General Comments

The manuscript by Fan et al. presents a detailed and technically sophisticated application of a fully distributed, physics-based hydrological model, the Water Balance Simulation Model (WaSiM) to investigate surface–subsurface hydrological interactions in the glacierized catchment of Martell Valley, South Tyrol, European Alps. This study addresses a significant gap in our understanding of groundwater dynamics in alpine cryospheric environments. I appreciate the authors' effort in undertaking this important and challenging task.

The strength of this manuscript is the use of extensive observational data, often lacking in high-mountain environments. The comprehensive implementation of WaSiM to simulate both surface and subsurface hydrological processes is impressive. However, the manuscript would benefit from a clear articulation of its novel scientific contributions in relation to existing studies. I suggest including a dedicated comparison with similar modelling efforts to better highlight what is gained or potentially lost by the chosen approach.

Additionally, the limitations of applying such a model in complex mountain terrain, particularly with respect to spatial resolution, assumptions regarding subsurface properties, and data coarseness, should be discussed in detail with a separate section addressing model uncertainties, assumptions (e.g. uniform aquifer thickness). Further elaboration on key methodological decisions, such as the use of dual melt approaches, justification for subsurface parameterization, and the model's ability to represent delayed responses in groundwater will also help strengthen the manuscript.

Answer: We thank the reviewer for the constructive and reasonable feedback on our manuscript, and their appreciation in taking on this challenging task of performing a fully-distributed physics-based surface-subsurface modeling of the high-elevation glaciated environment.

We fully agree that a clear articulation of our study's novel scientific contribution in relation to the existing studies, and a dedicated comparison with the similar modeling efforts, would be greatly beneficial for highlighting the strengths and limitations of the chosen approach. We also agree to add in-depth discussions on the aspects mentioned in the comments, such as modeling assumptions regarding the subsurface properties, justified choices on spatial resolution, subsurface parameterization, and key methodological decisions in the revised manuscript.

Specific Comments

Comment #1: The authors position their study as one of the first to develop and implement a physics-based modelling framework to simulate surface–subsurface interactions in a glacierized environment. I agree with this claim, as there are indeed only a limited number of comparable studies. However, I recommend that the authors include a dedicated section comparing their approach with similar studies to clarify what is gained (or lost) by using this approach.

Answer: We agree that a dedicated section on comparing our study with the existing handful of similar modeling efforts will be beneficial to strengthen the manuscript. We will add this subsection in the revision.

Comment #2: The authors provide information about the catchment surface conditions, e.g. 40% bare rock, 34% grassland, and so on. I recommend that the authors include similar information about the valley floor settings (sediment-filled region), as most of the movement occurs in this region. Additionally, providing information on the land cover (bare rock, grassland, forest) in the map (Figure 1) can help readers better understand the area and relevance for hydrology.

Answers: We agree to provide further information on the valley floor settings into the manuscript, and add the land cover information into the Figure 1.

Comment #3: In Section 4.1 and Table 3: the description of the snow/rain partitioning scheme using the parameters *TRS* (i.e. temperature at which half of the precipitation falls as snow [0 °C]) and *Ttrans* (i.e. half of the temperature range from snow to rain [+2 °C]) could benefit from further clarification or simplification. Some readers who are less familiar with hydrological modelling may find the current phrasing difficult to understand. I suggest explicitly stating that this defines a linear transition of snow fraction from 100% at –2 °C, 50% at 0 °C, and 0% at +2 °C. Additionally, the authors may explain the use of two different approaches (EB and T-index) for snow and ice melt, even though both processes are forms of cryospheric melt that are influenced by similar energy exchanges?

Answer: Thanks for the helpful suggestion. We will explicitly explain these two parameters by following the suggestion.

Ideally, the energy balance method should be adopted for both snow and ice melt processes. However, the energy balance method (with snow redistribution) is only available for snow melt in the model, so we adopt the extended temperature-based approach which enables to include the radiation information for simulating the ice melt. We agree to further elaborate on these method decisions in the revision.

Comment #4: 4.2. Unsaturated zone and groundwater model: In a mountain environment, especially when glacierized, the use of uniform aquifer properties ignores the highly heterogeneous nature of the subsurface system. Such assumptions may hold true in plains, where natural depositions are relatively uniform, but not in mountainous environments. The authors later in the results (L361–362, and on several other occasions) imply the heterogeneity of the subsurface properties. Also, why are vertical flows modelled but not horizontal flows in the unsaturated zone? Please explain. Are the authors sure about the absence of permafrost in the upper parts of the study area?

