Atmos. Chem. Phys. Discuss., 8, S11812–S11825, 2009
www.atmos-chem-phys-discuss.net/8/S11812/2009/

Atmospheric Chemistry and Physics Discussions

ACPD

8, S11812–S11825, 2009

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: 21 March 2009

The first noticeable feature of the paper is its length. The authors have had little success in synthesizing their thoughts and often got lost in side details. They also could have made better use of their 40 references.

We have moved a considerable amount of detail to the supplementary material to reduce the length of the paper. We hope this will make the remaining material more accessible for the reviewer.

Second, and related, the effort was arbitrarily focused on some parts of the error budget while nearly forgetting about the model error. What would be the point of the six pages of Section 2.5 if the transport error was to dominate the error budget?

As this reviewer points out, we have addressed the transport error in only the most approximate manner. Our finding that these transport errors are important points to the need for investigating this term in a more comprehensive fashion (perhaps with a TransCom-like effort, such as that being led by S. Maksyutov at the moment). If such a more detailed study were to verify that transport errors far outweighed all the other terms in the budget, then this comment might have some merit. However, even then, the results of this paper would still be useful, as they would point to the flux accuracies possible if the transport models were to be improved and the transport errors reduced. The quest to make use of OCO-like satellite data, if/when it arrives, will be one of identifying and removing systematic errors in the problem, and improving the transport models must be included in this effort.

Last, some parts of the data assimilation system were left crude, even though it was created more than two years ago. The first weak point is the adjoint model, the approximations of which make all the iteration-dependent results (Fig 9 and related conclusions) of little interest.

Since the supposed "weak points" of our adjoint are never detailed by the reviewer, we cannot directly respond to the comment. Below, we discuss discuss why we feel the adjoint is not responsible for the issues brought up in the detailed comments.

The second weak point is the diagonal covariance matrices that prevent the authors from achieving their "perfect experiment", despite their claim.

Below, we give our reasons why we did not explicitly specify correlations (non-zero off-diagonal terms in the covariance matrices) in this study. We feel we can still use the term "perfect model" to apply to tests done without these correlations, especially since we are using the term to indicate the absence of other error sources.

These issues dramatically limit the information content of the paper. A shorter, focused and more balanced version of this paper is needed.

ACPD

8, S11812–S11825, 2009

Interactive Comment

We agree that a shorter version of the paper would be more accessible. We have shortened it considerably in response to all the reviewers' comments.

Replies to Specific comments:

p.20053, l.20: the continental scale can also provide insight into flaws of the carbon models.

This is actually correct. The 22-region TransCom inversions did suggest that the annualized extratropical southern ocean sink was considerably smaller than the previous best estimates from the Takahashi flux product, a view that has gained credibility more recently with the latest ocean flux inversion work and carbon modeling studies of the Southern Ocean. We have reworded this to: "...but not at the regional scales where they would be most useful for identifying flaws in the carbon models."

p.20053, l.22: the mixing errors affect the continental scales as the regional ones.

We are not trying to claim otherwise here.

p.20055, l.3: "quantify" is very optimistic.

We think using 'quantify' is accurate. We have quantified the expected flux estimation uncertainty given when the OCO data are factored in, as well as error reductions computed from these. Of course, one is entitled to question how accurate our estimates of these quantities are... in that case, we could respond to the details.

p.20056, l.5: there are both improvements (those noted indeed) and steps backwards (the "perfect model" assumption in the reference run, and the loss of accuracy in the computation of the error reduction). Which effect dominates?

The reviewer is not particularly clear in what she/he means by "steps backward" here. What is meant by "the loss of accuracy in the computation of the error reduction", for example? The "perfect model" cases simply quantify the uncertainties due to the random errors in the problem. We are not sure why the reviewer considers this a "step

ACPD

8, S11812–S11825, 2009

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

backward".

p.20056, l.9: is the "full-physics" system used here the prototype for OCO level-2 data or will another algorithm be used?

