Response to RC1:

General comments

Comments 1: This manuscript is well written, and the work done appears quite meticulous and informative from a methodological point of view, but I am not fully convinced of the novelty of this manuscript. The manuscript shows that GXAJ-S-SF outperforms GXAJ. It is self-evidently almost certain that a more accurate hydrological partitioning can be achieved when two important physical processes snowmelt and freeze-thaw are included into the model. Thus, it is certainly expected that GXAJ-S-SF will outperform GXAJ in a region that experiences the S and SF processes. Furthermore, since GXAJ consistently underestimates the runoff, and since the physical processes modeled in SF can only increase the runoff but not decrease it, it is a foregone conclusion that upon calibrating SF you will arrive at a better fit for GXAJ-S-SF than GXAJ. As far as I can tell, there are no novel or interesting findings regarding hydrological processes in this manuscript, nor are there meaningful analyses about the utility and information content of the hydrological models used beyond the goodness-of-fit metrics NSE, RBE, RMSE. Therefore, I recommend that after major revisions addressing the concerns I have elaborated below, this manuscript could be suitable for publication as a technical note.

Response 1: Thank you for your detailed evaluation of our manuscript and your constructive feedback. We highly appreciate your comments on the novelty and contributions of the study, which have provided us with valuable guidance for improving the manuscript. Below, we address your concerns and clarify the innovative aspects and unique findings of our research.

We acknowledge your point that the inclusion of snowmelt and freeze-thaw processes (SFG) in the GXAJ-S-SF model makes performance improvements over the original GXAJ model somewhat expected in regions experiencing these processes. Indeed, prior studies (e.g., Ala-Aho et al., 2021) have shown the significant impact of SFG on runoff at small scales. However, as emphasized in the revised introduction and discussion sections (see lines 50-70, 135-155), there are still substantial knowledge gaps regarding the complex impacts of SFG processes on runoff in large basins, especially concerning the level of model complexity required to accurately simulate these processes. Our study systematically addresses these gaps, making the following key contributions:

- **Innovation in snow-freeze-thaw coupling:**

We developed a physically meaningful method to couple snowmelt and freeze-thaw processes within the hydrological model. This coupling allows for a quantitative analysis of how snowmelt and frozen ground influence runoff, soil moisture dynamics, and evapotranspiration. Importantly, the coupling approach is computationally efficient, requiring relatively few additional parameters while retaining strong physical interpretability.

- **Assessment of dominant hydrological processes in a large basin subject to SFG:**

Unlike previous studies that often focus on small-scale applications, our study systematically compares simplified models (e.g., without snow-SFG extensions or considering snow only) with extended models that account for both snow and SFG processes. These comparisons reveal that SFG processes significantly alter runoff dynamics at large basin scales, with quantitative results showing an increase in surface runoff by 39–77% and significant reductions in interflow and groundwater runoff compared to models during cold months that neglect SFG processes. This provides critical insights into the role of SFG in shaping hydrological regimes and highlights the necessity of incorporating such processes in predictive models for cold regions.

- **Integration of process understanding with practical applications:**

By analyzing the effects of varying frozen soil depths and their spatiotemporal distributions on hydrological processes, our study provides a comprehensive understanding of the complex feedback mechanisms of frozen ground on hydrological systems. These insights not only enhance the understanding of snow-SFG-hydrology interactions but also offer practical guidance for predicting hydrological responses to future climate change in cold mountainous regions.

In summary, while the improved performance of GXAJ-S-SF may appear intuitive, our study goes beyond performance metrics by systematically addressing the scientific question of how and to what extent SFG processes impact hydrological dynamics in large basins, an area with significant gaps in current knowledge. We have revised the manuscript to emphasize these contributions more explicitly, including updates to the title, introduction, discussion, and conclusion sections.

Comments 2: Would it not be more meaningful to compare GXAJ-S-SF to a different hydrological model that also includes snowmelt and seasonal freeze-thaw? For example one of those models you mentioned in L96 – 112. Even if an actual model comparison is not done, it would be useful to discuss the differences and similarities in model processes between GXAJ-S-SF and other similar models with snowmelt and freeze-thaw functions.

Response 2: Thank you very much for your valuable suggestions regarding our research! We have carefully followed your advice and expanded the discussion section, particularly in Section 4.1 "*Key limitations in hydrological models in relation to their process complexit*y". In this section, we have compared the processes of the current GXAJ-S-SF model with other hydrological models that include snowmelt and freeze-thaw functions.

Specifically, we elaborated on the similarities and differences in how these models simulate hydrological processes, highlighting the unique features and advantages of the GXAJ-S-SF model. These comparisons have also helped us to clarify the innovations and improvements that the GXAJ-S-SF model offers in representing hydrological processes.

Additionally, we discussed how the process complexity and computational efficiency of the GXAJ-S-SF model compare to other models. This provides further context for understanding the trade-offs between model simplicity and accuracy, as well as the potential for applying our model to other cold-region catchments. We believe this discussion addresses your concerns and adds value to the manuscript by offering a more comprehensive perspective.

We sincerely appreciate your constructive feedback, which has helped us improve the quality of our manuscript.

Comments 3: It would be more rigorous to re-run the models with different priors. For example, there could be a configuration of GXAJ, with soil property related parameters set at an "annual average effective value" taking into account that the soil is frozen for 9 months of the year. This hypothetical configuration of GXAJ could possibly produce results as good as GXAJ-S-SF, but it is possible that this configuration of GXAJ was not tested because the optimization algorithm was stuck in a local minimum. Given the highly nonlinear processes involved in this model, I think that calibrating from a single set of priors may be insufficient.

Response 3: Thank you for your valuable comments on our research. We greatly appreciate your feedback and have addressed the relevant issues with detailed revisions and additions as follows:

On Optimization Algorithm and Parameter Settings:

Regarding your concerns about model optimization and the use of "prior parameters," we understand the potential limitations of calibrating the model using a single prior parameter configuration. In our study, we employed the SCE-UA optimization algorithm to precisely calibrate key parameters and obtain the optimal solution for the model. To avoid the risk of local optima, we set a range of prior parameters and randomly selected different configurations within the allowed range, running the optimization algorithm multiple times. While this approach was not explicitly described in the original manuscript, we have now supplemented the relevant content to clarify that we considered different prior parameters during optimization (Section 2.2.3, lines 404-414). Through this approach, we enhanced the model's stability and minimized reliance on a single configuration, addressing the limitations you raised.