Answer: The mountainous environments can be highly heterogeneous as shown in the different groundwater heads responses in Figure 3. However, the detailed spatial subsurface properties are vastly unknown. Therefore, we could only include essential spatial heterogeneity in the model, i.e. by calibrating the hydraulic properties of different soil layers and soil types. Such compromised assumptions are adopted due to rare data available in the subsurface spatially. We agree to further discuss such assumptions regarding subsurface properties, data scarcity, and their limitations in the revision.

Here we adopt the Richard's equation to solve water flow in the unsaturated zone, which is 1-D vertical flow in the soil columns. In fact, most of the existing hydrological models (based on our knowledge) only simulate the 1-D vertical flow in the unsaturated zone but no horizontal flow. This could be due to the complexity of solving 3-D flow in the soil and the vertical flow is the primary process to recharge groundwater. The subsurface lateral flow is generated in a conceptual way in the model depending on the soil water content, local slope, and hydraulic conductivity.

In the highest parts of the Martell Valley, permafrost causes thawing and freezing in the underground. However, this is not included in our model due to lack of spatially distributed soil temperature data. We will further elaborate and discuss this model assumption and its related uncertainties in the hydrological simulations in the manuscript.

Comment #5: 4.3. Streamflow generation: The model computes surface runoff, subsurface lateral flow, and baseflow separately and routes them using a reservoir cascade. This lumped treatment may oversimplify the interactions and timing differences among flow components.

Answer: The surface runoff and the subsurface lateral flow are routed to the streamflow using a reservoir cascade approach with a flow travel time concept applied (Schulla, 2024). The baseflow is generated and transmitted through the adjacent groundwater and river cells in a physics-based approach. We will further clarify these model specifics and its uncertainty in the manuscript.

Comment #6: Since groundwater/subsurface water dynamics are central to the study, I suggest the authors justify the use of 1.3 m as subsurface thickness across the entire catchment. The region is dominated by bare rock, grasslands, forest, and glaciers. It is obvious that subsurface conditions and thickness are not uniform. I believe the thickness of 1.3 m is low for such regions, especially in valley floors, with former glacial depositions or landslides. The authors mention numerical reasons and limitations, I understand that. In that case, the authors need to do a sensitivity test to see the influence of variable thickness on subsurface flow and storage, and on the overall outcomes. Additionally, the use of 1.3 m contradicts Figure 3(a), as the hydraulic head of borewell ID 4478 is around the same depth.

Answer: The adopted uniform soil thickness across the catchment is a simplified assumption, which is calibrated based on the groundwater levels observed at the five groundwater piezometers. It is challenging to assign spatially varying subsurface depths in the ungauged area, e.g. the valley floors. We will further discuss this assumption, the subsurface properties, and data coarseness in detail in the manuscript.

We agree to add the result of the sensitivity test of the soil thickness to justify our calibrated results, and its influence on the subsurface flow, storage, and the overall outcomes.

For Figure 3(a), this is not a contradiction though. The lowest groundwater head is about 1.2 m while the adopted soil thickness (subsurface thickness) is 1.3 m. This is due to the specific setup required by the WaSiM software that the soil thickness must be the same as the aquifer depth, as the bottom part of the soil layers represent the saturated zone.

Comment #7: The model ran at 25×25 m resolution. Isn't the soil profile from the global database (Harmonized World Soil Database), which has a resolution of 1 km, too coarse, particularly in a topographically complex region? The use of such data in a 77 km^2 area may not represent local settings (like moraine or similar glacial deposits). Please clarify.

Answer: The adopted soil data are the ones that are available for Martell Valley, despite the rather coarse spatial resolution of 1km. It is challenging to get more detailed soil profile data for this area. We adopt a high spatial resolution of 25 x 25 m for the modeling task to consider the detailed topography impact on hydrology. We agree with the likely uncertainty in the hydrological simulations due to the soil data resolution, and we agree to discuss more on the subsurface heterogeneity and data coarseness in the revision.

Comment #8: The authors have adopted manual calibration following Fatichi et al. (2015) to avoid computational load and time. This is justifiable, but manual calibration may be inefficient and subjective when dealing with many parameters across modules. Please clarify how the authors overcame this and what standardized approaches were used.

