Atmos. Chem. Phys. Discuss., 12, C13675–C13692, 2013
www.atmos-chem-phys-discuss.net/12/C13675/2013/

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.

A. Fraser et al.

ac.fraser@ed.ac.uk

Received and published: 1 April 2013

We thank both of the reviewers for their helpful comments and careful reading of our manuscript. In the following response we repeat the original comments in italics.

Anonymous Referee 1

This study investigates the application of total column CH4 measurements from the GOSAT satellite instrument to global inverse modeling of sources and sinks. After several studies demonstrating the quality of GOSAT retrievals it is very useful to see how this translates into source/sink estimates. More studies of this kind will certainly follow, which will be able to take advantage of the numbers that are reported here.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Despite the significant constraints that are provided by the GOSAT data in terms of uncertainty reduction, they don't seem to bring important new insights. A latitudinal shift in rice paddy emissions is reported, but the robustness of this result remains unclear. GOSAT allows an important extension of the measurement coverage in the tropics, but how this translates into adjustments in emissions versus bias corrections remains unclear. A further effort in this direction would strengthen the value of the manuscript.

This is a first study to examine the consistency between the surface data and the GOSAT data. We chose to perform the analysis over relatively coarse areas to be consistent with the available information from the surface network. We think it would be premature to draw any conclusions on changes in emissions.

The reported OSSEs provide useful insights in the functioning of the inversion, but also raise questions as explained below. These and a few other issues will need to be addressed to make the step to ACP.

These comments are addressed below where they are raised in more detail.

GENERAL COMMENTS

Looking at figure 2 it is unclear why the reported posterior global emissions in Table 1 are lower than the prior. All regions except Boreal Eurasia show a significant increase in XCH4 when GOSAT data is used, which somehow is not reflected in the global emissions. The answer may be in the treatment of the initial condition in the inversion, but I couldn't find information describing what was done. Another option might be the lifetime of methane, but if I understand well it is not optimized. Further information is needed to explain the relation between Table 1 and figure 2 and how the initial concentrations are treated.

Unlike a 4DVAR inversion, the ensemble Kalman filter does not optimize the initial conditions. Differences between the model's initial conditions and the observations are

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

treated in the bias correction, which is described in Section 4.1 and in Appendix B. The inversion optimizes the bias between the model and the measurements over the entire time period of the inversion. While the model concentrations are smaller than the GOSAT observations, the difference is considered a bias and the emissions are not adjusted upwards to account for it. Appendix B describes how the initial bias was selected, and gives results of experiments performed adjusting this initial value and how many nodes are fitted, showing that the retrieved bias is robust. In Appendix A we show that the bias correction scheme retrieves a known bias from simulated data.

The reviewer is correct that we do not optimize for the lifetime of methane (the OH field). Our OH field is from a full-chemistry run of the GEOS-Chem model, and is consistent with observations of methyl chloroform (CH3CCl3, or MCF, see Figure 4 of Patra et al., ACP, 2011). We have added this information to the description of the model in Section 3.

It is unclear why an OSSE with perfect data representing a truth that is equal to the prior, nevertheless leads to the flux updates in the inversion as shown in figure 8. In this case, the prior model should be in the cost function minimum already so that there is nothing left to optimize.

In the original Figure 8, only the experiments with the prior flux being 20% larger than the true fluxes are shown, hence the differences between the posterior flux and the true fluxes. To avoid this misunderstanding we have changed Figure 8. In our new Figure 8, included in this response, we show the results of the experiments where the prior fluxes are the true fluxes and where the prior fluxes differ from the truth. We have also updated the results of the OSSE, as suggested by the second last comment from reviewer 2. We discovered an error in the code used to generate the simulated data used in the OSSEs, which resulted in the simulated data not being generated from the correct prior fluxes. This error has been fixed, and the new OSSE results show that when the prior fluxes are equal to the true fluxes and the simulated data is used as the observations, the posterior fluxes return the prior fluxes to within 0.02%.

ACPD
12, C13675–C13692, 2013

Interactive Comment

Else it is very surprising that a bias of several ppb leads to virtually no adjustment in the inverted fluxes. Even if the bias correction captures most of the bias, some adjustment is expected since the bias parameters have a limited uncertainty and therefore the inversion should find an optimal compromise between adjustments of the prior fluxes and bias parameters. It raises the question how well the optimized model fits the satellite data.

