Brief description of the modifications of the manuscript

The major modifications of our manuscript are

- More relevant literatures and explanations are added in the introduction to highlight the enhancement of our study.
- An additional scene with a series of wide prior parameter distribution cases are added in the revised manuscript. And more detailed discussions are added in section 4 to test the effects of prior distribution of parameters on the performance of estimation.
- Dual-parameter estimation cases are added to give more detailed discussions about multi-parameter estimation analyses.

Detailed responses

Anonymous Referee #1

General Comments

1. The authors apply the augmented ensemble Kalman filter to a 3-layer land surface model, called AVIM, to update both the soil moisture states and soil parameters simultaneously. The authors show that while the unconstrained 3-parameter estimation scheme improved the soil moisture prediction moderately with the parameters failing to converge to the true values, the constrained estimation scheme succeeded both in improving soil moisture prediction and in estimating soil parameters. This work nicely exhibits the values of combining parameter estimation with the model state update using the augmented Kalman filter technique in hydrological modeling and introduces a novel way to constrain parameter perturbation errors. The suggested constrained perturbation scheme has reduced the degree of freedom of the Kalman update scheme facilitating the convergence of the update. The topic of the research is relevant to the scope of HESS and the idea of constraining the parameter perturbation scheme is novel. This reviewer however has several comments on some technical parts of the method used in this work. I recommend the publication of this work after resolving the comments summarized below. Also, the manuscript contains noticeably many typos and grammar errors. I have listed some of them in the Specific Comments below (they do not include all the typos and errors, though).

The AVIM model was perturbed by adding zero-mean Gaussian errors to 1) soil parameters; 2) soil moisture initial conditions; 3) model forcing data (precipitation and short-wave). With given figures and tables, it is very difficult to tell if the perturbation scheme has created enough ensemble spread of soil moisture, which basically represents the model prediction error of soil moisture. Inferring from the comparison between "True", "Prior", and "Obs" soil moisture values in Figure 1, the perturbed model prediction of soil moisture seems to have smaller error than the observation until July, which implies insufficient model perturbation. Unlike the atmospheric models, most hydrologic models tend to suppress the large errors in soil moisture, which leads to substantially decreased initial soil moisture errors after several rainfall events. Consequently, soil moisture contents need to be perturbed continually in order to maintain the targeted level of mode prediction error. In relation to this comment, please clarify if the "Prior" soil moisture values in Figure 1 are the averages of ensemble soil moisture or a sample ensemble member from 100 ensembles.

A: We summarize this comment in following items and answer below one by one:
1.1 The manuscript does not explain clearly if the perturbation scheme on 1) soil parameters; 2) soil moisture initial conditions; 3) model forcing data (precipitation and short-wave) to the AVIM model can create enough ensemble spread of soil moisture.
1.2 Please clarify if the "Prior" soil moisture values in Figure 1 are the averages of ensemble soil moisture or a sample ensemble member from 100 ensembles.
1.3 Figure 1 shows that the "Prior" soil moisture seems to have smaller error than the soil moisture observation until July, which implies insufficient model perturbation.
1.4 Soil moisture contents need to be perturbed continually in order to maintain the targeted level of mode prediction error.

1.1 The manuscript does not explain clearly if the perturbation scheme on 1) soil parameters; 2) soil moisture initial conditions; 3) model forcing data (precipitation and short-wave) to the AVIM model can create enough ensemble spread of soil moisture.
A: We agree with this comment. The ensemble spreads of "prior" soil moisture was not given in the Figure 1 in the first manuscript. Also, there had some errors in the descriptions of how to get "prior" states of soil moisture in the third paragraph and Table 1 in section 3 of the first manuscript. Errors in initial soil moisture condition were imposed by replacing the "true" states with assumed imperfect values (i.e., 0.12 cm^3/cm^3) and adding zero mean Gaussian noise with a standard deviation of 50% of the assumed values. The "prior" soil moisture was the mean of ensemble "prior" integrations. We modified these statement errors in the revised manuscript and added ensemble spreads of "prior" soil moisture of both layers in Figure 1. The revised Figure 1 showed that, "prior" soil moisture states for both two layers had enough ensemble spreads by adding Gaussian perturbation to the initial soil moisture for both two layers and to precipitation and radiations once daily. Please see the revised manuscript for more on this comment.

