Response to RC1:

General comments

Comments 1: Most of my concerns from the original submission have been addressed satisfactorily. The current manuscript is good work, and I enjoyed reading it. I recommend that this manuscript be considered for publication after minor revisions to make some small but important improvements that would greatly improve the clarity of the manuscript.

Response 1: Thank you for your positive feedback and for recognizing our efforts in improving the manuscript. We appreciate your constructive comments, which have helped enhance its clarity and quality. Below, we provide detailed responses to your specific suggestions and outline the corresponding revisions to further improve clarity and precision.

Specific comments

Comments 2: Your response 11: You write that "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.". Would it not be clearer if you expressed this as "compared to the GXAJ model, no new parameters were added to the GXAJ- component of the GXAJ-S model"?

This would imply that all additional complexity in GXAJ-S is compartmentalized in the -S part only, which is my understanding of your work.

Response 2: Thank you for your insightful suggestion. We appreciate this clarification, as it more precisely conveys that the additional complexity in the GXAJ-S model is confined to the -S component, while the GXAJ component remains unchanged in terms of parameters. We have incorporated this revision into the manuscript (see lines 404–406):

"*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.* ***Compared to the GXAJ model, no new parameters were added to the GXAJ component of the GXAJ-S model.***"

Comments 3: Your response 10: It is still not clear to me which parameters you calibrated in your study, and which you used a fixed value for. In L384-386, you state that SNOW17 has 4 parameters that must be calibrated, and 6 that do not have to be. However, in Table 2 you list a “prior range” for 9 of these parameters, with the word “prior” implying that you calibrated all 9 of these to yield a posterior value. In Table 1 it appears that all of the listed parameters are calibrated because you use the phrase “prior estimate”. Please indicate in Table 1 and Table 2 which parameters you calibrated from the field data, and which parameters you used certain values for without any calibration.

Response 3: We sincerely appreciate the reviewer’s insightful comments regarding parameter calibration and fixed values in Tables 1 and 2. To improve clarity, we have explicitly distinguished parameters obtained from external sources (e.g., literature, remote sensing datasets, or soil properties) from those calibrated in our study. Specifically, we have:

Revised Table 1 to indicate which GXAJ model parameters were estimated from external sources and which were calibrated.

Updated Table 2 to clearly label SNOW17 parameters as either ‘calibrated’ or ‘fixed value.’

Table 1. *GXAJ model parameters and their descriptions.*

Module	*Parameter*	*Description*	***Source or Calibration***
Canopy interception	LAI_{max}	*Maximum LAI for the vegetation in a year*	***Derived from LDAS based on vegetation types***
	h_{lc}	*Height of vegetation (m)*	***Derived from LDAS based on vegetation types***
Channel precipitation	W_{ch}	*Channel width within a cell (km)*	***Estimated based on measured cross sections***
Evapotranspiration	W_{UM}	*Tension water capacity of upper layer (mm)*	***Estimated based on initial W_M***
	W_{LM}	*Tension water capacity of lower layer (mm)*	***Estimated based on initial W_M***
	C	*Evapotranspiration coefficient of deeper layer*	***Estimated based on LAI and h_{lc} of vegetation***
	K	*Ratio of potential evapotranspiration to pan evaporation*	***Calibrated (prior range: 0 – 1)***
Runoff generation	W_M	*Tension water capacity (mm)*	***Estimated using θ_{fc}, θ_{wp} and vadose zone thickness***
	θ_s	*Saturated moisture content*	***Obtained from literature based on soil types***
	θ_{fc}	*Field capacity*	***Obtained from literature based on soil types***

	θ_{wp}	*Wilting point*	***Obtained from literature based on soil types***
	S_M	*Free water capacity (mm)*	***Estimated using $\boldsymbol{\theta_s}$, $\boldsymbol{\theta_{fc}}$ and humus layer thickness***
	K_i	*Outflow coefficient of free water storage to interflow*	***Estimated based on soil properties***
	K_g	*Outflow coefficient of free water storage to groundwater*	***Estimated based on soil properties***
	C_i	*Recession constant of interflow storage*	***Calibrated (prior range: 0 – 1)***
Flow routing	C_g	*Recession constant of groundwater storage*	***Calibrated (prior range: 0 – 1)***
	C_s	*Recession constant in the lag and route technique*	***Calibrated (prior range: 0 – 1)***
	L_{ag}	*Lag time*	***Calibrated (prior range: ≥0)***

