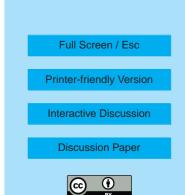**Atmospheric
Chemistry
and Physics
Discussions**

Interactive
Comment

Interactive comment on "Carbon source/sink information provided by column CO$_2$ measurements from the Orbiting Carbon Observatory" *by* D. F. Baker et al.

D. F. Baker et al.

Received and published: 20 March 2009

We appreciate the reviewer's suggestions for where we can shorten the paper, and have tried to follow these as best as possible (see below for specifics). We agree that the errors addressed in the paper are only a subset of all those likely to have been encountered by OCO, and we will try to make that point clearer in the text: the random error results should be considered "best-case" results, achievable only after all sizable systematic errors are identified and removed; the systematic error cases that we have examined in the paper, however inadequately, give only a rough idea of those likely to have been encountered in reality. Possibly a greater issue is whether the relatively low random X_{CO_2} retrieval errors that we have used here, taken from the linear error analysis of Boesch, et al, really capture the full magnitude of random errors likely to

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

have been encountered by OCO: we know, for example, that they did not contain any contribution from the residual image effects and "W-pattern" errors observed with the actual OCO CCD instrument.

As for whether the results should be applicable to other inverse approaches, we feel they ought to be, at least for those methods based around a least squares error functional and using similar assumptions for the measurement and a priori flux covariances. We can think of some approximations and errors with other methods that might prevent achieving these lower errors: a forward-running filter will have an initial transient that will take a while to die away, so errors will be higher during this time than what a smoother, running both directions in time, would give (our variational method is like a smoother in this regard); and ensemble filtering methods must make certain approximations that introduce additional errors to the optimal estimate provided by the full (non-ensemble) filter.

Replies to Detailed Comments:

p. 20058 l. 11-12: Shouldn't the a-priori correlations reflect assumptions on the fluxes, rather than properties of the data (at least in theory)?

For the final flux estimate to be a statistically meaningful combination of the prior flux estimate and the data, yes, they should. In a practical sense, a priori correlations smooth out the correction added onto the prior, limiting the spatial detail that can be corrected. By not explicitly adding spatial correlations in our setup here, we permit the method to correct the prior at the finest resolution it can (the grid-box resolution), at the risk of overfitting the data. If there is overfitting, however, this should be reflected in the error statistics that we calculate.

The spatial discretization of the model effectively imposes a correlation length on the estimate. If the true flux correlation length is less than this, then it would not make sense to add more via off-diagonal terms in P_0. Over land/ocean, Chevallier, et al (JGR, 2007) used values of 500/1000 km. Our 5 deg longitudinal grid-box dimension

effectively gave us this correlation over land, at least in the east-west direction, so we chose not to add more. Perhaps adding longer correlations explicitly over the oceans would have helped to converge the oceans faster than they did. When solving at finer resolutions, it would certainly be more important to add in the spatial correlations explicitly.

p. 20062 l. 1-6: I was wondering whether this choice of P_0 is actually appropriate for these tests. Shouldn't you use the same choice that you will use later with the real data? Otherwise, the same errors/biases of the data will lead to different errors in the fluxes between the tests and the real retrievals (same issue for the measurement uncertainly, p. 20068 l. 24ff).

The choice for P_0 described on lines 1-6, computed from the absolute value of the actual prior-truth flux difference, is the appropriate one to use in the "perfect model" experiments (1 2), in which we are trying to isolate just the impact of the random measurement errors. [We are trying here to get the same uncertainties that a method that computes a full rank covariance matrix would give. In our simulation setup, using a more realistic guess for P_0 would then introduce P_0-related errors into the estimate, as well.] We agree with the comment here that for all the other experiments, it would be more realistic to use the mistuned P_0 used in Experiment 3 (as well as the mistuned measurement uncertainty). Because we found in Experiment 3 that this mistuning slowed the convergence of our method considerably (by at least a factor of 3), in the interest of being able to finish our runs in a manageable amount of computer run time, we went back to using the perfectly-tuned P_0 and R values for the remaining experiments. Computational expediency was the sole reason for this choice. We argue that one ought to be able to add on, in some way, the flux errors due to the mistuning obtained in Experiment 3 to the results for the later experiments to get an approximation to what would have been obtained, had mistuned P_0 and R been used in those experiments.

p. 20072 l. 1-8: I'm a bit concerned about how specific the values of this error measure are with respect to the particular choices of prior flux and the flux generating the

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

pseudo-data. The error reduction will be large where the two involved models happen to be particularly different, which is however not a property of the satellite information but rather by chance. Wouldn't absolut error measures (like in Fig. 13) be more relevant throughout, given that biases are the most important issue?

