Review of the manuscript Impact of parameter updates on soil moisture assimilation in a 3D heterogeneous hillslope model by Natascha Brandhorst and Insa Neuweiler

Summary

The manuscript is well structured and written. Many studies have already been conducted on the topic, but I believe the present study provides some new insights. I have listed below only one general comment that would require additional analyses and some minor comments that should be considered to improve some descriptions and to strengthen the discussion. After that, the manuscript can be considered in my opinion for publication.

General main comment

[1] The effect of the DA in the validation points are generally poor. There is no specific information on the correlation length (L) of the generated random fields (L336) but I hypothesize that L is lower than the distance (d) between the location of the assimilated observation and validation observation. Thus, it should be interesting to see at which distance (d) DA improves the estimated soil moisture. There is no the need to run any other simulation but rather to calculate, e.g., RMSE at increasing d. It should be interesting to see and discuss if/when d > L the effect of DA is poor. On the one hand, this would help in defining where to install point-scale measurements. On the other hand, this would support the discussion of the low representativeness of point-scale measurements and the need for alternative soil moisture observations (e.g., L625).

Specific comments in order of appearance (Line number L)

L37. The problem of non-uniqueness due to insufficient observations is identified also in the present study. Moreover, however, the limited representativeness of point-scale soil moisture measurements is also highlighted. This could also be discussed in the conclusions of the present study.

L85. Due to the limited improvements of the DA when more realistic heterogenous cases are performed, I wonder if the simplified approaches proposed in literature (by the application of Miller scaling (Bauser et al., 2020) or global calibration coefficients (Shi et al., 2014)), might be preferable. I'm not asking to conduct any additional simulations or DA tests but these approaches could be further recall in the discussion and conclusions, i.e., what do the Authors think about using these approaches in the lights of the results obtained in the present study?

L233. A few details about the high performance computer and the computational resources used for these tests might be useful to highlight the effort for performing the simulations in the present study.

L240. Evaporation is prescribed and this might be one reason, in my opinion, of the instability of the simulations. If this is the case, I suggest the Authors extending the discussion based on that (e.g. at L385; at L431).

L233. Please specify here the thickness of each soil layer.

L246. If possible, please justify why you have used 181 days.

L250-252. The Authors well acknowledge that the experiments have been conducted eliminating some unwanted sources of uncertainty. Thus, it should be argued that, in real test cases, the results could be even worse than the one presented here. I would recall this aspect in the discussion and conclusion.

L291. How soil set-up is created (the random fields) is not well described. Information is reported only later (L335-338) but without information of the correlation length. I believe these are important details as different results can be obtained with different set-up. Thus, this information should also be presented at the beginning of the section.

L339. I suggest adding here a title "3.4 Performance metrices". I would then move the title "4. Results and discussion" before L354.

L420-421. Correlations between parameters and states at the validation locations and the observations are too small to induce an update of the former. This in my opinion could be related to the correlation length of the random field (L). See general comment #1 above. If this is the case, the discussion should be extended accordingly.

Figure 9. Not sure if I missed something, but I did not get why reference values are not plotted here. If possible, I would also add these lines (as red lines in figure 6). Discussion should be extended accordingly at L419.

L336-341. I would move NMD description to the method section after the description of the RMSE (i.e., L353). All together this sub-section should be named e.g., "3.4 Performance metrices"

L525. Point instead of colon?

L588. The term cumulative value could be in my opinion misleading. Please specify, e.g., variance of the field, cumulative density functions etc.

L614-615. Updating the saturated hydraulic conductivity turned out to be less important. This might be related to the upper boundary condition used (i.e., figure 2) while the results could be different when other hydrological conditions are prescribed. If this is the case, I would extend the discussion accordingly.

L620. Cumulative quantities. Please be more precise (see also comment above L588).

L625. Here I assume that you mean soil moisture observation from remote sensing and cosmic ray neutron probe. I would rephrase to be more precise.