Table 2. *SNOW17 model parameters and their descriptions.*

	Parameter	*Description*	***Calibration or Fixed Value***
Major parameters	SCF	*Snow correction factor, or gage catch deficiency adjustment factor*	***0.7 - 1.6 (calibrated)***
	$MFMAX$	*Maximum solar melt factor during non-rain periods, assumed to occur on June 21 (mm·°C-1·6hr-1)*	***0.5 - 2.0 (calibrated)***
	$MFMIN$	*Minimum solar melt factor during non-rain periods, assumed to occur on December 21 (mm·°C-1·6hr-1)*	***0.05 - 0.49 (calibrated)***
	$UADJ$	*The average wind function during rain-on-snow periods (mm·mb^{-1})*	***0.03 - 0.19 (calibrated)***
Minor parameters	NMF	*Maximum negative melt factor (mm·mb^{-1}·6hr^{-1})*	***0.45 (fixed value)***
	$TIPM$	*Antecedent temperature index parameter*	***0.9 (fixed value)***
	$PXTEMP$	*The temperature that separates rain from snow (°C)*	***0 (fixed value)***
	$MBASE$	*Base temperature for snowmelt computations during non-rain periods (°C)*	***0 (fixed value)***
	$PLWHC$	*Percent liquid water holding capacity for ripe snow (decimal fraction)*	***0.1 (fixed value)***
	$DAYGM$	*Constant daily amount of melt which takes place at the snow-soil interface whenever there is a snow cover (mm·day-1)*	***0.7 (fixed value)***

These modifications ensure transparency in parameter selection and calibration. We appreciate the reviewer's suggestion, which has helped enhance the clarity of our methodology.

Comments 4: Model comparison: I suggest including a table with the computational time (including both the calibration time and actual simulation time) of simulating a comparable scenario for all 3 models GXAJ, GXAJ-S, and GXAJ-S-SF. That would provide readers with a more complete information on which model to choose. The additional physical detail of GXAJ-SSF may not be necessary in some applications that prioritize fast computation over accuracy.

Response 4: Thank you for your valuable suggestion. We have included a table in the revised supplementary material (Table S1) comparing the computational time of GXAJ, GXAJ-S, and GXAJ-S-SF under a comparable scenario. This table reports both calibration and simulation times, providing readers with a clearer understanding of the trade-off between computational efficiency and model complexity.

We have also described this addition in the revised manuscript (Lines 574–578). Furthermore, we have specified the computing environment, including processor details, memory, operating system, programming language, and the number of calibration iterations. This ensures transparency and helps users make informed decisions based on their computational resources and modeling needs. Below is the added content:

"*To provide a more comprehensive comparison of the three models, we have included an evaluation of computational efficiency. Table S1 presents the calibration and simulation times for GXAJ, GXAJ-S, and GXAJ-S-SF. The results indicate that while GXAJ-S-SF provides improved physical representation, it requires longer computation time compared to GXAJ and GXAJ-S. This information is useful for users who may prioritize efficiency over accuracy in certain applications. Please let us know if any further details are needed.*

Table S1. *Computational time comparison for the GXAJ, GXAJ-S, and GXAJ-S-SF models.*

Model	*Calibration Time (hours)*	*Simulation Time (seconds)*
GXAJ	*1.5*	*9*
GXAJ-S	*6.1*	*39*
GXAJ-S-SF	*6.7*	*41*

 # All simulations were conducted in the following computing environment: *AMD Ryzen 5 3600X 6-Core Processor, 32GB DDR4 2133MHz RAM, Windows 10 operating system, and MATLAB R2023a for model implementation and execution. The computations were performed in single-threaded mode, with 400 iterations set for the calibration period.*"

Response to RC2:

General comments

Comments 1: Thank you for the authors' efforts in improving the manuscript. This version of the paper shows improvement compared to the previous one. However, I still believe that the current paper is not yet suitable for publication in HESS. In my opinion, the novelty of the proposed contributions is insufficient, and the reliability of the conclusions remains inadequate.

Response 1: We sincerely appreciate your time and effort in reviewing our manuscript and for providing constructive feedback. We are grateful for your acknowledgment of the improvements made in this revised version. However, we regret that our work has not yet fully met your expectations regarding novelty and the reliability of our conclusions.

To address your concerns, we have undertaken substantial and comprehensive revisions throughout the manuscript (include Abstract, Introduction, Results, Discussion and Conclusions). These include enhanced validation using multi-source data, deeper interpretation of hydrological processes in cold regions, and expanded discussions that place our work in context with existing studies. These revisions aim to better clarify the novelty of our methodology and reinforce the robustness of our conclusions.