The "full-physics" algorithm used here to compute the OCO X_{CO_2} retrieval uncertainties was originally intended to be the sole method used to process the OCO radiance to CO_2 concentrations. Because this method was computationally slow, a second technique was developed to process the vast majority of the data, with the "full-physics" algorithm being used to validate this second method. This second method did not provide a vertical weighting function for the X_{CO_2} measurement, so we found it more convenient to use the full-physics approach to provide both the measurement uncertainties and the vertical weighting. The uncertainty estimates from this approach are somewhat on the low side and can be viewed as best-case numbers to be aspired to, as systematic errors in the measurements are removed.

p.20057, l.16: strictly speaking, relating the definition of the state vector to the observation system is not correct.

It is not clear what the reviewer is trying to say with this comment. It is certainly correct to tailor what is estimated in the state vector of any problem to what is likely to be observable, given the data at hand.

p.20057, l.22: such an assumption is part of the prior information.

Yes, we agree, the assumption made about the diurnal cycle is effectively part of the prior, even if it does not explicitly show up in the prior, as with our treatment here. There should be an additional error term (not addressed here) in the a posteriori flux error uncertainty due to errors in the assumptions made about the diurnal cycle.

p.20058, l.17: is the CFL criterion satisfied with such a long time-step?

Some minor filtering of the winds at higher latitudes was required to satisfy the CFL criterion at this time step.

ACPD

8, S11812–S11825, 2009

p.20058, l.23: the remark is meaningless since the convergence rate also depends on the optimality system and on the minimization algorithm. The only relevant verification of the accuracy of an adjoint model is against its tangent-linear model. How different is the authors' computation of the authors' computation of $(Hx)^T(Hx)$ compared to $x^T H^T Hx$?

Meaningless? Why should one care if there are small errors in the adjoint if, when used in an assimilation, these errors do not appreciably affect the final estimate or the convergence rate? We have done the adjoint test the reviewer asks about, as well as point-to-point tests. The agreement is not exact, but good to within about 1% - 2%, point-to-point, for typical two-week runs. As we showed in Baker, et al (2006), however, this small error does not have a significant impact on the final estimated flux value. We have not tested whether this slows our convergence appreciably: we should do so, as well as test whether these errors cause the land/ocean tradeoff noted in the paper.

p.20059, l.11: it may be relevant to mention where the computational problem lies. It is not in the linear equations themselves, but in the estimation of the Jacobian matrix of the transport model.

The key word here is "direct", indicating a method that solves the linear equations with a method executing a predictable number of operations, as opposed to an "iterative" method, which may realize large savings in computations by computing an approximate solution more quickly. Since we already mention the filling of the Jacobian matrix in the previous sentence, we will just add "(non-iterative)" after "direct" in the sentence in question to make things clearer.

p.20059, l.15-18: the authors seem not to include the initial CO_2 concentrations in the state vector, which would not be correct.

While our method does include the initial CO_2 concentrations in the state vector in general (see Baker, et al, 2006), we have not considered these errors in this study. Since the errors in the initial CO_2 concentrations that influence the surface flux estimate

ACPD

8, S11812–S11825, 2009

Interactive Comment

die away quickly (after a few weeks), this has little impact on our flux results computed over a full year.

p.20059, l.28: the authors should make it explicit that they are talking about the posterior error covariance matrix.

We have added " a posteriori" here to clarify this.

p.20060, l.1: the variational and the ensemble methods do not necessarily achieve the same level of accuracy for a given computational effort.

This is true. We have reworded this discussion in response to Reviewer 2's comments, and this point is brought out in that revised discussion.

p.20061, l.10: interpolated.

Thank you for catching this... corrected.

p.20061, l.19: the reason given is not appropriate. Physically-based differences may lead to anything resembling or not the prior errors.

We must disagree with this comment. We think it is better to base the a priori errors on some actual physical basis (represented here by our best understanding of that basis, taken from carbon models). Perhaps the reviewer objects to basing this on the difference between two models. We would agree that such a difference should reflect only a single draw from the multi-dimensional probability density function that should be represented by the a priori covariance matrix. Determining the proper form for that true covariance matrix is quite difficult, as we are sure the reviewer knows. One could perhaps base the statistics on the distance between two points on the surface and fit this using the differences between many different carbon flux models (perhaps with different lengths for land and ocean). This would give a smoother prior than what we have used here; we prefer the detail in our prior as being more representative. In any case, we feel that it is more the error made in specifying the prior, rather than the precise form of the prior for the control case, that is more important for the tests we do in this paper. Our Experiment 3 is meant to explore this issue.

p.20062, l.4: the authors should be more explicit. Do they take the grid-point annual statistics of the differences or something cruder?

