Reviewer 2

Comment: This is a very interesting paper that addresses the elusive role of preferential flows by introducing an additional consideration to the bucket model, which distinguishes the role of preferential flows through shape functions. It eloquently poses the paper's main questions and refers to some of them at the end, thereby focusing the paper on specific aspects of the model and its ability to capture the role of preferential flows. In that respect, this paper has a substantial contribution to the field and should eventually be accepted for publication. However, this is not a "stand-alone paper" as, without reading the previous work by Turk 2024, there is no way to establish the relation between the "bucket ratios" and the model without going over Eqs. 1-26 in the said paper. The submitted paper continues the model from Turk 2024, starting from Eq. 33, and proceeds to examine the sensitivity to the SAS function optimization, specifically to the sensitivity to the preferential flow in light of the parameter (shape function) in the said equations. However, this sensitivity is hard to understand without the full model and figures 3 and 4 in Turk 2024. Further figures are needed from Bogena et al. 2015 (figures 1 and 2) and Turk 2024 (figure 1) to visualize and understand the layout of the field and its relation to the bucket model in Figure S1. I would argue that this paper is a continuation of Turk 2024, model-wise, and it can only be understood if one reads the 2024 paper prior to this paper, and as such, narrows the potential readership of this interesting and relevant paper.

Response: We thank Reviewer #2 for the positive feedback on our manuscript. We will address the concerns by providing additional background material in the Supplement. Below, we outline how we consider the issues raised by the reviewer and the changes we intend to make. Reviewer comments are shown in black, while our responses are shown in blue. Specifically, we will (i) briefly summarize the relevant parts of the model setup from Turk (2024), including the bucket ratios and key equations, and (ii) add figures adapted from Bogena et al. (2015) and Turk (2024) to illustrate the field layout and its relation to the bucket model. This will ensure that the present manuscript can be read independently without requiring the reader to consult the previous work.

Comment: Furthermore, coming from the transport in soil and rock community, the model seems very phenomenological in nature, which I guess is standard for "bucket models". While it makes a nice attempt to move away from the data-driven model and statistical approaches towards more mechanistic observation, there is no explicit attempt to present the causality between physical parameters to observations. This has to be done when we account for preferential flows, as these preferential flows are the outcome of specific conditions (permeability heterogeneity of the root area and groundwater, head differences within the groundwater stemming from the slope, and slope runoff on a local scale). However, the discussion on these aspects is limited to Section 4, and even there, it does not provide specific details. They are found to some extent in Bogena (2015) for Wüstebach, yet I couldn't find them for the HOAL in Turk (2024). Why is that?

This approach stays on the data side, and while the correlation it provides is an important step towards a more rigorous model, can't there be an effort to hypothesize about the mechanisms that form the varying dominance of preferential flows and see if the data and the missing parameters might corroborate this hypothesis?

Response: We acknowledge that our approach is phenomenological and will add a short discussion in Section 4 to link literature-based findings on soil texture, moisture dependence, and slope effects (e.g., Bogena et al., 2015 for Wüstebach, and additional references for HOAL) to the SAS parameterization used here. While our study does not explicitly test mechanistic functions, highlighting these connections will help frame how the SAS shape parameter may reflect underlying physical controls and point to directions for future mechanistic work.

Comment: To be more specific in my request, let's take the following example: Is there a possibility to search the literature and provide a functional form that relates how clayey soils under variation in saturation contribute to the emergence of preferential flows, and how this functional form relates to the shape function in this study? Can we compare the sensitivity of this functional form to these aspects (clay concentration, saturation) and the sensitivity of the shape function? Can it account for the model's sensitivity to the shape function? The same can be done for the differences in slope between the two data sets presented in this study. While this is not a mechanistic model, it indicates what parameters are needed in a mechanistic model, and the sensitivity to them. This will allow this community, and adjacent communities like soil and aquifers, to frame their mathematical approaches in the context of this excellent problem, especially given the excellent datasets presented here. Datasets that urge for a more rigorous and mathematical modeling approach, in space and time, that require some input on the mechanistic origin for these probability density functions controlling the current buckyeet model.

Response: Indeed it would be interesting to compare the SAS function shape to different catchment properties, which could allow an a priori parameterization of the SAS function without calibration. In our opinion, this would require testing the model in other catchments that cover a wide range of climates, soils, and topographies, followed by statistical analysis to identify the influential factors controlling SAS function shape. For instance, at the lysimeter scale, Asadollahi et al. (2020) showed that the SAS function shape was similar to the analytical solution of the advection–dispersion equation. However, applying such mechanistic or semi-mechanistic relationships at the catchment scale remains challenging in the context of a "bucket model," as processes are aggregated over the entire system. We will add a short discussion paragraph to highlight this as an open question and future research need, which could provide a bridge between empirical SAS modeling and mechanistic approaches.

