Response to Reviewer 3

We would like to thank Reviewer 3 for the comments and remarks which improved the manuscript. Please find our revisions and replies to the general and specific comments below. The line numbers in blue square brackets refer to the line numbers in the uploaded revised manuscript with tracked changes.

General comment / main concern:

R3: *My only main concern is about the experimental design. While the overall research question is about constructive large monoliths e.g., over 1 m³ (L39) that should obtain more accurate representation of soil processes, the experiments are conducted to smaller size blocks (1 x 0.5 x 0.35 m). In addition, only two tests are conducted and with same initial and boundary conditions (constant flow of 5 l min-1 after 24h from fully saturated soil). Finally, as different monoliths are compared, soil heterogeneity affect the interpretation of the results. So, how could we address the questions and reach any conclusions? In contrast, based on the research question and the overall general introduction of the study, I would have expected to see, e.g.: tests conducted on larger combined monoliths (e.g., two blocks of 1 x 0.5 x 0.35 m); tests conducted at the same monolith uncut and cut; tests conducted with different initial and boundary conditions. Field tests over the same area could have also been be performed for comparison. For this reason, I would strongly suggest to increase the number of experiments to try to derive some more general conclusions. In case these could be not anymore feasable, I suggest to rephrase several parts of the manuscript and weakening some statements as listed in the specific comments below.*

Answer: In general, we agree with what the reviewer states in this comment. Conducting experiments on larger monoliths, using different initial and boundary conditions, or the inclusion of field tests would have all been beneficial to support our hypothesis. However, we see no solution to the problem of soil heterogeneity other than repeating experiments with a high number of replicates. In this way, we may at least find out if soil heterogeneity is larger than procedural heterogeneity (which we did). Please be aware, that using the same soil monoliths uncut and cut will also introduce heterogeneity of unknown size due to the handling necessary during the resampling procedure.
Indeed, as R3 already mentions, conducting these additional experiments was/is not feasible. In fact, our experimental set-up presented in this manuscript (6 monoliths, 2 experimental sets, analysis of runoff characteristics, chemical properties, and flow velocity, 4 flow pathways) was already quite time-consuming and labour-intensive. While we agree that the suggested additional experiments are reasonable additions and something that may be done in future studies, we think that our results are sufficiently informative to draw the conclusion that combining monoliths is a sound procedure. However, we discuss this issue in more detail in the manuscript and rephrased several parts of the manuscript as suggested by R3 and the other reviewers (see Responses to Specific Comments).

Specific comments

R3: *Title not really meaningful, please consider rephrase. The two worlds should be field and lab but no experiments are actually conducted to answere that a combined monolith is the best of both. See general comment.*

Answer: We agree and rephrased the title to "*Technical Note: Combining undisturbed soil monoliths for hydrological indoor experiments*". [Lines 1-2]

R3: L7 I guess there are several major decision in soil hydrology depending on the objectives and expertise. The statement that a major decision in soil hydrological research is whether to conduct experiments outdoor or indoors is in my opinion a bold statement. Please consider rephrasing.

Answer: We agree that the decision whether to conduct the experiment under field or laboratory conditions is one of several options in soil hydrology. We rephrased the sentence: "*An important decision […]* ", to better reflect this. [Line 7]

See also response to Comment L19.

R3: L16. The study can be improved by adding some suggestions about which experiments should be conducted for a definite conclusions, e.g, block size, initial and boundary condition, comparison to field tests etc.

Answer: We deleted this sentence from the Abstract (following other reviewers' suggestions). We think that the wording (*our results suggest*…, *we propose*…) is defensive enough to clarify that there is still room for improvement. [Lines 17-19]
Nevertheless, we discuss this issue in more detail in section 4.2.: "*We used a comprehensive approach, analysing runoff characteristics, chemical loadings, and flow velocity and are, thus, confident in our conclusions. Nevertheless, there is room for further research, for instance applying different boundary and initial conditions, such as other soil types and flow rates, or a higher sample size and number of blocks to be combined. A comparison of large monoliths taken with heavy machinery with blocks of similar extent that are made up of combined smaller monoliths would be very interesting.*". [Lines 297-301]

R3: L19. Is this really a cardinal question? For some reasons one could always be fine with lab or field test.

Answer: We agree that one could always be fine with lab or field tests. However, each comes with specific advantages and drawbacks and they are, thus, not interchangeable and dependent on the research question and resources. The rather strong (original) wording of this sentence has also been driven by the fact that in soil hydrology (as opposed to other hydrological fields) the question of conducting indoor or outdoor studies is quite often an issue.
To account for this comment, we changed the wording from "*cardinal*" to "*important*". [Line 22]

R3: L68 The blocks were interchanged so that the left front of the left block faced the right front of the right block. Why?

