Response to RC1:

General comments 1:

The issues that I raised in the second round of review have been appropriately addressed. However, I agree with reviewer #2 that the novelty of this manuscript appears marginal, which I also mentioned during the first round of review. Perhaps this study is more likely to generate regional interest rather than general interest, given its large dependence on empirical methods and lack of novel generalizable physical insight. Nevertheless, the methodology presented in this manuscript may provide an alternative modeling method to existing ones, and its relative physical simplicity could be an advantage for specific cases if the benefits this simplicity are mechanistically justified and made clearer in the manuscript. Ultimately, I am undecided about whether this manuscript has a sufficient level of novelty or potential impact for HESS. Hence, I recommend minor revisions so that the authors may further clarify in the manuscript the value added by this study to the hydrological community.

Response 1:

We sincerely thank the reviewer for confirming that the concerns raised during the second round of review have been appropriately addressed, and for providing constructive feedback on the novelty and broader relevance of the study. In response, we have carefully revised the manuscript to further highlight the value of this work to the wider hydrological community and clarify its scientific contributions. See in particular our below responses to General comments 2 and 3, through which we now included key clarifications on how our study adds scientific insights for the hydrological community, beyond the regional perspective. In addition to these clarifications in the introduction (just before stating the objectives) and the discussion sections, we accordingly re-formulated (i) the abstract, (ii) other parts of the introduction (e.g., further explaining the general importance of understanding seasonally frozen ground (SFG) in large basins), (iii) methods, (iv) results and (v) conclusions (e.g., clarifying the modelling framework transferability and quantitative process-level insights of significance also beyond the study region).

General comments 2:

As I also mentioned in the first round of review, the physical insights that the study generates are not particularly novel, though it may be quantitatively more accurate than competing models in certain circumstances (e.g. the case study presented in the manuscript). A key issue with the manuscript is that the mechanistic reasons for why the predictions made by this model are good, or even better than other models, are not sufficiently explained. Given the mostly empirical nature of the model, it is possible that the good results may have arisen due to statistical fluke or overfitting to additional degrees of freedom. Hence, it is not clear whether the modeling methods introduced in this manuscript are generalizable to other regions or scenarios, and whether they represent a superior alternative over other existing models. It is recommended that the authors provide a convincing mechanistic explanation of why their fitting results (NSE) were superior to VIC and SWAT. Providing such an explanation would also help to allay concerns that some of the design choices of the introduced model seem to be rather arbitrary and non-generalizable, as I mentioned during the first round of review (e.g. my comment #6 regarding discrete soil layering in GXAJ).

Response 2: We thank the reviewer for raising this important point. In this study, the developed model integrates physically based processes (such as snow accumulation and melt, freeze–thaw dynamics, and soil water storage) with empirical simplifications to reduce parameter complexity and the reliance on extensive input data. We agree with the reviewer that the manuscript has retained a vagueness in explaining its novelty and main contributions. In response, we now clarify (in the introduction section) that a main novel aspect of the manuscript is how additional processes are accounted for in a three-step manner by a modular model design (one module per process; with the snow and frozen ground modules being grounded in well-established physical principles, e.g., SNOW17-based snowmelt equations and Stefan-based frost depth estimation). This allows for increasing the complexity while transparently checking the model performance of each step. In particular, any potential increases in model performance are then related to the dynamics created by the additional module (and the corresponding account for a new process).

In the introduction we also explain that, to the best of our knowledge, this has not been done earlier in large cold region basins. This is because previous comparisons have regarded models that differ in either structure (Gao et al., 2018; Li et al., 2018; Song et al., 2022), or structure as well as complexity (e.g., Ahmed et al., 2022; Gao et al., 2018; Guo et al., 2022) . In both cases, differences in model performance may then partly be due to fundamental structural or parametrization differences between models, introducing uncertainty in how performance may be linked to complexity (i.e., inclusion or omission of processes). Therefore, we believe that the current approach, which greatly reduces such structural uncertainty effects by study design, is useful and of general scientific interest in advancing the process understanding and prediction of large-basin runoff in cold regions, in addition to the presented insights for the considered basin.

