Author’s Response to Referee Comments by Anonymous Referee #2

We would like to sincerely thank the referee for the careful and constructive assessment of our manuscript. We are grateful for the time and effort devoted to the review, as well as for the valuable comments and suggestions provided. We have carefully considered all the points raised and agree that addressing them will further improve the clarity, presentation, and overall quality of the manuscript. Below, we respond to each comment in detail and indicate how we plan to incorporate the suggested changes should we be invited to submit a revised version.

General Assessment: In their study ‘Hydrological Modeling of Climate Change Impacts on Alpine Catchments: Shifting Snow Dynamics, Runoff Regimes, and the Frequency of Extremes’ Seelig et al. investigate the hydrological response of 15 nested headwater catchments in the Austrian alps to changing climatic conditions under 40 different emission scenarios for the near and far future. The investigated catchments are located at different elevation zones and feature varying geomorphological characteristics that influence aquifer storage and low flow buffering capacities. Implementing an exceptionally high amount of observation data they developed a nested conceptual model framework based on the GR4J-CemaNeige model, which was evaluated using a broad set of process specific performance indices. The main analysis investigates the buffering capacity of high alpine aquifers regarding regime changes, rainfall-runoff response and extreme events. The study supports the widely known shift of runoff regimes from snow to rain dominated discharge with stronger and earlier high flow extremes, stronger summer deficits and higher low-flow volumes (especially during winter). The novelty of the study lies within the explicit evaluation of nested systems along the elevation gradient, featuring varying storage capacities and highlighting the strong sensitivity of small high elevation catchments to changing hydrometeorological conditions.

The study is well written, clearly structured and features suitable well-designed figures and tables. The specific investigation of shallow sedimentary aquifers in the context of streamflow buffering represents a substantial contribution to scientific progress within the scope of Hydrology. The authors have applied a wide range of scientifically proven methods and discussed their results thoroughly, linking to a variety of appropriate and timely references. Only the description of the model setup und calibration is missing some information that would allow better reproducibility (see specific comments). The results are fully sufficient to support the interpretations and conclusions, while it remains sometimes a little difficult to toggle information form the large Supplementary material and the main manuscript. Further, the title could benefit from further highlighting the novelty of considering shallow-rock aquifers and their impact on runoff characteristics. However, it is quite long and

comprehensive already, so this is only a minor suggestion and should be up to the authors weather to rephrase.

After thorough evaluation of the Manuscript I agree with most of the comments previously raised by Giulia Mazzotti and instead of raising them repeatedly build up upon them. I would like to emphasize especially two of the major concerns.

Response: We sincerely thank you for your careful and constructive assessment of our manuscript and for your positive evaluation of its scientific contribution, methodological approach, and presentation. We also appreciate your observation that some aspects of the model setup and calibration should be described in greater detail to facilitate reproducibility. Together with the comments raised by Giulia Mazzotti, your feedback has helped us identify several aspects where the methodological description and the links between the main manuscript and the Supplementary Material could be made clearer. We have therefore carefully revisited these aspects and will incorporate additional information and clarifications in the revised manuscript, should we be invited to submit one.

Regarding your suggestion to further emphasize the role of shallow aquifers in the title, we agree that this aspect represents an important element of the study and appreciate your suggestion. At the same time, as you note, the current title is already rather comprehensive, and we would prefer not to make it substantially longer or more specific. Instead, we suggest to strengthen the emphasis on this aspect in the Abstract, Introduction, and Conclusions, making the novelty of explicitly considering aquifer-related runoff buffering along the elevation gradient more apparent. We believe this will communicate the central contribution of the study more clearly while retaining a concise and informative title.

We are grateful for your thoughtful and detailed assessment. Below, we address the concerns you raise, while also taking into account the related points discussed in our response to Giulia Mazzotti.

Major Comment 1: How representative are the high elevation rock glaciers (and their according aquifers) within the nested catchment concept in the Region?

I agree with Giulia that the specific emphasis on Rock Glaciers in the introduction and description of the catchments suggests a more detailed investigation of their effects on the hydrologic functioning whereas the results and discussion only slightly focus on the differentiation between catchments with or without rock glacier features. This might be information that is hidden within the specific catchment description but given the length and extent of the manuscript it is difficult for the reader to remember the catchment characteristic throughout the paper. Here it might be helpful to reframe the results highlighting characteristic-specific process responses rather than catchment specific responses. This would undoubtedly increase the transferability of the results.

