Referee #2

We thank the referee for his supportive review and for the useful comments. Below are our answers, including proposed action, *following each of the comments.* This includes the general comment regarding the treatment of interpretation as interpretations and not as results (comment #9).

1. The manuscript uses many abbreviations. This is a question of taste, of course, but I had to scroll back and forth a lot during reading. A table might be useful.

   Thanks for this suggestion. A list of abbreviations was added after the abstract (P.2, line 39)

2. *Figure 1 and Section 2.2: It did not become clear to me how the infiltration was carried out. I might have missed the description. From the text I understood that the infiltration was done by a constant head that was realized with the overflow tank. The resulting infiltration rates were not constant and depended on the soil conditions. In Figure 1, there is a peristaltic pump that seems to inject water with a constant rate. If water was pumped with constant rate, I would not understand how this could be, as the infiltration rate was measured and was not constant.*

   The referee probably missed the overflow pipe (w3 in Fig. 1).

3. *Line 145-146: I did not understand the sentence. What is meant by no pressure head above the unsaturated soil surface? Does this mean no ponding water? If there is water, the water always has a pressure head, which can be zero or positive or negative, but there cannot be no pressure head. I guess what is meant is no water.*

   Thanks for pointing out this clarity issue. We have rephrased the sentence and it now reads "*Six 72-hour column experiments were conducted. All experiments started with a dry system (i.e., after at least nine days of drainage), and no ponding on top*." (P.5, lines 136-137).

4. *Line 146: I also did not understand how the effluent supply can be higher than the infiltration. From figure 1 I understand that the effluent is the infiltrated water. Or were there separate supplies of effluent and clear water?*

   The water (effluent) supply is constant, and almost always at a rate higher than the soil's capacity. This leads initially to ponding, and then to overflow, as described in the text (P.5 lines 137-138)

5. *Line 191: The 'h' is missing after the times.*

   Thanks. Corrected (P.7, line 182)

6. *Lines 192-194 and Figure 2: I was here also confused about the infiltration rates. If I understand correctly, the infiltration rates were concluded from the volume in the tank. How often was this done? In Figure 2, the mean infiltration flux is given and I assume that this is the time average of the infiltration flux. At least this would explain why it is so constant over longer time spans in figure 2. What time intervals were chosen for averaging? And why? Why not showing the time series of the infiltration rate without averaging? This would allow to get a better picture of how much the infiltration rate is related to the water content in the upper soil column part.*

   Our intention was to use a pressure transducer to measure infiltration in high temporal resolution, but due to technical difficulties it did not function during some of the experiments, and we chose not to present different data for different experiments. The average is for an entire period, which is 24 h for FDO, 23 or 24 and 1 h for AI-LF1, and 24 h for AI-HF1. This is now clarified in the text: "*As a pressure transducer was not functioning for some of the experiments, the fluxes are averaged over time as presented in Fig. 2*." (P.7, lines 185-186)

7. *Related to the previous comment: If I understood the averaging of the infiltration correctly, I think it would be better to use the term 'mean infiltration' consistently in the text instead of calling it 'infiltration' (example lines 275, 279). This would prevent misunderstanding.*

   Thanks. Indeed, using mean infiltration is more appropriate and we have changed that at 20 different locations throughout the text.

8. *Section 3: Do I understand it right that the first 24 hours of the experiments FDO and AI-LF1 were the same? The large difference of water content in the upper soil at the beginning is thus the inherent uncertainty in the experimental results. I think this has to be considered when conditions in the experiments are compared. Smaller deviations have to be interpreted a bit carefully.*

   The observation is correct – FDO and AI-LF1 are identical in terms of operation in the first 24 h, other than small changes in the initial conditions. However, we disagree that the differences are high. For most of the time the water content in both experiments is 28%, with FDO increasing to 35% at about 18 h, with practically no difference in recharge throughout the period. The main difference is probably the earlier release of entrapped air in FDO. This is clearly explained in the text (P.8-9, lines 233-237).

9. *Lines 259-266: This paragraph contains an interpretation of the increasing water content and infiltration rates with time in the air injection experiments with low infiltration rates. The interpretations are plausible, however, they are just interpretations. There are other*

possibilities that could explain the findings, such as redistribution of water with the air injection that leads to connected patterns with higher water saturation and thus higher conductivity, facilitating higher infiltration. The increased water content could be local, and would then be rather accidentally. Also, the interpretation with the breaking of crusts reads a bit odd. I am not an expert on bio-clogging, but the time seems very fast to me. If breaking of clogging crusts was relevant for the increase of injection, the crusts must have formed before the first 24 hours. Is it realistic that this happens so fast? Is there evidence for crust building in the sand column (maybe visible at the surface)? I am also not so convinced by the preferential pathways. Why should a re-arrangement of grains (this is how I would interpret the 'creating wider pores' read) should lead to large pore clusters connected from top to bottom? It is possible, but not very likely. To clarify: The interpretations are plausible and I would not argue against them. But the preferential pathways and crusts are later in the manuscript treated as results. I think they need to be kept as possible processes that take place, not as given ones.

Thanks for this observation. Our empirical observations indicate that in SAT systems crusts do form (biological, but also physical due to compaction under very high hydraulic loads), but clearly this phenomenon was not specifically looked at in our experiment. To clarify that, we have added the following sentence:" *It is not certain that crusts were formed in this controlled experiment, but they do exist in SAT systems*." (P. 9, line 254). We have also changed the text accordingly in two other locations (P.10, line 279; P11, line 320).

10. *Line 278: If the water content in the whole column decreased, where did the water go? Could one observe a larger outflow from the column that matches the decreased water content?*

    Thanks for this observation. Drainage from the system was not measured in high enough temporal resolution to confirm enhanced drainage (which makes sense, considering the elevated air pressure), but clearly during air injection water drained from the system bottom, while the air did not allow further water recharge from top to compensate for that drainage. The following text (P.10, line 271-272) was added for clarification: "*The drop in water content is due to air impeding infiltration, while gravitational drainage continues (drainage measurements were not conducted in high enough resolution to confirm enhanced drainage due to air pressure)*."

11. *Line 350-351: This conclusion is here a bit misplaced and should better be moved to the end, when also the water quality was discussed.*

    Thanks for this suggestion. We have moved this conclusion (that also includes operational suggestion) from section 3.1 to the conclusions (P.19 lines 575-582)

12. *3.2: 'Effluent quality', not 'effluent quality'*

    Typo corrected. Thanks (P.12, line 354)

13. *A general comment: Would it for practical application not be important to study the range of influence of an air injection point? In the experiments, air influenced the water quality in the whole column, but as is later written in the discussion, will probably have an influence only to a certain depth. The column had a not too big cross section, so that the air flow was directed mainly vertically. In a real field a large area would have to be reached. This might involve a lot of injection points.*

    This is very accurate. Our intention in a follow-up experiment, in field conditions and with non-synthetic wastewater, is to experiment with both depth and separation of air injection laterals. Such data, used to calibrate a numerical model, would allow upscaling and optimization at full scale. This intention is clearly mentioned in the conclusions (P. 19, line 600-602)