**Response to RC4:**

*Original comment:* This manuscript presents a fully distributed surface–subsurface modelling application to the high-elevation, glacierized Martell Valley using WaSiM, supported by an unusually rich observational dataset including groundwater levels. I did not review the original submission, and therefore my evaluation is primarily based on the present version of the manuscript. I have nevertheless read the responses to the previous reviewers and checked the corresponding revisions.

Overall, I find the study interesting and potentially suitable for HESS. The field dataset is valuable, the modelling exercise is substantial, and the simultaneous use of snow, glacier, streamflow, and groundwater observations is a strength. The authors have also made a considerable effort to address the previous reviews. In particular, the revised manuscript is more explicit about the limitations of the model, the daily temporal resolution, the spatial transfer of subsurface parameters, and the fact that the conclusion regarding lateral subsurface flow is model-based rather than directly observed.

However, I still have several substantial concerns regarding the interpretation of the modelling results. In a number of places, limitations or structural assumptions of the model are interpreted as hydrological process insights. I therefore recommend major revisions, principally aimed at separating more clearly what is demonstrated by the observations, what is demonstrated within the particular WaSiM configuration adopted here, and what can genuinely be inferred about the real catchment.

Answer: Thank you for taking the time to review our manuscript and for the overall positive assessment. We have provided a point-to-point response to each comment below and revised the manuscript accordingly.

Major comments

Comment #1.1: Interpretation of the negligible subsurface lateral-flow contribution

This is presented as one of the main findings of the paper, but I do not think the current evidence supports the strength or novelty assigned to it.

The supplementary Monte Carlo analysis convincingly demonstrates that, within the adopted model structure and parameterization, simulations with dr = 0 and krec = 1 reproduce the observed shallow groundwater heads better than simulations allowing stronger modelled “interflow”. This is a useful model diagnostic. However, it does not demonstrate that lateral subsurface flow is physically negligible in the catchment. The authors now acknowledge this distinction in several places, but the interpretation remains stronger elsewhere, including statements that the result provides “new insights” into limited lateral subsurface flow in high-elevation landscapes. There are several reasons why I think greater caution is needed. First, the piezometers are located in relatively flat floodplain/alluvial settings rather than along hillslopes, precisely where unsaturated lateral flow would normally be expected to be most important. Second, the model has a very shallow lower subsurface boundary (1.3 m) and a simplified representation of subsurface heterogeneity. Third, “interflow” is a specific WaSiM flux formulation

and should not automatically be equated with all possible rapid lateral subsurface flow pathways in the real catchment. Preferential flow, local perched systems, fractured-rock pathways and other processes may not be represented by this model component.

Answer: Thanks for pointing out that we did pay enough attention to all instances that we refer to model insights and not physics. We have also checked the manuscript throughout to ensure this interpretation consistently. Indeed, our results show that the hydrological dynamics of the catchment can be modelled without the WaSiM interflow component, which does not mean that the lateral subsurface flow does not exist. Furthermore, we acknowledge that WaSiM adopts the Richard's equations to solve the unsaturated zone matrix flow; other preferential flow pathways (e.g., fracture-rock pathways, local perched systems) might exist in the field but are not explicitly represented in the model. We clarified this limitation at L 559-565 in Section 6.2 in the revised manuscript.

Comment #1.2: The authors' Monte Carlo analysis therefore supports the narrower statement that explicit WaSiM interflow is not required to reproduce the available daily groundwater and discharge observations under the selected model structure. I suggest consistently adopting this formulation throughout the Abstract, Results, Discussion, and Conclusions. The physically broader inference that lateral subsurface stormflow is of limited importance should either be removed or framed explicitly as a hypothesis emerging from this particular model configuration that remains to be tested independently.

Relatedly, I do not find this result particularly "surprising" or "unexpected". Given that the groundwater observations constrain predominantly shallow alluvial aquifers in relatively flat areas, limited simulated unsaturated lateral flow at these locations is not in itself unexpected. The genuinely interesting result is instead that the available streamflow and shallow-groundwater dynamics can be reproduced without requiring the WaSiM interflow component.

Answer: Thanks. We answer this together with comments #12 and #17 below.

