streamflow at increasing warming levels, Nature communications, 14, 1176-1176, 10.1038/s41467-023-36804-6, 2023.

Li, K., Tian, F., Khan, M. Y. A., Xu, R., He, Z., Yang, L., Lu, H., and Ma, Y.: A high-accuracy rainfall dataset by merging multiple satellites and dense gauges over the southern Tibetan Plateau for 2014-2019 warm seasons, Earth System Science Data, 13, 5455-5467, 10.5194/essd-13-5455-2021, 2021.

Lutz, A. F., Immerzeel, W. W., Shrestha, A. B., and Bierkens, M. F. P.: Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation, Nature Climate Change, 4, 587-592, 10.1038/nclimate2237, 2014.

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, 2021.

Nan, Y., He, Z., Tian, F., Wei, Z., and Tian, L.: 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, 4147-4167, 10.5194/hess-26-4147-2022, 2022.

Xu, R., Tian, F., Yang, L., Hu, H., Lu, H., and Hou, A.: Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network, Journal of Geophysical Research-Atmospheres, 122, 910-924, 10.1002/2016jd025418, 2017.

Jiang, Y., Yang, K., Yang, H., Lu, H., Chen, Y., Zhou, X., Sun, J., Yang, Y., and Wang, Y.: Characterizing basin-scale precipitation gradients in the Third Pole region using a high-resolution atmospheric simulation-based dataset, Hydrol. Earth Syst. Sci., 26, 4587–4601, https://doi.org/10.5194/hess-26-4587-2022, 2022.

Response to Reviewer #2

Comment 1:

This work aims to disentangle the long-standing question of hydrological sensitivities to climate change. This is done by feeding climate forcing into hydrological simulations. Specifically, the climate forcing is perturbed in order to assess the sensitivity of the hydrological response in a large mountainous basin. Of particular interest is the use of a detailed hydrological model capable of modelling not only streamflow but also snow cover and glacier mass balance, both key elements for medium-term water management and water resources optimisation. The authors point out that the novelty of the work reside in both the assessment of "patterns and drivers of local hydrological sensitivities" and to the case study. This is also due to the fact that there are contrasting results from similar analyses in the same region of interest. However, in my opinion, one of the most interesting features of this work is the adoption of a sort of multi-objective calibration based on streamflow, snow cover area, glacier mass balance and stream water isotopes, alongside with a spatial validation of the calibration. Therefore, I believe that this paper is worth publishing in HESS after answering the following questions.

Response:

Many thanks for your appreciation on our work. We have revised the paper thoroughly according to your comments.

Comment 2:

The model calibration is one of the weakest parts of the papers. There is no reference to the algorithm used as well as all the details regarding its applications which, in my humble opinion, should not only be mandatory to ensure the replicability of the paper, but are also important to understand the calibration results. For example, if the PSO scheme is adopted, the sensitivity analysis of the hydrological model parameters is not possible. Conversely, a Monte Carlo procedure could help to define the confidence band of the hydrological results. This should be explained according to the desired research objectives.

Response:

Many thanks for your comment. The calibration procedure was indeed described simplistically in the manuscript, because the main aim of the paper was to explore the hydrological sensitivity to perturbed climate, and only one set of parameters were adopted to derive the results. In specify, we adopted an automatic algorithm, the Python Surrogate Optimization Toolbox (pySOT, Eriksson et al., 2019) for model calibration. The pySOT algorithm uses radial basis functions (RBFs) as surrogate models to approximate the simulations, reducing the time for each model run. The symmetric Latin hypercube design (SLHD) method was used to generate parameter values, allowing an arbitrary number of design points. In each optimization run, the procedure stopped when a maximum number of allowed function evaluations was reached, which was set as 3000. In this study, we repeated the pySOT algorithm for 100 times, and a final parameter set was selected from the calibrated parameter sets manually based on the overall performance on multiple objectives. We acknowledge that the parameter calibration and uncertainty analysis are rather weak in this study, which is due to, as you said, the desired research objectives. We have added the details of calibration procedure in the Method section (L228~L236), and addressed the shortage in the Limitation section (L677~L683).

