Editor Comments
Public justification (visible to the public if the article is accepted and published):
I have read your manuscript, the two reviews and your responses with great interest and pleasure. Personally, I like the attempt to include signatures of preferential flow in transit time modelling and I concur with both reviewers that your work is based on comprehensive data and sound modelling. Yet, both reviewers provided a long list of detailed points to optimize the presentation of your work, to make the underlying modeling approach more transpartent to the readers and to discuss causality behind the correlation. I think these points can be addressed within a round of major revisions, along the path you outlined in your very detailed and constructure responses. Looking forward to receiving the revised manuscript,

Authors' General Comment:

Dear Prof. Erwin Zehe,

We thank you for the assessment of our manuscript in EGUSPHERE-2025-2597 and for your comments. We are grateful for the time and effort invested in reviewing our manuscript and for the reviewers' detailed suggestions, which have helped us improve the clarity and transparency of our work. We appreciate your positive feedback regarding our attempt to include signatures of preferential flow in transit time modelling. We made several line edits to enhance the precision and consistency of terminology throughout the manuscript including key revisions sections as follow:

Abstract: Revised to clearly reflect the study objectives, hypothesis formulation, and key findings in a concise and structured manner.

Introduction: The study objectives and hypotheses have been reformulated to enhance focus and stepwise analyses

Methods: The description of the model setup, including additional figures, model equations, and details on the SAS formulation and parameterization, has been expanded.

Figures 2 and 3: Now updated as Figure 3 and 4, both figures have been reorganized to reduce redundancy and improve clarity in presenting the stepwise analyses implemented.

Discussion: We have revised the discussion sections to explicitly address all three research questions for completeness. A new section titled "Implications and Limitations" has been added to address causal interpretation of our results.

Thank you for considering our revised manuscript for publication in Hydrology and Earth System Sciences.

Sincerely,

Point-by-Point Authors' Responses to the Editor and Reviewers Comments.-

In the following, we provide our detailed **point-by-point** responses to each of the comments. Reviewer comments are shown in **black**, while our responses are shown in **blue**. Line numbers refer to the revised version with track-changes.

Reviewer #1

Comment: Thank you for the opportunity to review the manuscript entitled "Catchment transit time sensitivity to the type of SAS function for unsaturated zone and groundwater." The authors present an interesting study comparing model performance across different setups that vary in the parameters used in the SAS function. This approach aims to deepen the understanding of catchment functioning, specifically the contributions of water from the unsaturated zone and groundwater to catchment transit times in two catchments. Overall, the manuscript addresses an important research gap by challenging an assumption in transit time distribution (TTD) modeling and has the potential to make a valuable contribution. However, I believe that the manuscript in its current form requires major revisions before it can be considered for publication. My main concerns are as follows:

Response: We thank Reviewer #1 for the thorough evaluation and constructive feedback on our manuscript. Below, we outlined the issues raised by the reviewer and the corresponding changes we made in the revised version. Reviewer comments are shown in black, while our responses are provided in blue.

Comment: The language used throughout many sections of the manuscript is often vague and imprecise, which unfortunately leads to several hydrologically inaccurate statements. For example, there is some conflation between residence time distribution and transit time distribution (line 290), the SAS function is described as specifying if rather than how young or old water leaves the storage (line 6), and terms commonly well-defined in hydrology, such as information content (line 10) and "process", are used incorrectly. A thorough revision of terminology is necessary, particularly in the abstract and introduction, to clarify these points and improve the manuscript's overall precision.

Response: We thank the reviewer for this comment. We revised the manuscript to improve the precision and consistency of terminology throughout. Specifically, we ensured a clear distinction between residence time distributions and transit time distributions (now explicitly defined in Section~2.3), clarified that the SAS function describes how—rather than if—water of different ages leaves a storage, and revised the usage of terms such as information content and process to align with established hydrological definitions. These changes were implemented throughout the abstract, introduction, and methods sections to enhance clarity and terminological consistency.

Comment: The hypotheses as currently formulated are not effectively integrated into the manuscript. First, the connection between the hypotheses and the discussion is missing and could be strengthened by revising the discussion section, which currently has some weaknesses (see one of my later comments). Second, the hypotheses are not formulated in a testable manner, which limits their usefulness in framing the study.

Response: We revised the formulation of the study hypothesis to make them more precise and testable. The revised text now reads (Lines 5 - 8):

"In this study, we hypothesized that the preferential release of young water in both the unsaturated zone and groundwater contributes measurably to the streamflow tracer signal. This effect can be represented through StorAge Selection (SAS) functions, which specify how young or old water leaves a storage.

We also revised the discussion section to explicitly link the results back to these hypothesis.

Comment: Although the results section is generally well-written, the key messages are often difficult to discern. A clearer focus and synthesis of the main findings would greatly enhance the manuscript's impact.

Response: The Results section was revised to emphasize the key messages by adding short synthesis statements at the end of relevant subsections, the main findings and their relevance to the study objectives.

Comment: I am uncertain about the comparability of the different model setups as currently presented. The authors rely on common performance metrics such as NSE for comparison, but their stated hypotheses suggest an information theory or information content approach. If the authors choose to maintain the current comparison method, they should provide a more detailed explanation and justification for this choice. NSE also has certain limitations, which should be taken into account in this context.

Response: We thank the reviewer for this comment. We revised the hypothesis to focus on the variability in tracer and streamflow data rather than on an information theory framework. Consistent with this, tracer transport models are generally evaluated based on tracer fit. We therefore assessed model performance using NSE, MAE, and correlation coefficients (r values), which together capture both variability and error characteristics across scenarios, while each metric also has its limitations. We agree that NSE alone can be misleading, and we clarified this in the methods while justifying our choice of combining NSE, MAE, and r as a balanced evaluation approach.

In addition, we highlighted in the discussion that weekly-resolution δ^2H measurements can lead to deceptively high NSE values, even when key groundwater age-selection parameters (e.g., preference for young vs. old water) remain poorly constrained (lines 507- 512).