As previously stated, classical physically based hydrological models, such as VIC, GBEHM, and WEB-DHM, provide detailed representations of hydrological processes. Their complex structures and numerous parameterization schemes often require large amounts of ground-based measurements (e.g., soil hydraulic properties, vegetation structure, snow thermal conductivity), which are rarely available in cold and data-scarce regions. This limitation can result in large simulation uncertainties. In the first part of the discussion (section 4.1), we now additionally clarify that, by comparison, our three-step approach implies that a limited number of additional parameters are introduced in each performance evaluation step, which enables the identification of well-functioning levels of model complexity while involving only a small number of parameters - five in the original GXAJ model and four in the snow module. This greatly reduces the risk of overfitting. As further discussed in section 4.1, we have also considered the risk of coincidental good performance by potentially overfitted models by evaluating in which way the addition of process-based modules alters the model behavior in multiple sub-catchments and over multiple seasons. We could then for instance see that, rather than increasing the sub-catchment and seasonal performance in random ways, the addition of the snow and SFG modules specifically increased cold-season performance in low-temperature (high-altitude) parts of the study area, which is consistent with the expected effects of the considered processes. This hence provides a logical explanation that helps readers understand (as asked-for by the reviewer) why the simulation performance demonstrated in our case study was strong (e.g., with high NSE) despite being based on few parameters as compared with e.g. VIC and SWAT applications.

General comments 3:

Furthermore, although the authors claim that the simplicity of the introduced GXAJ-based methods makes it advantageous compared to physical process-based models in geologically or topographically complex cold-region scenarios, the GXAJ-based methods appear to suffer from the same limitations, as discussed by the authors themselves in lines 769-784. Hence, the authors claim that "In data-limited regions such as the Yalong River basin, physical models may rely on data that are not available through direct measurements, such as ground temperature. This complicates parameterization processes and introduces uncertainties in the results." should be further explained and justified.

Response 3: We thank the reviewer for raising this important point regarding our claim about the practicality of GXAJ-based methods in data-scarce cold regions. We aimed at addressing this issue together with our addressing of the "General comment 2", see in particular our above answer in the second paragraph, starting with "*We now additionally explain that, by comparison, our three-step approach implies that a limited number of additional parameters are introduced in each performance evaluation step, which enables the identification of well-functioning levels of model complexity while involving only a small number of parameters - five in the original GXAJ model and four in the snow module*". This aspect hence contributes to explaining why the GXAJ model and its modular extensions are less prone to suffer from such limitations. This is further justified by the fact that we explicitly check that the addition of the snow and SFG modules resulted in a model behavior that is consistent with the expected effects of the considered processes (e.g., increased cold-season performance in low-temperature (high-altitude) parts of the study area) despite the data limitation issues. This is now also explained and justified in the beginning of the discussion (section 4.1), enhancing also the discussion of the generality of the approach in lines 795–817 of the revised manuscript.

Once again, we sincerely thank the reviewer for the constructive comment regarding the scientific value and applicability of the manuscript. This provided us with an excellent opportunity to clarify our reasoning, refine the discussion, and more clearly articulate the novelty and practical significance of our modeling approach in cold and data-scarce regions. We hope the revised manuscript better reflects these improvements.

References

Ahmed, N., Wang, G., Booij, M. J., Marhaento, H., Pordhan, F. A., Ali, S., Munir, S., and Hashmi, M. Z.-R.: Variations in hydrological variables using distributed hydrological model in permafrost environment, Ecol. Indic., 145, 109609, https://doi.org/10.1016/j.ecolind.2022.109609, 2022.

Gao, B., Yang, D., Qin, Y., Wang, Y., Li, H., Zhang, Y., and Zhang, T.: Change in frozen soils and its effect on regional hydrology, upper Heihe basin, northeastern Qinghai-Tibetan Plateau, Cryosphere, 12, 657–673, https://doi.org/10.5194/tc-12-657-2018, 2018.

Guo, L., Huang, K., Wang, G., and Lin, S.: Development and evaluation of temperature-induced variable source area runoff generation model, J. Hydrol., 610, 127894, https://doi.org/10.1016/j.jhydrol.2022.127894, 2022.

Li, Z., Yu, J., Xu, X., Sun, W., Pang, B., and Yue, J.: Multi-model ensemble hydrological simulation using a BP Neural Network for the upper Yalongjiang River Basin, China, Innovative Water Resources Management - Understanding and Balancing Interactions Between Humankind and Nature, 8th International Water Resources Management Conference of ICWRS, Gottingen, Web of Science ID: WOS:000459240300045, 335–341, https://doi.org/10.5194/piahs-379-335-2018, 2018.

Song, L., Wang, L., Zhou, J., Luo, D., and Li, X.: Divergent runoff impacts of permafrost and seasonally frozen ground at a large river basin of Tibetan Plateau during 1960-2019, Environ. Res. Lett., 17, 124038, https://doi.org/10.1088/1748-9326/aca4eb, 2022.