

We are pleased to send you the final version of our manuscript. We have taken into account the last comments and suggestions of Referee #1 and Referee #2. Our responses (**in blue**) to their comments (**in italic**) are detailed below. Note that the Code and data availability section has also been modified and includes now the DOI of the repository archiving the R scripts and data used for this study.

Response to Anonymous Referee #1

The comments from the previous review have largely been addressed in the responses and manuscript. Concerns regarding novelty of this study also have been adequately articulated and discussed. The manuscript generally reads well and contains detailed descriptions of theory and methods, making this work suitable for a wide audience. I still have minor comments, largely related to clarifications and writing clarity. Details are below.

Comments:

- It would be good to know the ET values that were used to correct the precipitation signal, and how these ET compare to other studies in the region (likely in SI). It seems like using effective precipitation aided the interpretation of this study, but it is hard to evaluate how realistic the Thornthwaite's ET values are.

Thornthwaite's potential evapotranspiration values used to correct the precipitation signal are accessible following the DOI indicated in the Code and data availability section. Note that Pfister et al. (2017) compared in a region encompassing our study area monthly potential evapotranspiration data obtained via both the FAO-reference Penman–Monteith and Thornthwaite formulas. Their results show that the difference is less than $\pm 5\%$. This information was added to the manuscript at the end of the second paragraph of section 3.1.3.

- It would be helpful to discuss some of the issues with tritium (and in general only using a single tracer). For instance, aggregation errors in tritium can bias estimated mean travel times young (Bethke & Johnson, 2008). Without additional tracers, the full transit time distribution remains unknown, and the mean can be biased.

We followed this advice. New elements have been added at the end of the discussion section in this respect.

Line Comments:

Abstract: The content looks good; however, there remains grammatical and language errors. I suggest further review on the language.

L9: manifold 'of' ...

L25: delete 'obtained'

L28: delete 'particularly well'

L29: improve to improving

L30: delete 'on a larger scale'

L38: I had to read this long sentence multiple times to understand. I suggest breaking into multiple smaller sentences.

L65: Suggest replacing 'thorny' with challenging or some other word.

L65: '... risk of sample contamination...'

L66: Consider removing 'rapid exchange with the atmosphere after emergence' as it is basically stated just prior.

L71: Suggest deleting 'first to carry out'

L91: Suggest revising this sentence for clarity: "...an important aspect that must be considered is the tritium measurement accuracy..."

We agree with all "Line comments" above and have modified the manuscript as suggested.

L167: "it is assumed not to dominate water transit times in the aquifer at large scale". Why is this the case? Is this in contrast to studies that show fast-flowing components dominate the flux-weighted transit time distribution in streams (e.g. Berghuijs & Kirchner, 2017)?

We agree that a reference to support this hypothesis is missing. We added at the end of the sentence the reference to Farlin et al. (2013a), who concluded on a rather small relevance of the fast flow component in the Luxembourg Sandstone aquifer (based on a time series analysis of spring water chemistry and stable isotopes).

L195: Suggest dropping 'consultancy firms' and just citing the reports as normal.

We followed this advice.

L410: To clarify, are the plotted spring tritium decay-corrected?

The tritium data are decay corrected to the date of sampling and then plotted as at the date of sampling, as it is standardly done.

Could the downward trend be explained by the sampling month, given there is large variation in the intra-annual tritium input based on season?

Tritium sampling occurred in all three sampling years approximately at the same season. Therefore, the observed long-term trends are unlikely to be biased by seasonal bias. In addition, the groundwaters are undoubtedly older than a few years. Hence, even if the samples were collected at different seasons, such seasonal variability in the tritium input would be smoothed out in the groundwater system integrating over time and space.

4.1.2. I suggest being consistent with terminology 'mean transit time' or 'age', but not interchanging.

We followed this advice and replaced "Age" with "Mean transit time" in the title of section 4.1.2.

L425: There are other ways to generate old groundwater ages that could explain the >35 year samples. For instance, diffusion of old-aged water from rock matrix has been shown important, especially in fractured rock systems (Bethke & Johnson, 2008; Rajaram, 2021). Without the use of other tracers, confidently rejecting these old-aged samples seems like a major assumption. Some further discussion (likely in Discussion Section) seems necessary. Similarly, Eq. 4 has assumptions, namely that spring discharge is sampling flowlines across the entire storage volume, which is not always the (Berghuijs & Kirchner, 2017). I feel some discussion on this is warranted.

See our response to Referee #1's second general comment.

Figure 8: Can the meaning of Hdry be added to caption?

The meaning of Hdry has been added to the caption of Figure 8.

L569: delete 'First of all'

We agree and have modified the manuscript accordingly.

Response to Anonymous Referee #2

List of a few formal revisions. The page/line numbers refer to the annotated pdf copy of the manuscript (file hess-2023-152-ATC1).

Page 5

Line 135: at this point a very short description of the bedding types and a few words about facies types would help the reader to understand the filtration modes through the vadose and phreatic zones, also linking the regional image provided by the former statements to the very local compositional properties.

We followed this advice. Two sentences were added here to shortly introduce the sedimentological origin of the Luxembourg Sandstone and the different subfacies related types.

See also suggestion on lines 148-150.

See our response to the "Lines 148-150" comment.

Line 141: replace "chemical" with "mineralogical", or simply delete "chemical"

We agree and have modified the manuscript accordingly.

Lines 143-144: please clarify if these are quartzarenites with 10% calcite cement and what do You mean by "carbonate matrix". Is it true matrix, that means micrite, or do You refer to cement? What is the 60% of calcite in the calcareous sandstones? Particles? Calcareous lithoclasts? In which proportion to quartz and other lithoclasts? (watch Your classification of these rocks).

We agree that this text is not accurate enough and have rephrased it accordingly.

Is there some difference in porosity l.s. between the two sandstone types? I would move here the sentence from lines 153-155.

See our response to the "Line 153-155" comment.

Lines 148-150: this sentence could be moved at line 135, to start the description giving an idea of the stratigraphy.

We followed this advice.

Lines 152-153: replace permeability with hydraulic conductivity (the dimensions You use are m/sec).

We agree and have modified the manuscript accordingly.

Line 153-155: this sentence should be moved to the lines before the statements about hydraulic conductivity.

We followed this advice.

Lines 161-163: could You please relate these fracture systems to some regional structure? You mention a fault, which fault?

We agree with this comment. We have modified the manuscript to link more clearly the fracture network and the regional geological structure.

Page 6

Lines 184-185: I still disagree with this assumption, that is reinforced by the new description of the aquifer heterogeneity. As a matter of fact, 350 m/h is a very fast velocity, unusual through fractured media with low to intermediate gradients. However, this is just my opinion, based on a general consideration, hence I don't ask for changes to this statement. Maybe some readers would agree with me, that's all.

Line 224: in this list and in the new Fig.2 I count 8 districts. Which is the ninth?

Correct, this is 8, not 9. We have modified the manuscript accordingly.