On Soil Characteristics and Related Parameter Settings:

We fully agree with your observation that the presence of frozen ground significantly alters soil moisture dynamics, which, in turn, affects the storage capacities of soil tension water and free water and their spatial-temporal distribution. In our study, we did not rely on a simplistic "annual effective value" to account for frozen ground effects. Instead, we dynamically adjusted the distribution of these underlying surface parameters by comparing the depth of frozen ground (characterized by spatial-temporal heterogeneity) with the corresponding soil layer thickness (as categorized in four specific cases in the methodology). This adjustment supported the model in reflecting key effects of frozen ground. Through this approach, we fully considered the spatial-temporal heterogeneity induced by frozen ground, thereby improving the accuracy of the simulation results. In the revised manuscript, we have further elaborated on this section (see Section 2.2.2 *Freeze-thaw process*), detailing how changes in frozen ground depth dynamically influence soil layer thickness and related parameters, enhancing the model's capacity to simulate frozen ground dynamics.

We believe that these revisions will better demonstrate the scientific validity and rationality of our approach, and we thank the reviewer for their insightful suggestions.

Comments 4: L409 – 417: “The accuracy in simulating the initial freeze and initial thaw dates was validated against ground temperature data from meteorological stations within the basin (Fig. S5), indirectly confirming the simulated soil freeze-thaw processes.”

Could you provide citations or a more detailed discussion to support the validity of this point? Since freezing and melting both start from the top, and since the temperature data for verification was measured at the ground surface, simulating the correct initial freeze and initial thaw dates does not help confirm that the model has simulated the freezing depth correctly over the 9 months with frozen soils.

Response 4: Thank you for your insightful comments. We have identified several studies supporting the use of surface temperature to validate the initial freeze and thaw dates, which provide a theoretical basis for applying this method in regions lacking observational data on frozen soil depth (Li et al., 2022). However, due to the absence of measured frozen soil depth data in our study area, we were unable to directly validate the simulated frozen soil depth. Consequently, this study primarily utilized available surface temperature data to verify the initial freeze and thaw dates, thereby indirectly supporting the reliability of the simulated freeze-thaw processes.

To further evaluate the model's reliability, we compared the spatial distribution of the maximum frozen soil depth simulated in this study with data from the Tibetan Plateau Permafrost Dataset (1961–2020) for the 2000s. The comparison revealed a high degree of consistency in both spatial distribution patterns and magnitude, with a correlation coefficient of 0.89. These results provided additional validation support, and we presented these comparative analyses in the revised manuscript (Section 3.1 "*Simulation of snow accumulation and freeze-thaw process*," lines 475-486).

We also recognize that the use of surface temperature and maximum frozen ground depth to verify the freeze-thaw process introduces some uncertainty (Li et al., 2022), Since the GXAJ-S and GXAJ-S-SF model variants used the same temperature, snow and frozen ground data in the present simulations, they can be expected to share similar data errors, However, due the non-linear nature of the modeled processes, such data errors may still not cancel completely when comparing different models. Nevertheless, observed differences in model performance between these models are mainly expected to reflect differences in model capabilities rather than differences in input datasets. Future work should focus on improving remote sensing data quality and exploring the long-term robustness of the model to further enhance performance and improve our understanding of the freeze-thaw processes in complex mountainous cold regions (Section 4.1 *"Key limitations in hydrological models in relation to their process complexity"* lines 642-659).

Comments 5: I think that the "modular approach" that you emphasize several times, including in the abstract and conclusion, is reinventing the wheel as it is just another name for loose coupling or one-way coupling, which is a basic hydrological concept.

Response 5: Thank you for your valuable comments and for providing us with the opportunity to clarify the term "modular approach." After carefully considering your feedback, we agreed that the term "modular" may overlap to some extent with concepts such as loosely coupled or unidirectional coupling, especially in the context of hydrological modeling. However, our use of "modular" aimed to emphasize the flexibility, scalability, and reusability of the model components. Specifically, the snow and frozen ground modules in our study were designed as independent components that could be enabled or disabled depending on environmental conditions (e.g., the presence of snow or seasonally frozen ground). These components were not hard-coded into the GXAJ model; instead, they were designed to be integrated into other hydrological frameworks without requiring significant modifications to the core structure of the model. This design approach enhanced the model's flexibility and adaptability, allowing researchers to extend or modify it to suit different cold-region environments.

To address your comments, we revised the manuscript to avoid potential misunderstandings caused by the term "modular approach." Instead, we described the design philosophy of the model components more accurately. In the abstract and conclusion, we replaced the term "modular approach" with "flexible and adaptable," ensuring that the core idea of the design is conveyed clearly and without ambiguity.

We hope this clarification and the corresponding revisions have adequately addressed your concerns.

Comments 6: After reading through the manuscript several times, I recognize that the bulk of the scientific contribution of this manuscript lies in the freeze-thaw process module in section 2.2.2. As shown in the results, it fits well with the measurements. However, I think some parts should be explained more clearly. What is the purpose of using two different representations of the soil layers in one model? Why not use the same layers for the computation of runoff, moisture and ET (Figure S1)? Does this mean that in the simulations, the humus layer could sometimes overlap with both the “upper soil” and part of the “lower soil”? And can the “upper soil” sometimes overlap with both the humus layer and the vadose zone? Does this not then imply that you need to interpolate some effective soil parameter values that may be inappropriate for the actual individual soil layers? How would this affect the runoff and discharge predictions? Furthermore, wouldn’t this mean that the parameter values you calibrate from field data do not have a proper physical meaning? I think that in order to reconcile the two different representations of the soil layers, it is inevitable that the calibrated parameter values are smoothed interpolations of the values that would actually describe each individual soil layer.

Response 6: Thank you for your positive feedback on the freeze-thaw process in section 2.2.2 and for your valuable suggestions. Regarding the use of two different soil layer representations in the model, we adopted this approach based on the design philosophy of the original GXAJ model. We did not provide sufficient explanation in the original manuscript, which caused some confusion, and we have now clarified this in the revised version (Section 2.2 *"Modeling approach"* lines 200-256, 327-379).

Specifically: When calculating runoff for a grid cell, soil saturation refers to the soil water content reaching the field capacity, not the saturation water content. The GXAJ model uses a saturation runoff mechanism, meaning that runoff only occurs when the soil's unsaturated zone reaches field capacity. Before this point, all incoming water is absorbed by the soil without generating runoff. In the GXAJ model, the tension water storage capacity (W_M, in mm) of a grid cell is determined by the watershed's topography, as well as soil, vegetation, and other surface conditions. We do not consider the uneven distribution of tension water content within the grid cell. To calculate the actual precipitation (P_e) available for runoff, we subtract the evaporation, canopy interception, and river precipitation from the measured rainfall during the calculation period, then check if upstream inflow replenishes the soil water content of the current grid cell.