The results from the OSSEs using different initial biases are different – but, as the reviewer rightly points out, not by very much. As we show in Appendix A, the inversion scheme is able to return the known bias in our experiments with simulated data. In Appendix B, we show that the bias correction scheme is robust to the number of nodes used and the first guess of the bias. We believe the similar flux results from the simulated data with different biases is a result of the bias correction scheme's skill in returning the bias.

SPECIFIC COMMENTS

Abstract: It is mentioned that that the largest emission changes of 75Tg/yr are found for Temperate Asia, but Table 1 lists only order 15 Tg/yr differences for this region.

The value cited in the abstract is a change in monthly emissions, while Table 1 lists annual differences. We now clarify this in the abstract:

"We find larger differences between regional prior and posterior fluxes, with the largest changes in monthly emissions (75 Tg yr -1) occurring in Temperate Eurasia."

Page 998: Are the bias correction parameters specified by month or is one single set used for the whole analyzed period?

One single set is used. We now state this explicitly in Section 4.1:

"For simplicity, we assume that the bias in GOSAT XCH4 data varies with latitude, following previous studies (Bergamaschi et al., 2009), and constant over the study period."

ACPD

12, C13675–C13692, 2013

Interactive Comment

Page 31007: "changes in methane emissions affect the surface concentrations before the total column measurements" It is not clear why this should be the case. It is mentioned elsewhere that the emissions have to be transported upwards, which takes time. However, the total column measurements, as provided by GOSAT, have sensitivity all the way down to the surface. Therefore there is no need for this signal to be transported upwards for detection by GOSAT.

This is true and our original wording was misleading. As stated by reviewer 2, surface concentrations and the total column are affected at the same time, but the changes in the surface concentration are larger and therefore detectable earlier on since the surface is only a part of the total column. We have re-phrased this section:

"The difference between the prior and posterior model was larger at the AGAGE sites since the surface concentration makes up only a portion of the total column. As a result changes in methane emissions are detectable earlier at the surface than in the total column."

Page 31008: "The error reductions for inversions using GOSAT data are at least twice the error reductions when only surface data" It should be mentioned that the use of a short assimilation window is not in the advantage of background surface measurements. It means that part of the information that is provided by the surface network is not used by the inversion. Else GOSAT measurement errors are likely to be correlated in time and space, not to mention biases that are not captured by the bias correction algorithm. No attempt is made to represent the uncertainty associated with this, which leads to over optimistic error reductions in comparison with the surface network.

We have experimented with including a lag window of three months, but we find that in most regions the changes are within the original error on the fluxes. (See Figure 2 included in this response.) Because of transport error in the model and unevenly distributed observations, in some regions it can be difficult to identify the origin and strength of the emissions correctly. In those regions, using a lag window can introduce

ACPD
12, C13675–C13692, 2013

Interactive Comment

likely non-physical changes in the seasonal variation of the fluxes. We chose to not use a long lag window for these reasons. In addition, GEOS-Chem driven with the posterior fluxes from the inversion with no lag window agrees better with data from AGAGE and TCCON than when using the fluxes from the three month lag window inversion. We have added our justification for not using a lag window to Section 4:

“ We do not use a lag window to estimate monthly methane fluxes: measurements of methane only affect fluxes in the month they were taken. Because of transport error in the model and unevenly distributed observations, in some regions it can be difficult to identify the origin and strength of the emissions correctly. In those regions, using a lag window can introduce likely non-physical changes in the seasonal variation of the fluxes.”

We do not take into account errors introduced to the fluxes by the bias correction. This error would be difficult to quantify, but we find (in Appendix B) that the inversions are robust to the bias correction. In any case, the focus of the paper is on the improvement in the inversions due to the more dense GOSAT observations versus the sparse surface network.

Anonymous Referee 2

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. poste-

ACPD

12, C13675–C13692, 2013

Interactive Comment

rior=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.

We address these comments below where they are expanded upon in more detail.

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.

We have reworded this sentence to be consistent with itself:

"In that study we found that the model reproduces the seasonal cycle of methane at the surface and in the free troposphere but overestimates the positive trend over the four year study period."

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.

The only differences between the model used in Parker et al. (2011) (where no bias was observed between the data and model, see Figure 3 of that paper) and the model used here are: (1) the wetland and rice emissions (updated from Bloom et al. (2010) to

ACPD

12, C13675–C13692, 2013

Interactive Comment

Bloom et al. (2012)) and (2) the biomass burning emissions (updated from GFED v2 to GFED v3). Wetland and rice emissions are the largest single source of methane and far outweigh the biomass burning source (260 Tg/year vs. 20 Tg/year). The Bloom et al. (2012) emissions have a much stronger Southern Hemisphere source, and a shift of the peak emission from just above to just below the equator (see Figure 4 of that paper), but with the same annual total source strength, by design.

