Summary

This manuscript uses low-flow flume experiments to test how channel organic matter (OM) loading affects aerobic respiration (AR) in the channel versus the hyporheic zone (HZ). The authors find a strong, rapid channel response and a muted, delayed HZ response, and propose that under weak hyporheic exchange the HZ acts as a relative cold spot for POM processing. The question is relevant, and the combination of empirical and modeling work in a well-controlled flume is a valuable approach. The high-resolution optode data are a particular strength, and I especially liked how the raw HZ DO images (Fig. 3) are paired with the integrated DO masses (Fig. 4). My main comments fall into three areas. First, more complete reporting of the data and analyses that link concentration changes to AR, particularly reaeration and hyporheic exchange, would strengthen the interpretation. Second, some research questions could be answered more directly, notably the relative contributions of the channel and HZ to AR (Q2), and a few conclusions extend beyond the experimental design, including the high-flow conceptual model. Third, a short discussion of which microbial communities may drive the observed respiration (e.g., leaf-associated communities) and how experimental conditions such as inoculation and the sequential design could influence their response would help clarify what the measured AR represents. Many of these points could likely be addressed with data and model outputs the authors already have. This work has the potential to make a valuable contribution to the field, and I hope the comments below are helpful in strengthening it.

Major comments

1.The study would benefit from more complete reporting of the data and analyses that link concentration changes to AR (e.g., reaeration and hyporheic exchange). Several key fluxes are discussed but not yet quantified, so the reported AR currently reflects net changes in DO and CO_2 rather than respiration rates in each compartment. Much of this could likely be addressed with data the authors already have.

- **a. Reaeration.** The Introduction (L66-67) emphasizes that reaeration must be quantified to estimate AR, but I could not find a reported reaeration coefficient (e.g., k_O2 or k600). Section 2.5 describes the N2 sparging test, and Fig. 7c shows a reaeration curve, so it seems the data exist. Reporting the fitted coefficient and its uncertainty, and moving this figure to the Results, would be very helpful. These values could then be used to correct the channel DO and CO_2 rates for gas exchange and to quantify the fluxes discussed at L405-408.
- **b. Hyporheic exchange.** The interpretation relies on weak channel-HZ coupling, so reporting exchange flux and residence times from hyporheicFoam would strengthen this considerably. For example, comparing residence times with reaction timescales (e.g., a Damköhler number) or reporting oxygen consumption rate constants would build quantitative evidence, as well as discussing how these compare to other hyporheic flumes would assist in transferability. The three permeability estimates span roughly an order of magnitude (L211-214), so it would help to clarify which value was used and how sensitive the exchange flux is to that choice. I also wondered whether the DW and UW zones might have diverged more under greater exchange.

2. The research questions are well motivated, but some could be more directly answered, and a few conclusions extend beyond the experimental design (e.g., Q2, the conceptual model).

- **a. Research question 2.** Q1 is addressed well through the three experiments. It was less clear that Q2 (the relative contributions of the channel and HZ to AR) is answered quantitatively. A simple budget or plot comparing channel and HZ AR for each experiment, with rates normalized (e.g., per volume or bed area) for comparison with published values, would directly support the conclusion that the channel "dominates" (L456-457). It would also be nice to explicitly revisit the second hypothesis (strong HZ feedback to the channel).
- **b. Conceptual model and flow conditions.** Fig. 8 is a helpful synthesis, but only one low-flow condition (0.011 m per s) was tested, so panel (a) is not directly supported by these experiments. If there is literature supporting it, you could include them, but this may be outside the scope of this paper to draw conclusions under high flow without having empirical numbers from the same flume experiment.

3. It would be helpful to discuss which microbial communities may drive the observed respiration and how experimental conditions could influence their response (e.g., leaf-associated communities, HZ response times, and inoculation). The interpretation assumes that AR reflects the equilibrated flume channel and corresponding HZ communities. A short discussion of other possibilities would strengthen the paper, even if the current data cannot fully resolve them.

- **a. Experimental design and microbial communities.** Running sequential experiments in a single flume is very reasonable given infrastructure constraints. A few comments and questions about the design:
    - How was it determined that the system had stabilized after the ~2 month inoculation (L114-117), e.g., stationary DO, CO_2, and DOC in the channel and HZ?
    - Were the leaves used to inoculate the flume the same species and from the same source as those used in the experiments? If they were collected at different times or seasons, how might differences in leaf quality (e.g., C:N, leachable DOC) have influenced the response?
    - Could the leaves have introduced a microbial community of their own, separate from the inoculated channel and HZ communities? It would be interesting to know whether microbes were already present on the leaves and how quickly leaf OM becomes available to them. A sentence or two acknowledging this would clarify what "channel AR" represents.
    - Relatedly, could the leaf pack be acting as a third respiration compartment, separate from the water column and the HZ? The channel response was both fast (within 30 min) and large: ~44 mmol O_2 in 24 h for POM200A. If this oxygen is not consumed in the HZ, it is presumably consumed on or above the streambed, so within the leaf pack, water column, or at the sediment surface. Discussing how much this matters for the interpretation would be helpful.

    - Each addition began once the channel had recovered, while the HZ was still re-equilibrating. Could carbon have accumulated in the HZ across runs and contributed to the muted response (e.g., L316)?
- **b. Response time of HZ communities.** How quickly might HZ communities be expected to respond to a 24 h pulse of leachate? If their response is inherently slower than the experiment, the muted HZ response could partly reflect timescale rather than a true cold spot, which could be worth mentioning in the discussion.