When calculating the sources of runoff (Figure S2(a)), the runoff in the grid cell is divided into three components: surface runoff, interflow, and groundwater runoff. The GXAJ model treats the upper soil layer in the unsaturated zone as the humus layer (determined by topography, soil, vegetation, and other surface conditions), with the bottom of the humus layer considered a "relatively impermeable layer." Some of the runoff generates interflow, while part continues to percolate, generating groundwater runoff. When the free water in the humus layer reaches saturation, surface runoff is produced. Similarly, we do not consider the uneven distribution of free water storage in the grid cell.

In summary, the GXAJ model (Yao et al., 2012) calculates the tension water storage capacity (W_M) in the unsaturated zone (Figure S1) and the free water storage capacity (S_M) in the humus layer to divide runoff into three components: R_s, R_i, and R_g. W_M determines whether runoff occurs and the amount of runoff (saturation excess runoff), while the free water content in the surface soil splits runoff into R_i and R_g. When the free water content reaches saturation, R_s is produced, as shown in Figure S2(a).

For evapotranspiration (Figure S2(b)), the GXAJ model uses a three-layer evapotranspiration model, dividing the soil (vadose zone) into upper, middle, and lower layers, with corresponding tension water storage capacities: W_{UM}, W_{LM} and W_{DM} (in mm). During actual evapotranspiration calculation, canopy interception is evaporated based on evapotranspiration capacity. When the intercepted water is less than the evapotranspiration capacity, the three-layer model is applied. The calculation principle is that the upper layer evaporates according to its evapotranspiration capacity. If the upper layer cannot supply enough water for evapotranspiration, the remaining capacity is drawn from the middle layer, with evapotranspiration in the middle layer proportional to the remaining capacity and inversely proportional to the middle layer's storage capacity. The ratio of middle layer evapotranspiration to the remaining capacity cannot be less than the deep layer evapotranspiration coefficient, *C*. If the middle layer cannot supply enough, the deep layer water will supply the deficit. The corresponding soil moisture and evapotranspiration are labeled as W^u, W^l, and W^d, and E^u, E^l, and E^d.

The original GXAJ model used different soil layers to better simulate the role of soil layers at different depths in hydrological processes. This means that the humus layer may overlap with the "upper soil" and part of the "lower soil," as you understand, but the specific situation may vary depending on the surface conditions. However, the soil surface parameters, such as tension water storage capacity and free water storage capacity, are derived from the physical properties of the soil (e.g., soil type and structure), as well as topography and vegetation, and thus have physical significance. Applying these concepts to a single-layer soil system would simplify the calculation, treating the entire soil layer as the unsaturated zone for runoff calculation and using a single-layer evapotranspiration model.

I hope this explanation resolves your doubts, and we will further improve the manuscript in the revision.

Specific comments

Comments 7: L192: Is saturation excess runoff a reliable way to partition snowmelt fluxes, which are fast and may often exceed the infiltration capacity?

Response 7: Thank you for the valuable feedback. The melting rate of snowmelt water is usually fast, and in the presence of a frozen soil surface, the soil's permeability is limited, which easily leads to surface runoff. We fully agree with this point, and it has been thoroughly considered in Sections 2.1 and 2.2.1 of our study. In our model, we specifically focused on the interaction between water movement and soil during the snowmelt process, considering the potential freezing of the soil surface. When the snowmelt water encounters the frozen soil layer, due to the low permeability of frozen soil, the snowmelt water cannot rapidly infiltrate into the soil, resulting in significant surface runoff. The detailed consideration of this process in the model ensures the rapid generation of runoff from snowmelt water. Additionally, even without considering the effects of frozen soil, the runoff generation mechanism in our study is based on saturation excess and infiltration excess (Figure 2), which also accounts for the possibility of infiltration excess runoff from snowmelt fluxes.

Comments 8: L277: If you divide the SFD by the cube root of ASD to get SFD*, then the units of SFD* are $[cm]^{2/3}$. What does that physically mean?

Response 8: Thank you for your attention. The empirical formula for frozen soil depth used in our study is derived from the research "*Influence of snow cover on soil freeze depth across China*" which utilized observational data from 378 meteorological stations across China (1980–2014). This study quantified the relationship between snow cover and the maximum seasonal freeze depth (MSFD), as well as the contribution of snow cover to MSFD. The results indicated that in areas with thin snow cover or short snow duration, the impact on freeze depth is minimal. However, in regions with thick snow and longer snow duration, snow cover significantly reduces the frozen soil depth. This relationship can be reasonably reflected by the formula dividing the freeze depth (SFD) by the cube root of snow depth (ASD).

From a physical perspective, the units of SFD* ($[cm]^{2/3}$) may not have strict physical meaning. However, the inverse relationship between SFD and the cube root of ASD reflects the empirical nature of this relationship, which has been validated with high accuracy across multiple stations. This makes it a practical and reliable method for describing the influence of snow cover on frozen soil depth, despite the lack of a physical explanation for the units.

Comments 9: L335: Please be consistent with terminology, do not interchangeably use primary parameters and major parameters.

Response 9: Thank you for the reviewer's correction. To avoid confusion between "primary parameters" and "major parameters," we have standardized the terminology to "major parameters" throughout the revised manuscript to ensure consistency in expression.

Comments 10: L346 – 349: It would be helpful to mark in Figure 2 which processes in SNOW17 were calibrated with measured data, and which were not.

Response 10: Thank you for your suggestion. In Figure 2, the SNOW17 model utilized remote sensing snow depth data to calibrate the simulated snow data. We have updated Figure 2 in the revised manuscript to clearly indicate the scope and method of calibration within the model.

Comments 11: L349 – 352: I am not sure what this actually means. You definitely need more parameters for GXAJ-S than GXAJ, because you are adding physical processes. Are you saying that just because the -S module is compartmentalized in a module that means that you do not add more parameters to GXAJ? I think that this is a confusing way to describe one-way coupling.

Response 11: Thank you for your valuable comments. We appreciate the opportunity to clarify this point. In our study, the goal was to improve model performance by introducing new physical processes while minimizing unnecessary adjustable parameters. While incorporating snowmelt (-S) and freeze-thaw (SF) processes required additional parameters to describe these physical processes, we ensured that no extra adjustable parameters were added to the model by adopting a fixed-parameter strategy during model construction. Below, we address your comments in more detail:

- Snowmelt process parameters: To enhance the effectiveness of the model improvement and avoid the possibility that the introduction of additional parameters could potentially improve simulation results, the SNOW17 model was initially run independently. Remote sensing snow depth data (considered as "measured values") were used as input, and the parameters were adjusted to align the model-simulated snow depth with the "measured values," thereby determining the snow parameters for the study area. This approach allowed the integration of the SNOW17 model with the GXAJ model to form the GXAJ-S model for calculating snowmelt runoff in grid cells, ensuring that no new parameters were added to the GXAJ-S model compared to the GXAJ model (Section 2.2.3 "*Model parameters and calibration*" lines 394-402).