Line 31: "In the light of…" or "In light of…"?

Response: We will revise the phrase to "In light of …"

Comment: The introduction is very good, and so are the questions which represent the study aim and objective.

Response: We thank the reviewer for the positive feedback on the introduction and the formulation of the study questions.

Comment: Figure 1. Can you provide insight on the streamflow? Maybe maps of the catchment with the locations of the sampling locations? That will help in understanding the significance and time lag of these measurements.

Response: We will add maps of both catchments showing the stream network and the locations of the sampling points to Figure 1 (or as a supplementary figure).

Comment: Figure S1: TT should be Tr by table S 1. One must define the parameters in black.

Response: We will correct Tr to TT in Table S1. In addition, we will clarify that T_T refers to the threshold temperature for snowmelt that needs to be calibrated.

Comment: "The darker blue box (Ss,a)" should be Ss,p Not all symbols are defined in Table S 1, please re-check. Maybe move Table S 1 next to Figure S 1 so it will be easier to go between them.

Response: We will correct the label to Ss,p and carefully check Table S1 to ensure that all symbols used in the figures and text are consistently defined.We will move the Table S 1 next to the Figure S 1.

Comment: Line 167: α and β, do not appear in Table S 1 or Figure S 1, please define them in the right context.

Response: We will revise the text and Table S1 to include definitions of α and β. Specifically, we will clarify that β is fixed as 1 for all fluxes, while α is fixed as 1 for preferential groundwater flow but allowed to be calibrated for the unsaturated zone preferential flow (SU,α).

Comment: The number of references to Figure S1 suggests that it should be included in the main paper rather than the supplementary material.

Response: We will move Figure S1 from the Supplement to the main manuscript.

Comment: Equations 1 and 2: How do you define Sr,max? Is it local in time (the maximum of the soil moisture for the given rain event) or is it simply the saturation value, namely the porosity? From the equation, its unclear how this term captures the preferential flow, specifically how it relates it to physical aspects, which are not saturation (heterogeneity and head differences)

Response: We will clarify that Sr,max is a calibrated parameter representing the maximum root-zone storage capacity, not porosity. The SAS function for age selection of preferential flow from the unsaturated root zone (Rf, Fig. S1) was formulated as a time-variable function of relative soil wetness to reflect changes in transport processes between wet and dry soil conditions. In particular, it accounts for the preferential release of younger water as soil wetness increases, consistent with previous studies showing that wetter soils promote faster flow paths with reduced mixing of water. The temporal variability in the SAS function was implemented through the time-dependent shape parameter $\alpha(t)$ (Eq. 2).

Comment: Line 223-226: Since there is no discussion on how preferential flow components like heterogeneity and head differences were implemented in the model, it is hard to understand if this is the right way to measure the model's sensitivity to it.

Response: We will clarify the model setup and revise the description of the second test to emphasize that our approach is based on age-ranked groundwater storage. Specifically, the section will be revised to:

"In the model setup, groundwater storage was formulated as consisting of an 'active' groundwater storage (Ss,a) and a hydrologically 'passive' storage volume (Ss,p, mm). The passive storage does not change over time if there are no deep infiltration losses (Zuber, 1986; Hrachowitz et al., 2016; Wang et al., 2023). While Ss,p does not contribute directly to runoff, it isotopically mixes with water in the active storage, thereby influencing the tracer signal of the outflow. The total groundwater storage (Ss,tot) was therefore considered as the sum of Ss,a and Ss,p, so that the age-ranked groundwater storage (Ss,tot) inherently reflected a mixture of these storage volumes. Next, we tested the sensitivity of the δ^2H signal to changes in the groundwater SAS function (Fig. 2b) by varying α across the same range—0.1, 0.7, 1.0, and 5.0—while fixing the previously optimized α_0 value for the root zone. This second test was designed as an age-ranked experiment to examine if (and how) preferential release of younger groundwater influences transit time distributions and tracer simulations. For clarification, the SAS formulation can only identify whether preferential release of young water occurs, but it does not resolve the underlying physical mechanisms that generate it."

Comment: Figure 2: Why is the CDF shown and not the PDF for the TTD? Indeed, one is the derivative of the other, yet the PDF allows us to see the tailing and the mean behavior more clearly. I'm also not sure if eCDF is the cumulative of , as it is not stated, and from the curvature, it's hard to establish it.