"The Spearman rank correlations (r) between observed and simulated δ^2H were lower in HOAL compared to Wüstebach, which can be attributed to differences in temporal resolution and the variability of isotope sampling (Fig. 1a, b). Although model performance metrics such as NSE or correlation coefficients quantify the agreement between simulated and observed isotope time series, they can result in seemingly good fits in the presence of sparse or irregular data sampling (Beven, 2006). In such cases, deceptively high NSE values may still occur even when key groundwater age-selection parameters (e.g., preference for young vs. old water) remain poorly constrained, thereby affecting transit time estimations (Stockinger et al., 2016)."

Comment: The discussion would benefit from a broader perspective. Addressing questions such as the following could improve the manuscript's relevance and clarity: What are the key findings for the research community? What are the broader implications of this work? How does this study advance process understanding in other, unstudied catchments? Under what conditions might one expect similar or different results?

Response: We thank the reviewer for this suggestion. We added a subsection on Implications and Limitations in the Discussion to provide a broader perspective (lines 550-554):

"While our study focused on a lumped catchment-scale framework, the results highlighted the need to advance toward more distributed models that can more directly link spatial heterogeneity in soils, slopes, and storage to preferential flow dynamics, and to advance process understanding. At the catchment scale, isotope-based modeling proved useful in capturing preferential flow in the unsaturated zone, but was limited doing so in groundwater due to the damping of the seasonal signal of water stable isotopes by large passive storage volumes assumed in the models. This suggests that in catchments with similarly damped water stable isotope signals, groundwater age selection will be difficult to constrain

Comment: While the authors refer to a previous study for a description of the model setups, I believe the manuscript would benefit from a brief summary of the underlying assumptions and structural details within the current text. This would enhance clarity for readers unfamiliar with the referenced work.

Response: A brief summary of the underlying model assumptions and structural details was added to improve clarity for readers unfamiliar with the referenced work. In addition, the relevant bucket-model equations for the water age balance (previously presented in Türk et al., 2024, Section "Integration of the rSAS function concept and the hydrological model") were included, together with a clearer description of how each bucket's tracer balance is formulated and parameterized within the SAS framework. (lines166-217)

Comment: Additionally, the configuration used to test passive groundwater storage requires further explanation. In particular, it would be helpful if the authors could elaborate on the rationale for choosing the specific storage volumes and clarify how mixing within this storage compartment was considered.

Response: We thank the reviewer for this comment. We expanded the description of the passive groundwater storage setup (Lines 306- 324) to clarify both the rationale for the selected volumes and how mixing was implemented. Specifically, we better explained in lines 308-314 that groundwater was represented as consisting of an active ($S_{s,a}$) and a passive ($S_{s,p}$) storage, where $S_{s,p}$ mixes isotopically with the active storage but does not contribute to runoff. The total storage (Ss,tot) therefore reflects both compartments, influencing the tracer signal of baseflow. We also clarified that (lines 315-316) the three passive storage volumes ($S_{s,p}$ = 500, 1000, and 5000 mm) were chosen to cover the range reported in comparable headwater catchments (Birkel et al., 2011; Benettin et al., 2015; Hrachowitz et al., 2021; Wang et al., 2023).

Some specific comments are provided below for consideration.

Comment: Title “sensitivity”: It is not clear whether the manuscript presents a sensitivity analysis. Please clarify if such an analysis was performed; otherwise, consider revising the title accordingly.

Title “type of SAS function”: Typically, SAS function types refer to, for example, gamma or beta functions. This may not be what you mean here please clarify the intended meaning.

Response:

The title was revised to better reflect the scope of the study. Because the analysis compared different SAS parameterizations rather than performing a formal sensitivity analysis, the term “sensitivity” was removed to avoid misinterpretation. We also clarified that “type of SAS function” refers to parameterizations favoring younger versus older water leaving storage, rather than to specific functional forms.

Revised title:

“Catchment transit time variability with different SAS function parameterizations for the unsaturated zone and groundwater”

Line 3: The terms fast and short are subjective and relative. Please avoid judgmental terms.

Response: We revised the sentence as to avoid judgmental terms :

Lines 1–3: “facilitate the rapid transmission of precipitation and solutes to streams, resulting in streamflow responses characterized by the release of young water (i.e., recent precipitation) from the catchment and correspondingly shorter transit time (on the order of days).”

Line 3: In the phrase “such preferential flow processes,” please clarify which specific processes from the previous statements are meant—do you mean preferential flow paths?

Responce: We revised the sentence and replaced “such preferential flow processes,” to” preferential flow paths” as:

Line 4 :“While preferential flow paths are documented in the both unsaturated zone and groundwater…”

Line 4: Please clarify the reference of the word “these.”

Responce: We removed “these” from the sentence.

Line 5: Use the term significant only where statistical significance (p-value) is reported; otherwise, rephrase.

Responce: The term “significant” was removed and replaced with “contributes measurably” to avoid implying statistical significance (Line 6).

Line 5: The statement “…by selecting specific SAS functions” seems self-evident because preferential discharge of young groundwater cannot be represented without selecting an SAS function. Please refine the hypothesis so that it is testable and cannot be answered with a simple “yes.”

Responce: We agreed with this comment and revised the hypothesis to make it more testable and conceptually focused. The revised text now reads (Lines 5-9):

“In this study, we hypothesized that the preferential release of young water in both the unsaturated zone and groundwater contributes measurably to the streamflow tracer signal. This effect can be represented through StorAge Selection (SAS) functions, which specify how young or old water leaves a storage.

Line 6: Replace "if" with "how."

Responce: We replaced "if" with "how."

Lines 6–8: It is unclear whether the functions were parameterized or if they were part of different model setups. Please clarify.

Responce: We agreed that the original wording was unclear and revised the sentence to explicitly indicate that the SAS functions—rather than different model setups—were parameterized. The revised text now reads (Lines 8-11):

"We systematically compared multiple parameterizations of the StorAge Selection (SAS) functions for the unsaturated zone and groundwater within a catchment-scale transport model using long-term measurements of hydrogen isotopes in water (δ^2H) from two headwater catchments: the Hydrological Open Air Laboratory (HOAL) in Austria and the Wüstebach catchment in Germany."