- Freeze-thaw process parameters: We utilized empirical parameters, which were also fixed in the model and not treated as additional adjustable parameters during the coupling process. By doing so, the freeze-thaw process was incorporated without introducing additional degrees of freedom (Section 2.2.3 "*Model parameters and calibration*" lines 402-407, Section 2.2.2).

By adopting this fixed-parameter approach, we ensured that improvements in model performance were achieved solely through the introduction of new physical processes, rather than by increasing model complexity with additional degrees of freedom. This clarification has been added to the revised manuscript to ensure transparency and a clearer understanding of the model parameterization strategy.

Once again, we thank the reviewer for highlighting this important aspect, which has helped us refine the clarity of our work.

Comments 12: L402 – 404: I think that the evidence of robustness is that the model *did not perform worse* during the validation period. Performing *better* during the validation period is not evidence of robustness. Conversely, performing *better* during the validation period suggests that you made some assumptions about the physical processes hard coded into the model, that were more valid during the validation period. Please discuss this in more detail if possible.

Response 12: Thank you for highlighting this important issue. We agree with the reviewer that better model performance during the validation period is not conclusive evidence of robustness. Instead, it may indicate that some assumptions about the physical processes hard-coded into the model align more closely with the conditions during the validation period.

We revisited the conceptualization of the snowmelt and freeze-thaw processes in the SNOW17 model and hypothesize that the improved performance during the validation period could be attributed to simpler hydrological conditions. Specifically, snow depths were generally lower during the validation period compared to the calibration period. Shallower snow conditions may reduce the complexity of snowmelt and freeze-thaw interactions, which are more prone to introducing uncertainties under deeper snow conditions. Consequently, the model was better able to capture these relatively simpler hydrological processes, resulting in improved performance during the validation period.

That said, we acknowledge that other factors might have contributed to the observed performance differences. For example, changes in climate conditions, precipitation patterns, or temperature distributions between the calibration and validation periods could also have influenced model behavior. While these possibilities cannot be entirely ruled out, we believe the model demonstrated reasonable stability during both periods, adapting to varying snow conditions and reliably simulating snowmelt and runoff processes.

In the revised manuscript, we have included a detailed discussion to explain the potential link between improved performance during the validation period and the simpler hydrological conditions (Section 3.1 "*Simulation of snow accumulation and freeze-thaw process*" lines 454-470). Future studies will aim to test the model under a wider range of conditions and across different regions to further evaluate its robustness and identify potential limitations.

We appreciate the reviewer's valuable insights and hope this explanation addresses the concern. If there are additional suggestions or areas you would like us to explore, we are happy to incorporate them.

Comments 13: Figure 4: What are the dashed lines?

Response 13: Thank you for your question. The dashed lines in Figure 4 represent the trend of snow depth variations (both model-simulated and remote sensing data) over the period from 2000 to 2018. We have revised the figure legend in the manuscript to provide a more detailed explanation to avoid confusion and ensure that the information presented is clearer and more intuitive for readers.

Comments 14: Figure 5: What is the dashed line?

Response 14: Thank you for your question. The dashed lines in Figure 4 represent the trend of snow depth variations (both model-simulated and remote sensing data) from 2000 to 2018. We have updated the manuscript to provide a more detailed explanation, ensuring clarity and avoiding any potential confusion for readers.

Comments 15: L427: You earlier defined an RBE, but not an RE.

Response 15: Thank you for your valuable feedback. We acknowledge that the definition of RE (Relative Error) in the original manuscript was incorrect, as it should have been abbreviated as RE instead of RBE. In the revised version, we have corrected its definition and description to ensure consistency in terminology and to prevent any confusion for readers.

Comments 16: L539 – 534: I think that the formation of a saturated layer above ground under these circumstances is possible only for very coarse soils that are inefficient at soil water redistribution. This is unlikely to be a general behavior. If you are referring to a specific soil type, please describe it. If you are claiming this as a general behavior, please provide references.

Response 16: Thank you for your valuable suggestion. The formation of a saturated layer above the frozen ground under freeze-thaw conditions is a well-documented phenomenon. This occurs primarily due to the low permeability of frozen soil, which prevents infiltration and causes meltwater or precipitation to accumulate at the base of the thawed layer (i.e., the top of the frozen soil layer). This effect is particularly pronounced when the thawed layer is thin, as water can easily accumulate at the freeze-thaw interface, leading to the formation of a saturated layer.

Numerous studies have reported this phenomenon, particularly highlighting how water accumulation at the freeze-thaw interface can result in a saturated layer above the frozen soil. We have included relevant references in the revised manuscript to further support this argument (Section 4.2 "*The impact of seasonal frozen ground/snow*" lines 693-700).

Comments 17: L540: If matric potential is the primary driver of moisture movement, then how does gravity cause a saturated layer to emerge at the frozen interface?

Response 17: Thank you for your thoughtful comment. We have carefully considered your suggestion and revisited the relevant statement. We realized that the main idea of this section is to emphasize the inhibitory effect of frozen soil on soil evapotranspiration, and the driving factors of soil water movement (such as matric potential) are not directly related to this mechanism. Therefore, to avoid confusion, we have removed the statement regarding "matric potential being the primary driver of moisture movement." We hope this revision addresses your concern and makes the explanation clearer (Section 4.2 "*The impact of seasonal frozen ground/snow*" lines 693-710).

Comments 18: L548 – 549: Which processes are you referring to, and what impacts? Are the processes you study not already naturally part of the local hydrological cycle and ecosystem?

Response 18: Thank you for your valuable feedback. The "processes" mentioned in the text, including freeze-thaw dynamics, soil water movement, and the effects of snow and seasonally frozen ground (SFG) on evapotranspiration, are indeed integral parts of the local hydrological cycle and ecosystem. However, we would like to emphasize that this study specifically focuses on how these processes-especially the freeze-thaw cycle and its effects on soil moisture and evaporation rates-can vary under different environmental conditions.

For instance, during the freezing period, frozen ground and snow cover suppress evapotranspiration. In contrast, during the thawing period, the formation of a saturated layer above the frozen soil may alter soil permeability and runoff patterns. These changes are particularly important in the context of climate change, as variations in the freeze and thaw periods could impact water resources and the stability of ecosystems over the long term.

In the revised manuscript, we have made these points more explicitly and further explained the potential impact of these processes (Section 4.2 "*The impact of seasonal frozen ground/snow*" lines 693-710).

Comments 19: L580 – L590: I feel that this is self-evident. It is a rehash of the widely known problem that data-calibrated hydrological models are often 'right for the wrong reasons'. It is a nice discussion that fits the work done, but does not contribute new knowledge.