Response: We will clarify in the text and figure caption that eCDF refers to the empirical cumulative distribution function. We selected cumulative distributions to show the individual curves more clearly, and for scenario comparison we found cumulative curves easier to interpret than PDFs.

Comment: A non-intuitive aspect is that preferential release of older water ($\alpha>1$) has the highest residence time. I assume that this is due to the interplay between the ground and root area flow, which affects the fast response water in Figure S1. Therefore, as α increases above 1, less water is directed to the fast response bucket; however, there is no indication of this. Where are the equations related to each "bucket"? How do you parameterize these "buckets"?

Another aspect is that Figures 2a, b, and 3 are basically identical. I understand that they are "illustrations", but if there is no difference among the illustrations, what is the point of showing 3 of them? Furthermore, the illustration of the different passive storage values does not provide any insight into how the analysis differs or how the results should change as a result.

Response: We thank the reviewer for this comment. We will improve the clarity of both the model description and the figures.

To address the first point, we will clarify in the methods section that preferential release of older water ($\alpha > 1$) increases mean residence times because less young water is directed toward the fast-response pathways (root-zone preferential flow), leading to a stronger weighting of slower groundwater contributions. We will also add the relevant bucket-model equations for water age balance equations and (currently included in Turk 2024) section "Integration of the rSAS function concept and the hydrological model", and provide a clearer description of how each bucket tracer balance through SAS formulation is parameterized.

To address the second point, we will streamline the figures to avoid redundancy. Specifically, we will keep only the necessary illustrations and clarify in the captions what information each figure is intended to convey. For passive storage values, we will highlight in the results section how varying Ss,p affects tracer damping and the tails of the TTD, which was the purpose of including those scenarios.

Comment: Line 248: 15 parameters are identified in Table S1, and I guess there are multiple locations at which they are measured, which spans the calibration to 55 and 190 parameters, but instead of guessing, please clarify this aspect in a text, or better yet, a figure outlining the measurements in the catchment area.

Response: We will clarify in the text that the 15 parameters listed in Table S1 were calibrated for each catchment. The number of parameter solutions (55 for HOAL and 190 for Wüstebach) refers to the sets of parameter combinations that satisfied the calibration criteria (DE < 1), not to the number of calibrated parameters. To avoid confusion, we will add a schematic figure of each catchment that outlines the main measurement locations (precipitation, streamflow, isotopes) and indicate how these data were used in the calibration.

Comment: Line 278: "In the HOAL catchment, the calibrated root-zone SAS shape parameter lower bound α0 = 0.14 indicated a strong preference for very young water through unsaturated root zone preferential flow pathways, suggesting that precipitation rapidly reached the stream with minimal mixing with stored water." While the correlation is clear, I'm not sure I understand the physical aspect of it. The unsaturated root zone is controlled by capillary forces if it is indeed unsaturated and stagnant, and by the connected paths of saturated areas (or preferential flows) under unsaturated flow. As such, the mechanism I can envision is that when starting with dry conditions and high infiltration, the invaded paths, which form the latter as preferential flows, must be occupied by younger water. However, under higher saturation, preferential flows are already established within the root zone; therefore, at lower infiltration rates, the discharge will consist of older water. Is this the physical mechanism suggested by this finding? If so, please refer to it; otherwise, what is the conclusion I need to draw from this finding, and how is it related to questions 1 and 2? This aspect is discussed in lines 364-372, where the difference between the clay soil of HOAL, which promote young water through the preferential flows, in contrast to the forest cover soil in Wüstebach which promote a more uniform infiltration, and therefore reduce the preferential flow effect and increases both the residence and the storage in the subsurface which is addressed later in the paper.

Response: We thank the reviewer for this comment. We agree that the physical interpretation of the calibrated α_0 value requires clarification. We will revise the text to make the connection to preferential flow mechanisms more explicit. Specifically, in the HOAL catchment the calibrated root-zone SAS shape parameter lower bound (α_0 = 0.14) indicates that under dry antecedent conditions with intense events, preferential flow paths are quickly activated, allowing younger water to reach the stream with limited mixing. In contrast, under wetter antecedent conditions, established preferential flow pathways facilitate more mixing compared to overland flow, leading to relatively older (older than recent precipitation) water contributions. This interpretation is consistent with the time-variable formulation of α (t) (Eq. 1), where the parameterization links soil wetness and precipitation intensity to the activation of preferential flows. We will also clarify that this finding directly relates to research question 1 (whether tracer data can identify preferential flow) and question 2 (how preferential flows influence transit time distributions).

Comment: Figure 6. Was there an attempt to relate the PDF (the derivative of the CDF) to the preferential flow specific length and time scales using the specific field data?