We use the monthly prior-truth flux differences at the grid-box scale, interpolated to the 7-day flux span. We have added "weekly" to this sentence to clarify this: "... of the actual weekly prior-truth flux difference..."

p.20062, l.11: the argument does not hold: what about aggregation errors?

We have addressed some error terms in this study, and have left others out, as the reviewer notes. Errors due to the assumed fossil fuel fields and the diurnal cycle of the land biosphere flux have not been addressed. The random part of the spatial aggregation errors have been addressed, as described in Section 2.5.2. We certainly agree that there will be errors made in attempting to model the diurnal cycle of the fluxes (especially over land), and that these will propagate through to errors in the estimated longer-term fluxes; it was not our intention to investigate these errors in this study.

p.20062, l.18: the flat mass-weighted average may be adequate for the apparent optical-path-difference OCO product.

It is not clear that this is the case. The vertical weighting of the X_{CO_2} measurement (what portion of the atmosphere the measurement is most sensitive to) is a function of the physics of the atmosphere and the details of the measurement approach, more than the details of the retrieval method. The apparent optical-path-difference should not have much more sensitivity to the stratosphere than the "full- physics" method, for example, since they are both using the same spectral bands with the same lack of sensitivity there. While detailed vertical weighting functions have not been computed yet for the apparent optical-path-difference approach, it seems more reasonable to assume they would be similar to those from the full-physics approach, rather than flat.

ACPD

8, S11812–S11825, 2009

Interactive Comment

p.20062, l.19: do those numbers correspond to measurement error or to observation error (with the transport and the representation errors included)?

These numbers reflect the errors added to the measurements and embodied in the assumed measurement error covariance matrix, R. Since the same constant errors were used across the full globe, it matters little whether one considers them to be straight instrument errors, or to include modeling errors (such as transport and representation errors).

p.20063, l.6: this argument may not hold for the model error.

We agree, since the model cannot represent variability inside its grid box. The error terms discussed in the text are instrument- and retrieval-related. The modeling errors can be thought of more as biases in terms of the cross-box average measurements.

p.20064, l.18: "striking" is a strong word for a feature that was expected.

Good point. We will replace "striking" with "noticeable".

p.20065, l.15: the distinction between along-track and track-to-box errors is rather artificial in this context. In both cases, it is a sampling problem.

True, they are both sampling errors. However, we find it useful to quantify them separately, since one term (the along-track term) can potentially be sampled, in the absence of clouds and high aerosols, while the other term (the cross-track one) cannot, since the instrument is not looking there. Further, we have estimates of both terms, separately, from Corbin, et al.

p.20066, l.10: the choice of the prior or wrong error statistics may also influence the correlated errors to a large extent.

If the comment is saying that incorrect assumptions about the overall measurement uncertainties are important, then we agree. Obviously, if all the random retrieval uncertainties from Boesch et al that we use here are too low, for example, then the flux

ACPD

8, S11812–S11825, 2009

Interactive Comment

uncertainties we obtain will also be too low.

p.20068, l.8: statistical correlations do not make the errors deviate from these so-called "extremes".

We are just trying to say here that there is a spectrum of error correlation lengths/times falling between zero (white noise) and infinity (pure biases).

p.20068, l.13: missing word.

We have added "that" after "unit" to clarify this.

p.20068, l.25 and following ones: what the authors do is not clear.

We have increased the measurement uncertainties in R in an attempt to account for the systematic errors added to the measurements (as advocated by Chevallier, et al, GRL, 2007, for example). To hopefully make this clearer, we have reworded "the assumed random measurement uncertainties" to "the uncertainties assumed in the measurement error covariance matrix, R"

p.20069, l.3: Why do the authors assume diagonal covariance matrices? Is this valid in space in the real world? Is this valid in time?