Comment 3:

I have some doubts about the objective functions. If I understand correctly, the model seems to have been calibrated using a single objective function composed of four equally weighted functions (NSE for streamflow at one site; NSE for isotope; RMSE for SCA; and RMSE for GMB). In this way, however, the authors mix different types of metrics. I try to be clear: the NSE varies between minus infinity and one, while the RMSE varies from 0 to plus infinity (theoretically, of course). As a consequence, different metrics have different impacts on the aggregated objective function. I therefore believe that the NSE (or RMSE) must be used for all the objects under consideration. Otherwise, one factor could be weighted more than the others. I would suggest to have a look at "Madsen, Henrik. (2003). Parameter estimation in distributed hydrological catchment modelling using multiobjective automatic calibration. Progress in Water Resources. 26. 205-216. 10.1016/S0309-1708(02)00092-1.,Section 2.2, Equation 3.

I also disagree with the definition of multi-object calibration. Basically, the authors use a single-objective calibration. In other words, they choose a specific solution in the multi-objective space. I believe that this strategy is fine if the aim of the paper is to focus solely on the uncertainty of climate forcing, although a more precise definition of the objective function and calibration is required, in my opinion.

Response:

Many thanks for your comments. The main reason that we mixed different types of metrics are as followed: (1) We found that NSE is not suitable for evaluation of objectives with strong essentially fluctuation (Schaefli et a., 2007), such as SCA and GMB. The NSE could be very low even if the simulation looks good. Meanwhile, error indexes such as RMSE and MAE were widely used for SCA/GMB simulation in previous studies (e.g., He et al., 2019; He et al., 2021; Lyu et al., 2023). (2) Although different metrics have different dimensional unit and range, previous results indicated that the values of these metrics are of the same order of magnitude when the model performance is acceptable (He et al., 2019; Ala-aho et al., 2017; Nan et al., 2021). (3) We did not simply adopt the parameter set which produced the best integrated objective function, but recorded all the parameters produced during the calibration procedure and manually selected one parameter set with best overall performance. Thus we believe that the choice of objective function had little influence on the main findings of this study. We have added the description of objective function selection in the revised manuscript (L222~L225).

We agree with you that the expression "multi-objective calibration" is inappropriate. What we mean here is that the model was calibrated toward datasets related to multiple objectives, but in the calibration practice a single objective function was adopted.

Comment 4:

Model calibration and evaluation section. I propose not only to carry out a spatial validation of the hydrological model, but also a temporal validation with a calibration period and then a test period. The hydrological response could then be carried out taking into account all 15 years from 2001 to 2015. At this stage it is indeed important to ensure the reliability of the hydrological model.

Response:

Many thanks for your suggestion. The main reason for such calibration scheme was the large basin area. In our previous studies in this region (Nan et al., 2021, 2022), we found that the model performance during validation period was highly correlated with that of calibration period. In specify, when the NSE during calibration period was in the range of 0.86~0.94, the NSE during validation period was also at a high level within the range of 0.77~0.92. This could be partly due to the strong linearity of precipitation-discharge relation in the large basin. But the model performance at internal stations had large uncertainties when the discharge at outlet station was simulated well. Consequently, we used the discharge data of internal stations to validate the model. Nonetheless, we have also collected the streamflow data at the outlet station during 1991~2000, and extended the simulation period, to conduct the temporal validation on streamflow simulation. We have clarified the calibration scheme (L241~L248) and added the result of validation period in the revised manuscript (L307~309).