Comment #12: Lines 401–408: I do not find the claimed "new insights" concerning the limited role of subsurface lateral flow particularly surprising. Lateral flow is generally expected to be important in steep terrain, whereas the piezometers used here are located in relatively flat areas. More importantly, the result demonstrates that the specific WaSiM representation of unsaturated-zone interflow is not required to reproduce these observations, rather than demonstrating that lateral subsurface flow is physically negligible in the catchment.

Comment #17: Lines 579–587: Please distinguish carefully between "interflow" as represented in WaSiM and lateral subsurface flow as a physical process. Demonstrating that the former is not needed to reproduce the available daily observations does not demonstrate that the latter is absent or generally unimportant in the catchment.

Answer: The comments #1.2, #12, and #17 are around the same topic and are addressed here. We agree to be more careful with the formulation and interpretation of the constrained subsurface lateral flow results and to be clear that the WaSiM interflow is not required to reproduce the daily groundwater and discharge observations in the study site. This statement is now adopted consistently in the Abstract at L 11-14, Results at L 397-406 in Section 5.3, Discussion at L559-

580 in Section 6.2, and Conclusion at L678-681. The physical implication is now framed as a hypothesis emerging from this model configuration that requires independent field-based testing.

We nevertheless consider this model-diagnostic result to be an important and novel contribution. To our knowledge, explicitly simulated unsaturated-zone interflow has rarely been evaluated in glacierized high Alpine catchments using a fully-distributed, spatially-explicit model jointly constrained by groundwater heads and streamflow. Recent field-based study by Tani et al. (2026) shows that storm runoff responses in steep hillslopes can be governed primarily by vertical unsaturated flow (rather than lateral), with rapid transmission through the soil profile followed by drainage through preferential pathways. The targeted hillslope field observations in our study site (Martell Valley) will be needed in the future to test the hypothesis of limited role of subsurface lateral flow directly.

We also clarify that the model experiment is spatially distributed over the high Alpine landscape. The WaSiM interflow is therefore calculated across the model domain according to local soil water conditions, hydraulic properties, and slope before being routed toward the stream network. Thus, disabling this component affects both hillslope and valley areas, and the simulated discharge integrates runoff generation from all contributing subcatchments.

Comment #2.1: Groundwater mismatch, topography, and use of the TWI

I am not convinced by the interpretation of the groundwater mismatch around borehole 4478. The manuscript argues that the high local TWI (Topographic Wetness Index) explains why the model has difficulty releasing water at this location. However, a physically based distributed model should itself reproduce the effects of local topographic controls to the extent that these controls are resolved by its discretization and governing equations. Explaining a model–observation discrepancy by saying that the topography produces high simulated storage risks becoming circular: the discrepancy may instead indicate that the local model representation, parameters, boundary conditions, or spatial discretization are inadequate.

Answer: Thanks. This comment is addressed together with Comment #13 below.

Comment #13: Lines 420–424: I am not convinced by the explanation that the high TWI at borehole 4478 accounts for the model's inability to reproduce the observed recession. A properly parameterized distributed model should itself reproduce the consequences of these topographic controls; the discrepancy may instead indicate limitations in the local model representation, parameters, or spatial discretization.

Answer: The comments #2.1 and #13 are on the same topic and are addressed here. We are not entirely convinced that this is circular reasoning: the model reproduces the effects of topographic controls in the limit of the model parameterization, i.e. here the scale. We can be more precise: it is not the topography (“nature”) but the chosen model representation of topography (“parameterization”). We have revised this part of the manuscript at L 409-411 accordingly to make clear that it is the model parameterization and not “nature” that leads to this effect. In the revision we clarify that TWI is one diagnostic of the modeled spatial groundwater behavior rather than the sole explanation for the discrepancy at borehole 4478. The relatively high TWI at this location reflects a modeled tendency toward wetter conditions, while we also acknowledge the

remaining discrepancy might also arise from unresolved fine-scale topography, subsurface heterogeneity, or spatial discretization (Riihimaki et al., 2021; Rinderer et al., 2014; Grabs et al., 2009). This revision is provided at L 419-423 in Result Section 5.3 in the manuscript.

