Anonymous referee #1

We thankfully acknowledge Referee #1's useful comments, which significantly helped us improve the presentation of the study. We report below our replies (denoted as AC, Authors' Comments) to the referee's comments, followed by brief descriptions of the changes we have made.

Brief description of the modifications to the manuscript

The major modifications of our manuscript are:

- the lagged regularized particle filter are improved in terms of kind of kernels, aggregation method of lagged weights, and the computational speed; and
- the experimental design of the study is improved to compare forecasts for varying lead times and process noises and independent observations with additional events.

Major Changes/Questions/Concerns:

1. *Abstract: Please mention the measurements assimilated. Also, please include some descriptive statistics/results so that the reader has a relative idea of the improvement when using your technique.*

   AC: We thank you for this comment. We reconstructed the abstract to describe the major findings of this study, including descriptive results and improvements.

2. *P.3385, l.23-26: Kalman-type filters do use a "linear correction step" but they are also "applicable to non-linear … state-space models" (e.g., ensemble Kalman filters). Please modify this sentence to state that SMC methods have the advantage of non-Gaussian state-space model, but remove the part about non-linear state-space models.*

   AC: You are right. Although there are some non-linear cases in the literature showing that SMC methods clearly outperform ensemble Kalman filters, it is still controversial in the practical applications for hydrology and water resources. We modified the sentence so that now it only refers to non-Gaussian models.

3. *Last two paragraphs in Section 1: Is there any way to merge/reduce the size of these paragraphs (e.g., remove redundancies)? For instance, P.3386, l.29-30 is essentially equivalent to P.3387, l. 13-15. Please pare down these paragraphs where possible for economy.*

   AC: We modified the paragraphs, removing redundancies.

4. *P. 3394, l.16-25: Does this mean the same measurements are assimilated more than once? I don't believe so, but I could not tell from Figure 2. Please clarify.*

   AC: The regularization step is not the re-assimilation of the same measurement. This approach gives small jittering in the resampling step to improve the diversity of particles. However, it is true that the state-space model is re-evaluated once more for the same time window in the regularization step. We modified Figure 2 to illustrate the concept of the regularized particle filter in detail.

5. *P. 3396, l.11-12: What is the remaining 8% of land characterized as?*

   AC: We re-calculated and changed the description about the percentile of land use. The new sentence is, "The land use consists of forest (76.7%), agricultural area (9.3%), residential area (7.5%), water body (2.0%), public area (2.7%), vacant land (1.2%) and road (0.6%), respectively." We apologize for the inaccurate information in the first manuscript.

6. *P. 3397, l.9-13: How are the fields generated (e.g., Kriging, nearest neighbor interpolation, etc.)? Please elaborate as to how the meteorological values are defined between the 13 observation stations.*

   AC: The nearest neighbour method was used to apply rainfall for each grid from 13 observation stations. We added the description in the end of the second paragraph of section 4.2.

7. *P. 3397, l.15-24: What are the boundary conditions for the groundwater? Constant head? Constant flux? Please elaborate with a sentence or two.*

   AC: We thank you for this comment. No flux boundary condition was specified at the catchment boundary for the groundwater flow. This description has been added in the middle of the fourth paragraph of Section 4.2.

8. *Equation (20): Why the use of additive error as opposed to multiplicative error (e.g.)? How might the selection of measurement error impact your results? Please elaborate with a sentence.*

   AC: We use error of the total soil moisture depth as an additive way in Eq. (18) and apply that for each grid in a multiplicative way (Eq. (21)). We think this procedure can provide proper perturbation, considering the spatial variations in a distributed hydrologic model. Similar noise definition for soil moisture has been applied for state updating of a distributed hydrologic model in the study of Kim et al. (2003). The multiplicative error is used for the perturbation of overland flow fields because the order of state variables is quite different in the upstream and the downstream, whereas the additive error is more reasonable in the case of soil moisture because it moves within limited ranges (e.g., from wilting point to porosity). In the case of measurement error, we follow the assumption widely implemented in the articles mentioned in the manuscript.

9. *P. 3402, l.14-16: "Narrow confidence intervals" are interpreted as an "enhancement of the probabilistic forecast." Is it possible that this is a result of too little a priori errors? Is there additional evidence you can provide that would strengthen your argument regarding an enhanced probabilistic forecast?*

   AC: In the case of real experiments in a catchment scale, it is hard to estimate and demonstrate which confidential intervals are more proper. In the revised manuscript, we eliminated this analysis and changed the experimental design to focus on more practical issues such as forecasts for varying lead times and stability of particle filters for the process noise. "Confidential intervals" are also discussed from different standpoints, including their sensitivity and applicability according to uncertain process noises as shown in Fig. 10.

10. *P. 3403, l. 19-20: Did you conduct any sensitivity analyses of "two step" versus "three step" versus "four step" ahead prediction? If so, did the number of steps influence the prediction capability? Please elaborate with a sentence or two.*

    AC: We redesigned our experiment and conducted sensitivity analysis for varying lead times up to a 24-hour-lead forecast. The results are summarized in Figs. 12 and 13. The comparison of two particle filters for varying lead times in a flood event is also illustrated in Fig. 11.