Additional considerations

1. CO_2:O_2 stoichiometry. How well the DO and CO_2 responses line up could reveal some of the underlying processes at play. For POM200A, the channel lost ~44 mmol DO and gained ~17 mmol CO_2 (L265-266), a ratio of ~0.4 rather than the ~1 expected for aerobic respiration. This could be an interesting result on its own, since paired O_2 and CO_2 estimates often diverge in rivers (e.g., Vachon et al., 2020; Shangguan et al., 2024), and a flume is a great setting to isolate why. Possible contributors include CO_2 evasion, carbonate buffering, and non-respiratory O_2 demand such as nitrification. Reporting channel DIC alongside CO_2(aq) would help the authors close the carbon budget.

2. Cold spot vs. CO_2 source. It would help to briefly reconcile the HZ as a "cold spot" (Abstract, Conclusions) with the Results statement that the HZ is "always an important source of CO_2" (L294-295). I think the point is that the HZ stores CO_2 but does not respond to channel loading or export it under low flow. Saying so explicitly would clarify how Fig. 8b should be read.

3. Pre-addition baselines. Are pre-addition baselines available for the channel and HZ? Referencing them would help separate responses to POM from background variability, such as the temperature-linked HZ DO fluctuations (L302-303).

Specific comments

Abstract

L15: "sediment with hyporheic zones" is confusing, as it suggests the HZ is a feature within the sediment rather than the zone of exchange itself. Consider defining the hyporheic zone at first use and keeping terminology consistent throughout (i.e., not using "sediment" as a substitute for "hyporheic zone," e.g., L40 and the Fig. 4 caption). Because the flume has no groundwater component, it would also help to give an operational definition in the Methods (e.g., here the HZ is formed by bedform-driven exchange of channel water).

L20: Define CO_2 (carbon dioxide) at first use. This happens throughout where DO is spelled out as dissolved oxygen before use, but carbon dioxide is not.

L24: This contradicts the Results. It says channel DO increased and CO_2 decreased, but the Results show the opposite (e.g., L232, L345).

Introduction

L32-39 and L130: CO_2 is not subscripted.

L33: Consider citing the original meta-analyses in addition to Harvey and Gooseff (2015).
L54: Remove “is.”
L66: This paragraph jumps around. Why is reaeration so important for AR? It needs a bit more explanation before moving to surface-subsurface fluxes. Consider separating out these ideas.
L75: A bit nitpicky, but is this flume large relative to other flume studies? Why is the word “large” used here?
L78-79: The hypotheses could each be one sentence tied directly back to your research questions (e.g., H1 and H2) to make it more explicit.

Methods

Has this flume been used in previous studies? Kaufman et al. (2017, 2021) and Norman and Cardenas (2014) are cited, but it would help to state this explicitly in the Methods.
L99-102: Low flow is central to the conclusions but is only briefly justified here. Why was this flow chosen?
L120, L265, L270: “m3” needs a superscript.
L125: “Mostly” dry Morus alba? What other leaves were used, and was drying standardized?
L144: The HOBO information can be shortened.
L151: Where are the GC validation results? Please include them in the supplement.
Section 2.2: Consider splitting the planar optode material into its own paragraph.
L180: “L-1” is set as a subscript instead of a superscript.
L189-193: The labeling appears reversed. Eq. 2a is the CO_2 solubility function of Weiss (1974), and Eq. 2b is the van’t Hoff form used for O_2. Please add sources for both (e.g., Weiss, 1974) The temperature coefficient in Eq. 2b also could be cited (e.g., Sander, 2015).
Section 2.5 needs citations. Is N2 sparging a common method for measuring gas exchange in flumes?
Please check the use of CO_2(aq) vs. CO_2 throughout

Results

L263, L272, etc.: The “Observations in the Channel/Hyporheic Zone” labels are not necessary unless they become subsections; the channel could be one paragraph and the HZ another. It would also help readers if all three experiments followed the same structure: (1) channel, then (2) HZ, each reporting DO at % saturation, CO_2, pH, and DOC. The authors report the change in channel DO saturation but the HZ in mmol which is difficult to compare; staying in the same units and comparing similar hours would help the reader.
L267: “POM200” should be POM200A.
L270: There is no need to restate the channel volume here.
L278: Typo in “POM-200A.”
L303: Consider “minima” instead of “valleys.” Temperature also is not shown, so a temperature panel in Fig. 4 would support this point.

Discussion

L353: Delete “are.”
L354: Delete “so.”

L375: Did you test the nutrient content of the leaves or the leachate, or is this assumed from Webster et al. (2016)? Channel nutrients are reported in Table C6, which would support the C-limitation argument if discussed.
L396: Connecting nitrate values from the supplement could support the denitrification claim.
L400-401: The text cites Fig. B2 for the boxplots and Fig. B3 for the PCA, but these appear as Figs. B1 and B2.

Figures

Figs. 6 and 7 read more as Results than Discussion, and could be rearranged to help the readers.
Fig. 5: The caption could be more detailed. What do yellow (channel) vs. red/blue (HZ) and the shaded background panels represent? The caption mentions stars, but the CO_2 panels use circles, and it says “porewater” although channel data are included.
Fig. 7: Please define (a) and (b). The figure seems out of order, and panels (a), (b), (d), and (e) are not referenced in the text. All panels should be referenced if included. Placing the reaeration results in the Discussion is also confusing (see general comments above).
Fig. 8: (a) and (b) are restated repeatedly in the caption and could be consolidated. This figure could also come earlier in the Discussion rather than right before the Conclusions.

Appendices

Fig. B1 (porewater DOC and CO_2 by depth) is not discussed in the text.
Fig. B2: Please describe what went into the PCA. It appears to be based only on the two sampling dates in Table C6 (pre/post POM200A), which is a small sample for this analysis.
Table C3: UW 60 cm DOC = 1.06 at hour 0 looks like a typo (10.6?).