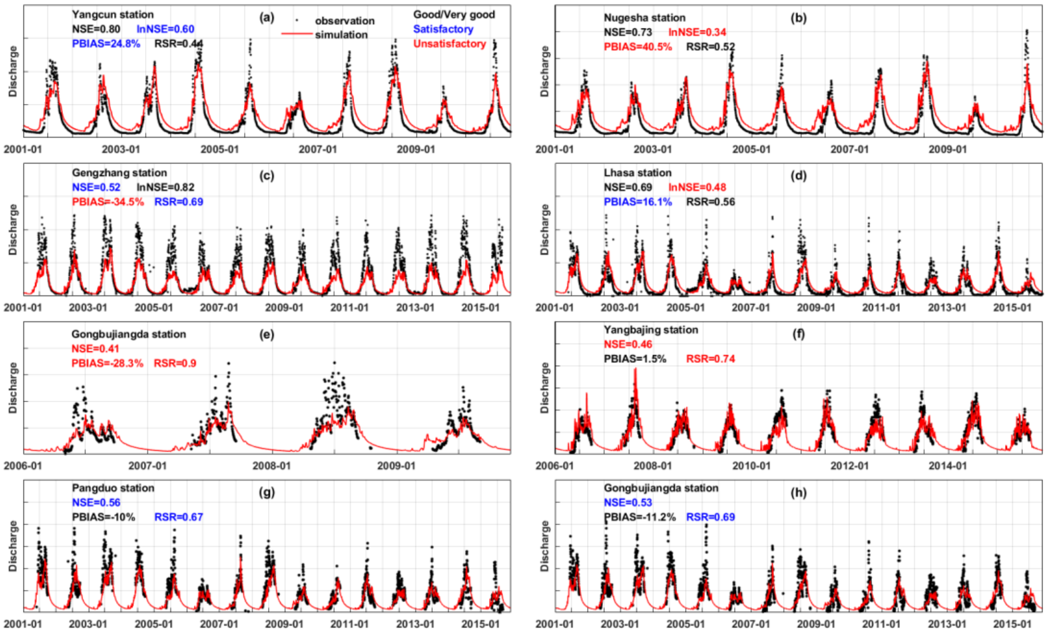
Comment 5:

Line 303: I propose to support the sentence "but were at acceptable levels with "D.N. Moriasi, J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel, T.L. Veith Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations".

The limitation of hydrological performance in terms of either maximum flow or low flow is consistent with several studies showing how the use of one metric in calibration can lead to less than ideal results for other metrics, as each is sensitive to particular characteristics of the time series and has its own limitations and trade-offs. See for example "Schaefli, B. and Gupta, H. V.: Do Nash values have value?, Hydrol. Process., 21, 2075-2080, 2007; "Gupta, H. et al.: Decomposition of mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80-91, 2009. "Mcmillan, et al. Five guidelines for the selection of hydrological signatures. Hydrol. Process., 31, 4757-4761, 2017. "; "Fenicia, F., et al.: Signature-domain calibration of hydrological models using approximate Bayesian computation: Empirical analysis of basic properties. Water Resour. Res., 54, 3958-3987." and "Majone, B. et al. Analysis of high streamflow extremes in climate change studies: How to calibrate hydrological models? Hydrol. Earth Syst. Sci. 2022, 26, 3863-3883". . I suggest that this consideration be taken into account when rewriting the section about the evaluation of model performance".

Response:

Thank you very much for providing these valuable references, which are really helpful for our deeper understanding on the calibration function selection issue. We have calculated the metrics PBIAS and RSR, and evaluated whether model performance is acceptable following the guidelines of Moriasi et al. (2007). We have added description about the model performance in the revised manuscript (L331~L332, L340~L342) and redrawn Figure 3 to show the acceptable level. We have also cited these useful literatures properly when describing the model performance evaluation in the revised manuscript (L342~L346).



Comment 6:

A few words should be devoted to the description of the concentration ratio and the concentration period, as is the case for NSE, LnNSE, RMSE ecc ecc.

Response:

Many thanks for your suggestion. We have added description (L290~L297) and equations (Eq. 8~10) of CR and CP in the revised manuscript.

Comment 7:

For the sake of clarity, I suggest deleting Figures 5a and 5b. They do not add any additional information respect to Figures 5c and 5d.

Response:

Many thanks for your suggestion. We have deleted Figures 5a and 5b in the revised manuscript.

Comment 8:

Figure 7: I do not understand figures from 7e to 7h and figures from 7m to 7p. It is not clear that the sum of the components is 100%. I suggest using a different type of figure style.

