Atmos. Chem. Phys. Discuss., 12, C12496–C12501, 2013
www.atmos-chem-phys-discuss.net/12/C12496/2013/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Estimating regional methane surface fluxes: the relative importance of surface and GOSAT mole fraction measurements" *by* A. Fraser et al.

Anonymous Referee #2

Received and published: 11 February 2013

This manuscript presents first inversion attempts using a GOSAT XCH4 retrieval. As this is a relatively new data stream and not much has yet been published with these data, the findings would be of general interest for people working both with flux inversions and satellite measurements and retrievals. The manuscript is well-written and generally easy to follow, and the figures are clear and easy to understand.

However there are some significant concerns. Firstly, the model set-up seems problematic, given the flux adjustment seen with the "perfect data" OSSE presented in Appendix A/Figure 8, and referenced throughout the paper as the theoretical upper limit of the inversion system. If the pseudo-data were created using the prior fluxes,

there should be no adjustment of the fluxes (i.e. posterior=prior), as there is no mismatch between the model and the measurements. This is something that needs to be explained or resolved before the paper can be published.

Also, the choice of AGAGE as "independent" validation data is problematic given the fact that coincident data were assimilated. These and other concerns are discussed further below.

p 30995 lines 8-11: This sentence is contradictory. It's not possible to reproduce the absolute concentration at the surface and in the free troposphere but at the same time overestimate the positive trend over the four year study period. Upon consulting the reference and looking at the data presented therein, I would rather say that the model did a fair job of representing the seasonal cycle in most cases, but overestimated the interhemispheric gradient (a common problem) and failed to match the trend. Granted, I'm not reviewing that (already published) paper, but the sentence included here contradicts itself nonetheless.

p 30996 lines 5-6: This description sounds as if the only low bias is in the tropics, while Temperate North America and Eurasia, North Africa, and Europe during part of the year also seem to have significant low biases. Are the biases in these regions (North Africa?) really due to lower wetland emissions? I'd rather list the regions where the match is good, as they seem more the exception here.

p 30997 line 8: performance is misspelt

p 30996 lines 15-17: Does this really mean that measurements made on the first day of the month essentially only influence fluxes from that same day? This seems an odd choice, and would mean that the fluxes towards the end of each month are very poorly constrained. If this is really the case, why isn't a lag window used? This seems especially important for XCH4, as the total column has a much bigger footprint in time and space than do surface measurements, and transport to the upper atmosphere (which takes time) needs to be taken into account in order to have the chemical losses well represented.

p 30997 lines 6-14: I must be misunderstanding something. In this idealized set up, where you're essentially trying to invert to find your prior, the error should be 0%, not 5-19%. The mismatch between your pseudo-observations and your prior forward run should be zero, so the inversion should not try to optimize the fluxes at all: you're already at the minimum of the cost function. Getting any deviation from this means that there's something wrong with your system, not that this is the theoretical upper performance limit. Or have I misunderstoond something? Were the pseudo-measurements perturbed in some way that isn't clearly explained in the text?

p 30997 lines 22-25: There seem to be two conflicting definitions of the representation error. Which is it? (I think the second definition makes more sense.)

p 30999 lines 13-14 and 19-21: In lines 13-14 you state that "when the variation of the fluxes within a region as a fraction of the total flux increases, eta is smaller." This seems consistent with the formula presented. In lines 19-21 you state that "The boreal regions have their maximum values in February or March, reflecting the large variation in fluxes within the regions at that time." These two statements are are inconsistent. Are the wetland fluxes (the only time-variant fluxes in the Boreal region in your model, I believe) really experiencing maximum variation in February? I'd expect it to be too cold still, and this high eta value reflects instead low variability and more measurements than in December-January.