We have carefully addressed your concerns regarding novelty and reliability in our detailed responses below. We highly value your insights and would greatly appreciate any further suggestions you may have to help improve the quality of our work. Once again, thank you for your thoughtful critique and for helping us enhance our manuscript.

Specific comments

Comments 2: The authors emphasize the development of a new hydrological model that considers snow and seasonally frozen ground. However, in my view, the approach of simply coupling empirical formula-based modules into a hydrological model is not sufficiently innovative. Extensive research has already been conducted on this issue in the Tibetan Plateau region, as the authors themselves have also mentioned. Moreover, several existing models, such as the VIC model (Cuo et al., 2015), CLM4.5 (Yang et al., 2018), GIPL2.0 (Qin et al., 2017a), WEB-DHM (Song et al., 2020), and GBEHM (Gao et al., 2018), provide more comprehensive descriptions of snow and seasonally frozen ground modules. The authors repeatedly highlight that the model developed in this study requires fewer input data, thereby emphasizing its applicability in data-scarce regions. However, the input data required by this model are essentially the same as those required by the aforementioned physically based models, primarily including topographic data, vegetation data, and meteorological input data. On the contrary, due to the simplified representation of physical processes in the model, more reference data are needed for parameter calibration. Therefore, I do not consider the simplicity of the model's physical description to be an advantage.

Response 2: Thank you for your valuable comments. We understand your concerns regarding the innovation and applicability of the model and would like to further clarify the main contributions of this study and the scientific rationale behind the modeling approach adopted. Indeed, we recognize that the advantages of the here considered modelling approach may be context-dependent. Below, we provide a detailed response to your concerns.

➢ **Coupling of Frozen Ground and Hydrological Processes**

Our study does not simply integrate empirical formulas into a hydrological model; rather, it systematically coupled seasonally frozen ground (SFG) processes with key hydrological components. The Stefan equation was used to calculate the spatiotemporal distribution of frozen depth, which directly influenced soil moisture/ice content in the vadose zone. This in turn changed the effective thickness of the vadose zone and humus soil layer (including effective tension water storage capacity and free water storage capacity), ultimately affecting multiple hydrological processes such as runoff generation, runoff distribution and evapotranspiration (see Section 2.2). The freeze-thaw process of frozen soil is affected by snow conditions. The improved model in this study takes these hydrological physical processes into account.

➢ **Comparison with Physically Based Models**

Indeed, we recognize that the advantages of the here considered modelling approach may be context-dependent. Considering the Yalong River basin case, we therefore now include an in-depth comparison of the performance of the investigated simplified (relative) models with the performance of physically based models. 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. This hence motivates our refined investigation example (see below) regarding how physical models perform in comparison with simplified models.

To investigate this issue, we referenced the application of the VIC model and SWAT model in our study area from 2007 to 2011 (Li et al., 2018) and compared it with our proposed model. The results show that the simulation accuracy of the VIC model (NSE = 0.75 during calibration and NSE = 0.65 during validation) and the SWAT model (NSE = 0.77 during calibration and NSE = 0.66 during validation) did not exceed that of our model over the same period (NSE = 0.87 during calibration and NSE = 0.74 during validation). This may hence be related to the uncertainties introduced by the parameterization of physical models in data limited regions, and suggests the need to expand observational efforts before expanding modelling efforts to further improve predictive capacity (see Discussion, Lines 649–668).

"*Although significant progress has been made in physical models that account for snow and freeze-thaw processes, their application in cold-region hydrology remains challenging. Due to the complex topography, heterogeneous vegetation cover, and uneven soil moisture distribution in cold regions, uncertainties in radiation and surface albedo estimation can lead to inaccuracies in surface energy balance simulations, introducing errors in ground temperature and soil heat flux estimations (Gao et al., 2018). Additionally, the spatial parameterization of physical models remains a significant challenge, and their structural and parameterization schemes require further refinement (Zhou et al., 2021). The diverse climatic and geographic conditions in cold regions further limit the applicability of many physical models across different study areas (Yong et al., 2023). Moreover, the complexity and uncertainty of cold-region hydrological processes increase the difficulty of model development and parameter calibration, which may negatively impact simulation accuracy (Gao et al., 2018; Qin et al., 2017). To further assess the performance of physical models in our study area, we compared the VIC model's simulation results from 2007 to 2011 (Li et al., 2018b) with those obtained using our simplified model. The results indicate that the VIC model exhibited NSE values of 0.75 and 0.65 for the calibration and validation periods, respectively, which did not exceed those of our model (0.87 for calibration and 0.74 for validation). This comparison illustrates that the data limitations in the Yalong River basin are likely to currently constrain the performance of physically based models. This hence suggests the need to expand observational efforts before expanding modelling efforts to further improve predictive capacity..*"