Line 10: The term "information content" implies the use of information theory; please clarify if that is the case.

Responce: Thank you for the comment. Our intention was to express that the δ^2H ratios in streamflow exhibited sufficient variability/sensitivity to identify preferential flow in the unsaturated zone. To clarify this, we revised the sentence as follows (Lines 11-12):

"The results indicated that δ^2H ratios in streamflow exhibited sufficient variability to confirm the preferential release of young water through preferential flow paths."

This clarification and consistent terminology were implemented throughout the manuscript to ensure accuracy and precision in language.

Lines 10–18: This section should be improved by clearly stating the main message of the results and explicitly explaining the relationship between age and specific SAS functions.

Response: We revised lines 10–18 to more clearly state the main message of the results and explicitly link the findings to the role of SAS functions. The revised text now read (Lines16-28):

" However, contrary to the unsaturated zone, the variability of δ^2H in streamflow was not sufficient to confirm the prefential release of young water from groundwater storage, as any seasonal variation of δ^2H in pore water was largely dampened by the catchments' substantial passive groundwater storage volumes with r values ranging between 0.54 and 0.60 in the HOAL, and ranging between 0.71 and 0.76 in the Wüstebach catchment. This interpretation was further supported by the fact that the observed attenuated δ^2H signal in streamflow could only be reproduced when the volume ratio between active and passive groundwater storage was less than 1 %, highlighting the dependence of SAS-based age selection on storage configuration. The damping effect, in combination with the groundwater SAS function parameterization influenced the estimation of the longer tails ($100 < T < 1000$ days) of the transit time distributions, making it challenging to quantify how much of the stream water is actually older than 100 days. The variability in streamflow TTD estimates arising from different groundwater SAS function shapes was considerable (±20 % for HOAL and ±23 % for Wüstebach), highlighting that TTD estimates are sensitive to how SAS functions are conceptualized and parameterized within the model. These findings underscore the need for complementary data sources, such as multiple tracers, high-frequency tracer analysis, and/or groundwater-level monitoring, to better constrain preferential flow processes and to reduce uncertainty in catchment-scale water transit time modelling."

Line 24: The term "dry period" depends on climatic context; please specify.

Response: We agree that the term "dry period" was too general and revised the sentence to specify that it refers to low-flow conditions resulting from below-average precipitation and/or reduced soil moisture availability. The revised sentence now reads (Lines 30-32):

"Groundwater plays a crucial role in the hydrological cycle and in sustaining streamflow during low-flow periods and influences the stream water age and quality (van der Velde et al., 2011; Hamilton, 2012; Kaandorp et al., 2018b)."

Line 25: Specify whether "reaching streams" refers to water after precipitation events or baseflow contributions.

Response: We revised the text (line 31) by replacing "quality of water reaching streams" also influencing the stream water age and quality."

Lines 28–30: As written, "this variability" (temporal) cannot be caused by spatial factors such as catchment topology. You may mean differences in temporal variability among catchments. Please reframe.

Response: We replaced "this variability" with "variation in flow timescales across catchments" (line 34) for clarity.

Line 33: The word "dramatically" is subjective; please remove.

Response: We removed the word "dramatically" from the sentence.

Line 38: A model, by definition, cannot detect anything: consider alternative terminology. Similarly, "quantify" may not be applicable in this context.

Response:(line 44) We revised the sentence to: "transport models can meaningfully represent preferential groundwater flow."

Lines 40–41: The terms "follow" and "system" are too vague: please specify.

Response: (lines 45-46) We revised the sentence for clarity. The revised version now reads:

"Water molecules entering at different locations within a catchment travel along distinct flow paths and exit the catchment at different times via streamflow or evaporation (transit time, TT)."

Line 43: Consider using "control volumes, such as catchments" instead of just "a catchment," since the approach could also be applied to a lysimeter or a stream reach.

Response: (line 48) We revised the text to read: "control volumes, such as catchments"

Line 46: A process cannot be quantified directly from a TTD; the TTD allows you to infer processes. Please check other occurrences where "process" is used in a similar way.

Response: We agree that a process cannot be quantified directly from a TTD. We revised the sentence to clarify (lines 51-52):

"Many studies have integrated hydrometeorological data and applied tracer-based modelling, using the transit time distribution (TTD) to infer flow processes and estimate transit times."

Line 48: Please clarify why "most" applies here.

Response: We removed the word "most."

Line 49: The message of this sentence is not sufficiently clear. Please rephrase.

Response: We rephrased the sentence for clarity as follows (lines 53 - 56):

"These studies have shown that streamflow typically consists of water from a broad spectrum of ages, with transit time distributions (TTDs) spanning from days to decades, thereby highlighting the importance of both rapid transmission of precipitation to the streams and its long-term storage in catchments."

Line 53: Keep terminology consistent. Use either function or model only when referring to different concepts. Here, it should be "SAS function." Also, note that the SAS function does not directly capture storage heterogeneity; please define precisely.

Response: We revised the sentence to (lines 59-61):

"The SAS function represents water age dynamics in hydrological systems by defining the relationship between the distribution of water ages stored within the system (residence time distribution, RTD) and the distribution of water ages leaving the system as outflows (transit time distribution, TTD)"

Line 55: Same adjustment as above. Use function instead of model if referring to the SAS function.

Response: We replaced (line 61) "model" with "function" when referring to the SAS function.

Line 57: Consider rephrasing, as it is unsurprising that time-variable TTDs reflect temporal TTD variability.

Response: We revised the sentence to (lines 63-65):

"Studies have shown that TTDs vary over time and that transport processes can differ under contrasting conditions, such as between wet and dry periods (Benettin et al., 2015b; Harman, 2015; Kaandorp et al., 2018a)."

Lines 59–63: Please re-check whether all cited studies actually applied SAS functions.

Response: We re-check (lines 65-70) the cited studies and ensure that only those that applied SAS functions are referenced in this context.

Lines 66–67: Please explain more clearly what is meant by "the age composition of groundwater flow to the stream."

