Response letter of hess-2019-324-RC2

Please find the responses to the comments.

Comments made by the reviewer were highly insightful. They allowed me to greatly improve the quality of the manuscript. I described the response to the comments.

Each comment made by the reviewers is written in *italic* font. I numbered each comment as (n.m) in which n is the reviewer number and m is the comment number. In the revised manuscript, changes are highlighted in yellow.

I trust that the revisions and responses are sufficient for my manuscript to be published in *Hydrology and Earth System Sciences*

Responses to the comments of Reviewer #2

The author investigated the data assimilation with a 3-D hyperresolutin land model named as ParFlow using ETKF on the various scenarios. Although I think that this manuscript is well written, I have some comments for publication.

Major comments

(2.1) 1. Ll. 316-321. Each ensemble member has different saturated hydraulic conductivity and rainfall rate using random numbers from lognormal distribution with mean = 0 and standard deviation = 0.15. Why does the author choose them? Does the author confirm their sensitivities? Please address the reason simply.

→ I chose them because this setting gives the sufficiently large error in precipitation considering the real-world applications but it does not introduce the strong non-Gaussianity to the precipitation data. This point was described in the discussion section of the original version of the paper:

> "Please note that the parameters of the lognormal distribution to model the uncertainty in rainfall were specified to make the rainfall PDF similar to the Gaussian distribution. I chose the lognormal distribution in order not to generate negative rainfall values and I intended not to introduce non-Gaussianity into the external forcing. The rainfall input which follows the Gaussian PDF was transformed into the non-Gaussian PDF of the background error by the strongly nonlinear dynamics of the topography-driven surface flow."

In the revised version of the paper, I have clarified this point in the section of experiment design when the lognormal distributed multiplicative error was introduced.

> "The two parameters of the lognormal distribution, commonly called μ and σ, were set to 0 and 0.15, respectively. These parameters were chosen to give the sufficiently large error in precipitation and saturated hydraulic conductivity. In addition, this setting makes the rainfall PDF similar to the Gaussian distribution, which is important to interpret the results of the experiments (see the discussion section)."

Although the multiplicative error used in this study was sufficiently large, the problem gets more difficult when the parameter was set to give larger uncertainty to rainfall. In addition, I could choose the parameter to give the biases and/or non-Gaussianity to the input rainfall, which may make the problem more difficult. Since these points are obvious and have been confirmed even in the conventional land surface models, I do not include the results with the different parameters. However, the specification of the prior uncertainty of rainfall and hydraulic conductivity must be important toward the real-world application, which has not been clarified in the previous version of the paper. In the revised version of the paper, I have clarified this point.

> "Although the prior uncertainty in rainfall and saturated hydraulic conductivity was arbitrary chosen in this study, the specification of the prior knowledge is not straightforward in the real-

world applications."

(2.2) 2. L. 365: RMSE is calculated by using all members, not an ensemble mean. Usually, I think the RSME is calculated by difference between an ensemble mean and truth. Although I guess the author's RMSE is better for the author's experiments, please explain why the author use all members for RMSE.

→ The NoDA has a very large spread since I gave the large uncertainty to the input rainfall and hydraulic conductivity. Despite the large spread, the ensemble members in the NoDA experiment are distributed around the truth since the model has no chaotic behavior. Therefore, I should evaluate not only if ensemble mean is consistent to the synthetic truth, but also if the ensemble spread is appropriately reduced. Therefore, here I evaluated the all members. This point was indeed unclear in the original version of the paper and I have included this point in the revised version of the paper.

> "I used all ensemble members to calculate RMSE because I should evaluate not only if the ensemble mean is consistent to the synthetic truth, but also the extremely large ensemble spread simulated in the NoDA experiment is appropriately reduced."

I have also used the other metrics to evaluate if the ensemble spread is appropriately reduced as the response to the other reviewer comment in the previous round of the review although the results have not been included in the main manuscript.

> "The other limitation of this study is that I could not thoroughly evaluate the skill of the ensemble data assimilation to quantify the uncertainty of its prediction. Following Abbazadeh et al. (2019), I calculated the 95% exceedance ratio and found that the ensemble forecast was systematically overconfident (not shown). In the synthetic experiments of this study, the number of rainfall events was small, and the timing and magnitude of rainfall were not diversified. Due to this limited amount of data, it is difficult to deeply discuss the accuracy of the quantified uncertainty by data assimilation. While the skill of lumped hydrological models was often evaluated by the probabilistic performance measures such as the 95% exceedance ratio (e.g., Abbazadeh et al. (2019)), the uncertainty quantification of the simulation of hyper-resolution land models is in its infancy. How surface lateral flows affect the accuracy of the uncertainty quantification by data assimilation should be investigated using more realistic data."

(2.3) 3. Ll. 461, 665-666: In my understanding, the ensemble Kalman filters (EnKFs) do not assume the Gaussian PDF and linearity. The EnKFs derive an optimal value under the Gaussian PDF and linearity. This does not mean assuming the Gaussian PDF and linearity.

→ I fully agree with this comment. This point was correctly described in the method section of the original version of the paper:

"It should be noted that the equations (10-13) give an optimal estimation only when the model and observation errors follow the Gaussian distribution. When the probabilistic distribution of the error in either model or observation has a non-Gaussian structure, results of the Kalman filter are suboptimal. This point is important to interpret the results of this study."