Response: We thank you for this constructive comment. We agree that the emphasis on rock glaciers in the Introduction and Study Site sections is not sufficiently reflected in the Results

and Discussion, where the findings are presented predominantly on a catchment-by-catchment basis. We acknowledge that a more explicit focus on specific hydrological process responses would improve the clarity and the transferability of our findings.

Should we be invited to submit a revised version, we will restructure the manuscript accordingly. We will condense the study-site description in the main text and move the detailed catchment descriptions to the Supplementary Material, presenting the relevant characteristics more systematically and concisely. We will revise the Results and Discussion to more explicitly link observed hydrological responses to specific catchment characteristics, including permafrost, rock glaciers (and other shallow aquifers), bedrock geology, and topography, rather than discussing the results primarily in relation to individual catchments. This will help distinguish characteristic-specific process responses from catchment-specific observations and clarify which findings may be transferable to other catchments. We will also explicitly discuss how representative the rock glaciers within the studied catchments are, accounting for the fact that their occurrence, distribution, and hydrological significance varies between mountain ranges.

We appreciate your suggestion, which will help us strengthen the process-based interpretation of our results and communicate their broader hydrological relevance.

Major Comment 2: What are the limitations arising from the use of a simple conceptual (and static) model like GR4J and not explicitly considering the Energy-Balance, Rock Glaciers (as active glaciers) and vegetation interactions?

As far as I understand from reading the manuscript, the specific catchment characterises (regarding soil structure, aquifer storage and land cover) are represented by the model parameters in the GR4J model. These model structures remain constant in time, thus not representing possible effects of permafrost/rock glacier thaw, or vegetation transitions. Thus, the results show how the catchments as they are today would respond to changing forcings but do not consider that we will likely see drastic changes of catchment characteristics. I think that this aspect could be further emphasised in the discussion of the results.

Response: We thank you for drawing attention to this limitation of the GR4J-CemaNeige modelling framework as employed in our study. As you rightly point out, our projections describe how catchments, given their calibrated present-day hydrological characteristics, respond to changing climatic forcing; they do not capture the potential evolution of catchment characteristics or the associated feedbacks between energy balance, cryospheric changes, vegetation dynamics, and subsurface storage.

In the Discussion and Conclusions, we will clarify this distinction explicitly and discuss its implications across our elevation gradients and study sites. Specifically, we will address how permafrost degradation may alter infiltration pathways, hydrological connectivity, subsurface storage, and runoff buffering from a process-oriented perspective. At the same time, vegetation expansion and canopy shifts may alter snow interception, evapotranspiration, soil

retention capacity, and groundwater recharge. These structural changes are particularly pertinent in high-elevation catchments dominated by permafrost degradation, whereas in lower-elevation catchments, vegetation transitions will exert a more direct influence on water partitioning.

We agree that making this limitation more explicit will provide a more appropriate framing of our results, distinguishing the hydrological sensitivities identified under present-day catchment characteristics from future responses that depend on their evolution. We will also emphasize the need for future research to investigate how these evolving catchment characteristics interact with changing climatic forcing, ideally through process-based modelling approaches that explicitly couple energy balance, cryospheric dynamics, vegetation changes, and subsurface hydrological processes.

We appreciate your constructive suggestion, which will help us better contextualize our findings and explicitly address the limitations of our modeling approach.

Comment on L37: Does a ‘retarded response’ mean a ‘delayed’ response here, meaning that we will see the shift a few years later? Or do you mean that the shift happens over longer time frames as it does in low elevations?

Response: Thank you for pointing this out. Here, by “retarded response” we mean that the hydrological shift occurs later in time in high-alpine regions compared to lower elevations. We agree that the original wording was ambiguous and will clarify this in the revised manuscript by explicitly referring to the later timing of the transition. Accordingly, we suggest revising the sentence to:

“While these large-scale shifts alter storage and release mechanisms globally, the hydrological catchment response differs significantly across altitudinal gradients, with transitions occurring earlier at lower elevations and later at higher elevations.”

We hope this revised wording makes clear that our primary meaning is a shift in the timing of the transition. We appreciate you bringing this ambiguity to our attention.

Comment on L49: L49: Again, a ‘retarding mechanism’ isn’t quite clear to me. Maybe rephrase.

Response: Thank you for raising this important point. We agree that the term “retarding mechanism” has been used ambiguously in the original manuscript, particularly in light of the clarification above. Here, we do not refer to a later timing or prolonged hydrological response, but specifically to the temporary storage and attenuation of precipitation by the snowpack during a precipitation event. We therefore suggest revising the sentence to make this distinction explicit:

“Second, as the fraction of extreme precipitation falling as snow declines, the temporary storage and attenuation of precipitation in the snowpack is reduced, allowing a larger fraction of the event to contribute directly to downstream runoff and thereby amplifying flood peaks.”