For the measurement error covariance matrix, R, the time and space errors are convolved, due to the orbital motion: the question is how correlated are the errors from one grid box to the next? (these are separated by 2 deg or about 210 km in the along-track direction for most of the orbit.) It is unlikely that significant random instrument and retrieval errors would be correlated across such large distances, especially after multiple measurements inside a grid box are averaged together. However, modeling errors could well be correlated. A surface-related retrieval bias could cause correlated measurement errors across a large desert, for example. CO_2 modeling errors caused by transport inaccuracies might also cause correlated measurement errors on these scales. We have treated both of these systematic error sources in separate experiments, admittedly without adding spatial or temporal correlations explicitly. One way

ACPD

8, S11812–S11825, 2009

Interactive Comment

of handling such correlations is simply to artificially increase the assumed measurement uncertainties in R, still without adding off-diagonal terms to account explicitly for the correlations. In that view, one could argue that the measurement uncertainties we used in our systematic error experiments might be too low, if such correlations are important.

For the a priori flux error covariance matrix, P_0, the 7-day flux spans that we used are long enough to justify neglecting time correlations (since this span is significantly longer than the diurnal and synoptic scales on which the largest flux errors are most likely to occur). In the space domain, the case is less clear. We have argued, in our response to Reviewer 2, that over land the assumption is likely not a bad one, since the 5 deg longitudinal box dimension used here already enforces about a 500 km correlation in that direction, similar to the 500 km spatial correlation length assumed in Chevallier, et al (JGR, 2007) and Roedenbeck, et al (2003), and so adding additional correlations through off-diagonal terms in P_0 is unnecessary. This argument fails over the ocean, however, where flux correlation lengths are longer. Since correlations in P_0 serve to smooth out the corrections made to the flux prior in the assimilation, by not adding these into our assimilation, we are permitting the assimilation to solve for the finest-scale detail that it can. The danger is that the assimilation may be overfitting, responding too strongly to variability in the data through exaggerated patterns in the final estimated fluxes. If this is the case, then these errors ought to be reflected in our calculated flux error statistics. Adding more appropriate spatial correlations in P_0 ought to result then in lower a posteriori errors (less over- shooting) than what we have obtained here. In that sense, our assumptions are conservative (i.e., if we would have added off-diagonal terms in P_0 to reflect the longer-scale correlations over the oceans, these should reduce unrealistic over-shooting there in response to retrieval and modeling errors, and our overall errors would be lower).

p.20069, l.8: This statement is not correct since the space-time correlations of the priortruth differences are not taken into account. The so-called "perfect model" experi-

ACPD

8, S11812–S11825, 2009

Interactive Comment

ment is rather the first "mistuned experiment".

If the reviewer's argument were to be correct, then we should be unable to retrieve the true fluxes exactly in the "perfect model" experiments, due to errors in our assumed P_0 causing flux errors similar to those in the mistuned experiments (as the reviewer hints at). However, this is not the case. In Baker, et al (2006) we used the same approach as used here, and were able to retrieve the true fluxes to almost arbitrarily high precision, in the absence of added measurement noise. (These experiments also show that the supposed adjoint errors that the reviewer points to are apparently of little importance to the convergence or the flux error results.)

p.20069, l.15: Same.

Again, we do not find these errors that the reviewer claims ought to be there.

p.20069, l.29: do the true fluxes have errors?

No... this was not worded well. We have removed "or their errors".

p.20071, l.9: the authors shift day and night, which may not be a good choice.

It was actually a useful choice, since it introduced a serious error in the representation of the diurnal cycle in transport, in addition to the near-day-long shift in the mean winds. Given our difficulty in modeling the mixing out of the PBL, this test could be seen to be important.

p.20071, l.12-end: at last, the transport error is considered, but in a crude manner compared to the details of the retrieval uncertainty.

Yes, we have not attempted to treat the systematic error cases as completely as the random errors: they are meant to give a rough idea of their importance. As noted in the conclusion, to do the transport error justice, a Transcom-like effort with multiple transport models involved would be required. This is clearly outside the scope of this study.

p.20071, l.15: why this factor? The results are about the tenth of the ppm. Is this realistic? Is there any reference that shows that the authors' model is so accurate?