Comment #2.2: The subsequent discussion of neighbouring grid cells also needs to be reframed. Spatial mismatch between a point observation and a model cell due to DEM uncertainty, spatial discretization, and geolocation uncertainty is well known in distributed modelling. Looking at neighbouring cells is reasonable, but selecting the cell producing the best match risks introducing an additional implicit calibration step. A more defensible approach would be to define beforehand a spatial support around each piezometer and report, for example, the ensemble mean/median and variance or range of simulations over all plausible cells. This would quantify representativeness uncertainty without effectively selecting the locally best-performing cell.

Answer: This comment is addressed together with Comment #14 below.

Comment #14: Lines 425–436: The mismatch between point observations and model grid cells due to spatial discretization and topographic uncertainty is well known and essentially unavoidable. Rather than identifying neighbouring cells that happen to provide the best agreement, a more rigorous treatment would be to define an objective group of plausible cells around each piezometer and report the mean (or median) together with the variance or range of the corresponding simulations. This would quantify spatial-support uncertainty without effectively introducing an additional local calibration step.

Answer: The comments #2.2 and #14 are on the same topic and are addressed here. We agree that choosing the best local support is a form of calibration. However, defining beforehand (without inverse modelling) the spatial support of the piezometer is highly challenging: given the local nature of the measurement and the heterogeneity of the subsurface, a simple spatial averaging approach may smooth out the observed dynamics and therefore provide a poor representation of the point measurement. One solution could be a stepwise refining of the spatial modelling support, another form of calibration (refine the grid until the observation is reproduced) but computationally highly demanding. We revised the manuscript to clarify that identifying the best-matching grid cell is used primarily to demonstrate that the model can reproduce the observed dynamics. Resolving the exact point-scale measurement would require further refining the spatial discretization and a more detailed representation of local subsurface heterogeneity.

That being said, we agree with the suggestion of reporting the median and spread of the simulated groundwater hydrographs for understanding the uncertainty. In the revision we have shown this quantification in Figure S14 in the Supplementary Material. At the same time, we retain that visually examining individual neighbouring-cell hydrographs is a crucial diagnostic approach, given that the strong local topographic heterogeneity in high Alpine catchment can produce markedly different groundwater responses over short distances. Therefore, it is important to not simply calculate the median. Accordingly, we find that the neighbourhood statistics provide the primary measure of representativeness uncertainty, while visually examining individual simulations is crucial for interpreting local model behaviour. We have revised the manuscript accordingly at L 424-431 in Result Section 5.3.

Comment #2.3: For the same reason, I am not persuaded by the recommendation that TWI should be used to guide future piezometer placement primarily so that observations correspond better to model grid cells. Monitoring design should principally aim at sampling relevant hydrogeological and hydrological conditions, not at placing instruments where a particular model is most likely to perform well. TWI could certainly be one criterion for stratifying monitoring locations, but the recommendation should be broadened and substantially softened.

Answer: Thanks for pointing out that we should be clearer stating that TWI is just one element for instrument placing. We made the implicit assumption that this is on top of other criteria to sample representative landscape units. TWI provides spatially explicit information on topographic controls on wetness and therefore offers an important basis for stratifying groundwater observations across contrasting hydro-topographic settings. We clarify in the revision that TWI should be used together with other site characteristics, including hydrogeology, geomorphology, soil properties, proximity to the stream network, and most importantly, the objectives of the monitoring campaign. In this framework, TWI remains a valuable criterion for guiding piezometer placement because it can help ensure that groundwater monitoring captures the range of topographically controlled wetness conditions represented across the catchment. The revision is provided at L 586-603 in Section 6.3. In addition, an iterative monitoring design informed by preliminary model simulations could be a useful strategy, particularly given that such model-guided planning has not yet been routinely incorporated into field campaign design.

Comment #3. Static versus dynamic stream network

The stream network is represented by fixed stream cells generated from the DEM. This is an important structural assumption in a high-elevation headwater catchment, where the active stream network can expand and contract strongly through time. Indeed, the authors later interpret their simulations in terms of winter network contraction and summer expansion, although the actual model stream-cell topology appears to remain static.