We have changed the sentence to highlight that the bias is at a maximum in the tropics:

"Unlike the comparisons in Parker et al. (2011), the new comparisons show a regional bias between the data and the model, peaking in the tropics, with GEOS-Chem generally underestimating the GOSAT data."

p 30996 line 8: performance is misspelt

fixed.

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.

The prior fluxes are specified on a either a monthly or yearly (depending on the source) time step, and the posterior fluxes are optimized on a monthly time step. In both the prior and posterior model there is a step change in fluxes at the start/end of each month. The inversions are performed at the end of the month – all the observations from that month contribute to the posterior flux weighted by the observation error.

See also our response to a similar comment on the use of a lag window by reviewer 1

ACPD
12, C13675–C13692, 2013

above.

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 misunderstood something? Were the pseudo-measurements perturbed in some way that isn't clearly explained in the text?

We have discovered an error in the code used to generate the simulated data – it was not being simulated from the prior fluxes, so in the presented experiments the true and prior fluxes were not actually the same. We have fixed this error (which only affected the OSSEs, and not the inversions using real GOSAT data) and now find that in the perfect case with prior and true fluxes being equal and assimilating perfect data the posterior fluxes in all regions are now within 0.02% of the prior fluxes. We have updated the text and figures accordingly.

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.)

The first was a definition and the second the implementation, and we apologize for any confusion caused. We have reworded the section to only use the second description and to be more clear.

"We describe the transport error as 0.5 % of the mixing ratio obtained by the flask measurement, and the representation error as the standard error of the monthly mean calculated from the observations made over that month (Wang et al., 2004)."

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

ACPD
12, C13675–C13692, 2013

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.

This was a mistake in the manuscript. The reviewer is right that the value reflects small variation of fluxes within the region in early spring. We have corrected the text:

"The boreal regions have their maximum values in February or March, reflecting the small variation in fluxes within the regions at that time."

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.

We choose to normalize by region precisely to avoid comparing regions to other regions. We define a simple metric to allow an assessment of the variation of information within a region with time and an understanding of what factors limit the inversions. Many factors are not explicitly included, for example: the size of the region and the longitudinal and latitudinal span of the region. To clarify this, we've added the following phrases to Section 4.2:

"Because the metric is only dependent on the number of observations and the variation

ACPD
12, C13675–C13692, 2013

within the region, and other factors that would influence the information content of the region are not explicitly included, values of eta within one region cannot be directly compared to values of eta in other regions.”

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.

See the discussion above about lag windows and Figure 2 in our responses to reviewer 1 above. The decrease in the posterior fluxes in South Africa when introducing a lag window is indeed balanced by an increase in fluxes elsewhere (e.g Tropical South America), but in general the fluxes are not changed by more than the initial posterior error.

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.

Yes, that is exactly what is meant, and we have expanded on this in Section 5.2:

“To avoid inconsistencies in the fluxes and resulting concentrations, we first “spin-up” the model for 4.5 years, from January 2005, using posterior fluxes from January to December 2010 and the appropriate GEOS-5 meteorology.”

ACPD
12, C13675–C13692, 2013

Interactive Comment

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

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

This is specified in Section 2, where the data are introduced, and we now repeat the information in Section 5.2 where they are used.

“Figure 6 shows daily mean and hemispherically-averaged GEOS-Chem (prior and INV3 posterior fluxes) and observations for two independent methane measurement networks (Sect. 2): AGAGE (surface mole fraction, June 2012 release) and TCCON (total column mole fraction, 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.

We did not mean to imply *statistical* significance in describing some of the changes in the Southern Hemisphere as e.g. “not significantly different”, but rather that the change was small. We reword this section to spell out the small changes and avoid this interpretation.

“For the Southern Hemisphere AGAGE comparisons the posterior standard deviation is increased by 0.2 ppb and the correlation coefficient is decreased by 0.03 while the biases are decreased by 9.1 ppb. At TCCON sites, the bias decreases by 1.7 ppb and the standard deviation decreases by 1.6 ppb, while the correlation coefficient increases by 0.03” *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?*

Figure 3 shows the individual time series of the TCCON observations and GEOS-Chem prior and posterior models that go into the paper's Figure 6d, as well as the average time series of the three stations. The prior (blue) and posterior (red) r-squared values are given on the figure. While GEOS-Chem at Darwin and Wollongong tends to underestimate the TCCON observations, at Lauder the model tends to overestimate the observations. When averaged, these fortuitously lead to a better agreement than at any one site.

