We thank the reviewer for their time in reviewing this manuscript. We carefully considered the comments, which point out some of the key challenges of modelling groundwater in high-elevation alpine catchments. We will provide point-by-point answers during the revision of the manuscript (should a revision be encouraged by the editor).

Hereafter, we provide a detailed discussion of the main comments of the reviewer:

(1) Details of high-elevation aquifers and subsurface heterogeneity:

The majority of the comments ask for a detailed/improved representation in the model of various types of aquifers (hillslope aquifers, stream-wide aquifers, comment on Line 230), of varying depths of different types of aquifers and their boundaries (comments on Lines 225, 256, 362 and 433), and of spatial subsurface heterogeneity distributed in the complex terrains. The reviewer asks for modelling the deep aquifers (e.g. several tens of or 100 m depth), which might exist in the Quaternary sediments in part of the catchment.

We would like to emphasize that we focus on modelling the shallow groundwater in this study. The requested details of the deep aquifers by the reviewer are not represented in our model (despite the fact that the model is potentially able to represent some of these details for simulation), because this detailed information is vastly not available. In fact, this information is even rarely available in the temperate climate and rather flat catchments, and not only in catchments with high elevation (>2000 m a.s.l.), with complex topography, and glacier coverage, such as the Martell Valley.

Not including the deep aquifers in the model does not mean losing the essence of the model at the headwater catchment scale. Shallow aquifers play important roles in sustaining winter streamflow, despite the relatively low magnitude in winter (Fig. 13 in Müller et al. 2022 for a proglacial outwash plain similar to Martell Valley). Modeling shallow groundwater with a hydrological model enables new understanding of the near subsurface hydrological processes and of how water partitioning is shaped by spatial and temporal surface water-groundwater interactions.

We include the essential heterogeneity in the subsurface by calibrating the hydraulic conductivity for each soil layer and for different soil types. Obtaining effective (calibrated) hydraulic conductivities is a prerequisite for a hydrologic model to reliably simulate observed streamflow, and this is a key challenge in physics-based models that often rely on pedotransfer functions to estimate hydraulic conductivity (Paschalis et al. 2022).

Heterogeneities occurring at a smaller scale (e.g. 25 m) are very difficult to characterize, as widely demonstrated by the vast literature in stochastic hydrogeology, starting from the detailed investigation of the famous sites of Borden and Columbus, for example (Rehfeldt et al. 1992; Woodbury, et al. 1991).

As pointed out by van Tiel et al. (2024), it is challenging to describe processes controlling cryosphere-groundwater interactions in alpine catchments due to a lack of observed data and modelling challenges. As a matter of fact, bucket models have been applied to the Martell Valley so far, in which the physically based description of groundwater fluxes is even more limited in comparison to our work (Schaffhauser et al, 2024; Puspitarini et al, 2020).

(2) Remote sensing data:

Remote sensing data (suggested by the reviewer) could potentially be used for estimating groundwater changes; we could think of GRACE data, but the low spatial resolution (>100km) is too coarse for high-elevation catchments characterized by a catchment area smaller than 100 km^2. The coarse resolution of the available remote sensing products for groundwater changes leads to high uncertainty, and it is hard to be adopted in a fully distributed hydrological model. Another limiting aspect of satellite-based gravimetry is the superimposition of gravity changes in glaciated high mountains. Here, changes in mass are caused by glacier melt, seasonal snow cover, and fluctuations in subsurface storage (Voigt et al. 2021).

(3) Physics-based modeling:

We respectfully disagree with the comment that our model is not physics-based. Part of the difference in perspective might stem from what we mean by "physics-based": We mean a model in which we describe groundwater movement according to the groundwater flow equation, while water movement in the unsaturated zone is described by Richard's equation. This approach is not fully 3-D as done by other software (e.g., HydroGeoSphere (Thornton et al, 2022)), but it still represents a step forward in comparison to bucket models, and it is, in our view, a good compromise between data availability and model complexity. This will be clearly specified in the revised version.

We acknowledge that WaSiM has an unusual specific requirement that the soil thickness must be kept the same as the aquifer bottom in the model, as the lower soil layers are taken as the saturated zone. Despite this numerical constraint, WaSiM shows to be useful to model the shallow aquifer (Fig.4 in this study), and it even outperforms the external coupling of surface (WaSiM) and subsurface (HydroGeoSphere) models for an alpine catchment, which show significant challenges in modelling the observed shallow groundwater heads within 2-5 m depth (see Fig. 4 in Thornton et al. 2022, 2021).

With the available information and data, we show in this study how to build this kind of fully distributed physics-based hydrological model for both surface and shallow subsurface processes in a high-elevation glaciated catchment, which is a missing piece in the literature.

(4) Adoption of a hydrological model (WaSiM) instead of a hydrogeological model:

In this manuscript, our research objective is to use a state-of-the-art hydrological model for groundwater modelling in such a high-elevation glaciated alpine catchment, with all its limits. We will make it clearer that this is not a groundwater model study. We adopt the hydrological model WaSiM, rather than a hydrogeological model (e.g. HydroGeoSphere, HGS), due to numerical and hydrological process understanding reasons.

The most relevant work on this topic among the handful modeling studies is the work of Thornton et al. (2022) who externally coupled the WaSiM (surface model) with HGS (subsurface model) for a high alpine catchment. That is, the snowmelt and rainfall generated from WaSiM are given as inputs to HGS. They show the challenges to neither get a good match of groundwater observations nor the streamflow (Figs. 3 and 4).