This assumption deserves explicit discussion. How does a prescribed stream network affect surface–groundwater exchange, especially given that groundwater exchange is only possible in relation to designated stream cells? Could the inferred seasonality in gaining/losing conditions, groundwater exfiltration, or the partitioning between surface runoff and groundwater flow be affected by the fact that hydrologically inactive and active portions of the network are structurally represented in the same way?

At minimum, this should be added as a limitation and the interpretation of Figures 8–9 should distinguish clearly between changes in simulated fluxes along a fixed potential channel network and actual expansion/retraction of the active stream network. It would also be useful to state whether WaSiM offers any possibility for dynamically activating/deactivating stream cells, or whether this could be considered in future work.

Answer: Thank you very much for pointing out this important omission in the paper. We have revised the manuscript to clarify that the stream network in the present WaSiM setup is prescribed from the DEM and remains static in the simulation, while seasonal variability is represented

through changes in the associated hydrological fluxes. In the Methods Section 4.3 at L 237-239, we now explicitly state that the stream grid cells define a fixed potential channel network throughout the simulation period. In the Results at L 510-511, we have clarified this point when interpreting Figures 8-9. The seasonal patterns shown in these figures are now described as temporal variations in simulated groundwater–stream exchange along the prescribed channel network. In the Discussion Section 6.4 at L 663-672, we have discussed more about this model limitation and highlight that dynamically updating stream networks in fully distributed models remains technically highly challenging because changes in channel topology require continuous updates of the spatial model structure and associated state variables. This issue is particularly relevant in glacierized environments, where evolving drainage pathways may accompany landscape and glacier changes (Fan et al., 2026). We therefore identify dynamic updating of stream cells as an important direction for future model development clarified this point in the revised version.

Comment #4. Claims regarding physically based integrated models

Several statements in Section 6.2 are too broad. In particular, the suggestion that even physically based models employing Richards' equation may fail to reproduce the rapid shallow-groundwater dynamics, followed by the implication that the present framework provides an advance in this respect, is not convincing as currently written.

The manuscript should explain precisely what WaSiM includes that other integrated surface–subsurface hydrological models do not. Richards-equation/Darcy-based integrated models can represent fast groundwater responses, saturation dynamics and groundwater feedbacks depending on their dimensionality, discretization, parameterization, boundary conditions and representation of preferential pathways. The observations presented here certainly demonstrate rapid shallow groundwater responses, but they do not by themselves demonstrate a general deficiency of physically based ISSHMs.

I recommend substantially narrowing this discussion and avoiding broad comparisons between "conventional", "conceptual", and "physically based" models unless the specific structural differences responsible for the claimed behaviour are clearly identified.

Answer: Thanks. We answer this together with comment#16 below.

Comment #16: Lines 571–576: This statement is not convincing. What does the model used in this study contain that other physically based integrated surface–subsurface hydrological models do not? Richards-equation/Darcy-based models can reproduce rapid groundwater responses depending on their structure, discretization, parameterization and boundary conditions. The authors should identify the specific structural advantage of WaSiM rather than making a broad statement about physically based models.

Answer: The comments #4 and #16 are on the same topic and are addressed here. We have revised the manuscript to enhance the statement and acknowledge that physics-based models can potentially reproduce rapid shallow groundwater responses when the relevant flow pathways are adequately represented through model structure, parameterization, spatial discretization, and boundary conditions. We have also clarified the structural advantage of the present WaSiM

application. In our WaSiM setup, surface runoff, unsaturated-zone lateral flow, shallow groundwater dynamics, and groundwater–streamflow exchange are represented within the same fully-distributed framework. This allows the relative contributions of these flow components to be evaluated jointly against both groundwater heads and streamflow observations.

An additional feature of WaSiM relevant to this study is its groundwater module (despite two-dimensional), which provides a physics-based representation of groundwater dynamics (Darcy equations) within the integrated modeling framework. This differs from many other distributed hydrological models, such as the fully-distributed ecohydrological T&C model (Fatichi et al., 2012), where drainage from the soil is transferred directly to the stream (no groundwater component), and semi-distributed mHM model (Kumar et al., 2013) and THREW-T model (Nan et al., 2025), where groundwater dynamics are represented using a linear store. The revisions are provided at L 544-558 in Section 6.2 in the manuscript.