Answer: Blocks were interchanged to better mimic the combination of two separately taken monoliths. If we would have combined the separated monoliths again directly at the cut without interchanging we were concerned that this would lead to a "perfect fit" (e.g., concerning macropore channels) that would not be possible to obtain when separately taken monoliths are used.

R3: L71. Discussion can be extended considering the results we should expect for different soil type

Answer: We added a discussion on different soil types at section 4.1.: "*In principle, we expect the combination method to apply to a large variety of soil types, but we cannot yet provide textural limits. However, combining two soil blocks requires removing the bordering at one side. Thus, structurally weak soils that could collapse without a frame are not suitable for this method.*". [Lines 195-197]

R3: L92. Why this flow? Is this representative of hydrological conditions? Alternative scenarios should have been performed. If not possible, discussion should be extended with suggestions about.

Answer: This flow rate was chosen based on previous experiments. We know from previous (unpublished) studies on similar soils that a large (though not extreme) rainfall event of 60 mm would produce a runoff within the range of 0.1 l min^{-1} m^{-2}. The chosen 5 l min^{-1} would therefore correspond to a source area of 500 m² that contributes to a concentrated flow at the field edge – which is probably at the lower end of realistic source areas. Other studies using artificial runoff (e.g., Guertault et al. 2021, Saleh et al. 2017) used much higher flow rates (12-99, and 99 l min^{-1}, respectively). This would not be possible to handle with our set-up. Nevertheless, we also chose a lower flow rate to encourage infiltration, as we speculated that the cut could act as a large macropore. Higher flow rates and, thus, higher flow velocities would promote surface runoff at the expense of infiltration.
Suggestions on experiments with different flow rates are included in the added paragraph referred to in the Response to Comment L16 (see above).

R3: L205. The fact that spatial heterogeneity affects more than cut monolith can not be considered as a proof of the validity of the combined procedure.

Answer: We agree that the high heterogeneity encountered is not a proof of the validity of the combination procedure. It is still possible that there is an effect of the combination, which would probably require a (maybe unfeasibly) high number of plots (replicates) to find out. Based on our data, even if there is an effect it would be small in comparison to the soil heterogeneity. Based on our findings, we only state that "… *there is no indication to reject the hypotheses*", which is true. Thus, we think that our conclusion that the advantages outweigh potential disadvantages of added data noise is supported.

R3: L213-214. Combining more monolith has been not proofed to be valid in the present study

Answer: We agree and rephrased and weakened this statement as follows: "*We conclude that the re-combination procedure did not lead to directional differences and, thus, had no adverse effect on runoff properties, suggesting that combining two (and probably more) blocks is a viable and practicable way to obtain single larger soil monoliths.".* [Lines 253-255]

R3: L230. What is Video 1?

Answer: Video 1 is part of the *Assets*. The *Assets* consist of a *Supplement* and a *Video supplement* and should both be downloadable at the Preprint Website (Copernicus Office). This short video shows a concentrated outflow at a single outlet, probably an earthworm channel.

R3: L244. During drying the cut seems to become relevant. So, it could be argue that also runoff could be affected depending on boundary conditions.

Answer: We agree; consequently, it is important that the monoliths are properly maintained, especially regarding regular watering. Experiments should not be carried out on dry-ish soil conditions. We added clarifications such as: "*We recommend that soil monoliths are kept outside in a sheltered but sunny location. Blocks need to be watered regularly for plant vitality and to avoid drying and an emergence of cracks in the soil structure that would affect runoff properties and impede the repair of the combined monoliths*", and

"*Furthermore, proper storage and maintenance, especially regular watering, are crucial to keeping the monoliths in good condition and are, in turn, dependent on the research aim, the duration of the*

experiment, climate, and resources (e.g., staff, storage space).", to the Discussion and Conclusion. [Lines 198-200; 306-308]

R3: *261. Same comment as for L213-214*

Answer: We think R3 is referring to the sentence "… *we recommend the use of combined monoliths* …". We admit that the term "*combined monoliths*" may be a bit ambiguous if two or more monoliths are meant. We suggest to leave this unchanged, but we added some more information in the following sentence to clarify this issue: "*Nevertheless, we encourage further research on this subject to better delimit the potential and possible limitations of this procedure, for instance using different experimental setups (e.g., number of monoliths), boundary conditions (e.g., flow rates, soil types, dimensions), or analysis methods (e.g., X-ray imaging).”*. [Lines 317-320]