Response: We revised the sentence (lines 72-73) to: "In many SAS function applications, the age distribution of baseflow (groundwater contribution to streamflow) is simplified by assuming uniform mixing of stored ages."

Line 69: SAS functions cannot be "measured"; you can parameterize them. Please revise.

Response: We revise the sentence (lines 74-75) to: "Noting that SAS functions are not straightforward to parameterize."

Line 79: Specify what the "release of young water" refers to.

Response: We revised the sentence (lines 83-85) to:

"Indeed, increasing evidence suggests that groundwater systems may not be completely mixed, and that the preferential release of relatively young water (recently recharged water) to streams may be a ubiquitous feature of groundwater in heterogeneous aquifers."

Line 80: Indicate where the "generally low longitudinal and transversal dispersivities" apply.

Response: We revised the sentence (line 86) to:

"… (ii) generally low longitudinal and transversal dispersivities in groundwater systems, leading to little mixing,"

Lines 81–83: This sentence appears disconnected from the previous one; please rephrase for cohesion.

Response: We rephrased (lines 88-91) the sentence for improved cohesion with the preceding text:

"This evidence suggests that groundwater systems are often not completely mixed, and the preferential release of young water may be common in heterogeneous aquifers. Therefore, SAS functions should be formulated to account for preferential release of young water and the nonlinearities in groundwater contributions and catchment responses."

Line 93: Clarify whether "long-term tracer observations" were conducted in streams or in groundwater.

Response: We updated the sentence to "long-term tracer observations in streamflow."

Line 96: The main objective could be presented in a way that connects more clearly to the underlying processes of interest rather than focusing solely on technical aspects.

Response: We revised the objective to more clearly emphasized the underlying hydrological processes of interest. The revised text now reads (lines 104-111):

"The objective of this study was to test whether stable water isotope δ^2H measurements in streamflow can be used within a simple, top-down, catchment-scale transport modeling framework to represent preferential flow in the unsaturated root zone and groundwater, and to assess their influence on transit time distributions. To this end, we evaluated how different parameterizations of SAS functions, which describe the release of younger versus older water from storage, affect simulated tracer signals at the stream outlet. By systematically comparing multiple parameterizations of SAS functions, we tested the hypothesis that the preferential release of young groundwater contributes measurably to the streamflow tracer signal and should therefore be represented in catchment-scale transport models. Additionally, we examined whether (and how) the extent and mixing assumptions of passive groundwater storage influence the interpretation of tracer signals and the estimation of transit times."

Line 103: If the term "information content" is used, ensure that it is correct. If information theory was not applied, please rephrase.

Response: We replaced "information content" with "variability" in line 112 to avoid misinterpretation.

Line 107: Please define what is meant by "interpretation" in this context.

Response: We clarified the wording by replacing "interpretation" with a more precise term: (line 116-117).

"Does explicitly accounting for the preferential release of young water in groundwater, through a SAS function, affect catchment-scale transit time distributions and simulated tracer signals in streamflow?"

Line 108: Clarify the meaning of "representation of preferential groundwater flow."

Response: We revised the research question to avoid the vague term "representation". The revised question now read (lines 118-120):

"How and to what extent do different groundwater mixing assumptions, in combination with varying passive storage volumes, affect the model's ability to reproduce streamflow tracer signals and the estimation of transit time distributions at the catchment scale?"

Line 122: Please state in which catchment the "predominant soil types" are found.

Response: We clarified this by specifying that the predominant soil types are found in the HOAL catchment (line 132).

Line 135: The term "catchment flow" is not defined. Do you mean streamflow?

Response: Yes, we mean streamflow, and we revised the text accordingly (line 145).

Line 160: It is important to briefly explain the underlying assumptions and general model setup, even if they are covered in a previous publication. Currently, it is unclear how the SAS functions are integrated into the model.

Response: We revised the sentence as follows to explain the underlying assumptions (lines 178-186):

"To route δ^2H fluxes through the model, we integrated the SAS approach (Rinaldo et al., 2015; Harman, 2015) into the hydrological model. In In this integrated framework, each storage defined within the hydrological model (e.g., S_r, S_f; Fig. 2) stores water of different ages at any given time t. These ages, denoted as T, trace back to past precipitation events and are ranked according to their input time. The age distribution of a storage at time t is referred to as $p_s(T, t)$, and its cumulative form, $S_T(T, t)$, represents the cumulative residence time distribution (RTD). The output fluxes O (mm d^{-1}) (e.g., Ea, Rs; Fig. 2) consist of subsets of water of specific ages drawn from the storage. Their age distributions are represented as $p_{Q,T}(T, t)$, and their cumulative form, $O_T(T, t)$, corresponds to the cumulative transit time distribution (TTD). The relationship between storage and output fluxes is defined by the SAS function ωO,m,j, which specifies the likelihood of selecting water parcels of different ages for release from storage, thereby translating the internal age structure of the storage into an age distribution of the output fluxes."

We added a description in the Methods section 2.3 explaining how each bucket's tracer balance is formulated and parameterized within the SAS framework (lines 186-217).

Lines 178–184: The SAS function is generally time-variable. Please clarify the novelty of the approach described. Explain why the precipitation value was set as it was and define Sr,max.

