General comments

a) The text content is generally well written but the theme's complexity and the structuring makes honestly the reading rather difficult, which can force the reader going back and forth the chapters to follow a storyline. I found very logical how the text was structured in the 'Study site and Methods' section, and I would really try to stick to this same structure when presenting 'Results', in the 'Discussion' section and even in the 'Conclusion' paragraph. My personal taste would be:

a.1) ADE parameters
a.2) TSM parameters (i.e. hydrologic exchange parameters)
  a.2.1. Identifiability of TSM parameters when $v = v_{variable}$
  a.2.2. Identifiability of TSM parameters when $v = v_{peak}$
a.3) Model iterations
  a.3.1. TSM first iterations
  a.3.2. TSM last iterations
  a.3.3. DYNIA
  a.3.4. Comparison with inverse modeling results (OTIS-P)
  a.3.5. Comparison with random sampling approaches (OTIS-MCAT)
a.4) Metrics and hydrologic interpretations of model results

According to this numbering, the 'Results' section starts with the a.3.1 jumps to a.3.3, then back to a.3.2, etc. The 'Discussion' section starts with a.4, then goes back to a.3.1 and a.3.2, follows with a.3.4, etc. This can make a difference between a quick and effective reading, and a tedious reading.

b) I would like to draw the attention about the use of the term 'parameters' in the manuscript. I personally have no experience with the identifiability of parameters in a TS model and it can bias my understanding of the term. Thus, this might be probably my wrong perception.
As a hydrologist, my understanding of **parameters** involved in the transport of solutes in a stream/river can be stream discharge, components of flow velocity, flow turbulence, grain size of streambed sediments, groundwater-stream (GW-SW) water exchange fluxes, stream channel topography, existence/absence of riparian vegetation, etc. In the manuscript, these parameters are treated as such, but then it introduces the term **TSM parameters** and **parameter sets**. In this manuscript's TSM model, each iteration simulates 115,000 parameter sets. An unexperienced reader would tend to think that > 100,000 parameters as the ones I mentioned above are involved in the transport of solutes in the stream. This should be obviously clarified.

c) The manuscript's title states 'to reduce the uncertainty' but I have missed more uncertainty analysis throughout the text.

Specific comments

Abstract

The abstract is concise and clearly written. How to fill the research gap is properly addressed.

1. Line 14: I would consider adding ‘adjacent groundwater bodies’ as a exchanging agent.
2. Line 24: … TSM parameters**, respectively**. The severe differences…
3. Line 26: … at **the** study site → The article ‘the’ makes reference to a determined object, but the study site has not yet been introduced. Consider changing ‘the’ for ‘our’ or ‘a study site in western Luxembourg’, for example.

Introduction

The introduction chapter needs to be revised in order to be fully understandable. The content is good, but an improved organization can elevate it.

4. Tend to list consecutive citations in chronological order (e.g. lines 72-73, 78-79, etc.).
5. Line 49: Many readers can disagree with such a definition of ‘hyporheic zone’. I don’t agree with the idea of a fully saturated thickness. I usually consider the definition given by Cardenas & Wilson, 2007 (Cardenas, M. B., & Wilson, J. L. (2007). Exchange across a sediment–water interface with ambient groundwater discharge. Journal of Hydrology, 346, 69– 80.), as a good approach.
6. Lines 79-82: This reads like a Research gap, and it is well stated, but it should fit at a later stage, before enumerating hypothesis and research questions.
7. Lines 103 and 113: I would not change paragraph in this sentence.
8. Line 105: ‘stream velocity’ → The stream itself does not move. Better using ‘streamflow velocity’.
9. Line 128: I would state the hypotheses in a new paragraph.

Study site and methods

This chapter is very well written according to me. Clear and concise. I would try to keep this structure in the following chapters.

10. Line 152: **Study site and ‘data’** → Better specify which data.
11. **Study site and data** → How long is the studied stream reach?
12. Lines 161-162: How was Q calculated? (± Analytical Error?)
13. Line 164: How was EC calibrated according to NaCl concentrations? Was EC temperature compensated? (± Analytical Error?)
14. Was there any difference in background EC between the injection and the measurement locations? Or was it assumed to be equal? Or is it one of the testing parameters? These aspects are actually interesting.
15. Line 181 (Equation 3): Why the second term of the second equation is negative? If $C_s < C$, the concentration of a certain solute in stream tends to get diluted. But according to eq. 3, $\partial C/\partial t$ would become larger…
16. 2.2: How the accuracy of the ADE and TSM results are assessed?
17. Line 207: I assume CDF is the Cumulative Distribution Function, but it has not been defined as such yet.
18. 2.3: How is the uncertainty of the model iterations’ results, DYNIA’s results and the comparison with OTIS-P and OTIS-MCAT assessed?
19. Line 270 (Equation 7): Introduce a tab in the equation’s label, so it can be in line with the rest of the equations.
20. 2.4: I assume that including FMED reduced the systematic error, but is it quantifiable?

