

## ***Interactive comment on “Diagnosing spatial error structures in CO<sub>2</sub> mole fractions and XCO<sub>2</sub> column mole fractions from atmospheric transport” by Thomas Lauvaux et al.***

### **Anonymous Referee #2**

Received and published: 10 May 2019

This manuscript describes a method of spatially filtering small ensembles of model forecasts. Ensembles of greenhouse gas (GHG) model simulations are useful for understanding transport errors whose covariances are needed when estimating surface fluxes using inverse modelling. Since GHG flux estimates are frequently needed on long (multi-year or decadal) time scales, large ensembles may be prohibitively expensive. Thus if a small ensemble is reliable, it could be useful for approximating transport error covariances if sampling errors arising from the small ensemble size can be filtered. Thus, the topic of this manuscript has important applications to GHG flux inversion systems. The application of the variance and correlation filtering methods from meteorology to GHG forecast ensembles is novel and, as shown in the manuscript, can

C1

reveal relationships (or lack thereof) between meteorological variables and CO<sub>2</sub>. However, conclusions regarding minimum ensemble sizes for estimating spatial variances or correlations should be carefully qualified. With small ensemble sizes, there is a lack of convergence from day to day that can be ameliorated with temporal averaging (or temporal filtering). This is akin to increasing the ensemble size with members from different days. Therefore, the authors are requested to review the manuscript and ensure all qualifications are presented when recommending ensemble sizes. Because I am not familiar with the iterative filtering methods in this work, my comments are mainly confined to the application of the filtering methods to carbon cycle science.

### Specific comments

1. While the idea of filtering small ensembles is very appealing, I am not comfortable with the criterion for success being solely based on the convergence of the schemes. The quality of the ensemble can be checked by issuing forecasts and comparing these to observations. If this is difficult to do, can the methods be checked using simulated observations?
2. P1, L12-14: “We conclude that. . .” This statement needs qualification. On P14, L7-10 the authors note that for daily convergence a larger number of members is needed. The smaller ensemble seemed to work, only if additional temporal averaging was done.
3. P6, L12: It is stated that a Schur filter will be used for correlations but in line 17, it is applied to covariances not correlations.
4. P6, eq. 8: It would be useful to define (in words) the symbol on the left side of (8). For example, is it an optimality criterion? Ménériet et al. (2015a) also does not define this symbol so it would be useful to add a few words here.
5. P7, eq.12: What is  $\bar{E}$ , in the denominator? The overbar is never defined. Since it does not appear in (13), it may be a typo.

C2

6. P7, L17-18: Using the stated assumption that the true covariance is approximated by the sample covariance (i.e.  $B_{ij}^* = \tilde{B}_{ij}$ ) will not yield (13) from (12). Instead, it will yield 1. If the true covariance  $B_{ij}^* = E(\tilde{B}_{ij})$ , then (13) can result.
7. P8, L14-15: I don't follow this argument. The trace of a covariance matrix (or sum of variances) equals the sum of its eigenvalues. Fast decreasing eigenvalues suggests you are talking about the spectrum of eigenvalues. The spectrum of eigenvalues of the correlation matrix does imply spatial filtering so a steep spectrum implies smoother fields. But what does the spectrum of eigenvalues of covariance matrix say about variance, as a function of ensemble size? Also, why should small ensemble sized in general have larger variances than larger ensemble sizes? Some references or mathematical derivation may be helpful here.
8. P8, L17: "dispersion" of what?
9. P9, L7-8: "Typically, ..." In Fig. 2, the 25-member ensemble has much greater temporal consistency of the optimal length scale. Since temporal correlations also suffer from sampling error, particularly with small ensemble sizes, how much can one infer about the temporal variability of the length scale with ensemble sizes of 5 or so?
10. P10, L7: "Better". Better than what? Presumably the statement refers to a comparison with Fig. 1.
11. Fig. 5 caption: What is the time? Presumably these are monthly averages.
12. P10, L16-17: This suggests you should be able to increase ensemble size using different days.
13. P10, L21-22: "XCO2 variance spatial patterns (Fig. 6c) exhibit distinct maximum values located in the southwestern part of the domain," I don't see this. I see

C3

maxima (orange and red regions) in the eastern part of the domain. Also, the red region south of Lake Huron is similar to that seen in CO2 at 5 km in Fig. 6b. This is somewhat reassuring.

14. P14, L9-10: "One important point here is the calibration step performed before filtering..." The calibration steps in Diaz-Isaac et al. (2018a) and Appendix A were done over 18 June -21 July 2008 which encompasses the period of optimal filtering (Figs. 2-4). How applicable would this calibrated ensemble be to a different time period? My guess is that the calibration might yield a different selection of members for a different time period as the meteorology changes. So would you need to redo the calibration continually with time? Some addition discussion of the calibration process as a function of time might be useful here or in section 4.2.
15. P18, L5: "Figure 10" should be "Figure 12"
16. Figure 12: Please add a legend to the bottom right panel of this figure, to label the curves. One can see which curve is which variable by comparing to the other panels, but a legend would be much more convenient for the reader.
17. P20, L4-6: As discussed in comment 1, the same qualifications need to be made here. Specifically, is this conclusion valid only in the context of additional temporal averaging?

#### Technical comments

1. P1, L4: "of which" should be "whose"
2. There are many instances of "fail at converging" which should be replaced by "fail to converge". i.e. P8L27, P8L30, P9L6, P9L15.

C4

3. P10, L9: “dependence to” should be “dependence on”
4. P12, L13: “similar. . .than” should be “similar. . .to”
5. P13, L10: “difficulty” or “inability”?

---

Interactive comment on Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2018-1113>, 2019.