11. *P. 3409, Table 1: Which station(s) are used in this comparison? Were the streamflow measurements used here the same ones that were assimilated? What about the other three (3) stations besides Katsura? Where these ever used in the experiments? Please clarify with a few sentences.*

    AC: We thank you for this comment. In the revised manuscript, we described the station and data period for all the simulation results to avoid confusion. Since model efficiency is estimated for the entire lead times, Table 1 is not needed; instead, Nash-Sutcliffe efficiency is plotted in Figs. 12 and 13.

12. *P. 3409, Table 1: How similar (e.g., total rainfall amount) was the rainfall during the calibration and verification periods? How might that impact the model performance? What is the calibration period was "dry" and the validation period was "wet." Please elaborate.*

    AC: All data periods simulated in this study were "wet periods", since this study was focused on flood forecasting. Soil parameters and lag time were roughly estimated for the data period of 2007, which was also a "wet" condition. If those are estimated for the dry period, accuracy of forecasts for wet periods may drop. Extended simulation periods and observed flows are shown in Table 1.

13. *P. 3414, Figure 5: All of the stream gauges are near the basin outlet. Where any upland stations available? Since these gauges essentially represent the same integrated hydrologic response, the information content in one gauge could be comparable to another gauge. How might the availability (or lack thereof) of upland gauge information impact the results? Please elaborate with a sentence or two.*

    AC: Among all stream gauges, it was not possible to get reliable discharge data except for Katsura and Kameoka due to high uncertainty and missing data. We added forecasts at the Kameoka station, located in the uppermost part, although Kameoka appeared to have still some uncertainty compared to Katsura.

14. *P. 3416, Figure 7: The rainfall record indicates precipitation events more or less every day. How might this strategy perform during drought periods when antecedent soil moisture conditions could have a significantly different effect on the hydrologic runoff response? Please elaborate with a sentence or two.*

    AC: Although this study was focused on flood forecasting, we could get improved forecasts even in drought periods (not shown in this paper). We think the lead time for the drought period can be extended when the spatial heterogeneity of the process noise is considered to estimate more appropriate antecedent soil moisture conditions in data assimilation.

15. *P. 3417, Figure 8: Again, are you comparing against the same gauge measurements that you assimilated? If so, then the results should always look better than the deterministic case. Any comparison that can show the efficacy of your technique should use independent observations for analysis. I found this to be the most confusing part in your paper, which could greatly benefit by a clarifying sentence or two.*

    AC: In the revised manuscript, we included comparisons of forecasts and observations at the Kameoka station, which were not used for data assimilation as shown in Fig. 13. As mentioned above, we could see improved forecasts via two particle filters for varying lead times at the "independent" Kameoka station shown in Fig. 13. However, the quantitative analysis between two particle filters was limited due to data uncertainty.

16. *P. 3418, Figures 9b and 9c: Which one is closer to the "truth"? Are there any ground-based soil moisture observations that can be used to answer this question? If not, then what is the value of this figure? Please explain.*

AC: Since this study was a real experiment in a catchment scale and there was, unfortunately, no ground-based soil moisture observation, there was no way to know what the truth was. Despite the importance of inverse modelling to estimate unknown and uncertain state variables, we concluded this would be beyond the scope of this study. We eliminated Fig. 9 and changed our focus to forecasts for varying lead times and various process noises, as mentioned above.

Minor Changes/Questions/Concerns:

1. *There are dozens of missing articles (i.e., "the", "an", and "a") throughout the manuscript. I am not going to comment on these particular grammatical mistakes because they are too large in number and not important to the merits of your technique. However, it may be worthwhile having an English-speaking technical editor briefly review your revised manuscript.*

   AC: We modified the revised manuscript according to the advice of an English-speaking editor.

2. *P.3384, l.13-15: Remove the sentence regarding MPI. It is a nice technical feature (as noted in the main text), but it is unnecessary for the abstract.*

   AC: We modified the sentence accordingly.

3. *P.3386, l.15-19: This is a good point. I recommend you reiterate this statement (to some effect) within the Conclusions as it is strong reminder as to the merits of your approach.*

   AC: Thank you for this comment. We included this in the Conclusions.

4. *Equation (1): Add space between the two "w_k" values in a similar manner as done with the two "v_k" values in Equation (2). This will make it easier for the reader.*

   AC: We corrected accordingly.

5. *P.3390, l.5: "particles" instead of "paraticles"*

   AC: We corrected accordingly.

6. *P. 3391, l7: "posterior" (i.e., one word) instead of "posteri or"*

   AC: We corrected accordingly.

Reference:

Kim, S., Tachikawa, Y., and Takara, K.: Applying a recursive update algorithm to a distributed hydrologic model, J. Hydrol. Eng., ASCE, pp. 336-344, 2007.