This revised wording distinguishes the temporal shift in hydrological response described above from the buffering effect of snow storage during individual precipitation events. We greatly appreciate this comment, which helps us make this key distinction much clearer.

Comment on L92ff: A short information on landcover within each catchment would be helpful to better understand how vegetation may also influence runoff response.

Response: We agree that providing an overview of the land-cover characteristics across our study catchments will offer valuable context for interpreting differences in their hydrological responses. It will also support our discussion regarding potential limitations of our modeling approach in catchments where substantial land-cover shifts are anticipated (as noted in our response to Major Comment 2).

In the revised manuscript, we will add this information to the Study Site description, highlighting key land-cover and vegetation characteristics relevant to runoff generation and evapotranspiration. Furthermore, we will explicitly address potential future land-cover transitions and their hydrogeological implications in the Discussion section. We appreciate this constructive suggestion.

Comment on Table 1: Is it possible to very briefly include how you (or in this case Wagner et al.) distinguished between active vs. relict Rock Glaciers?

Response: We agree that a brief explanation of how the rock glaciers were classified would be helpful and will add this information to the revised manuscript, should we be invited to submit one. We follow the classification of Wagner et al. (2020), which combines remote-sensing-based morphological characteristics with the regional permafrost model of Boeckli et al. (2012). Rock glaciers were classified as relict where permafrost was absent or only marginally indicated, whereas those located predominantly within the modeled permafrost area were considered to contain ice. We will briefly describe these classification criteria in the revised manuscript to make the information provided in Table 1 more transparent. Thank you for this suggestion.

References:

Boeckli, L.; Brenning, A.; Gruber, S.; Noetzli, J. (2012): Permafrost distribution in the European Alps. Calculation and evaluation of an index map and summary statistics. The Cryosphere 6 (4), 807–820. DOI: 10.5194/tc-6-807-2012.

Wagner, T.; Pleschberger, R.; Kainz, S.; Ribis, M.; Kellerer-Pirklbauer, A.; Krainer, K.; Philippitsch, R.; Winkler, G. (2020): The first consistent inventory of rock glaciers and their hydrological catchments of the Austrian Alps. AJES 113 (1), 1–23. DOI: 10.17738/ajes.2020.0001.

Comment on L213: The Spartacus dataset is arguably the best suited dataset for this study, but have you performed any quality checks against observations from within your small catchments? As arguably especially in small high elevation catchments, the offset might still be quite large.

Response: Thank you for raising this important point. We cross-checked the SPARTACUS dataset against available observations from weather stations within or close to our study catchments. Where local weather station observations were available, the comparisons generally showed good agreement with SPARTACUS data, though some local mismatches were observed.

We agree that these discrepancies and the uncertainties associated with gridded precipitation data, particularly in small, high-elevation catchments, warrant explicit discussion. In the revised manuscript, we will clarify these comparison checks in the Methods and expand our Discussion to address the potential impacts of meteorological forcing uncertainties on our modeling results. We appreciate this thoughtful question and the opportunity to make the quality assessment of our input data more transparent.

Comment on L229: How did you deal with cloudy days in the time series (as you are talking about a daily timeseries here) and did you use 500 x 500m estimates for the evaluation of small headwater catchments too?

Response: In our workflow, cloudy days were retained as missing values. Since the satellite-derived snow-covered fraction was used as a calibration target rather than as a model input, missing values could be retained as NAs without requiring any interpolation. Similarly, pixels flagged as suspicious or erroneous were excluded from the analysis from the outset. The available pixels were then aggregated within the five elevation bands per catchment, used to represent snow dynamics following the standard CemaNeige setup recommended by Valéry et al. (2014) and Riboust et al. (2019).

In addition to simulating snow water equivalent, which is the variable directly relevant for modeling water fluxes, CemaNeige also represents the snow-covered fraction for each elevation band. This provides the link that allows satellite-derived snow-cover products to be incorporated into the calibration process. We paid particular attention to the formulation of this relationship and followed the recommendation to use a hysteretic relationship between the elevation-band-averaged snow water equivalent and snow-covered fraction. Riboust et al.

(2019) evaluated alternative formulations across a large sample of catchments and several decades of observations and found this hysteretic formulation to provide the most stable model structure for applications under changing climatic conditions, which closely matches the context of our study. Furthermore, site-specific evaluations by Masten et al. (2026) in the Ötztal Alps confirmed that this hysteretic formulations performs best for capturing snow cover dynamics in high-alpine catchments within our study region.

