Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2016-234-RC1, 2016

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Can we detect regional methane anomalies? A comparison between three observing systems" *by* Cindy Cressot et al.

Anonymous Referee #1

Received and published: 25 April 2016

This manuscript investigates the spatio-temporal resolution of global methane inversions using different observing systems. This information is useful for a proper interpretation of inversion results obtained using existing observing systems, and for the design of new systems. An expected outcome is that larger regions are better resolved than smaller regions. Less expected is the finding that smaller regions are better resolved at the seasonal than at the inter-annual time scale. While trying to understand this, a couple of questions came up, as explained below, which I found have not been dealt with adequately yet. To make this study acceptable for publication this will have to be repaired, and/or explained more clearly.

GENERAL COMMENTS

Figure 3 and the tables depend on a detection criterion, like threshold SNR value which

should be exceeded to declare a region as detected or not. This criterion should be defined explicitly in the text. From one of the figure captions I found out that the criterion corresponds to SNR=1. This sounds like a rather loose criterion. Wouldn't something like a 95% confidence criterion be more appropriate? Whatever choice is made it should be stated and motivated clearly.

In this study the REFSURF inversion is used as reference, representing what the true variability would be like. As long as we don't know the true flux the results of an inversion may seem a defensible choice. However, this is only true as long as the validity of this approximation doesn't interfere with the conclusions that are derived from it in the end. Since this reference set of fluxes comes from an inversion of surface data itself, it suffers from the same flux detection limitations as the SURF inversion. Suppose that the setups of REFSURF and SURF were statistically equivalent, wouldn't you expect SNR=1? I mean if their posterior uncertainties are the same then REFSURF would be like a random instance of the posterior uncertainty of SURF. In this case what remains is equivalent to a comparison of the posterior uncertainties of the SURF, IASI, and GOSAT inversions.

The comment above has implications for the conclusions regarding the scale dependency of flux detection. For example, if REFSURF is not capable of resolving small-scale variability, it will generate noise (depending on the a priori constraints). If the use of GOSAT leads to better-constrained small-scale fluxes it may end up 'detecting' the noise of the REFSURF inversion, rather than the variability of the true fluxes at that scale. All we learn in the end is that GOSAT is better able to resolve small-scale fluxes than the surface network. That is something we could have concluded already looking only at their posterior uncertainties. Then what is the added value of the method that is used here?

The choice of reference period for calculating the signal should be explained better. It is chosen because 'it corresponds to a period of minimum atmospheric growth rate'. I guess this means that it has a minimal contribution from the long-term trend. However, shouldn't it be representative of the entire period also? The fact that you get anomalies that are predominantly positive suggests it is not. As a result your IAV signal will inevitably contain signal from a time scale outside the 3-year IAV window. It becomes even worse for the annual and seasonal time scale. I don't see how the method avoids signals from longer time-scales affecting the seasonal time scale. Wouldn't I have been better to take out variations on longer time-scales before computing seasonal anomalies?

The role of the prior flux uncertainty should be explained better. I wonder if some regions get already 'detected' without using any data, just because the prior uncertainty is already small enough to satisfy the detection criterion. It would explain why some regions are detected without the observing contributing any significant constraint (for example IASI detecting fluxes from NorthAmericanBoreal, when only data between 30S and 30N are used).

Another factor influencing the scale dependency of flux detection is the accuracy at which posterior fluxes are approximated. Give that only 10 Mont Carlo ensemble members are used this accuracy cannot be that high (see for example the appendix of Pandey et al, 2016 for a formula to compute the uncertainty of a Monte Carlo derived uncertainty for a given number of iterations). Although the limited ensemble size should not introduce a scale dependency, the number of iterations per inversion could do that, because, depending on the search algorithm used, the large scales may be solved first being the dominant eigenvectors of the optimization problem. No information is given about the number of iterations that is used, but in our experience the M1QN3 could converge slowly. Therefore additional information about the convergence of small-scale fluxes is needed.

Somewhere in the discussion a note of caution is required that the posterior uncertainties are derived without proper accounting for systematic errors in the satellite retrieval and transport model. Because of this, despite the use of real data, the detection statistics probably end up being rather optimistic. The use of Desroziers recipe for error

tuning does not account for the neglect of off diagonals in the R matrix.

SPECIFIC COMMENTS

Page 3, line 5: How appropriate is it here to use prior fluxes without IAV? It means that the prior is biased with regard to IAV, and as a result the posterior IAV will be low biased too (assuming that all other statistical assumptions are satisfied).

Page 4, line 1: I'm trying to understand the logic of the sqrt(2). How do you define interannual variability? Wouldn't it be the difference from one year to another? A sqrt(2) inflation rather suggests the variability of the 2-yearly flux in Tg/(2 year). How does that fit with the 3 year time windows? Apart from this I don't see why the assumption of uncorrelated errors would lead to a conservative estimate. I would rather think of posterior uncertainties as being negatively correlated because of limitations in independently resolving the yearly fluxes. Because of these complications I wonder how appropriate it is to address the 3 yearly time scale using a one year inversion anyway.

Page 4, line 6: 'The uncertainty in OH (5% after optimization)' Which scale does this refer to? If it is the global scale, then how about the uncertainty per latitude band?

Page 9, line 7: 'PBSURF signal is twice as often detected' Why is this? I guess it depends on the size of the regions in PBSURF compared to the scale of the regions that are evaluated. If the latter are smaller then wouldn't you rather expect that the large-region inversion suppresses the within-region variability? In that case they would become harder, rather than easier to detect. Some further discussion at this point would be helpful.

TECHNICAL COMMENTS

Page 9, line 7: "REFSURF" i.o. "SURF". Please check if there are other instances where REFSURF was meant.

Page 11, line 11: 'acknowledge' i.o. 'aknowledge'

Figure 5: What happens to the whisker-boxes at the 2Y time scale? I guess they become too compressed to see. If so, then please mention this somewhere (it shows up in other figures also).

Figure 5, 6 & 7, caption: ‘aggregation’ i.o. ‘agregation’

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-234, 2016.