Response 19: Thank you for your valuable feedback. We acknowledge your concern that this discussion might reiterate a well-known issue in hydrological modeling. To address this, we have revised and restructured the Discussion section (see revised manuscript) to better highlight the novel contributions of our study.

Specifically, we have refined this section ("*4.1 Key limitations in hydrological models in relation to their process complexity*") to emphasize the limitations of data-calibrated hydrological models and clarify how our approach addresses these challenges:

"*A limitation in the application of the GXAJ base model, which neglects impacts of snow and ice, is related to the fact that the parameters of its modules are determined based on historical basin characteristics. Although such models without frozen ground components can, through appropriate calibration or optimization of parameters, in some cases successfully reproduce historical hydrological processes in cold regions under stable conditions (Li et al., 2011; Zhang et al., 2017), they may not be suitable for evaluating the consequences of future changes as their* *calibrated values do not represent new conditions of the basin, and as the model lacks physical representation of key drivers of change. Our study demonstrates that incorporating the effects of seasonally frozen ground (SFG), snow, and other environmental factors into a basic model can provide robust and physically consistent results in simulating large-scale hydrological processes in cold regions, which can be particularly important for predicting hydrological impacts of future climate change scenarios. Key limitations in previous studies incorporating snow and seasonally frozen ground processes into hydrological models include taking 0°C as an assumed critical temperature for phase change, while neglecting the energy flux exchange between snow and soil layers, as for instance done in the VIC model (Liang et al., 1994). This potentially compromises simulation accuracy, particularly in regions where snow-frozen ground interactions are significant, such as areas with large seasonal variations in frozen ground depth. Similarly, while the SWAT model adjusts its parameters to accommodate permafrost conditions and introduces a new soil temperature module (Fabre et al., 2017), it does not account for the impact of snow depth on frozen ground and struggles to fully capture the complex and dynamic interactions between frozen ground and soil hydrological processes. Models such as the WEB-DHM (Qi et al., 2019), which employs enthalpy-based snow and frozen ground coupling, and the ATS model (Jafarov et al., 2018), have made substantial progress in simulating multi-year snow and frozen ground dynamics. However, their high complexity and demanding data requirements limit their applicability in large-scale hydrological simulations.*

In this context, while the here developed GXAJ-S-SF model does not rely on above-mentioned limitations of the VIC and SWAT models (as it accounts for the impact of snow depth variations on frozen ground and also captures multidimensional effects of snow-frozen ground coupling on hydrological processes) it still contains energy-related simplifications that makes it less dependent on extensive input data than e.g. the ATS model. In particular, the original lumped SNOW17 model (incorporated as a module in the GXAJ-S and GXAJ-S-SF model) was decentralized, and the energy exchange at the snow-atmosphere-soil interface was accurately simulated based on empirical relationships (He et al., 2011)."

Comments 20: L600 – 602: This argument is valid only if the modeled processes are linear. The processes you have modeled are potentially too nonlinear and have too many interactions for this argument to hold.

Response 20: Thank you for your insightful comment. We acknowledge that the argument regarding error propagation is more straightforward in linear systems and may not fully hold in highly nonlinear hydrological processes. To address this concern, we have revised the discussion (see Section 4.1, "Key limitations in hydrological models in relation to their process complexity", L642–L659) to clarify the potential limitations of our approach, specifically as follows:

"*In complex mountainous cold regions, observation remains a bottleneck (Gao et al., 2022). Due to limitations in measured data on frozen soil and snow depth in the considered Yalong River basin, satellite-based snow depth data and ground temperature station data were used in the present study for calibration and verification. In particular, errors in remote sensing snow depth data (Yan et al., 2022; Zou et al., 2014) can propagate to the model output. However, previous studies have specifically investigated the here used remote sensing dataset for the Yalong River basin showing that its accuracy is high (Wu et al., 2024), which suggests that model errors should be relatively low. We also recognize that the use of surface temperature and maximum frozen ground depth to verify the freeze-thaw process introduces some uncertainty (Li et al., 2022), Since the GXAJ-S and GXAJ-S-SF model variants used the same temperature, snow and frozen ground data in the present simulations, they can be expected to share similar data errors, However, due the non-linear nature of the modeled processes, such data errors may still not cancel completely when comparing different models. Nevertheless, observed differences in model performance between these models are mainly expected to reflect differences in model capabilities rather than differences in input datasets. Future work should focus on improving remote sensing data quality and exploring the long-term robustness of the model to further enhance performance and improve our understanding of the freeze-thaw processes in complex mountainous cold regions*."

We appreciate your valuable feedback, which has helped us refine our discussion and ensure a more rigorous interpretation of model uncertainties.

Comments 21: L603 – 604: I agree that remote sensing errors would probably not affected the core conclusions of this manuscript, but not for the reasons you provide in L600 – 602. As I explained in my general comments, I think your conclusions are mostly self-evident.

Response 21: Thank you for acknowledging that remote sensing errors are unlikely to affect the core conclusions of this study. Regarding your comment that our conclusions appear mostly self-evident, we would like to clarify that the hydrological impact of seasonally frozen ground (SFG) at a large basin scale remains insufficiently understood (Gao et al., 2022). The relative importance of SFG processes across different environments and their broader significance in larger basins are still not well understood (Ala-Aho et al., 2021). Our study provides direct evidence that SFG processes play a crucial role in shaping large-scale runoff dynamics in the Yalong River basin.

This conclusion is drawn from the comparative performance of the GXAJ-S-SF model, which explicitly accounts for both snow and SFG effects. Compared to the GXAJ-S and GXAJ models, which do not incorporate SFG, the GXAJ-S-SF model demonstrates significant improvements in simulating different runoff components. Notably, the original GXAJ model tends to slightly overestimate runoff during the calibration period and underestimate it during validation, with the underestimation being more pronounced in spring when snowmelt and frozen ground processes are dominant. The improved GXAJ-S-SF model reveals that increased surface runoff in cold months corresponds with a significant reduction in interflow and groundwater runoff. This insight contributes to a more comprehensive understanding of cold-region hydrological processes.

While incorporating snow and frozen ground effects leads to improvements in statistical performance metrics and runoff simulations-findings that may seem intuitive-our study is not solely focused on optimizing model accuracy. Instead, as highlighted in our response to Comment 1, our primary goal is to enhance the physical realism of hydrological modeling and deepen our understanding of cold-region hydrological processes. The consistency between our results and expected physical behavior reinforces the credibility of our approach and underscores the broader significance of SFG in shaping hydrological regimes.

We appreciate your feedback, which has helped us refine our discussion and emphasize the broader implications of our study.