These five snow-covered-fraction time series were then used as calibration targets together with discharge, using a composite objective function based on the KGE', with 75% weight assigned to discharge and 5% weight assigned to the snow-covered fraction of each of the five elevation bands (Riboust et al., 2019).

We specifically checked whether this approach was appropriate for small headwater catchments, where the number of available 500 × 500 m pixels per elevation band is necessarily limited. We carefully inspected the resulting snow-cover time series across all catchments, paying particular attention to these smaller headwaters. The satellite-derived snow-cover dynamics were consistent with visual inspections of raw satellite imagery, regional webcams operating in nearby valleys, and our own field observations. Because our monthly field campaigns (conducted year-round for discharge measurements and sensor maintenance) provided direct ground-truth insights into seasonal snowpack behavior, together with the nearby webcam feeds, we could confirm that the band-aggregated snow-covered fractions offer a physically plausible representation even in small headwater settings.

We will clarify the handling of cloudy days, the spatial aggregation across the elevation bands, the role of the snow-covered-fraction formulation in the calibration, and the quality checks performed in the revised manuscript, should we be invited to submit one. We will also provide further details on these checks in the Supplementary Material to make the procedure and its applicability to small headwater catchments more transparent. We very much appreciate this thoughtful question and would be pleased to address these aspects more explicitly, which would make this essential part of our calibration framework more transparent.

References:

Masten, M.; Seelig, S.; Vremec, M.; Seelig, M.; Winkler, G. (2026): An Enhanced Python Framework for Hydrological Modeling in Alpine Catchments: Snow Hysteresis and Glacier Ice Melt. Environmental Modelling & Software 197 (2), 106842. DOI: 10.1016/j.envsoft.2025.106842.

Riboust, P.; Thirel, G.; Le Moine, N.; Ribstein, P. (2019): Revisiting a Simple Degree-Day Model for Integrating Satellite Data. Implementation of Swe-Sca Hystereses. Journal of Hydrology and Hydromechanics 67 (1), 70–81. DOI: 10.2478/johh-2018-0004.

Valéry, A.; Andréassian, V.; Perrin, C. (2014): 'As simple as possible but not simpler'. What is useful in a temperature-based snow-accounting routine? Part 2 – Sensitivity analysis of the Cemaneige snow accounting routine on 380 catchments. Journal of Hydrology 517 (12), 1176–1187. DOI: 10.1016/j.jhydrol.2014.04.058.

Comment on L236f: This is a very brief description of the downscaling and bias correction. It might be helpful to actually explain the applied method and data briefly.

Response: Thank you for this helpful suggestion. We agree that the current description of the downscaling and bias correction is too brief and will expand it in the revised manuscript, should we be invited to submit one. The climate model data from the EURO-CORDEX EUR-11 archive (Jacob et al., 2014; comprising 8 runs for RCP2.6, 16 runs each for RCP4.5 and RCP8.5) were first interpolated to a 1 x 1 km grid using the observational grid as the target, and subsequently bias-corrected on a grid-cell basis using the Scaled Distribution Mapping method described in Chimani et al. (2016, 2019) and Switanek et al. (2017). We will revise and expand the manuscript accordingly to make the downscaling and bias-correction procedure, as well as the underlying data and methodological choices, fully transparent. We very much appreciate you pointing this out.

References:

Chimani, B.; Heinrich, G.; Hofstätter, M.; Kerschbaumer, M.; Kienberger, S.; Leuprecht, A.; Lexer, A.; Peßenteiner, S.; Poetsch, M. S.; Salzmann, M.; Spiekermann, R.; Switanek, M.; Truhetz, H. (2016): ÖKS 15 Klimaszenarien für Österreich. Daten - Methoden - Klimaanalyse. Hg. v. Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft (BMLFUW), Wien.

Chimani, B.; Matulla, C.; Eitzinger, J.; Hiebl, J.; Hofstätter, M.; Kubu, G.; Maraun, D.; Mendlik, T.; Schellander-Gorgas, T.; Thaler, S. (2019): STARC-Impact Guideline zur Nutzung der ÖKS15-Klimawandelsimulationen sowie der entsprechenden gegitterten Beobachtungsdatensätze, Wien.