➢ **Parameterization process**

We understand that the physical model can provide more explanations, and then its parameterization scheme is still a huge challenge. The enhanced model developed in this study integrates multiple key cold region hydrological processes while maintaining low parameter complexity (Section 2.2.3 Model parameters and calibration), making it particularly suitable for cold regions with complex hydrological and meteorological conditions and scarce data such as the Yalong River Basin.

➢ **Scientific Contribution and Innovation**

The novelty of this study lies not only in the coupling of snow, frozen ground, and hydrological processes (as mentioned above), but also in (see revised '*Abstract*', '*Results*' and '*Conclusion*' sections):

- Providing a quantitative analysis of the impact of snow/frozen ground on runoff partitioning and evapotranspiration (Section 3.3 '*Model differences in simulated runoff components and soil evapotranspiration*' and 4.2 '*The impact of seasonal frozen ground/snow on hydrological processes*').
- Demonstrating the complex interactions among snow cover, frozen ground, and the unsaturated zone (Section 3.1 '*Simulation of snow accumulation and freeze-thaw process*' and 4.2 '*The impact of seasonal frozen ground/snow on hydrological processes*').
- Offering a flexible and adaptable modeling framework that can be seamlessly integrated into hydrological models beyond GXAJ.

In summary, the enhanced modeling framework proposed in this study improves runoff simulations while providing new insights into the role of snow and frozen ground in shaping water balance components. The comparison with both the studied model set and more complex physically based models suggests that data limitations in the Yalong River Basin may currently constrain the performance of physically based models. This highlights the need for expanded observational efforts to improve predictive capabilities before extending physically based modeling approaches.

Thank you for your time and consideration.

Comments 3: Additionally, the authors primarily demonstrate the model's accuracy through the performance of streamflow simulations. However, this is far from sufficient for a study on hydrological processes in high-mountain basins, where multiple processes contribute to the overall dynamics. Given that the focus of the paper is on analyzing the impacts of snow and frozen ground on streamflow, detailed validation of these two critical intermediate processes is essential. However, such validation is currently lacking. For the snow module, the authors only use remotely sensed snow depth data for calibration and validation. However, the accuracy of these data remains uncertain, as it is well known that remote sensing of snow depth in the complex terrain of the Tibetan Plateau is subject to significant uncertainties. I recommend that the authors use more authoritative MODIS snow cover data to conduct a more comprehensive validation of the snow module results. As for the frozen ground module, the current validation is mainly limited to the start dates of freeze-thaw cycles (the results do not appear to be very satisfactory, and the authors have not provided quantitative metrics). There is a notable lack of validation for key physical variables, such as soil temperature and soil moisture (both ice and liquid water). I suggest that the authors collect in-situ measurements from the study region or validate their results against more authoritative remote sensing or reanalysis soil data to enhance the reliability of their findings.

Response 3: Thank you for your valuable comments on this study. In response to your concerns regarding the insufficient validation of the snow and frozen ground modules, we have made further improvements and validation efforts based on your suggestions:

➢ **Snow Module Validation:**

We acknowledge that remote sensing snow depth data may contain uncertainties, particularly in complex terrain. However, previous studies have specifically evaluated the dataset used in this study for the Yalong River Basin, demonstrating its high accuracy (Wu et al., 2024), which suggests that the model errors should be relatively low. To further enhance the validation, we have compared MODIS snow cover data with our model simulations. The results indicate that snow cover extended over up to half of the study area, with a high correlation coefficient (0.91) between the simulated and observed daily snow cover fractions. Figure S8 presents the spatial distribution of simulated snow depth alongside MODIS-derived snow cover on December 1, 2015, demonstrating strong consistency in coverage patterns (lines 669-701).

"*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, this study used multi-source remote sensing data and reanalysis data for calibration and verification from multiple perspectives. 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. This study further compared MODIS snow cover data with model simulations, revealing that snow cover extended over up to half of the study area, with daily snow cover fraction exhibiting a high correlation coefficient of 0.91 between the two datasets. Figure S8 illustrates the spatial distribution of simulated snow depth and MODIS-derived snow cover on December 1, 2015, demonstrating strong consistency in coverage patterns. We also recognize that the use of surface/soil 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*.