Response: No, we did not attempt to relate the PDF to preferential flow length and time scales using field data in this study, as this would require additional field measurements that go beyond the scope of the present work.

Comment: Line 316-320: In the spirit of my previous comment, the mechanism here is that it is solely controlled by the "pipeline of the preferential flows" within the groundwater, and the storage size represents the pipeline length and the extent or dominance of groundwater preferential flows? If this is the case, do we have indications for this for the HOAL and Wüstebach? Do we see that the first is more heterogeneous and therefore is more likely to lean towards preferential flows than the latter? Alternatively, it may stem from the difference in slope, where the first is steeper than the latter (as is clear from Figure 1 in Turk 2024 and Bogena 2015, respectively), which will exacerbate the inherent heterogeneity, leading to more preferential flows and shorter storage times for the HOAL case.

This is partially addressed in lines 345-350; however, the mechanism by which the dominance of preferential flows is related to the different conditions of each field is provided later in lines 364-372, favoring the first explanation. However, the second explanation is partially supported by the passive groundwater storage difference between these fields (lines 392-396). Are they both true? Equal?

Response: Our model formulation does not explicitly represent the physical "pipeline" of preferential groundwater flow, nor does it resolve heterogeneity or slope effects mechanistically. Rather, the sensitivity analysis with different passive storage volumes provides a phenomenological way of testing how much tracer damping could result from larger or smaller passive storage contributions.

We will clarify in the discussion that both mechanisms likely contribute: (i) soil/heterogeneity and slope control the activation and connectivity of preferential pathways, and (ii) passive groundwater storage controls the damping and persistence of tracer signals (and thus the residence times). Our data support the presence of both effects but do not allow us to quantify their relative contributions. We will revise the text to state this explicitly and to frame the two as complementary hypotheses rather than competing explanations, noting that disentangling them would require additional filed observations.

Comment: These comments on the mechanism of the preferential flows and their sources are reflected in the following (lines 412-416): "An alternative explanation, however, must also be considered: it is possible that such preferential groundwater flow processes are simply absent or negligible in the HOAL and Wüstebach catchments. The current data and model structure are insufficient to conclusively rule out either possibility. Ultimately, distinguishing between limitations in model sensitivity and the actual absence of preferential flow processes requires additional, spatially distributed tracer data and complementary hydrometric observations." As there are no hypotheses on the mechanism leading to preferential flows that stem from the physical aspects of the fields —permeability heterogeneity of the root area and groundwater, head differences within the groundwater stemming from the slope, and slope runoff on a local scale —one cannot draw a clear conclusion. These physical considerations are evident only at the correlational level from the data, as is clear from the Turk 2024 model, and not from a causality standpoint, which would require a more rigorous consideration. This is a major weakness of this study, and an attempt to formulate a hypothesis that will drive others to see what the actual causality is is missing. Nonetheless, this statistical approach is an important step toward a rigorous model, as it points to the conditions under which preferential flows can be more dominant. An example of how the statistical approach may point towards a more rigorous model is found in lines 430-438, where the distribution tail points to the ratio between active and passive storage.

Response: We thank the reviewer for this comment. We agree that our current framework cannot resolve the causal mechanisms of preferential flow processes, and that this represents a limitation of simpler, top-down catchment-scale, isotope-based transport models. We will expand the discussion (lines 412–416) to clarify this point and to formulate ideas that may guide future work. Specifically, we will note that (i) soil heterogeneity and slope likely govern the activation and connectivity of preferential flow pathways, (ii) differences in passive versus active storage modulate tracer damping and the apparent age distributions, and (iii) the relative importance of these processes remains unresolved with the current data.

"Our analysis does not allow us to establish causal mechanisms of preferential flow. However, based on previous field evidence, factors such as soil heterogeneity, slope, and passive storage volumes may play a role in modulating the strength of preferential flow signals. While these mechanisms remain hypotheses beyond the scope of top-down catchment-scale, isotope-based transport models, the SAS-based framework highlights conditions under which such effects may become detectable."

Comment: Discussion: I extremely like how the authors pose the three questions at the beginning of the paper, yet they only return to the first two at the end of the paper. Even if the conclusion is that it is inconclusive, this should appear for completeness.

Response: We thank the reviewer for this suggestion. We will revise the Discussion and Conclusions to explicitly return to all three research questions. For the third question, we will emphasize that passive storage volume and associated mixing assumptions introduced substantial variability in the estimation of transit time distributions, particularly in the long tails. However, the current data and model structure do not allow us to resolve this variability conclusively, and we will highlight this as a key limitation in catchment scale lumped transport model.

Line 399: add a space after groundwater.

Response: We will add a space after groundwater.