Additionally, we compared our study directly with that of Thornton et al. (2022) who coupled WaSiM externally with HydroGeoSphere, which is a three-dimensional hydrogeological model but requires a substantially high degree of calibration effort and hydrogeological data (often unavailable in glaciated high Alpine catchment) at L 514-527 in Section 6.1. Such comparison suggests that intermediate-complexity integrated models (e.g., WaSiM) may provide a useful compromise for process-based modelling in data-limited high Alpine environments.

Comment #5. Model dimensionality and hydrogeological simplifications

The model is repeatedly described as “fully integrated” and “fully distributed”, which is technically understandable, but readers should not interpret this as meaning that the subsurface hydrogeology is represented in comparable detail. The groundwater model is two-dimensional, the lower groundwater/soil boundary is fixed at 1.3 m, most groundwater parameters are spatially uniform, and the calibration observations all come from shallow aquifers in broadly similar geological settings. The same calibrated subsurface configuration is subsequently transferred to much of the ungauged catchment.

The authors acknowledge many of these limitations, including in response to the earlier reviews, which is welcome. However, the implications should be carried through more consistently into the Results. Catchment-wide maps of groundwater recharge, heads, exfiltration, and infiltration can appear much more physically constrained than they actually are. In particular, results outside S7–S9 should be presented explicitly as exploratory model outputs rather than inferred catchment-scale groundwater behaviour.

This is especially relevant to the conclusion that deeper groundwater contributions are unlikely to be substantial in the upper headwaters. Given that the model effectively excludes deeper groundwater flow paths, this conclusion cannot be strongly tested with the present setup.

Answer: Thank you for this comment. In the revised manuscript, we have improved the consistency of the interpretation across the Methods, Results, Discussion, and Conclusions, ensuring that the scope and limitations of this finding are described in a consistent manner throughout. The groundwater observations used for calibration represent shallow unconfined aquifers and are concentrated primarily in the upper Subcatchments S7–S9. Accordingly, we now

distinguish more clearly between groundwater responses that are directly constrained by the observations in these areas and the catchment-wide groundwater fields that represent spatially distributed model outputs under the adopted subsurface parameterization. In the Results, the findings outside S7-S9 about groundwater recharge, heads, exfiltration, and infiltration are now presented as exploratory spatial model outputs where observations are limited. The revision is provided at L 485-487 in Section 5.5.

The present model configuration is designed primarily to include shallow groundwater dynamics. The inference about the unlikely deep groundwater flow in the upper headwater catchments are now removed. Instead, we state explicitly that the available shallow groundwater observations and the present model configuration constrain the near-surface groundwater system, while deeper flow contributions remain outside the range directly tested by this setup. The revisions are provided at L 628-633 in Section 6.4 Limitations.

## Specific/minor comments

Comment #6: Lines 11–14: The statement concerning limited subsurface lateral flow and the conclusion that groundwater flow together with surface runoff is "dynamic enough to explain the observed daily streamflow dynamics in the computational settings" is not sufficiently clear. In particular, "in the computational settings" is vague and the physical meaning of this conclusion should be stated more precisely.

Answer: We have revised these lines in the Abstract accordingly and avoided the broader physical implication of the constrained subsurface lateral flow. It is revised at L 11-14 that "Joint evaluation against shallow groundwater heads and streamflow shows that the daily streamflow and shallow groundwater dynamics can be reproduced without requiring the explicit WaSiM interflow component, providing new constraints on the representation of unsaturated-zone flow pathways in integrated high Alpine hydrological models".

Comment #7: Lines 51–52: The sentence beginning "However, the streamflow generation process…" is grammatically incomplete and should be rewritten.

Answer: We have revised this sentence at L 52-53.

Comment #8: Lines 246–248: The stream grid cells are generated from the DEM and therefore correspond to a static definition of the drainage network. An increasing body of literature highlights that stream networks, especially in headwater catchments, can be highly dynamic. Please discuss how this assumption might affect the results, particularly groundwater–surface water exchange and runoff partitioning. It would also be useful to indicate whether a dynamic representation of the stream network could be considered in future developments.