1.2 Please clarify if the "Prior" soil moisture values in Figure 1 are the averages of ensemble soil moisture or a sample ensemble member from 100 ensembles.
A: We thank for this comment. The "prior" soil moisture values in Figure 1 of the original manuscript were a sample ensemble member from 100 ensembles. We changed them to the averages of ensemble soil moisture as showed in Figure 1 of the revised manuscript.

1.3 Figure 1 shows that the "Prior" soil moisture seems to have smaller error than the soil moisture observation until July, which implies insufficient model perturbation.
A: According to the Table 1 and Figure 1 in the manuscript, the "prior" soil moisture states had much more dry initial condition and much larger value of empirical parameter *b* than that of the "true" soil moisture states.

Dry errors in the initial conditions caused negative bias in soil moisture states. Positive error in empirical parameter *b* caused positive bias in soil moisture states. And that the effects of initial condition errors decrease with time, while the effects of errors in parameter *b* increased with time. Therefore, in a certain model integrations period (i.e., the first four months in Figure 1 of the revised manuscript), positive error from parameter *b* and negative error from initial conditions in the "prior" states may offset each other.

This is the mainly reason that caused the “prior” soil moisture seem to have smaller error than the soil moisture observations in this period.

1.4 Soil moisture contents need to be perturbed continually in order to maintain the targeted level of mode prediction error.
A: In our manuscript, we used error from three sources 1) soil parameters; 2) soil moisture initial conditions; 3) model forcing data to represent uncertainties in soil moisture contents. According to the answer of comment 1.1, those perturbations could create enough ensemble spread on soil moisture for both two layers in the whole experiment period (Figure 1 of the revised manuscript). Therefore, it is not needed to perturb soil moisture contents to get additional prediction error in model integrations.

2. It is stated that the precipitation was perturbed using zero-mean Gaussian noise with 20% of standard deviation. Due to the bounded nature of the precipitation (i.e., precipitation cannot have negative values), negative precipitation after perturbation would have been trimmed to zero. This can cause positively biased soil moisture. Is this part of the reason Prior_sm1 and Prior_sm2 in Figure 1 are positively biased in the drying down periods? Judging from the saturated hydraulic conductivity for the “Prior” run, which is about an order of magnitude greater than the “True” run, the significant positive bias is not readily understandable.
A: Yes, it may be one part of the reason. But, of course, this part is not the main contribution. Errors in parameter b, saturated hydraulic conductivity and soil moisture suction are the governing reasons for the significant positive bias in soil moisture.

Comparing to the “true” run (Table 1 in the manuscript), large positive biases were existed in parameter *b*, saturated hydraulic conductivity, and the magnitude of saturated soil moisture suction in the “prior” run. Based on the equations (2) and (3) in the manuscript, larger value of parameter b leaded to smaller unsaturated hydraulic conductivity and larger unsaturated soil water suction because the ratio of unsaturated to saturated soil moisture was always less than 1.0. Positive bias in saturated hydraulic conductivity could cause water flow between different soil layers more easily. Larger magnitude of saturated soil moisture suction made the drying down processes more slowly. Therefore, combined effects of errors in all those three parameters, there were significantly positive bias in soil moisture for both layers in the drying down periods.

3. According to the description of AVIM in the section 2.1, the model has three soil layers of 0.1, 0.9, and 3.6 m, respectively. However, it says in the deepest layer of 3.6 m, “both soil water flux and heat flux are assumed to be zero with constant soil moisture and temperature”. I request the authors to check this boundary condition carefully again. If the bottom layer has the no-flow boundary, the upper layers of 1 m in total would be quickly saturated after a few sizeable rainfall events. Most land surface models have a “free-drainage boundary condition” in the bottom layer.
A: We thank for this comment. There is our mistake in the manuscript. The AVIM model has the free drainage bottom boundary condition and zero heat flux condition at the bottom layer. The new sentences is “Free drainage assumptions is used for the bottom

layer."

4. Due to the way how the RRE (Relative Root Mean Squared Error), it is not possible to directly compare the errors between the cases with different update frequencies (i.e., 1-day, 10-day, : : : , 40-day updates). It seems to me that the relatively high RRE of the 10-day case originates from the degraded soil moisture updates by assimilating the observed soil moisture not because the enhanced performance of the parameter estimation. The overall improvements in term of RMSE (not of RRE) by 10-day updates should be summarized with the 1-day-update RMSE, so that the reader can compare their improvements directly.