Comments 22: L606 – 615: The benchmark model GXAJ you refer to is not a different model, but it is just GXAJ-S-SF without the snow and freezing capacities. This discussion is not meaningful because it is self-evident.

Response 22: Thank you for your comment. We acknowledge that the GXAJ model serves as the baseline version of GXAJ-S-SF, differing only in its exclusion of snow and frozen ground processes. However, we would like to emphasize that the significance of our comparison lies not in merely adding more processes but in demonstrating how these physically meaningful processes influence hydrological responses(see Section 4.1, “Key limitations in hydrological models in relation to their process complexity”, L660–L668).

“*Hydrological modeling typically prioritizes model fitness, which in theory can be improved by introducing more fitting parameters. However, this study highlights differences that are due to addition of process-based modules (regarding snow and frozen ground). This implies that improvements in model fit and differences in associated model output (e.g. runoff and evapotranspiration) reflect how the considered snow and/ or frozen ground processes more concretely alter hydrological flows. This therefore increases the understanding of underlying hydrological processes (Gao et al., 2022) in large-scale applications such as the Yalong River basin that additionally has a complex topography with large elevation differences yielding high spatio-temporal heterogeneity in snowmelt and freeze-thaw cycles of soil.*”

Additionally, to better highlight the novelty of our study, we have made substantial revisions throughout the manuscript. If the reviewer has any further questions or suggestions, we would be happy to discuss them.

Comments 23: L615 – 617: What is the modular approach being contrasted against? What results did you show that support this statement?

Response 23: Thank you for your insightful comment. In response to your suggestion (Comments 5), we have thoroughly revised the manuscript to clarify the concept of the "modular approach" and ensure that its context and contrasts are more explicitly defined. Additionally, we have refined the discussion to better support our statements, incorporating feedback from Comment 22 as well. The revised manuscript now provides a clearer comparison of the modular approach with alternative modeling strategies and explicitly presents the results that support our conclusions. We appreciate your valuable feedback, and we welcome any further discussions to ensure clarity and rigor in our presentation.

Comments 24: L630 – 634: This is a great point, and could be expanded to make the discussion more interesting.

Response 24: Thank you for your positive feedback and for highlighting this important aspect. We appreciate your suggestion and have expanded this section accordingly (see Section 4.2, “The impact of seasonal frozen ground/snow”, L733–L751). In the revised manuscript, we have further elaborated on the broader implications of our findings, particularly regarding the challenges of applying hydrological models in cold regions with complex geological and environmental conditions. We emphasize the necessity of considering local variability when extending model applications, while also recognizing the potential for improving process understanding and generalizability.

"*This study quantitatively analyzed the impact of seasonal snow and frozen ground on hydrological processes based on the hydrological model, and its validity was confirmed not only by measured runoff but also by multi-source data, especially the trends in snow and frozen soil changes. Although our developed model has great application potential in other cold regions, it should be used cautiously without prior understanding of the modeling system. Snow and frozen ground are just part of the factors affecting cold-region hydrology, with other factors intertwined with frozen ground having significant impacts. Geological conditions, in particular, greatly affect frozen ground but have large spatial heterogeneity and are challenging to measure. The empirical parameters of the SNOW17 model and Stefan equation have clear physical significance and have been validated by previous studies (Anderson, 2006; Ran et al., 2022; Zou et al., 2014). However, the soil and geology of mountainous basins are extremely complex and vary significantly across regions. This complexity introduces challenges in applying these models to different watersheds, requiring recalibration of their values. For instance, soil texture, moisture retention, and thermal properties can vary considerably, influencing the depth and dynamics of the seasonal frozen ground. Similarly, variations in topography, vegetation cover, and geological composition can impact runoff, infiltration, and evapotranspiration processes. Expanding the application of complex hydrological models therefore requires careful attention to local and regional variability in ambient conditions, but may also considerably increase the understanding of processes and the generalizability of the assumptions made*."

We sincerely appreciate your insightful comments, which have helped us strengthen the discussion and enhance the clarity of our study.

Response to RC2:

General comments

Comments 1: The authors present a modeling study on the hydrological impacts of snow and frozen ground dynamics in a topographically complex basin. The topic of cryospheric changes and their impacts on hydrology is both significant and timely. However, the authors should address several key issues in the current manuscript to enhance its quality before it can be considered further.

I think the novelty of this study is not sufficiently distinctive or well-highlighted. There have already been numerous modeling studies on snow and frozen ground dynamics in the Tibetan Plateau region, both the basin-scale and regional-scale studies are conducted. Moreover, the models employed in previous studies provided more advanced representations of snow and frozen ground processes, particularly in terms of frozen ground dynamics, compared to the model used in this study. Therefore, the authors need to consider how to better emphasize the unique contributions of this study in comparison to prior research.

Response 1: Thank you for your thoughtful evaluation and constructive feedback. We appreciate your comments regarding the novelty of our study, which prompted us to further clarify and emphasize its unique contributions.

In the revised manuscript, we have explicitly highlighted the following key aspects that distinguish our study:

- **Hydrological impacts of snow and frozen ground in large basins:**

While numerous studies have examined the role of snow and frozen ground in hydrological processes, most have focused on small-scale or regional-scale applications. However, at a large basin scale, key questions remain regarding the extent to which seasonally frozen ground (SFG) influences runoff and how model complexity affects hydrological simulations (lines 57-70). Our study systematically evaluates these effects by comparing models of different complexity levels. To emphasize this contribution, we have revised the title, discussion, and conclusion sections accordingly.

- **A simple and data-efficient snow and freeze-thaw coupling method:**

Unlike many existing models that require extensive parameterization of freeze-thaw dynamics, our study integrates snowmelt and freeze-thaw processes in a physically meaningful yet computationally efficient manner. The developed snow and freeze-thaw coupling module requires relatively few additional parameters and has low dependence on input data—an essential advantage for data-scarce cold regions. This approach makes it feasible for large-scale applications, particularly in regions lacking detailed soil freeze-thaw observations (see Section "4.1 Key limitations in hydrological models in relation to their process complexity").

- **Quantitative assessment of the impacts of SFG on hydrological processes in large basins:**

Our study provides a systematic comparison between simplified models (which exclude SFG processes or consider only snow processes) and extended models that incorporate both snow and SFG effects. The results reveal that SFG can significantly alter hydrological components, increasing surface runoff during cold months (by 39%–77% compared to models that ignore SFG) while reducing interflow and groundwater runoff. Compared to previous studies, our approach not only quantifies these effects at a large basin scale but also provides insights into how different model structures influence runoff predictions. These findings have been further emphasized in the revised Abstract, Introduction, Result and Discussion sections.