The factor was chosen somewhat arbitrarily. In retrospect, the values chosen were too low, as can be seen for certain regions when comparing the results of the experiments in which transport errors were added to those where both transport and aerosol bias errors were added: the larger measurement uncertainties assumed in the second case damped out some of the transport-related errors that were not damped out in the transport error-only case. This, however, serves to illustrate the impact of another mistuning error that will surely be present in any real assimilation.

p.20071, l.25: what would an alternative goal be?

Sometimes it is helpful to state the obvious at the beginning of a discussion. The goal would perhaps not be so obvious to all readers as it seems to be for this reviewer.

p.20072, l.3: why was the study restricted to one year?

Time and computational resource limitations were the prime reasons.

p.20073, l.2: should we have expected another behavior?

We do not necessarily think so. However, we could imagine that some readers might think that there would be a similar relative improvement everywhere (in terms of fractional error reduction), and this was not found to be the case.

p.20073, l.20: the increase of the land error illustrates the fact that the experiment is mistuned.

We do not agree that this is correct, and the comment does not give any details to support the assertion. Since the assimilation is optimizing weighted fluxes, there is no requirement that absolute fluxes be optimized, at least not when comparing one region against another.

p.20074, l.24: the authors had to do it but did not.

ACPD

8, S11812–S11825, 2009

Interactive Comment

Again, we refer the reviewer to the experiments done in Baker, et al (2006) to show that, in fact, we did force agreement between the assumed a priori flux error covariance matrix and the actual prior-truth flux errors.

p.20074, l.26: too many words.

We have removed a "using the" to correct this.

p.20075, l.12-19: the explanation is wrong: the convergence should not depend on the "true" error statistics, but only on the assigned ones.

The a posteriori covariance, P_x, does indeed depend both on the assumed error statistics and on the true ones (see equation (4)). So if the assumed error statistics are not in agreement with the true ones, the assumed P_x will not agree with the true one, and there would be no hope of converging in a single step. This, in general, is the case we always face in the real world. So we stand by our explanation and wording.

p.20076, l.8-10: is this realistic?

If the uncertainty reductions seem high, keep in mind that the X_{CO_2} retrieval errors assumed here are considerably lower than in some previous studies (Chevallier, et al, JGR, 2007; Baker, et al, 2006)

Section 3.4: It is difficult to read Figs. 10-11. Further, the results could be condensed.

We have taken Reviewer 2's suggestion of dropping one of the two figures, the seasonal one.

p.20078, l.16: mistyped word

Thank you for catching this. We have changed this to "systematic"

p.20079, l.19: not really.

We are not sure what the reviewer is disagreeing with here. Line 19 says "...while those with the smallest initial errors see the smallest initial errors see little to no im-

ACPD
8, S11812–S11825, 2009

provement." Perhaps the reviewer feels that there is indeed an improvement. We will be conservative and stay with our current wording.

p.20083, first paragraph: the conclusion is pointless. The authors should first improve the accuracy of their adjoint model.

The reviewer seems to be trying to argue that a discussion of the slowing of convergence due to the mistuning of P_0 is pointless, because of some supposed inaccuracies in our adjoint model. Even if the reviewer is correct that errors in our adjoint are responsible for the tradeoff between the convergence of the oceans and the land (we do not agree that this is so), this would not begin to explain the factor of three slowdown in convergence when P_0 is mistuned, nor our argument that the ocean fluxes may never get fully converged if the assimilation is ended before fully converged. In his/her comments above (p.20073, l.2:), the reviewer appeared to feel it obvious that the low-magnitude ocean fluxes should be converged last, after the larger land fluxes were fixed, presumably for good numerical reasons. Now he/she seems to be arguing that they are not converged because of some supposed problem with our adjoint. For our response to the suggestion that our adjoint is not up to the task, see above.

p.20083, l.18: "certainly" can be removed.

To be more conservative, we will drop "certainly" and reworded this to: "...would constrain the key sources and sinks of CO_2 well on a global scale."

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