Response:

We are sorry that we make you confused about the figures. Figures 7e-h and m-p present the relative contribution of each runoff component in the total runoff during different season in different climate perturbation scenarios. We believe that the pie graph might be a better figure style which can make it clearer that the contributions of each component add up to 100%. However, five pie graphs would be needed for each subplot if this figure style is adopted, leading to too many subfigures. What's more, another reviewer required us to add the error bar to denote the standard deviation, which is difficult to present by other styles. Consequently, we prefer reserving the current Figure 7. Nonetheless, we have added tables listing the numbers in Figure 7 in the Supplementary Information to make it clearer (Table S1~S4).

Comment 9:

I suggest using the correlation values (r) in the text to identify and comment on positive\negative correlations and strong weak correlations (see Ratner, Bruce The correlation coefficient: Its values range between +1/-1, or do they not?). The p-value only indicates that there is a relationship between two groups.

Response:

Many thanks for your comments. We have calculated the correlation coefficient, and reported them in the text when describing the positive\negative correlations (e.g., L550, L561, L569).

Comment 10:

Limitations section: I suggest adding that future work should address the problem of sensitivity analysis of hydrological models, multi-objective calibration and goal-oriented calibration.

Response:

Many thanks for your comments. We have addressed issues related to calibration procedure in the limitation section of revised manuscript (L677~L683, L689~L692).

References

Ala-Aho, P., Tetzlaff, D., McNamara, J. P., Laudon, H., & Soulsby, C. (2017). Using isotopes to constrain water flux and age estimates in snow-influenced catchments using the STARR (Spatially distributed Tracer-Aided Rainfall–Runoff) model. Hydrology and Earth System Sciences, 21(10), 5089-5110.

Eriksson, D., Bindel, D., & Shoemaker, C. A. (2019). pySOT and POAP: An event-driven asynchronous framework for surrogate optimization. arXiv preprint arXiv:1908.00420.

He, Z. H., Pomeroy, J. W., Fang, X., and Peterson, A.: Sensitivity analysis of hydrological processes to perturbed climate in a southern boreal forest basin, Journal of Hydrology, 601, 10.1016/j.jhydrol.2021.126706, 2021.

He, Z., Unger-Shayesteh, K., Vorogushyn, S., Weise, S. M., Kalashnikova, O., Gafurov, A., Duethmann, D., Barandun, M., and Merz, B.: Constraining hydrological model parameters using water isotopic compositions in a glacierized basin, Central Asia, Journal of Hydrology, 571, 332-348, 10.1016/j.jhydrol.2019.01.048, 2019.

Lyu, H., Tian, F., Zhang, K., & Nan, Y. (2023). Water-energy-food nexus in the Yarlung Tsangpo-Brahmaputra River Basin: Impact of mainstream hydropower development. Journal of Hydrology: Regional Studies, 45, 101293.

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, 2021.

Nan, Y., He, Z., Tian, F., Wei, Z., and Tian, L.: 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, 4147-4167, 10.5194/hess-26-4147-2022, 2022.

Response to Reviewer #3

Comment 1:

The work is a good, well performed, exercise of modeling. After the model calibration, the authors did some scenarios simulations to assess the effects of temperature increase and precipitation change. This should be the purpose when modeling climate changes effects.

Response:

Thank you very much for your appreciation on our work. We have revised the paper thoroughly according to your comments.