Response: We added that the dual dependence of α(t) on both soil moisture and precipitation intensity extends previous SAS applications by providing a more flexible representation of preferential flow dynamics in the unsaturated zone of the HOAL catchment. We also explicitly defined Sr,max as the calibrated maximum root-zone storage capacity, which controls the scaling of relative soil wetness in the α(t) formulation. In addition, the precipitation threshold (P_thresh) was clarified as a calibration parameter representing the intensity above which infiltration-excess preferential flow pathways are activated. The revised section now reads (lines 238-250):

"In the HOAL catchment, previous studies have highlighted the non-linearity of preferential flow generation in the unsaturated zone, where both precipitation intensity and soil wetness control the activation of preferential flow pathways (Széles et al., 2020; Vreugdenhil et al., 2022). Overland flow occurs when precipitation exceeds a certain threshold, routing recent precipitation directly to the stream with minimal interaction with stored water (Türk et al., 2024). To account for the combined roles of soil wetness and precipitation intensity in the activation of preferential flow and the release of young water in HOAL catchment, we parameterized the SAS function for preferential flow from the unsaturated root zone (Rf , Fig. 2) using a time-variable shape parameter α(t) defined as a function of both soil wetness state and precipitation intensity. Specifically, α(t) was formulated as a function of relative soil wetness (scaled by the maximum root-zone storage capacity, Sr,max) and precipitation intensity (PI , mm d−1), with a threshold parameter (P_{thresh}) controlling the onset of precipitation-driven preferential flow. This causal formulation was implemented to ensure that α(t) dynamically responds to both wetness conditions and event-scale precipitation forcing, allowing the model to capture the non-linear activation of preferential flow observed in the catchment. The dual dependence of α(t) on soil wetness and precipitation intensity, therefore, extends previous SAS applications by providing a more flexible representation of unsaturated zone preferential flow dynamics in HOAL."

Lines 195–199: Please explain why streamflow, log streamflow, flow duration curve, runoff coefficient, and δ^2H were selected as variables.

Response: We revised the text to clarify the rationale for selecting these performance metrics. Specifically, the original sentence was revised as follows (lines 266-272):

“For model parameter optimization, we used the Differential Evolution algorithm (Storn and Price, 1997) and an objective function that combined five performance criteria related to streamflow and δ^2H dynamics. The objective function included the Nash-Sutcliffe efficiencies (NSE) of streamflow (to evaluate overall discharge dynamics), logarithmic streamflow (to match low-flow conditions), the flow duration curve (to capture the distribution of flows over time), the runoff coefficient averaged over three months (to ensure water balance consistency), and the NSE of the δ^2H signal in streamflow (to constrain tracer dynamics) (Table S2). These individual performance metrics were aggregated into the Euclidean distance DE, with equal weights assigned to streamflow and the δ^2H signature, according to:”

Line 199: Avoid subjective terms such as “perfect.”

Response: We removed the term “perfect” from the sentence line 275

Line 203: Clarify what is meant by “each combination.”

Response: We removed the term “each combination” from the sentence for clarity line 275.

Line 211: The need for stepwise calibration and the exact order of steps should be made explicit.

Response: To clarify, we did not apply a stepwise calibration but instead used one calibration of the hydrological parameters. The stepwise analysis was then performed afterwards by varying the SAS function parameters (α_0 for the unsaturated root zone and α for the groundwater compartments) while keeping the calibrated hydrological parameters fixed. We revised the text (lines 301-305) accordingly to avoid confusion.

Lines 218–221: This information may be better placed earlier in the manuscript for clarity.

Response: We moved this information earlier in the manuscript to improve clarity. (lines 285-300)

Lines 228–237: This section should more clearly describe how different passive storage volumes were implemented in the earlier-described model setup and why these volumes were selected. As it stands, the model configuration is not fully reproducible.

Response: We revised the section to include the following clarification (lines 309-314):

“In the model setup, groundwater storage was represented by an active storage component ($S_{s,a}$) and a hydrologically passive storage component ($S_{s,p}$, mm). The passive storage $S_{s,p}$ does not contribute directly to the quantity of baseflow, but it isotopically mixes with water of the active storage. For the SAS function, the total groundwater storage was therefore defined as the sum of $S_{s,a}$ and $S_{s,p}$ ($S_{s,tot} = S_{s,a} + S_{s,p}$). As a result, the age-ranked groundwater storage ($S_{s,tot}$) inherently reflects a mixture of these storage volumes in the age composition of baseflow (Q_s, Fig. 2).

We also justified the choice of passive storage volumes as follows (lines 315-317):

“We applied three different passive storage volumes ($S_{s,p}$ = 500 mm, 1000 mm, and 5000 mm) to cover the ranges reported for comparable headwater catchments (Birkel et al., 2011a; Benettin et al., 2015a; Hrachowitz et al., 2021)”

Lines 240–246: Please summarize the key takeaway for the reader.

Response: We added a summary sentence at the end of the section to highlight the main message (lines334 -339):

"Overall, δ^2H in precipitation exhibited large variability in both catchments; however, this signal was attenuated in streamflow. In the HOAL catchment, event-based streamflow δ^2H samples reflected how precipitation inputs were rapidly transmitted to the stream, whereas weekly samples alone would have masked this variability. This highlights the importance of event-based sampling for detecting preferential flow signals, which may remain obscured with weekly data alone. This applies to the Wüstebach catchment, where only weekly streamflow δ^2H measurements were available, which may have prevented the detection of such rapid responses."

Line 248: Define what is meant by "feasible parameter solutions."

Response: We revised the sentence to clarify that it meant parameter solutions with acceptable model performance (DE < 1) (line 341-342):

"Model calibration resulted in 55 feasible (acceptable model performance with DE < 1, Fig. S2) parameter solutions) for HOAL and 190 feasible parameter solutions for Wüstebach."

Lines 266–276: Clarify whether these results relate to the stepwise analysis, and connect them to Figure 4. The calculated fractions shown in Figure 4 should also be introduced earlier.

Responce: We clarified that these results are not related to the stepwise analysis but are based on the initial model calibration conducted prior to the stepwise experiments. We also introduced the calculated fractions earlier in the text and explicitly linked them to Figure 4. In addition, the Methods section (Section 2.3.1) was updated to describe how these fractions were calculated to ensure clarity and reproducibility. Specifically, we added (lines 281-283):

"In addition, the young-water fraction of daily streamflow ($F_Q(T < 90)$) was calculated as the sum of streamflow fractions with transit times up to 90 days. Its monthly variability was then analyzed in relation to the corresponding monthly variability of soil wetness (Sr/Sr,max) for both catchments."

Furthermore, the Results section was revised to begin with (lines 359-362):

"Figure 5 presents the transit time distributions (TTDs) estimated from the initial model calibration, conducted prior to the stepwise experiments. For TTD estimations, we used the model-calibrated parameter set that yielded the lowest DE. The results presented hereafter are conditional on the underlying model assumptions and should be interpreted in light of the associated uncertainties."