Answer: We tackle this task by calibrating the model from top to bottom and module by module. The sequential calibration of module by module is a commonly adopted logical procedure to calibrate such a fully-distributed physics-based hydrological model. We first perform manual sensitivity tests on the key parameters in each module. We then perturb and calibrate the identified sensitive parameters manually in detail to the observed variables.

The module by module calibration offers a good diagnostic power, as it isolates which modules (e.g., snowmelt, glacier melt) are causing discrepancies between observation and simulation. This allows an incremental validation, as each module can be tested and validated before integrating with the next. By doing so, errors in specific processes (e.g., snow or glacier melt) can be addressed without compromising other well-performing modules. Through the simplification of the calibration strategy, the parameter interactions are reduced, which leads to more stable model results.

Despite the calibration is sequential, we rerun the whole model (including all modules) each time when a parameter in a module is perturbed, and we focus on the model performance to the observed variables of that module. For example, when we calibrate a parameter in the snow module, we run the whole hydrological model and all temporal and spatial hydroclimatic outputs are produced, but we focus on the model performance compared to the snow water equivalent at the observed stations and spatial snow coverage. In this way, the hydrological processes between the modules are interconnected and the consistency is ensured.

We assigned the default values to the insensitive parameters given in the WaSiM user manual. Besides that, some parameter values are adopted from the literature. We agree that manual calibration is not an optimal solution. However in this study, our aim is not to reproduce the exact catchment, but to understand the hydrological processes with a reasonable parameter set, and the simulated dynamics or relative changes should be reasonably consistent. We agree to articulate our calibration procedure in more detail in the revision.

Comment #9: The authors point out an 8% underestimation of the annual glacier MB. An additional explanation about this is required. Since one glacier is used for calibration, the model may not have captured the heterogeneity of glacier responses in the catchment, as glacier MB is also affected by aspect, elevation, and other non-climatic parameters.

Answer: We agree to provide more explanation of the underestimated annual glacier MB. As only one glacier's mass balance in the catchment is available, this could lead to uncertainty in simulating the glacier responses in the whole catchment. We will discuss the influential factors mentioned by the reviewers in detail, such as the aspect, elevation, and other likely factors.

Comment #10: In Figure 2(c), please also include information about the plot within the circle. Though it is mentioned in the text, I believe some explanation is also required in the caption.

Answer: Agree and we will add this information in the caption.

Comment #11: L358: Such low discharge can also be contributed by frozen streams in winter at lower elevations, apart from the reasons mentioned.

Answer: Agree and we will add this information in the text.

Comment #12: L369–370: Is it due to the use of 1.3 m as the subsurface thickness and homogeneous subsurface properties in the model? I believe the groundwater level should show a lag time in response to surface conditions. It may also be due to the model time step (i.e. daily).

Answer: In fact, this result is obtained from the observed groundwater level and river level data, The groundwater from the deep circulation could have a long lag time in response to the surface conditions. However, the shallow groundwater observed here within 1.2 m depth shows to be as responsive as the river water level to peak melts and rainfall events in the headwater subcatchment. A lag less than 24 hours cannot be shown in the daily simulation though.

Comment #13: L385–394: I believe that the delayed response, especially in the early melting period, is valid. Additionally, the statement challenging the commonly adopted hydrological modelling approach about the role of soil needs more evidence.

Answer: Here could be a misunderstanding due to the wording. In fact, the observed shallow groundwater heads show very fast response to the early melt (nearly no delay) – this is an observed phenomenon. The simulated delayed response for months is thus incorrect when the subsurface lateral flow is allowed in the model. Additionally, the simulated groundwater hydrographs with delays depart from the observed groundwater hydrographs. That is, the observed groundwater hydrographs can only be simulated when the subsurface lateral flow is forced to 0 in this study. This finding could relate to site characteristics. We will look for more evidence on supporting this statement.

Comment #14: Section 5.6 can be made shorter; I see several repetitions of points already mentioned earlier.

Answer: We will try to shorten the length of this section. The points that have been discussed following the results will be shortened in this section.

Reference

Schulla, J.: Model Description WaSiM (Water balance Simulation Model), http://www.wasim.ch/downloads/doku/wasim/wasim_2024_en.pdf, 2024.