p 30999 lines 23-25: I understand that this is a relative metric, and indicates only the confidence level of a given month compared to other months for that region, but I think it might be hard to interpret due to the normalization on a region-by-region basis. For instance, how does an eta value of one in Boreal Eurasia compare to an eta value of one in Tropical South America? A careless reader could assume that the uncertainty on the two months (July 2009 for the former and February 2010 for the latter) is the same, but the normalization by region makes it difficult to interpret. Would it make

sense to normalize it by the maximum value for any region? I'm not sure if this is the solution, but perhaps it could at least be clarified.

p 31003 lines 18-20: If increasing the lag window to three month causes such a significant difference in your regional fluxes (some change in the shape of the seasonal cycle, but, more importantly, a shift throughout the year), why do you limit yourself to a one month lag? It seems that the shorter assimilation window causes a consistent high offset, at least for this region, which is surely balanced by a low offset elsewhere. If you had increased the lag window and seen no change, I could understand the argument justifying the choice of such a short assimilation window, but this directly contradicts this choice. Please explain.

p 31004 line 10: Should this really say January 2005? If so, I'm confused. In section 5.1 it's stated that the study period is July 2009 through December 2010, when the data were available. Do you mean that to spin it up you took one year of the posterior fluxes (e.g. January 2010 through December 2010) and repeated them as a flux climatology from January 2005 with time-varying meteorology, and then only analyzed the results from June 1, 2009, onward? If so, this needs to be better explained.

p 31004 line 13: Please specify the version of TCCON data you are using (i.e. GGG2009 or GGG2012).

p 31004 lines 23-25: I agree that there is no statistically significant change in the SH AGAGE comparison (correlation coefficient decreases by 0.03). But is the change in the NH AGAGE comparision (correlation coeeficient increase by 0.09) any more statistically significant? Neither seems convincing to me. Is there some statistical argument why a change in the r^2 value of 0.09 is significant but a change of 0.03 is insignificant? What is the cut off here? As for the TCCON comparisons, I agree that an increase in the r^2 value for the NH from 0.29 to 0.48 (an increase of 0.19) may be significant, but the agreement is still very poor. I am actually more impressed that the r^2 value for the TCCON SH is so high - it is not obvious from the figure, but this is agreement is actually decent.

p 31005 lines 26-30 and figures 6 7: I am really surprised that averaging over the southern hemisphere TCCON stations (Darwin, Wollongong, Lauder), which show lower r^2 values in Figure 7 (0.3, 0.5 and 0.6 respectively) could result in a correlation coefficient greater than 0.7, as shown in Figure 7. Is this really correct? Also, the word "completely" must be removed from line 29. This implies that having an r^2 value a bit more than 0.5 would "completely reproduce the variability in the observations", rather than vaguely matching the seasonal cycle.

p 31007 lines 18-20: Changes in fluxes do not influence the surface concentrations "before" the total column. Rather, changes in the surface mixing ratio values play only a part in the total column value, but they're certainly changing at the same time. Surface measurements have higher variability than do column measurements as a result.

p 31007 lines 20-21: This is a general (and significant) problem with the use of the AGAGE station data for "independent" comparison. I understand that there is a difference between the continuous and flask measurements, but data that are assimilated at the same location and at overlapping times means that this cannot be considered an independent comparison.

p 31008 line 8: has -> have x2 (data are plural)

page 31010 lines 12-13: This is a bold conclusion given the experiments carried out. What if the spatial structure in your prior was incorrect (e.g. 20% more emissions in Temperate Eurasia and 20% less in Europe, or similar)? This is a more subtle effect than the global emissions being off by 20% across all regions but with perfect prior knowledge of the geographical distribution of the fluxes (which is comparatively easy for the model to fix, and does not depend on transport). A better (and more realistic) experiment would be to try create pseudo-data with a perfectly reasonable independent model set-up (by using a different wetland model with a different distribution of fluxes, for example), and seeing how well your model can retrieve this given an inconsistent

prior.

Figure 8, and Appendix A: Again, as mentioned above, I do not understand why the "ideal" case results in any flux error - the prior should be identical to the posterior in this case.

Interactive comment on Atmos. Chem. Phys. Discuss., 12, 30989, 2012.