Answer: We have revised the manuscript to clarify that the stream grid cells are generated from the DEM and remain fixed throughout the simulation, while seasonal variability is represented through changes in hydrological fluxes along this prescribed channel network in Methods Section

4.3 at L 237-239. In the Results at L 510-511, we clarify the seasonal patterns in groundwater–stream exchange as temporal flux variations along the static potential stream network. In the Discussion in Section 6.4 at L663-672, we explicitly clarify the limitation of this structural assumption for the simulated spatial extent and magnitude of groundwater–surface water exchange and for runoff partitioning, particularly during periods when the active stream network may expand or contract in reality. We also identify dynamic activation or updating of stream cells as a relevant direction for future model development in high Alpine catchments, which remains a key challenge of updating drainage networks in evolving glacierized landscapes (Fan et al., 2026).

Comment #9: Lines 265: What is the "colmation of the soil"?

Answer: "Colmation" is a term used in the WaSiM model: it represents the water exchange rate between the adjacent saturated soil cells and the stream network cells. It has a unit of m/s. This has been clarified in the caption of Table 4.

Comment #10: Lines 365–386: Section 5.2 discusses observed river and groundwater responses together with simulated rainfall and melt forcing, but not the corresponding simulated groundwater and streamflow responses. It is therefore unclear why the hydrological simulations are not discussed in this section. Please either restructure the section or clarify its purpose.

Answer: The purpose of Section 5.2 is to examine how the observed river and shallow groundwater systems respond to the estimated hydrological inputs during snowmelt, glacier melt, and rainfall events. We therefore intentionally use the observed river water levels and groundwater heads as independent response variables, while the simulated rainfall, snowmelt, and glacier melt fluxes are used to characterize the timing and magnitude of the water inputs to the subcatchment. This allows the timing of the observed hydrological responses to be examined directly without relying on the model simulations for both the forcing and the response variables. The corresponding simulated groundwater and streamflow dynamics are subsequently evaluated against observations in Sections 5.3 and 5.4. We have clarified this purpose at L 359-365 in Section 5.2 in the revision.

Comment #11: Lines 395–396 and elsewhere: Please add units for MAE and RMSE; these quantities are not dimensionless.

Answer: Thanks for pointing this out. We have revised the manuscript to report units for all MAE and RMSE values. The units now correspond to those of the evaluated variable: mm for Snow Water Equivalent (SWE), m for groundwater heads, and $m^3/s$ for river discharge.

Comment #12: Lines 401–408: I do not find the claimed "new insights" concerning the limited role of subsurface lateral flow particularly surprising. Lateral flow is generally expected to be important in steep terrain, whereas the piezometers used here are located in relatively flat areas. More importantly, the result demonstrates that the specific WaSiM representation of unsaturated-

zone interflow is not required to reproduce these observations, rather than demonstrating that lateral subsurface flow is physically negligible in the catchment.

Answer: This comment has been addressed together with Comment #1.2. The revision is provided at L 397-406 in Section 5.3.

Comment #13: Lines 420–424: I am not convinced by the explanation that the high TWI at borehole 4478 accounts for the model's inability to reproduce the observed recession. A properly parameterized distributed model should itself reproduce the consequences of these topographic controls; the discrepancy may instead indicate limitations in the local model representation, parameters, or spatial discretization.

Answer: This comment has been addressed together with Comment #2.1. We have revised the text at L 407-423 in Section 5.3 accordingly.

Comment #14: Lines 425–436: The mismatch between point observations and model grid cells due to spatial discretization and topographic uncertainty is well known and essentially unavoidable. Rather than identifying neighbouring cells that happen to provide the best agreement, a more rigorous treatment would be to define an objective group of plausible cells around each piezometer and report the mean (or median) together with the variance or range of the corresponding simulations. This would quantify spatial-support uncertainty without effectively introducing an additional local calibration step.

Answer: This comment has been addressed together with Comment #2.2. We have revised the text at L 424-431 in Section 5.3 accordingly.

Comment #15: Lines 532–534: I would avoid describing the result as “unexpected”. Given the location of the groundwater observations and the adopted model structure, the more interesting result is that WaSiM can reproduce the available groundwater and streamflow observations without requiring its interflow component.