Jacob, D.; Petersen, J.; Eggert, B.; Alias, A.; Christensen, O. B.; Bouwer, L. M.; Braun, A.; Colette, A.; Déqué, M.; Georgievski, G.; Georgopoulou, E.; Gobiet, A.; Menut, L.; Nikulin, G.; Haensler, A.; Hempelmann, N.; Jones, C.; Keuler, K.; Kovats, S.; Kröner, N.; Kotlarski, S.; Kriegsmann, A.; Martin, E.; van Meijgaard, E.; Moseley, C.; Pfeifer, S.; Preuschmann, S.; Radermacher, C.; Radtke, K.; Rechid, D.; Rounsevell, M.; Samuelsson, P.; Somot, S.; Soussana, J.-F.; Teichmann, C.; Valentini, R.; Vautard, R.; Weber, B.; Yiou, P. (2014): EURO-CORDEX. New high-resolution climate change projections for European impact research. Reg Environ Change 14 (2), 563–578. DOI: 10.1007/s10113-013-0499-2.

Switanek, M. B.; Troch, P. A.; Castro, C. L.; Leuprecht, A.; Chang, H.-I.; Mukherjee, R.; Demaria, E. M. C. (2017): Scaled distribution mapping. A bias correction method that preserves raw climate model projected changes. Hydrol. Earth Syst. Sci. 21 (6), 2649–2666. DOI: 10.5194/hess-21-2649-2017.

Comment on L250ff: It is not quite clear to me weather the model accounts for water losses along the catchment chain eg. induced through evapotranspiration or leakage? Also, do I understand correct, that the presence/absence and activity of rock glaciers is solely captured by the parametrization of the GR4J storage compartment – thus remains unchanged throughout the modelling period?

Response: Thank you for raising this important point. We agree that both aspects of the model structure should be described more clearly.

Regarding water losses and fluxes along the catchment chain, GR4J accounts for evapotranspiration through its soil moisture accounting component, driven by potential evapotranspiration from the gridded ARET dataset. The actual evaporation rate is determined as a function of the current water level in the soil moisture accounting store to calculate the rate of water losses through evapotranspiration from the store, bounded by its calibrated capacity x_1 (Perrin et al., 2003).Catchment leakage and delayed subsurface contributions are subsequently governed by the model's routing component, where subsurface storage dynamics are controlled by the routing store capacity x_3 (Perrin et al., 2003). In the revised manuscript, we will expand Section 3 to detail these store-based evaporation mechanics and subsurface routing functions.

Regarding the representation of rock glaciers (as well as other catchment features that may evolve over time), you are entirely correct: they are captured implicitly through the calibrated model parameterization, which remains unchanged throughout the modeling period. This also means that our projections do not account for potential changes in subsurface storage capacity, for example if permafrost ice melts and previously ice-filled pore space becomes available for groundwater storage, or if thaw alters subsurface connectivity and flow paths. We did not explicitly incorporate such changes because there is currently no straightforward way to parameterize them with sufficient accuracy and testability across our study sites.

As also discussed in our response to Major Comment 2, this means that the projections represent the response of catchments with their calibrated present-day hydrological characteristics under changing climatic forcing. We agree that this important limitation should be made more explicit and will highlight the resulting uncertainties in the Discussion and Conclusions.

Thank you again for raising this crucial point, which will help us clarify both the representation of water fluxes in the model and the limitations associated with its static representation of catchment storage characteristics.

Reference:

Perrin, C.; Michel, C.; Andréassian, V. (2003): Improvement of a parsimonious model for streamflow simulation. Journal of Hydrology 279, 275–289. DOI: 10.1016/S0022-1694(03)00225-7.

Comment on L277: How was the KGE modified?

Response: We apologize for the imprecise terminology in the original draft. The wording "modified Kling-Gupta efficiency" was misleading, as we did not modify the formula ourselves.

Instead, we used the established formulation proposed by Kling et al. (2012) (often denoted as KGE'), which differs from the original Gupta et al. (2009) formulation in calculating the variability ratio from the coefficients of variation of the simulated and observed time series. By reducing cross-correlation between the bias and variability ratios, this formulation limits the effect of compensating errors. It has therefore been recommended over the original metric in comparative model evaluations (e.g., Cinkus et al., 2023).

In the revised manuscript, we will remove "modified" to avoid ambiguity and simply cite Kling et al. (2012) when defining the KGE' objective function. We will also include the exact mathematical formulation in the Supplementary Material for full clarity. Thank you for pointing this out.

References:

Cinkus, G.; Mazzilli, N.; Jourde, H.; Wunsch, A.; Liesch, T.; Ravbar, N.; Chen, Z.; Goldscheider, N. (2023): When best is the enemy of good – critical evaluation of performance criteria in hydrological models. Hydrol. Earth Syst. Sci. 27 (13), 2397–2411. DOI: 10.5194/hess-27-2397-2023.