We agree that absolute error measures are generally a better way to quantify the constraint provided by a given set of measurements since, with sufficiently loose priors, they reflect mainly the measurement information. However, the fractional error reduction statistic also has some key advantages for our problem. When a funding agency asks what value a satellite's data can provide, it is usually couched in terms of the improvement over existing knowledge that will be gained. The error reduction puts improvement over the land and the ocean on a similar scale, whereas their absolute errors may differ by an order of magnitude. Finally, our field in general has taken to using the error reduction statistic as its favorite metric in OSSE studies of this sort, despite its inadequacies; for comparison to previous results, it needs to be included. We actually favor providing both the absolute error and error reduction plots when presenting results, but when space is tight, we go with the error reduction statistic. We have at least included plots of the a priori uncertainties assumed here, so that areas in which these are either quite high or quite low can be identified.

In this context: Wouldn't it make sense anyway to use at most a very smooth prior (not a detailed one from models), as the fine-scale flux structure may be hoped to be contained in the data?

Using a smooth prior makes sense to the extent that the fine-scale flux details from a model are usually more uncertain than the coarser-scale patterns. However, if the fine-scale detail from a model is physically- based and fairly robust (e.g. the difference between fluxes from a lake and a forest in an area with many lakes) then it ought to be included. The data may not be of sufficient resolution to resolve such detail on its own, and leaving it out would degrade the ability to model the data accurately a priori. We should note that, of the two flux models used in this study, we used the smoother one

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

(CASA/Takahashi) for our prior, and the more detailed one (LPJ/NCAR ocean) for the truth.

p. 20072-20073: This is a long discussion on the convergence. I'm however not convinced about the value of showing and discussing results that have not yet reached the cost function minimum, and would suggest to omit them (which would shorten the paper, incl. Fig 8). The authors could still make the point that land fluxes alone could be obtained in an shorter minimization, by actually performing the experiment they outline in p. 20074 l. 1ff, and just citing the number of iterations they needed, compared to the standard run.

We have followed the reviewer's suggestion and moved the details of the convergence, including Fig. 8, to the supplementary material, in the revised version of the paper.

p. 20083 l 15ff: The results are given for using OCO data only. In the current discussion, often a combination of satellite and in-situ data is envisaged, with the hope to reduce biases. If you do not agree to such a combination, can you comment on why not?

We agree, of course, that the OCO data would have been used in combination with in situ data to constrain the surface sources and sinks. We have done a previous study (Baker, et al, 2006) that suggested that the in situ data provided only a very weak constraint on the weekly grid-box-scale fluxes such as we examine here. If that is true, it would make little difference to the a posteriori uncertainties due to random errors on whether the in situ data were included in this study or not. To keep this study as clean as possible, we left them out. Of course, as you hint at, outside the parts of the globe with dense data (western Europe, North America) the in situ data will perhaps prove most useful in identifying systematic errors in the satellite data. This should reduce the magnitude of the systematic errors in the satellite data, but does not invalidate the results obtained here for the particular levels of bias assumed.

Figs 10 and 11: These figures seem to convey the same message, except for minor

details. I suggest to move one of them to the supplement (incl. simplifications in the text).

One gives the results for weekly fluxes, the other for seasonal fluxes. We will take the suggestion and move the seasonal results to the supplement, and simplify the text.

Replies to Smaller comments:

p. 20053 ll. 2-16: A few references would be in order.

We add some in the revision.

p. 20054 l. 14: briefly explain terms like 'inclination'

We have replaced 'inclination' with 'orbital inclination'.

p. 20055 l. 4: 'We use a tracer transport model...' - also mention the inverse framework, otherwise it sounds like a forward study

We have reworded the first sentence as follows to clarify this: "In this study, we use an inverse method to quantify how well X_{CO_2} measurements from OCO will help estimate sources and sinks of CO_2 at the surface."

p. 20058 l. 27: "X_{CO_2} measurements" should only be used for real data; use a term like "pseudo-data" or "synthetic data" instead.

We have changed this to: "...using simulated X_{CO_2} measurements that are modeled..."

p. 20059 l. 4-5: mention that the cited studies use in-situ data, not satellite data.

We have changed this to: "...used in typical past time-dependent CO_2 inversions of in-situ data"

p. 20059 l. 20: Roedenbeck (2005) involved daily (not monthly) fluxes.

Thanks for catching this. We will drop "monthly" here.

p. 20060 l. 1ff: The authors discuss that their numerical method (iterative minimization)

was less appropriate than another numerical method (ensemble Kalman smoother).