Comment 2:

Correctly the authors listed also the limitations of the work (chapter 4.3). At L572-577 they state that the simulation of climatic change was at least too abrupt. By sure when arriving at +5° the glaciers will not be the same than now. They already are reduced since the simulation period (2015). But furthermore, I would remark that the temperature increase due to the climate change is far from being linear. We currently measure mainly longer warm periods (sometimes with higher temperature maximums) and shorter cold periods, and this leads to an increased mean temperature. Simulating a linear increase is of course easier than introduce a climate (stochastic?) model, but the effects on hydrology are surely different. Remarkably on ice or snow melt, but also, for instance, on evapotranspiration estimates. A second deep source of uncertainties is due to lack of data over such a large basin. "The CMFD was made through fusion of ground-based observations with several gridded datasets from remote sensing and reanalysis" (He, J., Yang, K., Tang, W. et al. The first high-resolution meteorological forcing dataset for land process studies over China. Sci Data 7, 25 (2020). https://doi.org/10.1038/s41597-020-0369-y). The CMFD, like all large-scale datasets, is the product itself of a model that incorporates from sparse ground data (about 735 CMA stations over 9.5x106km²) and remote sensing data combined by some algorithms and interpolated. As well known, eventually, the uncertainties of a cascade of models become rapidly great.

In my opinion these limitations severely reduce the usefulness of the work.

Response:

Many thanks for pointing out the major limitations of our study. We acknowledge that the current climate perturbation method is rather simple, but we believe that adopting GCM to analyze runoff change would also lead to significant uncertainties. For instance, according to Cui et al. (2023), the projected change in precipitation for the warming of 1.5°C had a very large uncertainty, ranging from -13.68% to 7.20%, based on 22 CMIP6 models. The climate perturbation function has the special function on investigating the roles of particular precipitation or temperature perturbations in shaping the responses of specific hydrological processes, and was widely adopted in previous studies (e.g., He et al., 2021; Su et al., 2023; Rasouli et al., 2014). So we believe that our analysis is helpful for understanding the separate effect of changing temperature and precipitation on runoff, and the spatial pattern of runoff change.

We acknowledge that the data scarcity problem leads to large uncertainty in our results. Actually, as you said, all the existing datasets have large uncertainty on the Tibetan Plateau, because of the sparse ground data. We have tried to adopt multiple datasets to constrain the model and validate the simulations on multiple objectives. However, given the uncertainties exist in all the dataset, it is indeed hard to say whether this procedure could improve confidence of the results. We have addressed this issue in the revised manuscript (L305~L307, L669~L672).

Comment 3:
L55 – the melting processes of frozen water are determined by the energy budget.
Response:
Many thanks for your correction. We have rephrased this sentence in the revised paper.

Comment 4:
L75 - I presume that you mean that the contribution of melt ranges from 0.86% to 40.59%, or it was estimated as 0.86% ± 40.59%. The use of a minus sign is not very clear due to the very wide range of values. Also at L45 and L77.
Response:
Many thanks for your careful reading. What we mean by the minus sign is the range, which is indeed misleading. We have changed it to "~" in the revised paper.

Comment 5:
L140 – the drainage area is 2 x10^5 km²? The measure unity misses.
Response:
Many thanks for your careful reading. We have added the measurement unit in the revised paper.

Comment 6:
L165 – 1.875° at 30°N is a rectangle of about 180x208 km, 3.7x104 km². This means just a little more than five full pixels on a 2x10^5 km² basin.
Response:
Many thanks for your comment. We adopted the output of an isotope enabled General Circulation Model (iGCM) to drive the model, the spatial resolution of which is indeed coarse. However, our previous work (Nan et al., 2022) indicated that the bias of iGCM on the TP is highly related to elevation, and we have conducted bias correction according to the elevation of each REW. Although the bias correction procedure was still simple, it could provide a rather reasonable isotope input data.

Comment 7:
L180 – eight subzones. If I understand correctly the subzones should be 12 (6 surface x 2 subsurface)
Response:
We are sorry for not describing the model clearly. The six subzones in the surface layer are parallel to those in the subsurface layer. A subzone refers to either a zone in the surface layer or a zone in the subsurface layer, rather than the combination of two subzones in different layers. Consequently, there are eight subzones in each REW.

Comment 8:
L188 – does the model consider even the melt caused by (liquid) rainfall? It could be rather

important mostly (but not only) in the changing rainfall scenarios.

Response:

Many thanks for your comment. We are sorry that we don't quite understand what you mean by "melt caused by liquid rainfall". If you mean the rainfall occurring in glacier covered areas, it was considered in the model but was regarded as rainfall runoff rather than meltwater.