Gupta, H. V.; Kling, H.; Yilmaz, K. K.; Martinez, G. F. (2009): Decomposition of the mean squared error and NSE performance criteria. Implications for improving hydrological modelling. Journal of Hydrology 377 (1-2), 80–91. DOI: 10.1016/j.jhydrol.2009.08.003.

Kling, H.; Fuchs, M.; Paulin, M. (2012): Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. Journal of Hydrology 424-425 (4), 264–277. DOI: 10.1016/j.jhydrol.2012.01.011.

Comment on L278: What were the time periods used for calibration and validation? And what forcing data did you use during this step, the SPARTACUS data? If so, did you also validate the model performance using the RCP data?

Response: Thank you very much for raising this point. We acknowledge that the original description was not sufficiently clear. Giulia Mazzotti raised a closely related issue, and we therefore recognize that the distinction between the use of observational data and the subsequent climate-projection simulations was not adequately reported.

For the parameter optimization, we used SPARTACUS precipitation and temperature and ARET potential evapotranspiration as meteorological inputs during calibration only, and compared the resulting simulations with the available observed discharge and MODIS-derived snow-cover time series. The periods considered were determined by the availability

and temporal coverage of these observational datasets for the individual catchments and variables. Once the model parameters were established, the climate-impact simulations were performed entirely with the individual climate projections, including their respective historical periods (i.e., projections into the past). This preserves the internal consistency of each climate-model realization and allows historical-to-future changes to be assessed within each individual projection, rather than comparing future projections against a common observational reference. We validated model performance using the RCP scenarios to quantify the resulting uncertainty across the ensemble of projections but did not report them in the original submission. We agree that these results should also be presented to improve the transparency and reproducibility of our results.

We realize that this workflow was not sufficiently transparent in the original manuscript and we will clarify these aspects, should we be invited to submit a revised version. We very much appreciate this question, as it has helped us identify an important aspect of the modelling framework that should be communicated more clearly.

Comment on L280: This refers back to my comment on line 229: I don't quite get how you were able to compare observed snow cover data (as a spatial fraction) with the bucket type SWE data of the model.

Response: We agree that the connection between the spatially distributed MODIS satellite observations and the modeled SWE was not sufficiently transparent in the original manuscript.

To clarify how these datasets are compared, the modeling and evaluation workflow operates as follows across all catchments:

(1) Spatial Disaggregation: Each catchment is discretized into five equal-area elevation bands, and CemaNeige simulates band-average snow accumulation, melt, and snow-covered area (SCA) separately within each band.
(2) SWE-to-SCA Conversion: For each elevation band, the mean modeled SWE is converted into a corresponding mean modeled SCA using the linear hysteretic formulation. Riboust et al. (2019) demonstrated across a large sample of catchments that this hysteretic structure provides the most robust and stable framework under changing climate conditions, a finding further supported in our region by site-specific testing in the Ötztal Alps (Masten et al., 2026).
(3) Band-Averaged Comparison: To ensure spatial consistency, MODIS pixel-based observations are aggregated to compute a mean observed SCA for each elevation band. The mean modeled SCA time series for a given band is then compared directly against this mean observed SCA time series for the corresponding band.

In this way, satellite-derived snow cover fractions averaged over each elevation band serve as an observational constraint on modeled elevation band-average SWE via the SCA–SWEhysteresis function.

In the revised manuscript, we will clarify this link in the Methods section and expand the Supplementary Material to provide detailed descriptions and evaluations. We appreciate this comment, as explicitly laying out this connection strengthens the description of our snow-model calibration.

References:

Masten, M.; Seelig, S.; Vremec, M.; Seelig, M.; Winkler, G. (2026): An Enhanced Python Framework for Hydrological Modeling in Alpine Catchments: Snow Hysteresis and Glacier Ice Melt. Environmental Modelling & Software 197 (2), 106842. DOI: 10.1016/j.envsoft.2025.106842.

Riboust, P.; Thirel, G.; Le Moine, N.; Ribstein, P. (2019): Revisiting a Simple Degree-Day Model for Integrating Satellite Data. Implementation of Swe-Sca Hystereses. Journal of Hydrology and Hydromechanics 67 (1), 70–81. DOI: 10.2478/johh-2018-0004.

Valéry, A.; Andréassian, V.; Perrin, C. (2014): 'As simple as possible but not simpler'. What is useful in a temperature-based snow-accounting routine? Part 2 – Sensitivity analysis of the Cemaneige snow accounting routine on 380 catchments. Journal of Hydrology 517 (12), 1176–1187. DOI: 10.1016/j.jhydrol.2014.04.058.

