Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2018-1113-RC1, 2019

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Diagnosing spatial error structures in CO_2 mole fractions and XCO_2 column mole fractions from atmospheric transport" *by* Thomas Lauvaux et al.

Benjamin Ménétrier (Referee)

benjamin.menetrier@irit.fr

Received and published: 25 February 2019

1 General comments

In this article, the authors use covariance filtering methods recently developed in the NWP data assimilation context and adapt them in the atmospheric transport framework. Their goal is to get a better estimation of the forecast error spatial structures for in situ carbon dioxide mole fractions and total column dry air mole fractions. I think that it is a very good idea to import methods from one domain to another, and I congratulate the authors for this effort.

I am not a specialist of atmospheric chemistry, so I will not be able to assess the relevance of the experimental setup used in this paper. In my comments, I will focus on the covariance filtering aspects, since the authors uses the theory that I have developed for my PhD thesis.

I think that some parts of the paper need major clarifications regarding the use and implementation of filtering methods. Detailed remarks are listed hereafter.

2 Specific comments

1. I think that there is a mistake in the definition of the criterion $\mathcal{L}(\mathbf{F})$ line 14. The criterion should be a scalar, so the transpose sign should be put after the first parentheses: $\mathcal{L}(\mathbf{F}) = \mathbb{E}[(\mathbf{x}^\star - \mathbf{F}\widetilde{\mathbf{x}})^\mathrm{T}(\mathbf{x}^\star - \mathbf{F}\widetilde{\mathbf{x}})]$. As a consequence, the partial derivative of $\mathcal{L}$ with respect to F_{ij} is: $\frac{\partial \mathcal{L}}{\partial F_{ij}} = -2\mathbb{E}[(x_i^\star - \widehat{x}_i)\widetilde{x}_j]$. Setting this partial derivative to zero yields: $\mathbb{E}[(x_i^\star - \widehat{x}_i)\widetilde{x}_j] = 0$ for all i and j, so: $\mathbb{E}[(\mathbf{x}^\star - \widehat{\mathbf{x}})\widetilde{\mathbf{x}}^\mathrm{T}] = 0$, which is the result given at line 17. I am not sure whether the total variation of the criterion $\delta\mathcal{L}$ is really relevant here.

2. I don't really understand why the authors assume that the dichotomy algorithm fails at converging if the filtering length-scale becomes larger than 750 km. Indeed, as shown in Ménétrier *et al.* (2015a), appendix C, the filtering length-scale verifying the optimality criterion for a homogeneous and isotropic filter always exists and is unique. As explained in section 8 of this paper, the value of this optimal filtering length-scale can be related to two ratios:

    - the signal-to-noise ratio for the variance (related to the true variance values and the sample size),
    - the ratio between the variance signal and noise spatial variations length-scales (the noise length-scale being related to the forecast error correlation

length-scale, see Raynaud *et al.* (2009))

If the optimal filtering becomes very large, leading to a constant value for filtered variances, it does not mean that the dichotomy algorithm is failing. It is a sign that the noise amplitude is too large compared to the signal variations and/or that the variance signal and noise length-scales are mixed. As a consequence, the filter prefers to filter out all spatial variations and keeps the mean value only (which might be the best option). I suggest a rewriting of the discussion about the convergence failures.

3. The covariance filtering theory leads to fractions of polynomials with factors $(N-3)$ at the denominator, so 4 members is the absolute minimum ensemble size for this theory. I think that a 5-member ensemble is still too small to get significant results, and I would suggest to use ensembles of size 8-10 at least.

4. For the localization of correlation functions, the Schur filters given by equations (9) and (11) are not positive definite functions, as mentioned in Ménétrier *et al.* (2015b). Thus, a fit of the raw localization function with a positive definite function is required before applying the localization. It seems that the raw localization function was used in this paper, which is problematic. Also, it is not clear how the spatial and angular average was performed to estimate the statistical expectations of equations (9) and (11) via an ergodicity assumption.

5. For Figures 8, 9 and 10, the small domains overlap, which makes difficult to see the functions shapes. It would be better to have separate boxes as in Figure 13 of Ménétrier *et al.* (2015b). Moreover, it would be more helpful to plot the localization functions, rather than the localized correlation functions.