Nonlinear Processes in Geophysics Discussions

Nonlin. Processes Geophys. Discuss., referee comment RC2
https://doi.org/10.5194/npg-2022-5-RC2, 2022

Comment on npg-2022-5

Anonymous Referee #2

Referee comment on "Applying prior correlations for ensemble-based spatial localization" by Chu-Chun Chang and Eugenia Kalnay, Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2022-5-RC2, 2022

The authors propose a localization scheme for prior correlations and compare it with the traditional localization scheme based on distance dependence. This localization scheme is of interest for the implementation of ensemble data assimilation methods. The manuscript is quite well written and meets the submission requirements of the journal NPG. Nevertheless, the manuscript has the following issues that need further clarification and improvement.

Specific comments:

- This new localization scheme for YK18 relies heavily on the statistical formulation of Equation (5), so is there any similarity between this formulation and Anderson's work, and what are their similarities and differences? Please elaborate explicitly.
- Does the statistical result of Equation (5) depend on the number of samples? If so, how much does this sample dependence affect the final results?
- Equation (5) counts the correlation coefficients between the model grid points and the observed points, but we know that the observed variables are hardly fixed in their positions at different moments. This situation is especially prominent when assimilating satellite data in NWP. Since the position of the observed data is difficult to be fixed, the observation operator *H* is actually difficult to be fixed as well. Then how should the correlation coefficients between the model grid points and the observed points, which are calculated by Eq. (5), be applied to other moments?
- Similarly, the model in the validation experiment given so far is very simple, with only one variable. For a true NWP model, there are perhaps multiple model state variables such as *U, V, P, T, Q,* on the same model grid point. And due to the use of different grid schemes, these variables may not appear at the same location of the grid. So how to use Eq. (5) for statistics in this case and apply it to the real situation?

- The authors elaborate that one of the advantages of YK18 is that it is more computationally efficient. However, it can be seen from their analysis that in fact YH18 should essentially provide some new calculations of localization correlation matrices as well, so why does it make the improvement of computational efficiency?
- As for the "a faster spin-up" proposed in the manuscript, I do not quite understand it. The purpose of our data assimilation is to give a more accurate initial field and then drive the model to forecast. The spin-up seems to be more appropriate in the simulation of climate models.
- It seems that Section 2.3 of GDL appearing in Page7 should be Section 2.2.

REFERENCES:

- Anderson, J. L. (2007). Exploring the need for localization in ensemble data

assimilation using a hierarchical ensemble filter. Phys. D 230, 99–111. doi:

10.1016/j.physd.2006.02.011

- Anderson, J. L., and Lei, L. (2013). Empirical localization of observation impact

in ensemble Kalman filters. Mon. Weather Rev. 141, 4140–4153. doi: 10.1175/MWR-D-12-00330.1