

Review

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 and 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 from 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:

1st 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.

2nd 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.

Minor, Specific Comments

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?

L49: Again, a 'retarding mechanism' isn't quite clear to me. Maybe rephrase.

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

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?

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.

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?

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.

L250ff: It is not quite clear to me whether 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?

L277: How was the KGE modified?

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?

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.

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.

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

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.

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?

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 RCP2.6scenario? → 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.

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

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 be 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...

L526: there is a double 'this' here