Interactive comment on "On the localization in strongly coupled ensemble data assimilation using a two-scale Lorenz model" *by* Zheqi Shen et al.

Anonymous Referee #2

Received and published: 24 January 2019

General comments

This manuscript considers (distance-based) localization in the ensemble adjustment Kalman filter (EAKF) for coupled data assimilation problems, with a particular focus on a localization scheme in a two-scale Lorenz model. Overall, the manuscript is clearly written and reasonably organized, and also contains some interesting findings and insights, e.g., results with respect to the combinations of different levels of coupling in assimilation, as reported in Table 1 and Figure 8. I am in support of the publication of the current work, after some minor issues are addressed.

Specific comments

1. Between lines 10 – 15, page 7. The authors mentioned that P_{xx} is a diagonal-constant matrix. Does this mean that the off-diagonal elements in P_{xx} are all zero? If so, it seems to me that very "strong" localization is applied in the EAKF. To see this, let's use notations similar to those in Eq. (7) of the manuscript, but here I dropped the index n of ensemble members. Without loss of generality, and regardless of which type of EnKF is used, in general one would have the following update formula

$$\Delta x_m = \sum_{s=1}^{S} K_{m,s}\Delta y_s \,, m = 1, 2, , M; s = 1, 2, , S,$$

where m and s are the indices of model state variables and observation elements in the filter analysis scheme, respectively; M and S are the total numbers of model variables and observations; and $K_{m,s}$ is the element of the Kalman gain matrix on the mth row and sth column. In EAKF, when localization is conducted, it's equivalent to introducing some tapering coefficients $P_{m,s}$ to the update, such that

$$\Delta x_m = \sum_{s=1}^{S} P_{m,s} K_{m,s}\Delta y_s \,.$$

For the authors' specific problem in consideration, one has $M = S$. So "P_{xx} is a diagonal-constant matrix" means that $P_{m,s} = 0$ if $m \neq s$, or in other words, the model variable x_m is only updated using the observation at the same location as x_m. In reality, it may be possible that observations at nearby locations also contain certain information of x_m, so a "weaker" localization scheme may be useful. Although, I do see that, in this case, adding more observations in the update scheme may make localization much more complicated.

My suggestion here is thus to clarify the situation, and discuss the implication when P_{xx} (likewise, P_{zz}) is chosen to be a diagonal matrix. (No action required for the side remark in the sequel) In general, it should be desirable to make the localization scheme more general and more flexible. For this purpose, the authors may wish to have a look at the idea behind the recently proposed correlation based adaptive localization.

Technical corrections (minor issues)

1. Line 22, page 6. In "...inversely proportional to the distances...", "proportional" does not sound accurate.

2. First line, page 7. In "...the prior ensemble member", consider adding "n-th" before "prior".

3. Page 7 mentions "correlation covariance" in a few places. I guess it should be "cross covariance" instead.

4. First line, page 8. In "could beyond...", add "be" before "beyond".

5. Line 21, page 9. Double check the notation $\rho_z Z$.

6. In Eq. (10), define the operator $\otimes$ before using it. It does not seem to be a standard tensor product (between two vectors).

7. Last line of Section 3, page 10. In "...L-variable has equal effect...", it seems "has" should be "have".

8. In the definition of CE (page 10), why $\overline{x^{true}}$ should be squared in the denominator.

C3

9. Line 12, page 11. In "It is possible that the smaller MS-RMSE with SCDA in figure 7b is due to ...", it seems to me "figure 7b" should be "figure 6b" instead. Similarly, Line 20, in "whose results are shown in Figure 9d–f", maybe "Figure 9d–f" should be "Figure 8d–f".

10. Line 13, page 12. In "when N $\leq$ 320", should "$\leq$" should be "=" instead?

11. Line 24, page 12. In "...limited ensemble size", add "a" before "limited". Line 25, add "of" after "the presence".

12. Lines 5 – 6, page 13. Replace "factors" by "factor", and change "a update" to "an update".