A: According to this comment, we added the comparison of RMSE between the cases with different update frequencies in Table 3 and gave relative statements in the last paragraph of section 4.3 in the revised manuscript. Please see the revised manuscript for more about this comment.

5. I recommend the authors to replace the term 'idealized twin experiment' with the 'identical twin experiment', and 'constraint-based'/'post-constrant' with simply 'constrained' (e.g., 'constraint-based simultaneous state-parameter estimation' -> 'constrained state-parameter estimation').

A: We thank for this comment. The bad statement in our manuscript is changed according to this comment.

Specific Comments

Page 1434, lines 14-17 This sentence is very difficult to understand, even after reading through the entire manuscript. Please rewrite it with more plain terms. Also, significant -> significantly.

A: The bad statement in our manuscript is changed. The new sentence is "Increasing the number of estimated parameters, which cause more degree of freedom and imbalances in the EnKF assimilation processes, lead to a significant decline of the performance of parameter estimation."

Page 1434, line 19 temporal-sparse -> temporally sparse.

A: We corrected according to the comment.

Page 1434, line 22 ranging from 1-day to 40-days -> ranging from 1 day to 40 days.

A: We corrected according to the comment.

Page 1435, line 6 instrument -> tool.

A: We corrected according to the comment.

Page 1435, line 18 increasing amount of research attention -> increasing attention.

A: We corrected according to the comment.

Page 1435, line 22 clarify what "flow-dependent" means.
A: We corrected according to the comment. The new sentence is "it provides a flow-dependent background error covariance which obtained at each update and adjusts the background to newly available observations optimally."

Page 1436, line 5 There are two main weaknesses exist -> There exist two main weaknesses.
A: We corrected according to the comment.

Page 1440, line 10 Actually -> In fact
A: We corrected according to the comment.

Page 1440, line 18 the parameters which do not contained -> the parameters, which are not contained.
A: We corrected according to the comment.

Page 1440, line 19 post-constrained update -> constrained update.
A: We corrected according to the comment.

Page 1442, line 9 2-yr spinup -> 2-year spinup. Did you run the model twice for the period of 1 Jan. 1998 – 31 Dec. 1998?
A: We corrected according to the comment. No, we didn't run the model twice for the period of 1998. the 2 year spinup period is from 1 Jan. 1996 – 31 Dec. 1997.

Page 1444, line 2 chose -> chosen
A: We corrected according to the comment.

Page 1444, line 3 Cosby's paper -> Cosby et al. (1984).
A: We corrected according to the comment.

Page 1444, line 18 one-daily -> once daily or once-a-day
A: We corrected according to the comment.

Page 1446, line 6 been -> being (or just leave it out).
A: We corrected according to the comment.

Page 1446, line 18 It is shows -> It shows.
A: We corrected according to the comment.

Page 1447, line 8 can not converge -> do not converge.
A: We corrected according to the comment.

Page 1447, line 10 Different from -> Unlike.
A: We corrected according to the comment.

Page 1447, line 16 added on -> added to. Clarify what "inherent balance relationships"

means.
A: We corrected according to the comment. Here, the "inherent balance relationships" means "statistical relationships between parameters k_{sat}, ψ_{sat}, and b ". We changed this confused statement in the revised manuscript.

Page 1447, line 20 (figures are not shown) -> this result should be presented in the revised manuscript.
A: Dual-parameter estimation cases are added to give more detailed discussions about multi-parameter estimation analyses in the revised manuscript.

Page 1448, line 3 remove "been".
A: We corrected according to the comment.

Page 1448, line 13 potentialities -> the potentials.
A: We corrected according to the comment.

Page 1448, line 21 remove "been".
A: We corrected according to the comment.

Page 1448, line 29 guess -> hypothesis.
A: We corrected according to the comment.

Page 1449, line 9 what does "concerned" idealized twin experiments mean?
A: It means "related". We removed this bad statement in the revised manuscript.

Page 1450, line 1 40-days -> 40-day.
A: We corrected according to the comment.

Page 1450, line 15 different -> difficult?
A: We corrected according to the comment.