- **Combining hydrological process understanding with practical applications:**

By analyzing frozen soil depth and its spatiotemporal impact on hydrological processes, this study enhances our understanding of the complex feedback mechanisms between frozen ground and hydrology. These insights not only improve our knowledge of snow, freeze-thaw, and hydrological interactions but also provide valuable references for predicting hydrological changes in cold mountainous regions under future climate change scenarios. Additionally, the model's relatively simple parameterization and efficient data requirements make it suitable for practical applications in data-limited regions, where many existing models may face constraints due to high data demands.

We have incorporated these refinements throughout the manuscript to better highlight the contributions of this study. Once again, we sincerely appreciate your valuable comments, which have helped us improve the clarity and impact of our work.

Specific comments

Comments 2: In Figure 3c, it is evident that a significant portion of the study area is covered by permafrost. However, the Stefan model mentioned in the methodology is designed to model seasonal frozen ground. Did the authors separately account for the dynamics of permafrost in their study? If not, this could be a critical limitation that needs to be addressed or clarified.

Response 2: Thank you for highlighting this important aspect. As shown in Figure 3c, the study area is primarily dominated by seasonal frozen ground (SFG), while permafrost accounts for less than 10% and is sparsely distributed along the edges of the study area. This distribution is consistent with the characteristics of the region, though some of the permafrost extent in the dataset may be subject to uncertainty, particularly in boundary areas.

Since our study focuses on the hydrological impacts of seasonal frozen ground, we employed an improved Stefan model that is specifically designed to simulate SFG processes. This model is well-suited for the dominant frozen ground type in our study area. The simulation results demonstrate high accuracy in these regions, confirming the model's effectiveness in capturing key hydrological processes.

Regarding the limited permafrost areas, we acknowledge that our model does not explicitly account for permafrost dynamics. However, given its sparse distribution and small coverage, the influence of permafrost on basin-scale runoff processes is expected to be minimal. Additionally, in permafrost regions, the active layer undergoes seasonal freeze-thaw cycles, which are conceptually similar to SFG processes. As a result, our approach remains applicable to the majority of the study area, and any potential impact of permafrost on overall simulation results is likely negligible.

Comments 3: Line 272-275: How was this threshold 30cm determined? Was a sensitivity analysis conducted to assess the impact of this threshold on the results? Providing such an analysis would help evaluate the robustness of the study's findings.

Response 3: Thank you for the valuable question! The 30 cm threshold mentioned in our study is based on findings from previous research. Numerous studies have explored different snow depth thresholds. For example, Brooks et al. (1995, 1999) and Cline (1995) suggested that when snow depth reaches 30–40 cm, air temperature has little influence on ground temperature. Building on this, Hill (2015) developed a conceptual model indicating that thick snow cover (>30 cm) effectively insulates the ground, preventing deep freezing and enabling groundwater recharge. Conversely, for thin snow cover (<30 cm), the ground remains seasonally frozen during the snowmelt period, which delays groundwater recharge and shifts hydrological responses later into the summer.

We adopted Hill's (2015) 30 cm threshold for snow depth based on these well-established findings. Although we did not conduct a detailed sensitivity analysis in this study, previous research has demonstrated that this threshold is widely applicable across different cold regions. Additionally, the simulated hydrological response in our study aligns well with observed trends, further supporting the appropriateness of this threshold.
To address this point more explicitly, we have revised the manuscript to provide a more detailed explanation of the threshold selection and added relevant references. We sincerely appreciate the reviewer's insightful comment!

Comments 4: In Table 3, the authors utilized several data products from other studies. However, the accuracy of these datasets, particularly the snow depth data, which is critical for this study, has not been clarified.

Response 4: Thank you for pointing out this important aspect. We acknowledge that the accuracy of the datasets used in our study, particularly the snow depth data, is critical for ensuring reliable hydrological simulations. In the revised manuscript, we have provided additional details on the accuracy and validation methods of the datasets used, with relevant references to further clarify these aspects.

In particular, we have expanded the discussion on the uncertainty introduced by snow depth and frozen ground datasets and their potential impact on our results (see revised manuscript, Lines 642–659):

"*In complex mountainous cold regions, observation remains a bottleneck (Gao et al., 2022). Due to limitations in measured data on frozen soil and snow depth in the considered Yalong River basin, satellite-based snow depth data and ground temperature station data were used in the present study for calibration and verification. In particular, errors in remote sensing snow depth data (Yan et al., 2022; Zou et al., 2014) can propagate to the model output. However, previous studies have specifically investigated the here used remote sensing dataset for the Yalong River basin showing that its accuracy is high (Wu et al., 2024), which suggests that model errors should be relatively low. We also recognize that the use of surface temperature and maximum frozen ground depth to* *verify the freeze-thaw process introduces some uncertainty (Li et al., 2022), Since the GXAJ-S and GXAJ-S-SF model variants used the same temperature, snow and frozen ground data in the present simulations, they can be expected to share similar data errors, However, due the non-linear nature of the modeled processes, such data errors may still not cancel completely when comparing different models. Nevertheless, observed differences in model performance between these models are mainly expected to reflect differences in model capabilities rather than differences in input datasets. Future work should focus on improving remote sensing data quality and exploring the long-term robustness of the model to further enhance performance and improve our understanding of the freeze-thaw processes in complex mountainous cold regions.*"

We sincerely appreciate the reviewer's insightful suggestion, which has helped us improve the robustness and clarity of our study.

Comments 5: Line 309-404: In points with high snow depth, there are significant discrepancies between the model results and the remote sensing data. The authors should investigate the underlying causes of these differences.

Response 5: Thank you for your valuable comment. We acknowledge that discrepancies exist between the model results and remote sensing data in regions with high snow depth. One potential reason is that hydrological processes in areas with deep snow are more complex. The model employs a simplified parameterization approach to simulate snow accumulation and melt processes, which may not fully capture the intricate dynamics of snow accumulation-melt cycles. These factors may also explain why the model performance during the validation period (when snow depth is relatively shallow) is better than during the calibration period (when snow depth is deeper), as shown in Figure 4. However, despite these differences in areas with deep snow, the calibration and validation results still show relatively low RMSE and BIAS values, indicating that the model performs well overall in simulating snow depth dynamics. Furthermore, compared to the original model, the improved model exhibits significant improvements in simulating snowmelt runoff.

We acknowledge that snowmelt is a complex hydrological process, and given the limited data availability, we have made efforts to utilize existing observations and remote sensing data to simulate the snowmelt process with the highest possible accuracy. While certain limitations and uncertainties are inevitable, we believe that such efforts are meaningful and valuable, especially in cold regions where data scarcity poses significant challenges.