Comment on L303: Many of the used evaluation criteria are based on daily, seasonal or yearly timesteps. Given the large variety of sub-catchment size it is likely that especially the response times of small headwater is too small to be captured properly by eg. a daily based flashiness index. I realise that this is something you cannot properly evaluate throughout the study as forcing data is on daily timesteps but it might be worth noting in the discussion of the results.

Response: Thank you for this insightful comment. We fully agree that daily metrics, such as the flashiness index, cannot fully capture sub-daily response times in small headwater catchments. Because our meteorological forcing data are restricted to a daily resolution, fast hydrological processes and sub-daily dynamics remain beyond the scope of the current modeling framework.

In the revised manuscript, we will explicitly highlight this temporal resolution constraint in the Discussion. Specifically, we will address how daily metrics may fail to fully capture rapid catchment responses and point to higher-temporal-resolution forcing datasets as an important priority for future research. We appreciate this suggestion, which adds valuable nuance to our discussion of model limitations and evaluation metrics.

We appreciate this thoughtful comment, as it highlights an important limitation of the temporal resolution of our analysis and provides a clear direction for improving the assessment of hydrological extremes in future work.

Comment on Figure 3-5: The figures could benefit from increases label sizes, they are really hard to read in the print version

Response: Thank you! We will increase the label sizes across Figures 3–5 to ensure full readability in print.

Comment on L350: I very much appreciate the process based and hydrological plausibility check of the model setup and calibrated parameter set. this is done not often enough in hydrological modelling and is worth pointing out towards the quality of the study and paper.

Response: Thank you very much for this positive and encouraging comment. We share your view that assessing the physical and hydrological consistency of the calibrated model is an important step that deserves greater attention in hydrological modelling. We are therefore particularly pleased that you see this as strengthening the overall quality of the study.

Comment on L353: I wonder whether active rock glaciers are modelled as actual additional source of water, that potentially becomes less in the future or are they just treated as aquifers that buffer the hydrological response? If this source of water is not explicitly accounted for I guess the exchange parameter is making up for it, but since parameter values are constant in time isn't that skewing the water balance?

Response: Thank you for raising this point. The contribution of rock glacier ice melt is not explicitly modeled as an additional water source because estimating this flux is extremely difficult, and reasonable guidance for quantifying it remains lacking. Existing studies across mid-latitude mountain regions consistently suggest that rock glacier ice melt constitutes a very small fraction of rock glacier discharge (typically $\leq$ 1–3%), making its contribution negligibly small compared to total runoff (e.g., Krainer et al., 2011, 2015; Pruessner et al., 2021, 2022).

Given these constraints and the high conceptual uncertainty, we chose to neglect this flux rather than speculate on whether parameters (like the exchange coefficient x_2) compensate for it. Because rock glacier ice melt is omitted from the model structure, our framework does not simulate potential future declines as subsurface ice depletes, which represents a clear limitation of our modeling setup. In the revised manuscript, we will explicitly acknowledge this limitation in the Discussion and contextualize the expected minor magnitude of this flux using published literature from mid-latitude mountain settings.

We appreciate this comment, as highlighting this limitation provides important context for interpreting projected water balances in catchments containing rock glaciers.

References:

Krainer, K.; Chinellato, G.; Tonidandel, D.; Lang, K. (2011): Analysis of the contribution of permafrost ice to the hydrological water regime. PermaNET Project, Action Report 7.3, Grenoble.

Krainer, K.; Bressan, D.; Dietre, B.; Haas, J. N.; Hajdas, I.; Lang, K.; Mair, V.; Nickus, U.; Reidl, D.; Thies, H.; Tonidandel, D. (2015): A 10,300-year-old permafrost core from the active rock glacier Lazaun, southern Ötztal Alps (South Tyrol, northern Italy). Quaternary Research 83 (2), 324–335. DOI: 10.1016/j.yqres.2014.12.005.

Pruessner, L.; Huss, M.; Phillips, M.; Farinotti, D. (2021): A Framework for Modeling Rock Glaciers and Permafrost at the Basin-Scale in High Alpine Catchments. J. Adv. Model. Earth Syst. 13 (4), 269. DOI: 10.1029/2020MS002361.

Pruessner, L.; Huss, M.; Farinotti, D. (2022): Temperature evolution and runoff contribution of three rock glaciers in Switzerland under future climate forcing. Permafrost Periglac. Process. 100, 102. DOI: 10.1002/ppp.2149.