Line 290: The explanation ("This was due to…") belongs to the Discussion. Also, confirm whether "residence time" here should be "transit time," and note that this is the first mention of the term in the manuscript.

Responce: We revised the sentence accordingly by removing the explanatory text and rephrasing it as (line 398-399):

"This is consistent with root-zone storage residence times being predominantly shorter than 300 days."

In addition, we added a definition of residence time at its first occurrence in the Introduction (line 61):"The distribution of water ages stored within the system (residence time distribution, RTD)."

Line 313: Remove the subjective phrase "and somewhat surprisingly."

Response: We removed the phrase "and somewhat surprisingly" from the sentence line 421 .

Line 315: Specify which variability is being reduced—e.g., variability in streamflow?

Response: We revised the sentence to clarify that it refers to the variability in the δ²H signal in streamflow (line 423).

Lines 351–354: Consistency with previous findings does not necessarily justify the model configurations or research hypotheses; identical results can occur in the presence of shared erroneous assumptions. Please refine this reasoning.

Response: We revised the reasoning to emphasize that consistency with previous studies provides support but not proof of correctness. The revised text now read (lines 459-464):

"The consistency of our results with prior tracer-based modeling and SAS applications in both HOAL (Széles et al., 2020; Türk et al., 2024) and Wüstebach (Stockinger et al., 2019; Hrachowitz et al., 2021) provides additional confidence in the applied model configurations. However, we acknowledge that such consistency alone cannot exclude the possibility of shared assumptions. Therefore, we use these results as supporting evidence that the model setups are reasonable for testing the research hypotheses, while recognizing the need for further validation with complementary data and approaches."

Lines 357–364: Please state clearly what the novelty of this section is.

Response: We revised the text to highlight the novelty of this section by adding the following sentence at the end (lines 473-476):

"The results, quantitatively demonstrated through the SAS-based modeling framework, show that streamflow isotope data can reflect the activation of preferential flow in the unsaturated zone. While previous studies have identified such processes through field observations (Vreugdenhil et al., 2022; Pavlin et al., 2021; Wiekenkamp et al.,2016), our results demonstrate that they can also be captured and interpreted using catchment-scale tracer modeling."

Line 369: Clarify whether the described crust formation was observed directly, or if it is inferred.

Response: We revised the sentence to clarify that the crust formation was observed (line482-484):

"…formation of soil crusts cracking of the clay-rich topsoil during the dry summer months, creating direct preferential pathways that accelerate water transmission through the catchment (Exner-Kittridge et al., 2016)."

Line 373: Make the link to the previous sentence explicit.

Response: We rephrased the sentence to make the connection explicit (lines 487-490):

"Despite contrasting site characteristics, both catchments showed responses consistent with previous studies that documented the role of macropores and preferential flow pathways in the unsaturated zone, where water frequently bypasses matrix storage and exchange processes (Zehe et al., 2006; Angermann et al., 2017; Sprenger et al., 2016; Klaus et al., 2013; Loritz et al., 2017)."

Line 376: Clarify who assumes this.

Response: We thank the reviewer for this comment. We removed the sentence to avoid the ambiguous term "assume".

Lines 419–444: There is considerable repetition of the results in this section. Consider condensing.

Response: We agree with this comment and condensed the section accordingly lines 514-523

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 addressed the concerns by providing additional background material in the Supplement. Below, we outlined issues raised by the reviewer and the corresponding changes we made. Reviewer comments are shown in black, and our responses are shown in blue. Specifically, we (i) briefly summarized the relevant components of the model setup from Türk (2024), including the bucket ratios and key equations in method Section (line 178-217). We also added figures S1 and S2 adapted from Bogena et al. (2015) and Türk (2024) to illustrate the field layout and its relationship to the bucket model in the Supplementary Information. These additions ensure that the manuscript can be read independently, without requiring the reader to consult 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. We added a short discussion section (Implications, Limitations, and Outlook) linking literature-based findings on soil texture, moisture dependence, and slope effects for Wüstebach, along with additional references for HOAL (lines 556-567). While our model does not explicitly test mechanistic relationships, highlighting these connections helps to clarify how the SAS shape parameter may reflect underlying physical controls and points toward directions for future mechanistic research.

Specifically, we added the following lines 556-567:

"While the SAS formulation identifies the statistical signatures of preferential flow through tracer-based modeling, it does not explicitly resolve the physical mechanisms that induce preferential flow. Therefore, isotope data alone, when used within a catchment-scale lumped model framework, may be insufficient to distinguish between a true absence of preferential flow and a limited model sensitivity to detect it. In the

HOAL catchment, Exner-Kittridge et al. (2016) showed that alternating contributions from shallow and deep aquifers throughout the year were the main cause of the seasonal variability in nitrate concentrations in streamflow. These alternating contributions, together with extensive tile drainage and heterogeneous clay-rich soils, create rapid and spatially variable flow pathways (Exner-Kittridge et al., 2016; Pavlin et al., 2021). Such features, combined with overland flow of the HOAL catchment (Blöschl et al., 2016), may favor the activation of preferential flow. For the Wüstebach catchment, field studies indicated that soil and groundwater dynamics are coupled. Bogena et al. (2014) and Graf et al. (2014) showed that soil water variability decreases with depth due to lower porosity and root water uptake, while groundwater fluctuations closely follow soil moisture dynamics (Bogena et al., 2014), reflecting high infiltration and storage capacity in forest soils.

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 bucket 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 added 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.

We added a discussion paragraph (lines 568–577) to further emphasize the limitations of catchment-scale models.

"Although these processes remain beyond the explicit resolution of top-down, isotope-based transport models, the SAS framework enables delineation of the hydrological conditions under which preferential flow effects become detectable. Future research should combine stable isotopes with complementary tracers (e.g., tritium, chloride, or major ions) and higher-frequency sampling to enhance the diagnostic power of tracer-aided models. Furthermore, linking SAS function shapes to measurable catchment attributes could enable a priori parameterization, thereby reducing dependence on calibration. At the lysimeter pedon scale, Asadollahi et al. (2020) showed that SAS functions can approximate the analytical solution of the advection–dispersion equation; however, extending such mechanistic relationships to the catchment scale remains challenging within lumped "bucket" model frameworks. Addressing these challenges represents a key step toward integrating empirical SAS modeling with a process-based understanding of preferential flow and subsurface mixing"

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

Response: We revised 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 added maps of both catchments showing the stream network and the locations of the sampling points to a supplementary Figure S1 and S2.

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

Response: We corrected "Tr" to "TT" in Table S1. In addition, we clarified that TT refers to the threshold temperature for snowmelt, which is a calibrated parameter. We also added a description of all parameters in black font to the equation table (Table S2) for clarity.

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 corrected the label to Ss,p and carefully checked Table S1 to ensure that all symbols used in the figures and text are consistently defined.

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: α and β are not included in Figure S1 because they are part of the transport formulation rather than the hydrological model. Their definitions are provided in the Table S1. We revised the text to include explicit definitions of α and β. Specifically, we clarified that β is fixed at 1 for all fluxes, while α is fixed at 1 for preferential groundwater flow but is calibrated for unsaturated zone preferential flow (SU,α) for the model calibration.

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 moved 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 clarified that Sr,max is a calibrated parameter representing the maximum root-zone storage capacity, rather than porosity. Accordingly, the SAS function shape parameter representing preferential flow from the unsaturated root zone (Rf, Fig. 3) was formulated as a time-variable function of relative soil wetness (Sr/Sr,max), where Sr is the water volume in the root zone at time t, and Sr,max is the calibrated maximum root-zone storage capacity. Lines (233 -237) now read as:

"The SAS function shape parameter representing preferential flow from the unsaturated root zone (Rf , Fig. 2) was formulated as a time-variable function of relative soil wetness (Sr /Sr,max), where Sr is the water volume in the root zone at time t, and Sr,max is the maximum root-zone storage capacity (calibrated parameter). The Equation 8 adopts an increasing probability of young water release with increasing soil wetness through the time-dependent shape parameter α(t), reflecting changes in transport processes between wet and dry soil conditions"

For the HOAL catchment, Equation (1) further includes precipitation intensity as a second controlling variable for the SAS shape parameter. This dual dependence on soil wetness and precipitation intensity reflects the observed nonlinearity of preferential flow activation, where infiltration-excess conditions during high-intensity rainfall events rapidly trigger preferential pathways, even under relatively dry antecedent soil states.

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 acknowledge that preferential flow is influenced by multiple physical controls, including soil heterogeneity and hydraulic head differences, in addition to soil wetness. In our model, these processes are represented conceptually rather than mechanistically. Therefore, the SAS formulation can identify if and whether preferential release of young water occurs but does not resolve the underlying physical mechanisms that generate it. We have added this clarification to the Discussion section (lines 556- 557) to emphasize the conceptual nature of the SAS-based approach and to delineate its limitations in capturing the physical drivers of preferential flow.

"While the SAS formulation identifies the statistical signatures of preferential flow through tracer-based modeling, it does not explicitly resolve the physical mechanisms that induce preferential flow."

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 clarified in the text and figure caption that eCDF refers to the empirical cumulative distribution function. Empirical cumulative distributions were selected because they allow clearer visualization of individual curves, and for scenario comparisons, we found cumulative curves easier to interpret than probability density functions (PDFs).

Comment: A non-intuitive aspect is that preferential release of older water (α>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 clarified in the Methods section that preferential release of older water (α > 1) leading to a stronger weighting of older water contributions (lines 222–226). We also added the following explanation to the text:

"The time variability of the SAS function shape was determined by the age-ranked storage and the shape parameter α, which was bounded between 0 and 1 to represent a preference for younger storage, and greater than 1 to represent a preference for older storage. Preferential release of older water (α > 1) decreases the mean residence time of stored water, as older water is removed from storage. Conversely, preferential release of younger water (α < 1) increases the mean residence time of stored water, as older water remains stored for longer periods."

Additionally, we expanded Section 2.3 of the Methods to describe how each bucket's tracer balance is formulated and parameterized within the SAS framework, provided explanation of how tracer dynamics are represented through SAS parameterization (lines 174-217).

To address the second point, we streamlined the figure3 and 4 to avoid redundancy. Specifically, we retained only the essential illustrations and clarified in the captions what information each figure is intended to convey (Figure 3). For passive storage values, we highlighted in the Figure 4 how varying $S_{s,p}$ affects tracer damping of tracer signal in streamflow.

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 clarified in the text that the 15 parameters listed in Table S1 were calibrated separately for each catchment. The numbers 55 (for HOAL) and 190 (for Wüstebach) refer to the parameter sets that met the calibration performance criterion (DE < 1), rather than to the number of calibrated parameters. In addition, we added Figure S1 outlining the measurements in the catchment area.

Lines 341 -342

"Model calibration resulted in 55 feasible (acceptable model performance where DE < 1) parameter solutions for HOAL (Fig.S 3) and 190 feasible parameter solutions for the Wüstebach catchment (Fig. S 4)."

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 revised 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 added lines 375-381

"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. These reflected the SAS formulation (Eq. 7) on the dual dependence of α(t) on soil wetness and precipitation intensity. Under high-intensity precipitation, α(t) takes a value of 0.14, indicating that rapid activation of preferential pathways occurs, allowing precipitation inputs to reach the stream with minimal mixing with stored water. In contrast, under wetter antecedent conditions, α(t) increases toward 1, indicating greater mixing within the root zone and contributions of relatively older (i.e., older than recent precipitation inputs) water to streamflow."

We also clarified 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).

