I appreciate the authors' responses in answering my questions in the previous iteration. The revised manuscript addresses most of my concerns. In my opinion, the manuscript is nearly ready for publication, with a few minor issues below that need addressing.

Line 336: I am curious about why the locally optimal proposal cannot be applied for the non-linear observation operator? Even if the linear assumption is invalid, shouldn't it still be applicable? (like we can still apply the ensemble Kalman filter to assimilate non-linear observation even if it's not optimal) Please clarify if there's anything that I might be missing here.

Line 344: "The time averaged RMSE results are plotted in 5." -> Figure 5.

Line 344-345: "The effect of the non-linear observation operator can be clearly seen, where the time averaged RMSE for the linear case is lower in all set-ups".

This statement is inaccurate. There are a lot of factors determining the magnitude of RMSE. Assimilating a linear observation (i.e., with linear observation operator) with very large observation error standard deviation can also lead to a large RMSE. When the observation error standard deviation is fixed, using different observation operator leads to different shape of likelihood function, and therefore different magnitude of RMSE. I recommend rephrasing this sentence or just removing it.

Line 449: what are the units for the standard deviation (especially for ps)?

Line 459: "10 hPa" noises in ps seems to be quite large

Figure 9: I am still skeptical about the RMSE values in these experiments. Even in a model run without DA (the lower left panel), the RMSE are unrealistically large, suggesting the large RMSE is not related to the observing system. I suspect that this is a result of the experiment setup that the ps standard deviation in Q is set to 100 (hPa) and the observation error is set to 10 (hPa). I think setting the standard deviation as 100 (Pa) = 1 (hPa) and 10 (Pa) = 0.1 (hPa) could lead to a more realistic representation.