Comment on L402: In the Model, ET is a function of temperature and the production store, right? Do you have any idea, why we can see a future ET decrease for the RCP2.6scenarioo? → related to that in L597ff: you mention that you assume lower ET rates in steep high elevation catchments as a result of a lower production store (thus assuming a water- rather than energy-limitation of the process, correct?) This would sure be worth mentioning here or somewhere else in the main text.

Response: Thank you very much for raising this point. The apparent decrease in ET under RCP2.6 may result from how the results are presented in the figure, as ET actually increases significantly across all catchments, including under RCP2.6. We will revise the figure to make this clearer and avoid any possible misinterpretation.

Regarding the connection to L597ff, the formulation of GR4J indeed links actual ET to potential ET and the available water in the production store. Thus, in the high-elevation catchments, the smaller production-store capacity can constrain actual ET when water availability is limiting. In this sense, the parameterization implicitly represents a stronger water limitation of ET rather than an explicit energy limitation. We will make this relationship clearer in the revised manuscript, should we be invited to submit a revised version.

Giulia Mazzotti also suggested revisiting the figures to improve the clarity of the presentation. We recognize that this was not sufficiently well addressed in the original manuscript and will place particular emphasis on improving the figures and ensuring that the main results and their interpretation are presented unambiguously. We appreciate you drawing our attention to this point.

Comment on L441: The Reißeck group seems to behave different than the other catchments, do you have any idea why?

Response: We acknowledge that the Reißeck Group displays a distinct hydrological behavior compared to the other catchments. We observed this departure in our results as well, and in the revised manuscript, we will expand our Discussion to address potential site-specific factors that may contribute to this difference.

We suspect that its distinct response may be related, at least partly, to its particular topographic and geographic position. The Reißeck Group is located approximately 20 km south of the main Alpine ridge, whereas the other study sites are situated much closer to the main ridge. It is also in a relatively exposed position above the Drautal, which opens towards the Klagenfurt Basin to the south. These characteristics may result in a stronger influence of Mediterranean precipitation patterns compared with the other study sites. This interpretation is consistent with the findings of Haslinger et al. (2021), who demonstrated that the relative influence of Atlantic and Mediterranean precipitation patterns varies across the Austrian Alps and changes with the prevailing atmospheric conditions.

We would, however, like to present this interpretation cautiously, as our analysis does not allow us to demonstrate a causal link between the geographic position of the Reißeck Group and its distinct hydrological response. Should we be invited to submit a revised version, we will therefore highlight the particular characteristics of this study site and discuss this possible explanation as a preliminary interpretation, while clearly acknowledging the associated uncertainty and avoiding over-interpretation.

Thank you again for drawing our attention to this interesting feature of the results. It provides a valuable opportunity to discuss the distinctive characteristics of the Reißeck Group more explicitly.

Reference:

Haslinger, K.; Hofstätter, M.; Schöner, W.; Blöschl, G. (2021): Changing summer precipitation variability in the Alpine region. On the role of scale dependent atmospheric drivers. Climate Dynamics 57 (3-4), 1009–1021. DOI: 10.1007/s00382-021-05753-5.

Comment on L460ff: Given the model structure, where snowmelt is triggered by temperature as a threshold only, the patterns of change for snow indices doesn't surprise. So this section could potentially by shortened in order to have a slightly more compact manuscript. Also, it should be discussed, how the consideration of the actual energy balance could influence the results. This continues down to L475 where you point out, that we see a strong temperature-driven influence on the changing metrics. If Temperature and Precipitation are the only considered changing forcings – this isn't a surprise…

Response: Thank you for this insightful suggestion. We agree that because CemaNeige employs a temperature-index structure, some projected snow-index shifts follow directly from

the temperature forcing. In the revised manuscript, we will shorten and streamline this section, with a stronger focus on the hydrologically relevant implications rather than describing each snow metric in detail. We will also explicitly discuss the inherent limitations of temperature-index snow models. Furthermore, we will reframe the discussion around L475 to avoid overstating temperature as a primary driver, placing greater emphasis on how varying catchment characteristics modulate the hydrological response to these forcings.

We appreciate this feedback, as it helps make the manuscript more compact while sharpening the discussion of model limitations.

Comment on L526: there is a double ‘this’ here

Response: Will be corrected, thank you!

We would like to thank you again for your careful, constructive, and very helpful review of our manuscript. Your comments have helped us identify several aspects that require clearer explanation and presentation, and we believe that addressing these points will substantially improve the manuscript. We greatly appreciate the time and effort you have devoted to providing such detailed and thoughtful feedback, and we hope to have the opportunity to submit a revised version that addresses your suggestions accordingly.

Kind regards,

Simon Seelig

On behalf of all authors