Response to referee #2 – second response:

I strongly urge the authors to reconsider their statement: "We thus show that v_T, the mean travel time of the isotope and the tracer, is distinct from the mean travel time of the “water” itself, as determined from Darcy’s law." On first reading this appears to be saying "the mean travel time of the water is distinct from the mean travel time of the water". I understand this isn't their intended meaning, but this framing is at best obscure and might in fact be read as deliberately provocative.

RESPONSE: The quoted statement is from our Reply to the referee’s first comment. The specific wording in the original manuscript, and further reinforced in the revised manuscript, is slightly different. In the manuscript itself, we state: “By estimating the porosity and cross-sectional area of flow through the column, and for a fixed Q, $\bar{v}_w$ can be calculated using Darcy's law, and the apparent [*this word added in the revised version*] mean travel time of water through the column can be estimated by dividing the column length by $\bar{v}_w$.” The discussion regarding v_T and v_W first appears in Section 3.2, and the definitions of these two terms are stated clearly and simply. The mean velocity of a tracer, or isotope, can indeed be different from the apparent mean water velocity. There is nothing obscure or intentionally provocative in this – rather, this is a fundamental point that is unfortunately not always recognized when interpreting and quantifying dynamics of fluid flow and chemical transport.
***DONE*:** In the revised manuscript, we added the word “apparent”, as noted above, for clarity.

First of all, Darcy's Law (as I have always understood it) is a statement about the relationship between a pressure gradient and bulk water flux (volume per area per time). The present paper never reports, calculates, or relies on a pressure gradient, and so they don't appear to be making use of Darcy's Law in a way that I can recognize. Consider: the quantity v_W obtained from v_W=Q/(nA) can be calculated regardless of whether the flow in the porous media is laminar (and so Darcy's law would be expected to hold) or turbulent (in which case it would not). v_W is therefore quite independent of Darcy's Law.

RESPONSE: As we discuss in the manuscript (Section 2, Methods), we in fact prescribed the volumetric flow rate, Q, in the experiments, and state the values. The volumetric flow rate relies

directly on the pressure gradient, so that the expression v_W=Q/(nA) is indeed Darcy's law and applicable to the experiments we report. Yes, we agree that one can in principle (mis)apply Darcy's law to situations of turbulent flow, but this is not the case in our experiments, and in most others reported in the literature. In other words, the v_W that we calculate, and that is usually reported in other studies in the literature, is indeed representative of and based on Darcy's law; the v_W we report is definitely not "quite independent of Darcy's Law".
***DONE*:** In the revised manuscript, we added (line 134) the definition of Q as the volumetric flow rate (which appeared later), and added "*Q*" to the heading in Table 1, for added clarity.

Second, the authors seem to want to have their cake and eat it too, when it comes to the relationship between v_T and v_W. Consider these two statements:
"In particular, we emphasize that the fundamental formulation of the ADE *requires* that the velocity term in the equation correspond to – i.e., be identical to – v_W. Every textbook development of the ADE immediately invokes the mean linear water velocity, v_W, based on Darcy's law."
-- fine, so when the ADE applies we would expect v_W=v_T. Deviations from that would indeed be surprising, but that is not what was observed here (since the ADE does not apply to the data presented).

RESPONSE: We agree and indeed suggest in the manuscript, similar to the referee, that the ADE does apply when v_W=v_T. However, it should be noted that literature over the last 25+ years shows numerous examples of experiments (and numerical simulations) in a wide range of porous (and fractured) media, over a wide range of spatial and temporal scales, that exhibit non-Fickian (non-ADE) behavior --- in other words, deviations between v_W and v_T are actually very common. The difficulty is that the ADE is frequently assumed to hold, when in fact it does not for the situation and measurements under analysis.

"In fact, we do not invoke v_W in the CTRW, and thus there is no "mismatch" or inconsistency between these two parameters in the CTRW framework" and "the continuous time random walk framework (CTRW) formulation discussed below is essentially founded on v_T"
-- fine, so when the ADE does not apply and we have to use CTRW, we would not expect v_W=v_T, as their meaning diverges. v_T is a parameter of the CTRW conceptual framework,

and within that framework it is conceptually distinct from v_W. They coincide only when the CTRW reduces to the ADE.
In the present dataset the ADE does not apply ("the ADE cannot match the measurements, particularly the long tailing behavior") and so (by the author's logic) we should not expect v_W=v_T, and **indeed this is the case**.
So where is the mystery here?

RESPONSE: The referee again agrees with us, but then concludes that there is no mystery here. In response, we emphasize that it is misleading to state that “v_T is a parameter of the CTRW conceptual framework” – it is a general concept (see Sect. 3.2, which describes a Gedanken experiment) that is relevant in all transport studies and modelling efforts. Moreover, further to the Response above, if ADE behavior is actually not exhibited in such a wide range of porous (and fractured), then why is the ADE almost “automatically” assumed to be the correct assessment of the transport behavior and then applied? As discussed in the manuscript, application of the ADE in aquifer and catchment studies remains ubiquitous. And as discussed in the manuscript, we demonstrate how this can lead to serious over-estimation of aquifer storage thickness.

It seems like the primary issue point being made in this paper is about how in porous media sufficiently heterogeneous as to be non-fickian the quantity v_W=Q/(nA) cannot be naively interpreted as the 'mean velocity' of the water. Instead, the presence of long tails adds some important nuance and complexity to the very notion of "mean velocity".
That might be a useful point to make, but the weird distinction the authors draw between the "velocity of the water" and the "velocity of the water isotopes" rather obscures it. Also, I would note again that the phenomenon at issue appears to be the case *regardless of what tracer is used*, so I'm still not sure why isotopes are being singled out.

RESPONSE: As the referee states, we agree that the presence of long tails adds important complexity to the very notion of mean velocity. This important complexity and the fact that it needs to be recognized, both conceptionally and for the sake of different estimates of aquifer properties – e.g., thickness – is indeed a main motivation for this manuscript. However, we respectfully disagree with the statement: “the weird distinction the authors draw between the "velocity of the water" and the "velocity of the water isotopes"”. There is nothing weird, except for the realization that this distinction is often unrecognized, so that modelling and interpretation efforts are biased and incorrect.

Our point, as demonstrated and discussed theoretically, experimentally, and using model simulations, is that there is a distinction between the apparent mean water velocity and the velocity of any measurable tracer moving with the water. While this point would likely be accepted by most researchers when discussing chemical tracers, we argue that there is a misconception when comparing the apparent mean water velocity and the mean velocity of water isotopes. Water isotopes are commonly thought to *exactly represent* the apparent mean water velocity; but "tagging" of individual water molecules identifies them similarly to chemical tracers, so that the misconception noted here is often overlooked. This misconception then propagates to incorrect use of the ADE and estimates of catchment properties such as aquifer storage thickness.

PLEASE NOTE THAT VARIOUS REVISIONS TO THE MANUSCRIPT ARE NOTED IN EACH OF THE TWO RESPONSES TO REFEREE #2.