Responses to comments by referee 1

1. In terms of data collection, repeat photogrammetry was taken, and channel networks were analyzed through manual delineation of the tops of the channels only. Was there any effort to track the volume of the channel network? If multiple images were taken, they could be processed using Structure-from-motion to get topography. If not, was there anything else done to try and collect topographic data? Or monitor the sediment export?

   → We agree with the referee that measurements of the network topography would be valuable in the context of our experiment. Unfortunately, the grains of our experimental aquifer are homogeneous, making the surface of the aquifer insufficiently textured for photogrammetry. As an alternative, we tried to measure topography by projecting fringes on the network. However, the channel banks were too steep for this method to work, and led to large errors. Topographic measurements would have required significant changes in the experimental setup, which we plan to address in future work.

   Changes in the manuscript: line 228 (line number in the marked-up manuscript), "However, to reconstruct the water table, we choose to neglect the network topography and set it to zero. While this method correctly captures the shape of the water table across most of the experimental domain, it overpredicts the discharge by a factor of about two near the channel tips. Measuring the topography would help us to resolve this discrepancy. Unfortunately, because of their homogeneous color, our grains lack the texture required to use photogrammetry. We are instead currently testing a fringe projection method to extract the topography".

   → At the experiment outlet, water and sediments flowed into an overflowing tank. As the water level in the tank was kept constant, water coming from the experiment automatically flowed out. Conversely, sediments sank to its bottom and stayed in the tank. Thus, by measuring the weight of this tank in time - knowing the sediment and water densities - we estimated the mass of sediments flowing out. Unfortunately, the variations in water levels and important changes in water discharge induced significant errors in sediment mass estimates. In future work, we hope to build a more robust technique to measure the eroded sediment flux.

   Changes in the manuscript: line 213, "Establishing the exact nature of this relationship requires additional experiments. Moreover, precise measurements of the sediment flux would improve our ability to monitor the erosion intensity during network growth, which we currently assess only through visual observations."

2. The discharge measurements were made by measuring the mass of water coming out of the basin over a set period of time (line 86). What time period did you use to measure discharge? Did that water also contain sediment? If so, how was that accounted for in the mass/time measurements?

   → Discharge measurement were performed over a time interval of 1 to 5 min. Indeed, the water flowing out of the aquifer contained sediments. However, as both the average sediment flux and the grain density are low (around 20 $g.h^{-1}$ and 1500 $kg.m^{-3}$), we consider that its impact on the discharge measurements is negligible.

Changes in the manuscript: lines 88 to 91, "collecting in a beaker the mixture of water and sediments", "over time intervals ranging from 1 to 5 minutes" and "Because the sediment discharge is relatively low (about 20 g.h^{-1}) compared with the water discharge (at least 6 kg.h^{-1}), the sediment mass is negligible. Weighing the beaker thus yields a reliable estimate of the water discharge (Romon, 2025)."

3. The experiment ran for 35 days. Did you run it continuously? Overnight? How often did you measure things like discharge?

→ 25 days is the effective duration of the experiment, which we ran continuously as much as possible - including overnights and weekends. The only time is was stopped was because of a technical problem with the camera.

→ We measured discharge right before changing the water table height, then once more a couple minutes later. We sometimes did some extra measurements to add more data points, but these were not regular.

Changes in the manuscript: line 125 "measured the discharge of water leaving the aquifer, increased the recharge by a small amount (typically 0.1 L.min^{-1}), then measured discharge once more (Fig. 3a)." and line 127 "- during which the experiment ran continuously -".

4. On line 116-117, you note that you waited until the network had achieved a stable morphology before increasing the aquifer recharge rate. Please explain how you determined when the network was no longer eroding and had reached a stable morphology. Was it just through visual observation? Did you compare photos between different time periods?

→ As the referee mentioned, we used image comparison by subtracting photos from different time periods to identify any changes - or lack there of - at the channel tips. At the same time, we observed the network directly - where moving sediment were easy to see.