Comment 9:

Table 3 – I presume that DDFG is the degree-day factor for ICE (glacier) melt.

Response:

Thanks for your correction.

Comment 10:

Table 3 – -4.28C is a rather cold temperature for melting threshold. Of course, this happens because of the daily average, but it is better to explain it.

Response:

Thanks for your comment. We have explained it in the revised manuscript (L318~L320).

Comment 11:

L245 and following. This could be the main problem of the work. The temperature increase due to the climate change is far from being linear. See my general comments on modeling limitations.

Response:

Many thanks for your pointing this out. We need to clarify that although the temperature scenarios were set linearly, we did not assume the process of temperature changing. We only analyzed how runoff would change if the temperature is X°C higher than the current condition. The influence of gradual process of climate change on runoff has not been analyzed yet, which was indeed a limitation of this work and was already addressed in the limitation section (L646~L649)

Comment 12:

L258 - Did you combine also temperature and precipitation perturbations in a single simulation?

Response:

Thanks for your question. We did not combine temperature and precipitation perturbations in a single simulation, to focus on the separate effect of temperature and precipitation on runoff change.

Comment 13:

L269 – Figure 2. Streamflow discharge appears well simulated, NSE = 0.82 is more than satisfactory in such a long period. SCA is more confused, but still well acceptable, as the isotopes simulation. The average glacier mass balance comparison with observed values (L280) has no meaning. Even a linear mass loss could simulate those 11 points.

Response:

Many thanks for your comment. Although the 11 measurements of glacier mass balance are rather easy to simulate, the glacier calibration mainly aims to provide basic constraints for the amount of glacier meltwater. For instance, if the simulated glacier meltwater is twice the current simulation, we can expect that the sensitivity analysis result would be totally different.

Comment 14:
L316 – decreasing rate
Response:
Thanks for your correction.

Comment 15:
L333 – it is hard to believe that evaporation is not limited by water condition. Did you ever try to compare it with potential evaporation?
Response:
Many thanks for your comment. We calculated the evaporation and found it was much lower than potential evaporation (228mm/yr compared to 889mm/yr). It is actually inaccurate to say that the evaporation is not limited by water condition. It should be expressed as "the energy limitation played a more important role than water limitation on evaporation". This is mainly because the YTR basin is a relative wet region on the TP, and was also identified as an energy-limited area by Wang et al. (2022). We have changed the expression in the revised manuscript (L380~L381).

Comment 16:
L362 – possibly in April the fraction of runoff due to snow melt increases. But SCA is going to decrease …
Response:
Many thanks for your comment. We also attribute the runoff increase to the increasing snow melt, and we have added it in the revised manuscript (L411~L412).

Comment 17:
L385-386 – please control the measure unit (km^3/yr and km^3/s).
Response:
Many thanks for your correction. We made a mistake here, and the measurement unit should be km^3/yr.

Comment 18:
L388 – the glacier surface is (almost) impermeable but runoff from bédières rather quickly falls to the bottom through crevasses, moulins, wells. So, melt runoff flows mostly on the bottom of the glacier.
Response:
Many thanks for your comment. Although glacier meltwater could fall to the bottom of glaciers, very little of them infiltrates in deep groundwater storage. Consequently, we assume glacier meltwater generates runoff quickly through the surface pathway (but not necessarily through glacier surface), similar to other glacio-hydrological modeling works (e.g., He et al., 2019; Gao et al., 2017).

Comment 19:

L498 – in my experience the melt component due to (liquid) rainfall can amount up to 15% of the total melt. In the REWs with larger GAR the increase of rainfall can lead to a greater runoff, if this component is considered in the model. Fig. 9d means that over a glacierized area a precipitation increase corresponds to a runoff decrease. Projecting the interpolation line, we can suppose that for a 35% of GAR there are no runoff changes when precipitation increases to 120%. The 35% of glacierized area literally eats the rainfall increase over the whole basin. Rather surprising!

