



EGUsphere, referee comment RC1  
<https://doi.org/10.5194/egusphere-2022-434-RC1>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on egusphere-2022-434

Lili Lei (Referee)

---

Referee comment on "Guidance on how to improve vertical covariance localization based on a 1000-member ensemble" by Tobias Necker et al., EGU Sphere, <https://doi.org/10.5194/egusphere-2022-434-RC1>, 2022

---

### Summary

This manuscript uses a convection-permitting 1000-member ensemble simulation to examine the vertical localization. An empirical optimal localization is proposed, which minimizes the sampling error of correlations estimated from a 40-member ensemble comparing to those from the 1000-member ensemble. Vertical correlations and localization functions for different state variable and cross variables are systematically examined. Results show that different vertical localization functions are required for different variables and vertical height. Combination of the empirical optimal localization with adaptive sampling error correction is also investigated. The manuscript is well-written, and could be a valuable contribution for the data assimilation community. I have several comments as below.

- l22, It is more appropriate to say members of  $O(100)$ , since Canadian Center has 2 groups of 128-member ensembles.
- l100-105, the description of BCs is confusing. Is the GEFS 20-member analysis used 50 times to get 1000 BCs? Or climatological GEFS is sampled for 1000 BCs?
- l197, there are increased correlations below 800hPa. Are there any physical explanations for this?
- Eq. (5), is  $s$  the index for 40-member groups with  $S=25$ ? Eq. (5) is similar to Eq. (1) in Lei and Anderson (2014).
- If the sample correlation  $r_{40}$  tends to overestimate the true correlation  $r_{1000}$ , the EOL computed from Eq. (7) should be no larger than 1.0? As discussed by Lei and Anderson (2014), ELF can count inflation compared to GGF. But I am not sure about the EOL in Eq. (7), which could be true since the true correlation is known. Can the authors provide some derivations on this statement?
- Figs 2 and 7, how about the sample correlations estimated for cross variables?

- Figs 3-5, the UU EOL seems have values larger than 1.0. What is the exact value at the reference level? Why EOL estimates localization larger than 1.0 when sample correlations are close to 1.0? Intuitively, when sample correlations are close to the true correlations as 1.0, localization value goes to 1.
- I258-260, this discussion is based on sampling errors in correlations. But for cycling data assimilation experiments, too strong taper for cross variables may result in too weak corrections.
- I290, the "error reduction" is for estimated correlation, not for prior/posterior errors by using the EOL. Also it would be helpful to have some discussions about the estimated localization and localization applied for cycling data assimilation in the section of conclusions and discussions.
- Section 3.1, a curious question, if the direct-variable EOL is applied rather than cross-variable EOLs, i.e., UU is applied for UU, UV, UT, and UQ, how about the correlation error reduction?
- I396-399, it would be helpful to add some dynamical explanations for these results.
- I407, the computational efficiency issue can be treated by model-space localization (Lei et al. 2018).