In the revised manuscript, we have further investigated and discussed the potential causes of these discrepancies, including the limitations of both remote sensing data and the model itself (see revised manuscript, Lines 454-470, and Section “Discussion”). Additionally, we have expanded the uncertainty analysis section to provide a more comprehensive discussion on this issue and suggest directions for future improvements (Section “4.1 Key limitations in hydrological models in relation to their process complexity”,L642-659).

Comments 6: Line 412:415: For the simulation of frozen ground processes, verifying only the accuracy of the initial freeze and initial thaw dates is far from sufficient. It is also essential to validate the simulated soil temperature and soil moisture (including liquid water content and soil ice content). These variables are key to understanding how freeze-thaw processes influence basin hydrology. Therefore, the authors should provide validation results for these variables to demonstrate the reliability of the study's findings.

Response 6: Thank you for your insightful comment. We acknowledge the importance of validating soil temperature and soil moisture (including liquid water content and soil ice content) to further demonstrate the reliability of the simulated freeze-thaw processes. However, due to the lack of measured frozen soil depth, soil temperature, and soil moisture data—an issue common in most cold regions-we are currently unable to directly validate these variables.

To indirectly support the reliability of the simulated freeze-thaw processes, we used available surface temperature data to verify the initial freeze and thaw dates, a method that has been supported by multiple studies. In addition, to further assess the model’s reliability, we compared the spatial distribution of the maximum frozen soil depth simulated in this study with data from the Tibetan Plateau Permafrost Dataset (1961–2020) for the 2000s. The results showed a high degree of consistency in both spatial distribution patterns and magnitude, with a correlation coefficient of 0.89, providing additional validation support. These comparative analyses have been presented in the revised manuscript (Section 3.1, "*Simulation of snow accumulation and freeze-thaw process*" Lines 471-486).

We also recognize that verifying the freeze-thaw process using surface temperature and maximum frozen ground depth introduces some uncertainties (Li et al., 2022). Since both the GXAJ-S and

GXAJ-S-SF models used the same input datasets for temperature, snow, and frozen ground, they are expected to share similar data uncertainties. However, due to the nonlinear nature of the modeled processes, these errors may not completely cancel when comparing different models. Nevertheless, observed differences in model performance between these models are mainly expected to reflect differences in model capabilities rather than differences in input datasets.

We agree that future work should focus on improving remote sensing data quality and exploring the long-term robustness of the model to further enhance its performance and improve our understanding of freeze-thaw processes in complex mountainous cold regions. We have incorporated this discussion in Section 4.1 ("*Key limitations in hydrological models in relation to their process complexity*" Lines 642-659).

Comments 7: The 'Results' section is too brief and lacks depth in describing the characteristics of snow and frozen ground changes and their hydrological effects. For instance, there is insufficient discussion on how frozen ground processes alter soil temperature and moisture conditions, thereby influencing hydrological processes, as well as how snow changes directly impact runoff. Additionally, the manuscript does not adequately address how snow affects frozen ground processes and thereby indirectly impacts hydrology. Moreover, compared to the analysis of the differences in runoff simulations using various modules, I believe it would be more meaningful to explore how the synergistic changes in snow and frozen ground under climate change influence runoff in the study area during the past decades.

Response 7: We sincerely appreciate the reviewer's insightful comments and suggestions. We fully agree with your concerns and have significantly expanded the "Results" section to provide a more comprehensive description of snow and frozen ground dynamics and their hydrological effects.

In the revised manuscript, beyond the original content-covering the simulation of snow accumulation, freeze-thaw processes, runoff, and the hydrological impacts of snow and frozen ground (on runoff components and evapotranspiration)-we have made the following additions:

- Seasonal variations in snow and frozen ground to provide a more detailed temporal perspective.
- The influence of snow depth on frozen ground dynamics, highlighting the snowpack's insulating effect on soil freezing and thawing.

- The impact of snow and frozen ground on soil moisture, further illustrating their indirect influence on hydrological processes.

Additionally, we have strengthened the discussion on the contribution of snowmelt to runoff, ensuring that the hydrological impacts of these processes are more thoroughly analyzed. To address these revisions, we have made comprehensive modifications throughout the manuscript, including refinements to the research objectives, methodology, results, and uncertainty discussions, to better emphasize the contributions of this study.

Regarding your suggestion to explore the synergistic effects of snow and frozen ground changes on runoff under climate change over past decades, we fully acknowledge the importance of this research direction. However, the primary goal of this study is not only to analyze runoff differences under different model configurations but also to develop and evaluate an enhanced hydrological model for cold regions. This model integrates snowmelt and freeze-thaw processes through a modular and computationally efficient design, making it widely applicable. Furthermore, we quantitatively assess the contribution of snowmelt to runoff and the effects of frozen ground on soil conditions, runoff components, and evapotranspiration. While our findings contribute to a better understanding of the complex hydrological processes in cold regions, the study does not aim to systematically analyze the long-term impacts of climate change on runoff. We have clarified this research objective in the revised manuscript.

Additionally, due to the lack of long-term observed runoff data in the study area, our current analysis is limited to hydrological simulations from 2000 to 2018. A longer-term investigation would require more extensive observational data, which remains a direction for future research.

Once again, we sincerely appreciate your valuable feedback. We have carefully revised the manuscript to clarify and expand relevant discussions, ensuring a more thorough response to your suggestions.

Comments 8: In the 'Discussion' section, the authors should focus on how their findings represent an advancement over previous research and then call back to the research questions outlined in the Introduction. Rather than including an extensive literature review, the discussion should emphasize the novel contributions of this study and its implications for the field.

Response 8: We appreciate the reviewer's suggestion to refine the discussion by emphasizing the novel contributions of our study rather than providing an extensive literature review. In the revised manuscript, we have significantly restructured the Discussion section to focus on the advancements our study brings to the field and how these findings address the research questions outlined in the Introduction.

Specifically, we have divided the Discussion into two subsections:

- **Key limitations in hydrological models in relation to their process complexity** – Here, we critically analyze the limitations of existing models (e.g., VIC, SWAT, and WEB-DHM) in representing snow and seasonally frozen ground (SFG) processes and highlight how our newly developed GXAJ-S-SF model overcomes these challenges by improving physical realism while maintaining computational efficiency.
- **The impact of seasonal frozen ground and snow** – This section emphasizes how our model enhances the understanding of key hydrological processes influenced by SFG, including its effects on runoff components, evapotranspiration, and seasonal water storage dynamics. The discussion also connects back to the research questions by demonstrating how incorporating SFG improves hydrological simulations in cold regions and offers insights into climate change impacts on water resources.

By restructuring the Discussion in this manner, we ensure that the focus remains on our study's key contributions and implications, rather than reiterating a broad literature review. We hope this revision better aligns with the reviewer's expectations.

Thank you for this valuable feedback.