Changes in the manuscript: line 124 "we compared photographs from different time periods and waited until an absence of observable changes indicated that the network had reached a stable morphology."

5. For the modeling, the assumption was made that the channel slope was not important as it was only 2% and thus the water table at the outlet was used to set the elevation of the water table throughout the channel network, making the water table slope essentially zero in the channel network (line 152-155). Looking at Figure 5a, it looks like there is about a 3cm drop over the experimental domain (150cm x 150cm), which is a 2% slope on the water table. Thus, the channel slope that is neglected is the same as the water table slope. It is thus not negligible. The result is that anywhere there is a channel, the water table gradient in the model will be artificially high adjacent to it because the water table elevation drops to zero there (same as the outlet). This will generate higher velocities along the edges of the channel because of the artificially high hydraulic gradient. This seems rather important. In lines 175-177, you note that the difference between the water table heights in the channels and the modeled heights is quite high, indicating that the channel slope cannot be neglected. Why did you neglect it then? Can you please explain how you handle this discrepancy in a

way that does not impact the results and ensuing interpretation of the water table gradients in the vicinity of the channel network?

→ Indeed, the results in our manuscript indicate that the slope of the channels in the experimental network are non-negligible compared to the variations in water table elevation, highlighting the importance of measuring topography in any future experiments. Still, the piezometric data shows that simplifying the network elevation to $h = 0$ does not impact our reconstruction of the water table outside of the immediate vicinity of the channels.

→ However, as noted in this comment, this simplification artificially increases the gradient of the water table near the channel tips, causing a significant over-estimation of the groundwater velocity. Thus, we agree with the referee that the values of groundwater velocity presented in the article (in particular on figure 6 d-f) are not robust enough to be exploited.

→ To evaluate the error that our simplification of network topography induces, we discuss in a second appendix (B) the case of a simpler, one-dimensional system meant to represent a small section of our experiment in the vicinity of a channel tip. Using the analytical solution for the water table height in this one-dimensional configuration, we find that our simplified boundary condition ($h = 0$) results in an overestimation of the groundwater flux by a factor of about two.

Changes in the manuscript: We detail the computation in appendix B of the manuscript, lines 254 to 268, and figure B1.

→ Although our simplification of the network topography ($h = 0$) induces a non-negligible error on the groundwater flux, it allows an estimate of the right order of magnitude. Thus, we choose to represent the groundwater flux, $q = -Kh\nabla h$, in section 4 of the manuscript, instead of the velocity.

Changes in the manuscript: line 186 "The difference between our boundary condition and the actual water table height inside the drainage network (approximately 1 cm) results in an overestimation of the groundwater flux by a factor of about two (see appendix B)" and line 189 "In short, our numerical method accurately reproduces the water table, except in the immediate vicinity of the drainage network". Most importantly, we have rewritten the last paragraph of section 4 (lines 195-204) and replaced figure 6 d-f. Correspondingly, we have changed the caption of figure 6, lines 226 and 237 of the section Conclusions and Discussions.

6. The paper states that 6 experimental runs were conducted (line 104), yet only one of the runs is described here. While the approach of focusing on one experiment to highlight a particular process is reasonable, there are no data presented at all from the other 5 experiments. Could you combine some of the data from those experiments to bolster the data being presented here from a single experiment? I was left wanting some confirmation of the observations from this single experiment with more experiments, and since you ran more experiments, perhaps you can include some information on whether they support or contradict the observations seen in the one experiment presented here.

→ Unlike what the manuscript suggested, the 5 additional experiments served as preliminary tests. For some of them, we had not yet installed complete monitoring (piezometric data, discharge, regular photos...), and all were mostly used to test out the setup, so that a robust experiment (the one presented in the article) could be run.