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.

Agreed, we have removed the word “completely”.

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.

This is true and our original wording was misleading. See also our response to a similar comment by reviewer 1.

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.

ACPD
12, C13675–C13692, 2013

Ideally we would have more ESRL/GASLAB sites so that we could assimilate some and save the others for independent comparison, but this is not the case over much of the globe. There is added value in the AGAGE continuous measurements because they are more frequent and are not sampled from preferential wind directions, in contrast to some ESRL/GASLAB sites. We agree that since we are assimilating data at the same locations, the AGAGE data is not purely independent, and so we have renamed Section 5.2 to “Agreement with ground-based data” and replace “independent surface data” with “other surface data”/“data that has not been assimilated” in the text of Section 2 and Section 5.2.

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

fixed.

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. 20Temperate 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.

As mentioned above, we have updated the OSSE results. As suggested by the reviewer, we have run an additional experiment where we have perturbed the prior fluxes in each region by a random number between -20 and 20%. These results are now reported in the main paper, in detail in Appendix B, and in Figure 8. Results from this experiment are similar to those from the +20% experiment.

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.*

As addressed above, this was due to an error in the code used to generate the simulated data, which has now been fixed.

References

A.A. Bloom, P.I. Palmer, A. Fraser, D.S. Reay, and C. Frankenberg (2010); Large-Scale Controls of Methanogenesis Inferred from Methane and Gravity Spaceborne Data; Science; 327.

A.A. Bloom, P.I. Palmer, A. Fraser, and D.S. Reay (2012); Seasonal variability of tropical wetland CH4 emissions: the role of the methanogen-available carbon pool; Biogeosciences; 9.

R.H. Parker, H. Boesch, A. Cogan, A. Fraser, L. Feng, P. Palmer, J. Messerschmidt, N. Deutscher, D. Griffith, J. Notholt, P. Wennberg, and D. Wunch (2011); Methane Observations from the Greenhouse gases Observing SATellite: Comparison to Ground-based TCCON data and Model Calculations; Geophysical Research Letters; 38; doi:10.1029/2011GL047871.

P.K. Patra, S. Houweling, M. Krol, P. Bousquet, D. Belikov, D. Bergmann, H. Bian, P. Cameron-Smith, M.P. Chipperfield, K. Corbin, A. Fortems-Cheiney,ÂăA. Fraser, E. Gloor, P. Hess, A. Ito, S. R. Kawa, R.M. Law, Z. Loh, S. Maksyutov, L. Meng, P.I. Palmer, R.G. Prinn, M. Rigby, R. Saito, and C. Wilson (2011); TransCom model simulations of CH4 and related species: linking transport, surface flux and chemical loss with CH4 variability in the troposphere and lower stratosphere; Atmospheric Chemistry and Physics; 11.

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

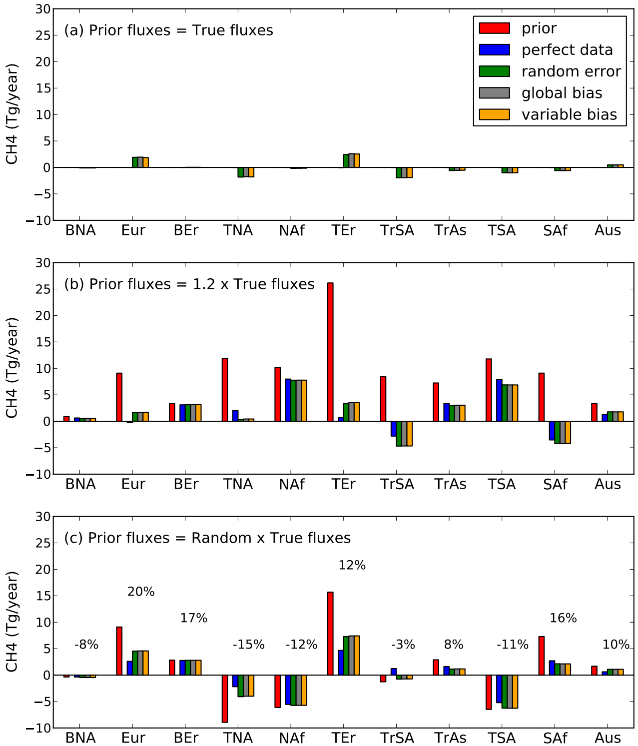

Fig. 1. New Figure 8, showing results of all the experiments adjusting the prior fluxes.