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."

➢ **Frozen Ground Module Validation:**

For the frozen ground module, we have further refined the validation process. In addition to validating the start dates of the freeze-thaw cycles and providing quantitative metrics, we have also incorporated ERA5 soil temperature data and spatial distribution data of maximum frozen soil depth. The comparison of these data shows that the simulated frozen ground depth from the model aligns well with the remote sensing/reanalysis data in terms of both time series and spatial distribution, with detailed results provided in lines 476-497.

"*This study systematically validated the simulation results of frozen soil depth based on the Stefan empirical formula through multi-source data comparison. Fig. 5 presents the frozen depth derived from ERA5 reanalysis data using four soil temperature layers (0–7 cm, 7–28 cm, 28–100 cm, and 100–289 cm; freezing occurs when layer temperatures fall below 0°C). The seasonal freeze-thaw depths calculated by the Stefan formula exhibit high consistency with ERA5-derived results in both freeze-thaw timing and variation trends. Notably, the ERA5-based frozen depths display a stepwise variation pattern, with the maximum freezing depth terminating at the 100 cm layer, likely attributable to the freezing inhibition effect caused by higher temperatures in the deep soil layer (100–289 cm). The simulations indicate that the freezing process initiates in late September, reaches the maximum depth of 1.4 m by late March of the following year, and completes thawing by late May. This temporal pattern aligns closely with ground temperature observations from basin meteorological stations (Fig. S6; mean errors of ≤5 days for initial freezing dates and ≤10 days for initial thawing dates).*

To further evaluate the model's spatial performance, the 2000–2018 mean maximum frozen depth distribution was compared with contemporaneous data from the National Tibetan Plateau Data Center (Table 3; Fig. S7). The Stefan formula-based simulations, incorporating station-based temperature interpolation, demonstrate smoother spatial transitions—a characteristic linked to model parameterization. Both datasets reveal a gradient pattern of deeper frozen depths in upstream valley regions and shallower depths in downstream areas, with a spatial correlation coefficient of 0.89. Furthermore, the observed decreasing trend in frozen depth during 2000–2018 corresponds with accelerated snowmelt patterns (Fig. 4), highlighting the coupled response of the cryosphere to climate change."

- **Data Sources and Validation Reliability:**

Due to the lack of in-situ measurements in our study area, we have employed multiple remote sensing and reanalysis datasets to validate the snow and frozen ground processes from various perspectives. This multi-source validation approach strengthens the reliability of the model results while accounting for potential uncertainties from different data sources. We believe this significantly enhances the credibility of our findings. In future work, we will continue to improve the quality of remote sensing data and assess the long-term stability of the model.

We appreciate your constructive feedback, which has helped us further refine the study. The manuscript has been revised accordingly, and we believe these improvements contribute to the robustness and clarity of our findings.

Comments 4: Finally, numerous studies have already investigated the hydrological effects of snow and frozen ground at large basin scales, ranging from basin-scale (e.g., Cuo et al., 2015; Qin et al, 2017b; Song et al., 2022; Wang et al., 2023a, 2023b) to the entire Tibetan Plateau. Therefore, I recommend that the authors compare some of their conclusions with those of previous studies, rather than simply stating that research in this area is lacking. Furthermore, I find the current conclusions to be insufficiently in-depth. For example, the critical processes of how changes in soil ice and liquid water in frozen ground affect streamflow are not thoroughly discussed.

Response 4: Thank you for your thoughtful comments. We appreciate your suggestions regarding the need for comparison with previous studies and a more in-depth discussion of key processes.

To address this, we have made substantial revisions to the '*Discussion*' section, where we now explicitly compare our refined conclusions with existing studies. This comparison highlights both the similarities and differences between our study and previous research, thereby strengthening the contextual relevance of our findings.

Additionally, we have integrated the refined findings into the revised '*Conclusion*' section. This includes the key findings described in the '*Results*' and '*Discussion*' regarding the influence of snow depth on frozen ground depth and duration, the seasonal impact of freeze-thaw cycles on runoff generation, the suppressive effect of snow and frozen ground on evapotranspiration during cold months, the seasonal role of snowmelt, and key findings regarding the future hydrological significance of snow and frozen ground. These comparisons aim to position our results within a broader literature context while emphasizing both the consistency with previous studies and the novel insights provided by our research.