Changes in the manuscript: line 109 "To investigate the formation of drainage networks in our laboratory aquifer, we ran several preliminary experimental runs, each lasting from a few days to a couple of weeks", and line 113 "Over time".

→ Several of the preliminary experiments led to the formation of drainage networks. Thus, we have added an appendix (A) to the article where we present some of our observations from these experimental runs, as well as pictures of the networks.

Changes in the manuscript: Appendix A includes figure A1 and lines 246 to 252.

7. Lastly, there is no discussion section in the manuscript, tieing your observations back to the literature and to the natural world. The second paragraph in the conclusions section highlights this a little bit. I think the paper would be stronger if you could spend more time relating the results from here back to other experiments or the natural world.

→ We agree with the referee that the article was lacking some discussion and links with natural world cases. Therefore, we titled the last section "Conclusions and Discussions", added information about the perspectives our experiment offers, and linked our findings to studies of natural networks.

→ How might the results here inform channel network evolution models in heterogeneous aquifers? To understand the impact aquifer heterogeneity could have on groundwater flow and channel growth, we could build an experimental aquifer with various layers of different grain types. In such a context, we expect the variations of hydraulic conductivity to influence the groundwater flow, leading perhaps to different groundwater velocities in each layer.

→ How might these results combine with surface flow in natural systems to set the pace of channel network evolution? In these experiments and in the few field studies we have conducted (to be published), we have considered areas where overland flow is negligible compared to infiltration.

Changes in the manuscript: line 216 "in areas where infiltration dominates over overland flow".

→ What does it mean in terms of landscape evolution if seepage erosion reaches a steady-state and essentially stops unless aquifer recharge increases? This suggests that many natural networks might currently be in steady-state and are no longer growing significantly (not considering surface erosion, glacier melt, or extreme weather events). Moreover, current river networks might have a morphology due to past, stronger groundwater flow, no longer representative of today's aquifer recharge.

Changes in the manuscript: To answer question 7, we have rewritten several paragraphs of the Conclusion and Discussion section: lines 217 to 224, lines 237 to 239.

A few minor notes:

1. The term “ramified” is not one I was familiar with to describe dendritic or branching rivers. Consider using a different term, like branching. (This may be a regional issue – I am based in the USA.)

   → In light of this comment, we have changed the term “ramified” to “branching” in the entire article (2 iterations, lines 24 and 55).

2. What kind of plastic were you using? Did it have any cohesion?

   → As stated in the article, we used Guyson guyblast plastic media US type 2 (sizes ranging from $d = 500$ to $1000\ \mu$m and density $\rho = 1500$ kg.m^2). Although we have not measured its cohesion, this plastic sand is angular and of irregular shape which makes it more cohesive than spheres.

3. On figures 2, 5, and 6, please include scale for the images or explain what area is covered in the figure captions.

   → We have added a scale on figures 2 and 5, and a clarification in the caption of figure 6 : “Each reconstruction of the water table and of the associated groundwater flux spans over the entire experimental aquifer (150 × 150 cm)”

4. Line 212 should be “split” not “splitted”

   → We have corrected this spelling error line 132.

5. Line 168 remove word “with” or “to” (we compare it to the piezometric. . .)

   → We have removed the word “with” line 178.

6. Line 191 should be “velocity” not “vel”

   → We have remove the word “vel” line 206.

Responses to comments by referee 2

1. Line 122 — “splitted” should read “split”

   → We corrected the spelling line 132.

2. The value and dimensions of alpha in the caption of Fig 4 do not look right.

   → There was indeed an error in the caption of Fig. 4, where the linear function was written incorrectly. Thus, we replaced the previous caption with : “Blue dashed line: linear fit to the data $A = \alpha Q$ with $\alpha = 3.8 \cdot 10^3$ s.m^{-1}”.

3. After Eq(2) in addition to defining K might also be a good idea to define R.

   → To clarify Eq. (2), we have added: “and R is the recharge rate” line 154.