ACPD
12, C13675–C13692, 2013

Interactive Comment

ACPD

12, C13675–C13692, 2013

Interactive Comment

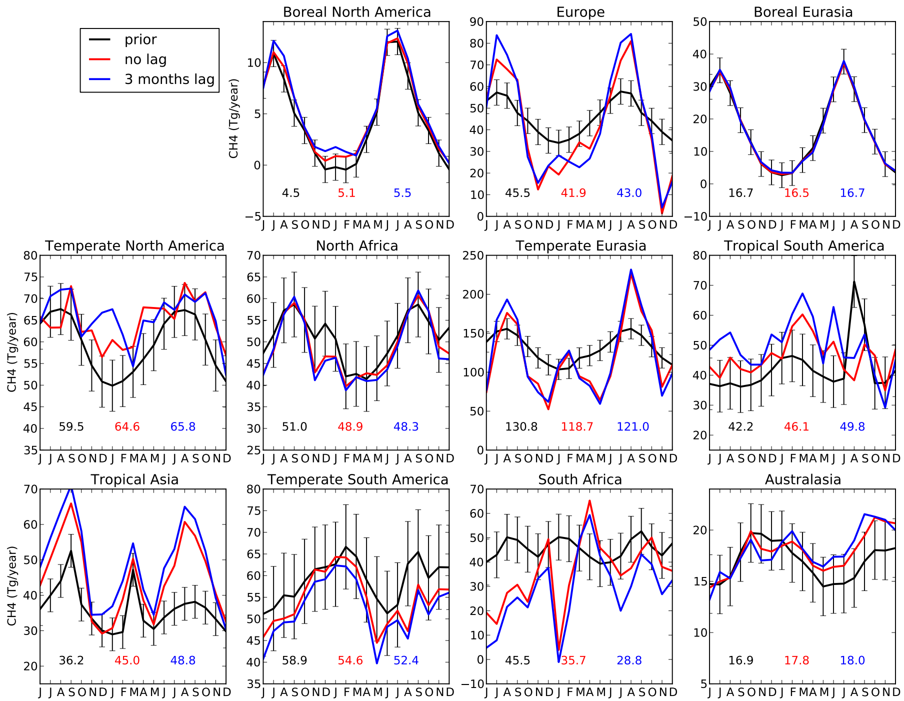

Fig. 2. Effect of lag window on inversion results. The inset numbers are the average flux in the region from the prior (black), no lag (red), and 3 month lag (blue) fluxes.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

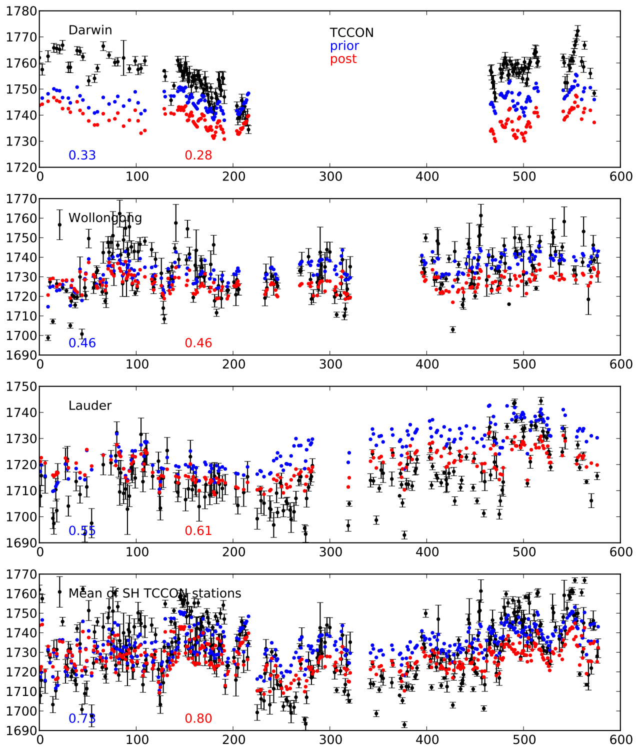

Fig. 3. Measured and modelled TCCON XCH4 at Southern Hemisphere stations. Numbers on the figures are r-squared values of the correlation between observations and the model.

ACPD
12, C13675–C13692, 2013

Interactive Comment

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper