First, we note that a fourth referee (first round of reviews) provided a *highly positive appraisal* of the manuscript and recommended only minor revisions which we addressed previously. No further review comments were uploaded by this referee.

Response to Referee #1, Report #3:

Nothing to suggest. Please see the recommendation to the editor

RESPONSE: We thank the referee for the positive recommendation for publication.

Response to Referee #2, Report #1:

The central claim of this paper, as stated in the abstract, is that "For both water isotopes and inert chemical tracers, the measured mean tracer velocity is not necessarily equal to the apparent mean water velocity". However it does not make sense to claim both of the following statements are true simultaneously:

1. The mean velocity of the tracer v_T is different from the mean velocity of the water v_w
2. The tracer is water molecules or moves identically to water molecules

The only way I can make sense of this paper is if v_T is not the mean tracer velocity, but is rather a scaling parameter of the CTRW model that has dimensions of velocity (but is not the mean). However during our review discussion the authors rejected this suggestion, and claim that v_T is in fact the mean tracer velocity. Therefore I cannot see how this paper is conceptually coherent.

RESPONSE: The referee appears to have accepted our responses and clarifications to all but one of the comments raised in the two previous replies. Here, the referee first notes a central clear claim ("…the measured mean tracer velocity is not necessarily equal to the ***apparent*** mean water velocity") and then lists two points that are claimed to be contradictory. But in listing these two points, the referee ignores the preceding sentence with the "central claim". We do *not* claim statement #2 above – rather, we claim that the average velocity of the ***tagged*** water molecules (i.e., the water isotopes) is not identical to the ***apparent*** mean water velocity.

At this point, the referee focuses on movement of water molecules, whether pure H_2O or deuterium isotopes. The referee misses the key recognition, as we state clearly in the manuscript, that the very act of tagging water molecules yields a measurement representative of their migration as a chemical tracer. This is stated clearly, for example, in
(i) Section 3.2, below Figure 2 describing the Gedanken experiment: "The act of tagging water effectively changes a water molecule to act as a "non-water" tracer, in the context of breakthrough measurements; in other words, the measured velocity represents the mean velocity of a tracer (be it a water isotope or an inert chemical), rather than the apparent mean velocity of the water."
and

(ii) Conclusions: “Notably, water isotopes exhibit the same transport behavior as tracers; the very act of tagging water molecules, implicit in the measurement of any water isotope, yields a measurement of their migration as a chemical tracer, which is not identical to the bulk water flow.”
DONE: In the revised manuscript, we have also noted this point in the Abstract, to add further clarity.

Response to Referee #3, Report #2:

Overall Evaluation

This manuscript presents a carefully designed laboratory study that addresses a fundamental and widely held assumption in catchment hydrology: that water isotopes (e.g., D_2O) behave identically to the bulk water molecules (H_2O) and can thus be used directly to infer the mean water velocity and transit times. The experimental work is robust, comparing breakthrough curves of deuterium-enriched water and bromide tracer under various flow conditions and porous media configurations. The key finding—that water isotopes exhibit transport behavior indistinguishable from that of an inert chemical tracer, leading to a mean tracer velocity ($\bar{v}T$) that is measurably slower than the apparent mean water velocity ($\bar{v}w$)—is significant and has direct implications for the interpretation of field-scale isotope data.

The use of the CTRW framework to quantify the non-Fickian transport and the clear demonstration of the inequality $\bar{v}w \neq \bar{v}T$ are particular strengths. The discussion linking these laboratory findings to potential overestimations of aquifer storage thickness in catchment studies is timely and relevant. The manuscript is generally well-written and structured. However, to strengthen the impact and clarity of the work, several aspects require further elaboration and clarification.

RESPONSE: We thank the referee for the positive evaluation and recommendation for publication. We appreciate the highly constructive comments, which we address below point-by-point. We are pleased to incorporate additional explanations for all comments in the revised manuscript.

Specific Comments

(1) The manuscript rightly concludes that "the very act of tagging water molecules... yields a measurement of their migration as a chemical tracer." This is a crucial point. It should be explicitly discussed whether and how the rapid equilibrium isotope exchange between H_2O, HDO, and D_2O (mentioned in the methods) influences this "tagging" concept. Does this equilibrium exchange mean that the "tag" is effectively transferred between molecules, potentially altering the perceived transport behavior? A brief discussion on how this isotopic exchange is accounted for in the interpretation of the D_2O breakthrough curves would strengthen the argument that it behaves as a conservative solute and not as a perfect proxy for the bulk water movement.