However, in the results section, I described that EnKF assumed the Gaussian PDF and linearity. In the revised version of the paper, I have modified this point by describing:

"As I mentioned in section 2.2, in the ETKF, the estimation of the state and parameter variables is optimal if and only if the model's error has the Gaussian PDF and the relationship between observed variables and unobserved variables is linear."

"Please note that the low representativeness of the soil moisture observations in the case of the HIGH_K reference shown in section 3.1 is due to the limitation of the Kalman filter that the error PDFs need to follow the Gaussian distribution to get the optimal estimation so that the increase of the ensemble size cannot solve this issue."

(2.4) 4. L. 616: "there are large errors in the area around 500<=x, y <=1500" I have trouble with this sentence. I cannot confirm the large errors in Fig. 9.

→ Since Figure 9 shows the improvement rate by data assimilation, the large errors cannot be directly found in Figure 9. This point was indeed unclear in the original version of the paper. I have clarified this point by simply mentioning that the large errors are not shown in the figures.

"In the noOF configuration, there are large errors in the area around 500<=x, y <=1500 (not shown) since the increase of soil moisture in this area is caused by the topography-driven surface flow which is neglected in the noOF configuration."

(2.5) 5. Ll. 688-699: Assimilating just one observation improves the analysis errors in whole domain despite the nonlinear equations. This means that the model has long-range spatial correlations. Therefore, I guess the ETKF works well without the localization. Also, the author mentioned that the localization scale depends on the model parameter. In order to confirm those, the author should investigate the spatial correlations.

→ I realized that this description was misleading. Because the negative impact of the non-linear and non-Gaussian on the state estimation can be found only in the edge of the area where topography-driven surface flow reaches (Figures 9b and 9c), I suggested omitting to update the state variables there as the heuristic approach. It is not identical to the localization method in the context of data assimilation, in which the spurious correlation sampled by an ensemble is eliminated by restricting the impact of assimilating observation. Therefore, what I suggested is not directly related to the spatial correlations. As the reviewer mentioned, the model has long-range spatial correlations except for the edge of flooding area. In the revised version of the paper, I have clarified this point.

> "As a possible heuristic approach to avoid the negative impact of the non-Gaussian background PDF, I can omit to update the state variables in the edge of the area where topography-driven surface flow reaches. The numerical experiments clearly indicate that the negative impact of the non-linear physics and non-Gaussian PDF is found only in the edge of flooding areas so that it is beneficial to simply omit to update the state variables in this area. It is similar but not conceptually identical to the localization method, in which the spurious correlation sampled by an ensemble is eliminated by spatially restricting the impact of assimilating observation (e.g., Rasmussen et al. 2015; Anderson 2007; Bishop and Hodyss 2009). "

(2.6) 6. Figure 8: In the OF configuration of Fig. 8 (a) and the noOF and OF configurationsof Fig. 8 (b), the DA_obs1 and DA_obs9 experiments have almost the same RMSE although the DA_obs9 experiments have 9 times observation information. Why?

→ This point was not mentioned in the paragraph where Figure 8 appears. However, I mentioned it later in the original version of the paper.

> "Figure 9b shows that the increase of the number of observations substantially improves the soil moisture simulation in the region which is affected by topography-driven surface flow (see also Figure 6c). However, the skill to simulate soil moisture is severely degraded in the lower-left corner of the domain, which causes the stalled improvement from the OF_DA_obs1 experiment to the OF_DA_obs9 experiment shown in Figure 8a."

In the revised version of the paper, I have provided the guidance for the readers to make it easier to find this statement.

> "The OF_DA_obs361 experiment has the smallest RMSE so that a denser observing network is beneficial to estimate soil moisture, although there is the stalled improvement from the OF_DA_obs1 experiment to the OF_DA_obs9 experiment (the reason for it will be explained later)."

(2.7) 7. Figure S4: A green line looks like to splits into a single outlier and the others. If so, Ithink this is ensemble clustering (EC, Anderson 2010, Amezcua et al. 2012). The EC is frequently generated by ensemble square loot filters included the ETKF and may be related to the non-Gaussian PDF. Therefore, please refer to the EC in section 4.

→ Thanks for the comment. Figures S1-S4 indeed show the ensemble clustering, which strengthens the conclusion that the non-Gaussian PDF has an important role in hyperresolution land data assimilation. I have included this discussion in the revised version of the paper.

> "In addition, I found ensemble clustering in which the ensemble members are split into a single outlier and the others (see Figures S1-S4). The previous studies found that this ensemble clustering is generated by the non-Gaussian PDF (Anderson 2010; Amezcua et al. 2012). Ensemble clustering shown in the analysis timeseries also implies that the non-Gaussian PDF plays an important role in the data assimilation of the hyperresolution land model."

Minor comments

(2.8) 1. Equation numbers are confused. For instance, Eq. 4 is written on the lines 141 and 148. Please correct the all equation numbers.

→ The numbering of the equations in the original version of the paper was indeed confusing. I have fixed this point following the reviewer's instruction.

References

Anderson, J. L.: A non-Gaussian ensemble filter update for data assimilation, Mon. Wea. Rev., 138, 4186-4198, 2010.

Amezcua, J., Ide, K., Bishop, C. H., and Kalnay, E.: Ensemble clustering in deterministic ensemble Kalman filters, Tellus, 64A, 1-12, 2012.