We did not intend to suggest that our variational method is less appropriate than the ensemble methods. We merely wanted to note that there is another class of estimation methods that could potentially be used for this problem, with certain advantages. While the ensemble filtering methods do have the advantage of not requiring an adjoint (at least, when running only in forward-mode as a filter), they have other disadvantages (some of which the reviewer noted) as well. We will reword the text to better reflect our feelings.

Though the Kalman method indeed has the advantage of not requiring an adjoint, it is actually much less efficient, because it involves an effective number of model runs that is not only determined by the number of ensemble members (which is roughly equal to the number of iterations of the iterative method), but is multiplied by the number of Kalman steps within the assimilation window. For a satellite inversion, the assimilation window certainly needs to be long, given by the time it takes to transport surface flux signals into the higher atmosphere which is slower than horizontal transport. A Kalman smoother with an assimilation window as short as 10 weeks (while Law, Atmos. Chem. Phys., 4, 477, 2004 suggests at least half a year even for surface data) would already need 10 times more CPU time (assuming a Kalman step of 1 week).

Since the variational method requires at least a forward and an adjoint run per iteration (equalling something closer to four forward runs between them, computationally), the situation for the Kalman smoother is not quite as bad as this, but the point remains.

Moreover, it would intrinsically yield an approximation only, while the iterative method can be converged to full accuracy (provided the adjoint is exact).

This is a key point: there is nothing to force the ensemble mean ever closer to the data, as there is with the iterating variational methods.

p. 20062 l. 10-12: It seems that rectification between the diurnal cycle of the fluxes

and that of the transport leads to another error source that should be mentioned.

Yes. To bring this out, we have reworded this to: "Similarly, the diurnal cycle of flux is not modeled here, since the OCO data, taken at a single local time per day, cannot resolve it; to the extent that the early-afternoon OCO data are biased with respect to the daily mean X_{CO_2}, the resulting CO_2 flux estimates may be biased as well, and this will not be reflected in the results shown here."

p. 20069 l. 20-21: Explicitly define the terms "precision" and "accuracy", because not everyone may share the same concept (or even be aware of common connotations).

We have reworded this to: "... for quantifying the precision of the estimate (the standard deviation of errors about the mean estimate), though not necessarily its accuracy (the standard deviation of errors about the truth), since they do not quantify the impact of systematic errors."

p. 20075 par. 2: As a comment, larger a-priori sigmas are generally expected to lead to slowlier convergence, because the effective number of degrees of freedom is larger.

We agree that a looser prior results in more degrees of freedom and a slower convergence. However, the a-priori sigmas assumed in Experiment 3 (the mistuning experiment) were, if anything, a bit tighter than the true ones (compare Fig. 2e to 2d) so we do not feel this argument can explain our slower convergence. As argued in the paragraph in question, we feel it is the difference in the assumed priors that is more the issue.

p. 20076 l. 21-26: There does not seem to be much additional information in the second aerosol test. I suggest to only show one.

To save space, we have removed the first aerosol test completely, leaving only the second (higher-level) case.

p. 20080 l. 8ff: How do the a-priori uncertainties compare between this study and Chevallier et al.? Can these differences explain the differences in results?

Chevallier, et al, (JGR, 2007) used ocean/land prior uncertainties of about 0.4/4.0 gC m^{-2} day^{-1}, which, in the units used in Figure 2, equal about 1.7/17.0 * 10^{-8} kgCO$_2$ m^{-2} s^{-1}. These are a factor of four or more times looser than what we have used (based on the actual difference between our two flux models). If the uncertainty reductions were strongly influenced by the a prior uncertainties assumed, then the Chevallier reductions should be even greater than ours, yet ours are larger. This leads us to believe that the reductions depend more on the measurement uncertainties assumed, as mentioned in the paragraph in question.

p. 20083 l. 5ff: I was surprised that you fear the system to become computationally infeasible. Wouldn't e.g. 200 iterations always suffice but still be fully managable?

We have shown that, with a realistically-mistuned P_0, near-full convergence already takes about 200 iterations for this fairly coarse resolution (5 deg longitude, 2 deg latitude). At finer resolutions, with more degrees of freedom, we should expect to need more iterations to achieve full convergence. And, given the much greater computational expense when running at finer resolution, we could easily envision the problem becoming computationally limited.

p. 20084 l. 5ff: The ongoing TransCom experiment on exactly this aim (lead by S. Maksyutov) needs to be mentioned here.

Yes, we should have added the ongoing effort led by S. Maksyutov to the two TransCom efforts mentioned, and we have now done so.

Replies to Minor corrections:

p. 20059 l. 23: "solving"

p. 20061 l. 10: "interpolated"

p. 20081 l. 20: "expected to be"

Thanks for catching these...

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Interactive comment on Atmos. Chem. Phys. Discuss., 8, 20051, 2008.