Lines 480-487:

"In HOAL, the low α_0 value indicated rapid, direct water transmission through preferential flow paths driven by intense rainfall, consistent with previous hydrometric analyses and field observations (Pavlin et al., 2021; Vreugdenhil et al.,2022), which allow young water to reach the stream with limited mixing. This was further facilitated by the formation of soil crusts and cracking of the clay-rich topsoil during the dry summer months, creating direct preferential pathways that accelerate water transmission through the catchment (Exner-Kittridge et al., 2016). In contrast, the higher α_0 value in Wüstebach480 (α_0 = 0.98) indicated that under wetter antecedent conditions, established preferential flow pathways promoted greater sub surface mixing, leading to relatively older water contributions. This likely reflects the influence of forest cover in Wüstebach, where enhanced infiltration promotes deeper and more uniform mixing (Wiekenkamp et al., 2016) than in HOAL."

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?

Do we see that the first is more heterogeneous and therefore is more likely to lean towards preferential flows than the latter?

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. Therefore, the SAS formulation can identify if and whether preferential release of young water occurs but does not resolve the underlying physical mechanisms that generate it.

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 generally represents a limitation of simpler, top-down catchment-scale, isotope-based transport models. We agree that our study primarily identifies preferential flow behavior from tracer-based modeling and does not explicitly resolve the underlying physical mechanisms that control it. We clarified in the Discussion that our findings should be interpreted as statistical rather than causal. However, we now also provide a hypothesis to guide future work, suggesting that the observed preferential flow behavior may result from a combination of (i) soil heterogeneity and macroporosity in the unsaturated zone, (ii) slope-driven hydraulic gradients that enhance lateral subsurface connectivity, and (iii) differences in active and passive storage volumes that control tracer damping. This addition links the SAS-based statistical interpretation to potential physical processes that can be investigated in future studies.

We expanded to clarify this point and to formulate ideas that may guide future work.

(Lines 556-577)

"While the SAS formulation identifies the statistical signatures of preferential flow through tracer-based modeling, it does not explicitly resolve the physical mechanisms that induce preferential flow. Therefore, isotope data alone, when used within a catchment-scale lumped model framework, may be insufficient to distinguish between a true absence of preferential flow and a limited model sensitivity to detect it. In the HOAL catchment Exner-Kittridge et al. (2016) showed that alternating contributions from shallow and deep aquifers throughout the year were the main cause of the seasonal variability in nitrate concentrations in streamflow. These alternating contributions, together with extensive tile drainage and heterogeneous clay-rich soils, create rapid and spatially variable flow pathways (Exner-Kittridge et al., 2016; Pavlin et al., 2021). Such features, combined with the steeper topography of the HOAL catchment (Blöschl et al., 2016), may favor the activation of preferential flow. For the Wüstebach catchment, field studies indicated that soil and groundwater dynamics are coupled. Bogena et al. (2015) and Graf et al. (2014) showed that soil water variability decreases with depth due to lower porosity and root water uptake, while groundwater fluctuations closely follow soil moisture dynamics (Bogena et al., 2015), reflecting high infiltration and storage capacity in forest soils. This behavior contrasts with the results obtained in the HOAL catchment, suggesting a more uniform subsurface mixing in Wüstebach. Although these processes remain beyond the explicit resolution of top-down, isotope-based transport models, the SAS framework enables delineation of the hydrological conditions under which preferential flow effects become detectable. The current isotope data and statistical model structure are insufficient to distinguish between a true lack of preferential flow and limited model sensitivity to detect it. Future research should combine stable isotopes with complementary tracers (e.g., tritium, chloride, or major ions) and higher-frequency sampling to enhance the diagnostic power of tracer-aided models. Furthermore, linking SAS function shapes to measurable catchment attributes could enable a priori parameterization, thereby reducing dependence on calibration. At lysimeter pedon scale, Asadollahi et al. (2020) showed that SAS functions can approximate the analytical solution of the advection–dispersion equation; however, extending such mechanistic relationships to the catchment scale remains challenging within lumped "bucket" model frameworks. Addressing these challenges represents a key step toward integrating empirical SAS modeling with a process-based understanding of preferential flow and subsurface mixing."

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 have revised the Discussion and Conclusions sections to explicitly address all three research questions for completeness. For the third question, we now emphasized that passive storage volume and associated mixing assumptions introduce substantial variability in the estimation of transit time distributions, particularly in their longer tails (Lines 525-542). However, the available isotope data and the current lumped model structure do not allow this variability to be resolved conclusively. We now highlight this as a key limitation of catchment-scale, tracer-aided transport modeling (Lines 544-577).

Line 399: add a space after groundwater.

Response: We added a space after groundwater.

Added References:

Asadollahi, M., Stumpp, C., Rinaldo, A., and Benettin, P.: Transport and water age dynamics in soils: A comparative study of spatially integrated and spatially explicit models, Water Resour. Res., 56 e2019WR025539, https://doi.org/10.1029/2019WR025539, 2020.

Blöschl, G., Blaschke, A. P., Broer, M., Bucher, C., Carr, G., Chen, X., Eder, A., Exner-Kittridge, M., Farnleitner, A., Flores-Orozco, A., Haas, P., Hogan, P., Kazemi Amiri, A., Oismüller, M., Parajka, J., Silasari, R., Stadler, P., Strauss, P., Vreugdenhil, M., and Zessner, M.: The Hydrological Open Air Laboratory (HOAL) in Petzenkirchen: a hypothesis-driven observatory, Hydrology and Earth SystemSciences, 20, 227–255, https://doi.org/10.5194/hess-20-227-2016, 2016.

Bogena, H. R., Bol, R., Borchard, N., Brüggemann, N., Diekkrüger, B., Drüe, C., Groh, J., Gottselig, N., Huisman, J. A., Lücke, A., et al.: A terrestrial observatory approach to the integrated investigation of the effects of deforestation on water, energy, and matter fluxes, Science China Earth Sciences, 58, 61–75, 2015.

Exner-Kittridge, M., Strauss, P., Blöschl, G., Eder, A., Saracevic, E., and Zessner, M.: The seasonal dynamics of the stream sources and input flow paths of water and nitrogen of an Austrian headwater agricultural catchment, Science of the Total Environment, 542, 935–945, https://doi.org/10.1016/j.scitotenv.2015.10.151, 2016.