Firstly, the HGS is highly computationally intensive, challenging to calibrate, solves Richards equations in 3-D, and needs a good quality of geological model and soil depth map, where this information is hardly available and simple assumptions have to be applied for high alpine studies. Most importantly, they pointed out that "Whilst the soil zone is very small compared to the unconsolidated and consolidated geological formations volumetrically, it likely exerts a disproportionately strong hydrological influence via its influence on the partitioning of liquid water at the surface" (p.7 in Thornton et al. 2022). This is why we adopt a hydrological model over a hydrogeological model; we aim to get good partitioning of liquid water. Interestingly, by switching off the subsurface lateral flow in the hydrological model, we find good match for both groundwater and streamflow observations. Therefore, we need the hydrological model for such process understanding.

Compared to existing studies (Schaffhauser et al, 2024; Thornton et al, 2022, 2021; Puspitarini et al, 2020), the further added value is that (i) we build an integrated hydrological model with WaSiM, which avoids the uncertainty introduced by external coupling between a surface model and a subsurface model, (ii) we model the surface and shallow subsurface hydrological processes in a high-elevation alpine catchment, with glacier coverage, which enables to quantify the hydrological impact of glacier melt, and (iii) we model the subsurface (unsaturated and saturated zones) in a physics-based approach, given that most hydrological models represent the subsurface in a conceptual way.

(5) Baseflow and subsurface lateral flow:

Regarding the rather low baseflow mentioned by the reviewer, our groundwater piezometers are mainly installed in the upper headwater subcatchments, where the groundwater baseflow is potentially lower than the locations close to the catchment outlets and the lower part of the catchment (see the conceptual diagram in Fig. 1 in van Tiel, et al. 2023). In fact, we observe in the snow accumulation period a decrease in the groundwater table, which is relatively well captured by the model for ID4479, while the model underestimates the drop in the GW level for ID4478. This shows that the shallow aquifer described in the model may locally overestimate the storage. However, the shallow piezometers do not fall dry, and the experimental evidence does not exclude a connection between the shallow aquifer and the river during the snow accumulation period. Therefore, we cannot conclude, as pointed out by the reviewer, that the baseflow is uniquely generated in the winter months by a deeper circulation path. Additionally, the water stable isotope data collected in winters 22/23 and 23/24 for both the groundwater and the river display consistent values. Such results are not shown in the manuscript because they do not have a high temporal resolution, but they support the hypothesis of a connection between the shallow aquifer and the river during the winter period. We will circulate this point more clearly in the manuscript.

Regarding the subsurface lateral flow, it may not be negligible in the hillslope processes mentioned by the reviewer. In our study, the piezometers are installed in a rather flat area, where the subsurface lateral flow is found to play a minor role by calibrating the model to the observed groundwater heads. This finding could relate to the site characteristics and the observed locations. We will further clarify and discuss these points in the manuscript.

Despite all the limits, this research is among the first on performing such a detailed physics-based modelling study for both surface and shallow subsurface hydrological processes. We agree to better articulate and tone down some of the statements as pointed out in the revised manuscript.

References

Müller, T., Lane, S. N., & Schaefli, B. (2022). Towards a hydrogeomorphological understanding of proglacial catchments: an assessment of groundwater storage and release in an Alpine catchment. Hydrology and Earth System Sciences, 26(23), 6029-6054.

Paschalis, A., Bonetti, S., Guo, Y., & Fatichi, S. (2022). On the uncertainty induced by pedotransfer functions in terrestrial biosphere modeling. Water Resources Research, 58(9), e2021WR031871.

Puspitarini, H. D., François, B., Zaramella, M., Brown, C., & Borga, M. (2020). The impact of glacier shrinkage on energy production from hydropower-solar complementarity in alpine river basins. Science of The Total Environment, 719, 137488.

Rehfeldt, K. R., Boggs, J. M., & Gelhar, L. W. (1992). Field study of dispersion in a heterogeneous aquifer: 3. Geostatistical analysis of hydraulic conductivity. Water Resources Research, 28(12), 3309-3324.

Schaffhauser, T., Tuo, Y., Hofmeister, F., Chiogna, G., Huang, J., Merk, F., & Disse, M. (2024). SWAT-GL: A new glacier routine for the hydrological model SWAT. JAWRA Journal of the American Water Resources Association, 60(3), 755-766.

Thornton, J. M., Therrien, R., Mariéthoz, G., Linde, N., & Brunner, P. (2022). Simulating fully-integrated hydrological dynamics in complex alpine headwaters: potential and challenges. Water Resources Research, 58(4), e2020WR029390.

Thornton, J. M., Brauchli, T., Mariethoz, G., & Brunner, P. (2021). Efficient multi-objective calibration and uncertainty analysis of distributed snow simulations in rugged alpine terrain. Journal of Hydrology, 598, 126241.

van Tiel, M., Aubry-Wake, C., Somers, L., Andermann, C., Avanzi, F., Baraer, M., ... & Yapiyev, V. (2024). Cryosphere–groundwater connectivity is a missing link in the mountain water cycle. Nature Water, 2(7), 624-637.

Voigt, C., Schulz, K., Koch, F., Wetzel, K. F., Timmen, L., Rehm, T., ... & Flechtner, F. (2021). Introduction of a superconducting gravimeter as novel hydrological sensor for the Alpine research catchment Zugspitze. Hydrology and Earth System Sciences, 25(9), 5047-5064.

Woodbury, A. D., & Sudicky, E. A. (1991). The geostatistical characteristics of the Borden aquifer. Water Resources Research, 27(4), 533-546.