RESPONSE: We explain in the Methods section, last paragraph, that "In excess of H_2O, D_2O rapidly converts to HDO in equilibrium ($D_2O + H_2O \rightleftarrows 2HDO$), which is linearly correlated to the measurable ArD^+ ion in the plasma." We note that analytical method measures elemental (not molecular) concentrations, so that the signal mirrors the *total* amount of the analyte(s), in this case the total ArD^+ in the sample. We also then note that, as a result, the equilibrium exchange of deuterium in the water is accounted for implicitly, with the BTCs shown in Fig. 1 representing effective measurements and the overall transport behavior.

DONE: In the revised manuscript, we have (i) added an explanation that we measure the total amount of deuterium in the samples (lines 163-165, revised manuscript), and (ii) in Section 3.1, end of the first paragraph (lines 189-194, revised manuscript), we expand the discussion to note that how this isotopic exchange is accounted for in the interpretation of the D_2O breakthrough curves.

(2) The study uses two packing configurations (homogeneous and heterogeneous) and two flow rates. While this is a good start, a more detailed justification for the specific grain sizes, the nature of the heterogeneity created, and the selected flow rates would be beneficial. For instance, what specific pore-scale structures or inclusion characteristics were targeted to induce the observed anomalous transport?

RESPONSE: We explained in Section 2, paragraph 2, our reasons for choosing the configurations and sand types. The setups are based on previous results from Elhanati et al. (2023); investigating similar types of experimental setups offers continuity and prior support that the transport is inherently non-Fickian. It is important to recognize, as shown repeatedly in the literature, that even pore-scale heterogeneity in macroscopically uniform sand columns, over a range of flow rates and travel distances, can induce non-Fickian transport.

DONE: In the revised manuscript, we have added this information at the end of Section 2, paragraph 2, lines 124-127.

(3) The manuscript states that the breakthrough curves for Br^- and D_2O are "similar." This similarity is visually apparent in the figures but could be further substantiated quantitatively. It is recommended to include a quantitative metric, such as the calculation of the temporal moments (e.g., mean arrival time, variance) or a statistical goodness-of-fit measure between the Br^- and D_2O curves for key experiments. This would provide a more objective and robust basis for the central claim of identical transport behavior.

RESPONSE: We calculated (but had not provided the metrics) previously, to support our claims regarding similarity of the BTCs shown in Figure 1. Thank you for suggesting this.

DONE: In the revised manuscript, we now include Table 2, following Figure 1, which provides the suggested metrics (mean travel times, standard deviations and correlation coefficients to compare each pair of BTCs), and reference to the new Table on line 197.

(4) The discussion in Section 3.4 is insightful but could be expanded. The extrapolation from a 1D laboratory column to a 3D catchment is a significant step. The authors should more explicitly discuss the potential limitations and necessary conditions for this scaling. For example, how might the relative impact of $\bar{v}w$ vs. $\bar{v}T$ differ in a 3D system with more complex flow paths, recharge dynamics, and the presence of unsaturated zones?

Acknowledging these complexities would provide a more nuanced perspective on the generalizability of the laboratory findings.

RESPONSE: We agree that expanding from a 1D laboratory setup to a 3D catchment is significant. We therefore tried to be "modest" in the analysis and in clearly stating limitations in the analysis, as they appear in the first paragraph of Section 3.4. Clearly, the presence of highly complex flow paths, recharge dynamics, and partially water-saturated regions in catchments will impact the relative difference between $\bar{v}_w$ and $\bar{v}_T$. This aspect remains to be investigated in detail, but it can be speculated that the added complexity might lead to an even wider range of sources for tracer retention, as suggested by some field-scale analyses of anomalous transport (e.g., Goeppert et al., 2020; Dentz et al., 2023).
DONE: In the revised manuscript, we have added these comments in Section 3.4, end of paragraph 2, lines 397-402.

(5) The manuscript notes the slightly different molecular weights (and hence diffusion coefficients) of H_2O and D_2O but dismisses this as a significant factor for the short-term experiments. Given that the study fundamentally challenges the assumption of identical behavior, a more thorough discussion of this point is warranted.

RESPONSE: A brief reference to molecular weights appears in Section 3.4, paragraph 5, Lines 418-421. In light of the referee's comment, we note that over relatively long time scales, the impact of different rates of diffusion may also be significant: diffusion on one hand can enhance tracer trapping and thus extend retention times in low-permeability zones, while on the other hand lead to increased uniformity of the tracer plume concentrations. We note that the specific impact of molecular diffusion in these scenarios remains to be investigated.
DONE: In the revised manuscript, we have added these comments in Section 3.4, end of paragraph 2, lines 421-426.