Response:

Many thanks for your comment. If we understand correctly, "the melt component due to rainfall" refers to the rainfall occurring in the glacier covered regions, which is regarded as rainfall overland runoff in the model. The meltwater in the model only refers to water melting from solid snow/glacier. In REWs with larger GAR, the contribution of rainfall in total runoff would be small, thus the runoff increasing in response to increasing precipitation would be accordingly small.

The inference that 35% GAR eats the rainfall increase over the whole basin is indeed surprising, but the deviation from data point to the regression line in Fig. 9d is rather large for high GAR, indicating that the confidence of regression line is likely low when GAR is large. This is because there are few REW samples with large GAR based on the current spatial discretization. The runoff change in response to precipitation increase depends not only the balance among precipitation, evaporation and glacier melt, but also the spatial scale. More works are needed to better understand this, and we have addressed this issue in the revised manuscript (L564~L568).

Comment 20:

L510 – the equation is useless, the correlation is high. But on fig. 10 also the "decreasing" REWs seems to have a correlation.

Response:

Many thanks for your comment. The seeming correlation between GAR and MAP is mainly because of the relatively narrow range under the dividing line. The distribution of "decreasing" REWs is actually rather dispersive.

Comment 21:

L513-517 – for "increasing" REWs the correlation between GAR, MAP and the increasing runoff is poorly understandable by the blue color intensity. Could you find a way to represent it with a number?

Response:

Many thanks for your comment. We have added a table presenting the GAR, MAP and runoff change in the Supplementary Information (Table S5).

Comment 22:

L535 – glacier meltwater increases with temperature, but glaciers are shrinking and, soon or late, this leads to a runoff maximum followed by a decrease. It has already happened to many glaciers.

Response:

Many thanks for your comment. Our model also simulated the change of glacier area, and results indicated that the glacier covered area decreased by 10% during the simulation period. However, this was not stressed in the paper, because the glacier area shrinking was not validated by measurement data. This study adopted the climate perturbation method and did not consider the gradual climate change process, thus the turning trend of glacier melt runoff was not reflected. We have addressed this issue in the limitations of revised manuscript (L653~L655).

Comment 23:

L542-543 – and for the other regions the future will be the same, moved a little further.

Response:

Many thanks for your comment. We have added this in the revised manuscript (L619~L622).

References

Cui, T., Li, Y., Yang, L., Nan, Y., Li, K., Tudaji, M., ... & Tian, F. (2023). Non-monotonic changes in Asian Water Towers' streamflow at increasing warming levels. Nature Communications, 14(1), 1176.

Gao, H., Ding, Y., Zhao, Q., Hrachowitz, M., & Savenije, H. H. (2017). The importance of aspect for modelling the hydrological response in a glacier catchment in Central Asia. Hydrological processes, 31(16), 2842-2859.

He, Z., Pomeroy, J. W., Fang, X., & Peterson, A. (2021). Sensitivity analysis of hydrological processes to perturbed climate in a southern boreal forest basin. Journal of Hydrology, 601, 126706.

He, Z., Unger-Shayesteh, K., Vorogushyn, S., Weise, S. M., Kalashnikova, O., Gafurov, A., ... & Merz, B. (2019). Constraining hydrological model parameters using water isotopic compositions in a glacierized basin, Central Asia. Journal of hydrology, 571, 332-348.

Rasouli, K., Pomeroy, J. W., Janowicz, J. R., Carey, S. K., & Williams, T. J. (2014). Hydrological sensitivity of a northern mountain basin to climate change. Hydrological Processes, 28(14), 4191-4208.

Su, T., Miao, C., Duan, Q., Gou, J., Guo, X., & Zhao, X. (2023). Hydrological response to climate change and human activities in the Three-River Source Region. Hydrology and Earth System Sciences, 27(7), 1477-1492.

Wang, L., Han, S., Tian, F., Li, K., Li, Y., Tudaji, M., ... & Yang, Y. (2022). The evaporation on the Tibetan Plateau stops increasing in the recent two decades. Journal of Geophysical Research: Atmospheres, 127(23), e2022JD037377.