Results

This chapter is properly written but the sub-chapters could be re-organized to build-up a smoother storyline. See the last part of General comment (a).

21. Lines 279-282: This reads more like in the 'Methods' section, right?
22. Why was v_{peak} and no other velocity chosen for the fixed velocity scenario? Why not median velocity which could be more a representative velocity value?
    What does v_{peak} mean?
    - Peak velocity during tracer experiments?
    - Mean velocity during the experiment E3 which had the largest discharge?
    - Peak velocity during the entire monitoring period?
23. Line 292: Figure 3c-f →Is Figure 3f informative in the tailing of the BTC?
24. Line 297: Add a comma after > 0.05)**,**
25. Line 342: 8.8 l/s for E3 → I would not consider this discharge as a low q_s, since you state that the exchanging flow is > 1/3 of the total Q of the stream.
26. Line 343: 15 hrs for E3 → I would not consider this as a long residence time. The flux velocity is roughly 4.6 m/hr, which according to me can be slow for dead zones, but not for hyporheic flux, or GW-SW exchange.
27. Lines 345-348: are these simulations physically possible at all, knowing the actual stream discharge? As for T_{sto}, do 0.8 (in E1) and 3.3 (in E3) m/s make sense?
28. Lines 353-354: Again… double check the applicability of these results into the actual flow regime.
    Comments 26, 27 and 28: Since you are doing a comparison between methods, rather than using terms like 'high' or 'low', it is better to use relative terms such as 'higher' or 'lower', or even 'distinct'.
29. Line 349: 'whether' does not sounds appropriate in this context. I would rather use 'regardless'?

Discussion

Same as in the previous chapter, the sub-chapters could be re-organized to build-up a smoother storyline. See the last part of General comment (a). Aren't 4.1 and 4.3 more related to each other, and 4.2 a unique sub-chapter. I think starting with 4.2 and then following with 4.1 and 4.3 would improve the chapter's sequence.

30. 4.1: Are the uncertainties specified?
31. Lines 402 and 406-407: Aren't these contradictory statements? Identifiability is contradictory under the same scenario (in both cases $v = v_{peak}$).
32. Line 438: 100,000 instead of 100'000.
33. Line 446-447: Seems like a very generalist closing statement.
34. Lines 460-462: Aren't these contradictory statements? The assumption $v = v_{peak}$ might not be representative of the advection role, but can encompass the effect of advection in the entire BTC.
35. Lines 480-503: Figure 8 is actually a very good review from other authors' findings. I quite like it. The piece of text relating to this figure is actually an important part of your discussion, but again it reads like a succession of sentences with vague organization, and it forces the reader to continuously go back and forth between the text and the figure. I propose the following. Instead of travelling through all the cited authors and comparing your results with their results, why can't it be presented in concordance with the different BTC's features, e.g.:

- Rising Limb
- Peak
- Falling Limb
- Tail

One can travel along the BT curve and compare your data with literature data. The reading would be more graphical and intuitive, and it could probably help to safe some word spacing.

By the way, it is interesting to know whether there is a method or a simple threshold value to distinguish between the end of the falling limb and the beginning of the tail?

36. Lines 502-503: '… Kelleher et al., 2013 also indicate …" → whereas it has a strong influence based on your results, right?
37. Lines 503-504: 'Different sensitivity…' → This reads like the staring of the next paragraph.

Conclusion

The conclusion reads well, but again, I would try to expose the ideas trying to follow the same sequence of findings shown in the results.

38. Line 523: '… that the BTC…' -→ Revise wording.

Tables and Figures

39. Line 739 (Table 2): $v = v_{peak}$ → Try to keep the same notation as in the text (i.e. Italic and subscript).
40. Line 777 (Figure 2): The meaning of colored dots for (m-o) plots not specified.
41. Line 778 (Figure 2): Try to keep the same notation as in the text (i.e. Italic and subscript).
42. Figure 3: Here you show the BTC in E1 when the streamflow is smaller and the identifiability is probably less dynamic. I would have liked to see differences in dynamism between the different tracer experiments.

    This does not mean that you need to re-do the figures for other tracer experiments, but probably comment a bit on it in the text.
43. Figure 4: I would probably combine Figs. 3 and 4 in the same figure. Both are showing exactly the same info (Experiment E1 and $v = v_{variable}$) but only for different iteration states, and can be potentially confusing.
44. Figure 7: In the Y-axis of plots g-I and j-l, please use similar notation. Use either 0.01 or 10^{-2}.
45. Lines 866-867 (Figure 7): 'm','n','o' seem to be leftovers.
46. Line 868 (Figure 7): '… and equal to v_{peak}**, respectively**'.
47. Figure 8: Plots g-l are presented first and explained later. They can both combine in the same piece of text.