Answer: We have removed the term “unexpected” and revised the discussion to emphasize the model-specific result: the shallow groundwater level and streamflow observations can be reproduced without requiring the WaSiM interflow component at L 559-565 in Section 6.2.

Comment #16: Lines 571–576: This statement is not convincing. What does the model used in this study contain that other physically based integrated surface–subsurface hydrological models do not? Richards-equation/Darcy-based models can reproduce rapid groundwater responses depending on their structure, discretization, parameterization and boundary conditions. The authors should identify the specific structural advantage of WaSiM rather than making a broad statement about physically based models.

Answer: This comment has been addressed together with Comment #4. The corresponding revisions are provided at L 544-558 in Section 6.2 in the manuscript.

Comment #17: Lines 579–587: Please distinguish carefully between “interflow” as represented in WaSiM and lateral subsurface flow as a physical process. Demonstrating that the former is not needed to reproduce the available daily observations does not demonstrate that the latter is absent or generally unimportant in the catchment.

Answer: This comment has been addressed together with Comment #1.2. The corresponding revisions are provided at L559-565 in Section 6.2 in the revised manuscript.

Assessment of the revision

The authors have responded seriously to the previous reviewers and many of their requests have resulted in clear improvements. I particularly appreciate the more explicit acknowledgement that the inferred absence of interflow is conditional on model structure and temporal scale, the expanded Monte Carlo sensitivity analysis, the more cautious treatment of low-flow performance, and the clearer discussion of the shallow groundwater limitation. Nevertheless, in my view the central process interpretation still needs another round of revision. The manuscript will be considerably stronger if the authors present the “zero-interflow” result primarily as a diagnostic finding about what the observations require from this particular model, rather than as direct evidence that lateral subsurface flow is generally unimportant in the Martell Valley or other high-alpine catchments. I therefore recommend major revisions.

Answer: Thanks for recognizing the value of our work. We have responded to each comment above accordingly, especially paying attention to interpreting the interflow result as suggested.

## References

Fan, X., Schaefli, B., Hofmeister, F., & Chiogna, G. (2026). When the past matters: How model initialization can lead to surprises in long-term simulations in glaciated environments. Hydrological Processes, 40(4), e70507. https://doi.org/10.1002/hyp.70507

Fatichi, S., Ivanov, V. Y., & Caporali, E. (2012). A mechanistic ecohydrological model to investigate complex interactions in cold and warm water-controlled environments. 1 Theoretical framework and plot-scale analysis. Journal of Advances in Modeling Earth Systems, 4, M05002. https://doi.org/10.1029/2011MS000086

Grabs, T., Seibert, J., Bishop, K., & Laudon, H. (2009). Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model. Journal of Hydrology, 373(1–2), 15–23. https://doi.org/10.1016/j.jhydrol.2009.03.031

Kumar, R., Samaniego, L., and Attinger, S. (2013). Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations, Water Resources Research, 49, 360–379, https://doi.org/10.1029/2012WR012195.

Nan, Y., Tian, F., McDonnell, J., Ni, G., et al. (2025). Glacier meltwater has limited contributions to the total runoff in the major rivers draining the Tibetan Plateau, npj Climate and Atmospheric Science, 8, 155, https://doi.org/10.1038/s41612-025-01060-6

Riihimäki, H., Kemppinen, J., Kopecký, M., & Luoto, M. (2021). Topographic Wetness Index as a proxy for soil moisture: The importance of flow-routing algorithm and grid resolution. Water Resources Research, 57(10), e2021WR029871. https://doi.org/10.1029/2021WR029871

Rinderer, M., van Meerveld, H. J., & Seibert, J. (2014). Topographic controls on shallow groundwater levels in a steep, prealpine catchment: When are the TWI assumptions valid? Water Resources Research, 50(7), 6067–6080. https://doi.org/10.1002/2013WR015009

Tani, M., Sidle, R. C., Sayama, T., Matsushi, Y., & Kojima, N. (2026). Vertical Unsaturated Flow Governs Storm Runoff Responses: Observational and Theoretical Foundations. Hydrological Processes, 40(7), e70627. https://doi.org/10.1002/hyp.70627

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. https://doi.org/10.1029/2020WR029390