

Interactive comment on “Evaluating a 3-D transport model of atmospheric CO₂ using ground-based, aircraft, and space-borne data” by L. Feng et al.

Anonymous Referee #1

Received and published: 27 September 2010

This paper uses GEOS-4 and -5 meteorology fields to drive the GEOS-Chem atmospheric transport model, forced by reasonable CO₂ source/sink estimates, to model atmospheric CO₂ concentrations across 2003–2006. An initial set of CO₂ sources/sinks are modified, using an ensemble Kalman filter, to agree with a subset of GLOBALVIEW-CO₂ in situ CO₂ observations (most taken near the surface) across that span. The optimized fluxes are compared briefly to results of the TransCom3 annual mean flux experiment. Then the trend and seasonal variations in the corresponding a posteriori CO₂ fields are compared to CO₂ observations that were not used in the flux inversion, including aircraft and AIRS data taken above the boundary layer. This comparison is used to assess the extent of systematic errors in the transport model, especially

C8004

regarding the model's vertical mixing.

There is value in model analyses of this sort: certainly a study of this sort ought to be done before using a model to interpret column-averaged CO₂ retrievals from satellite, for example. One might want to bring other tracers, like SF₆, into the analysis to be more rigorous, but this is a good place to start. I think the authors have done a reasonably good job with their analysis here. The presentation is generally clear and I like the fact that the authors took the trouble to compare with comparing aircraft measurements. There are a couple of points I would like to have clarified. More importantly, I believe there are a couple problems with the flux optimization approach that I think may well have a significant impact on the results. (Rather than taking these criticisms in a negative way, I hope the authors may find them useful in helping to understand some otherwise hard-to-explain aspects of their results, such as the strong uptake by the northern land areas, and the differences in the trends.) Finally, I make some suggestions about where the authors might go a bit further in interpreting their results, and I suggest an avenue for future work.

I would like the authors to clarify which GLOBALVIEW sites were used in the flux inversion – it is too difficult to tell from Figure 1. Were any of the biweekly aircraft profiles used? This is very important for interpreting the results: if only surface sites were used, then the comparison with the aircraft data after the inversion can be interpreted more easily in terms of vertical mixing errors, whereas if there are vertical profiles in the flux inversion, then interpreting the a posteriori fits depends more on where one looks. Secondly, the global mean flux results obtained from the flux inversion are larger for 2004–2006 than a simple fossil fuel minus atmospheric increase calculation would give. I assume this is because an additional anthropogenic input due to biofuel burning has also been added, correct? If so, please give the annual totals assumed for this term. Further, when comparing your results to those of previous analyses (Transcom) that did not use this term, it might be useful to subtract off the contribution from this term, or at least mention the impact in your discussion.

C8005

The a priori land biospheric fluxes used in this study (from the CASA model) are balanced over the course of a year, or multiple years (that is, the overall magnitude of respiration is chosen to balance photosynthetic uptake), whereas we know that, on a global scale, the land biosphere actually has been taking up carbon in recent decades, however. The ocean fluxes have no inter-annual variability. And the fossil fuel fluxes assumed here have no seasonal or diurnal variability (which is fine, since a good global model for these does not yet exist). The atmospheric inversion done in this paper is thus critical for correcting the deficiencies in these prior fluxes, so that, when run through the transport model, they give somewhat realistic concentrations that may be compared to atmospheric CO₂ measurements. The inversion approach used here is similar to that used in the TransCom3 interannual variability experiment (Baker, et al, 2006) in that the same spatial patterns are assumed for the flux corrections for the same 22 emission regions; what is different is the inversion method used (an ensemble Kalman filter instead of a single "batch" matrix inversion). The fluxes that need to be corrected by this inversion have an impact on atmospheric concentration gradients that takes years to fade away. There are still significant gradients in the meridional direction after three years of mixing, and gradients between the troposphere and stratosphere take even longer to smooth out. If errors are made in the flux optimization step, these errors will corrupt the trends, seasonal cycles, and vertical gradients discussed in this paper, and the conclusions about the model transport may well be incorrect.

To do this flux correction properly, then, it seems to me that a data span longer than the three (or four) years used here would have been more appropriate. If there are significant errors in the concentration field assumed at the beginning of the span (and from the discussion in the text on page 18031 line 16-on, it would seem like there are, since only a global offset was added to correct initial errors), then some method of correcting errors in fluxes before the start of the measurement span is required, otherwise these errors will corrupt the fluxes estimated during the span; since this does not seem to have been done here, this initial condition error may well penetrate into the span for a year or two, at least. Examining a longer span of data (starting in

C8006

2000, say) would make the comparisons done for 2004-on less sensitive to errors in the initial CO₂ field (or, in other words, to errors in the fluxes assumed before the start of the span). Using a longer data span would also have provided a tighter constraint on the long-term trend: I believe the errors in the trends given in Table 4 would decrease if this was done.

The use of the ensemble Kalman filter (EnKF) for this study causes the largest problems, however, I think. The EnKF is an approximate estimation method that is best used on problems that are too large to use exact methods on. For the problem examined here (with $4 \times 12 \times 22 = 1056$ flux variables being solved for), an exact inversion approach could have been used, such as the batch inversion method used in the Transcom study. First, in that case the exact covariance matrix could have been solved for, rather than the ensemble approximation given here. More importantly, though, in that case the impact of fluxes on concentrations could be carried out across the full 4-year span considered. By choosing to use only an 8-month window in the filter (or smoother) used here, the impact of fluxes on concentrations further away than 8 months has been severed. If flux from, say, a tropical forest, is convected upwards and not seen at any surface sites until longer than 8 months later, that flux is unobserved in this approach and can take on any value needed to satisfy the global trend at no cost in the inversion, while compensating errors will be made in other places. In the exact (batch matrix inversion) approach, the transport is carried out for three years or more, so that the impact from all regions is felt at the sites, and thus constrains them. (Using a longer data span would also help here.) The authors also chose to go with an ensemble implementation of the Kalman smoother, introducing further approximations (even if they use an ensemble that is as large as the number of degrees of freedom in the problem). Given the difficulty of getting the mean fluxes correct in these sorts of atmospheric inversions, one should want to minimize errors due to method as much as possible. Why not just use a simple, single matrix inversion to optimize the fluxes? I am worried that the large uptake obtained in the northern land biosphere may have more to do with the flux inversion method used than with the transport model, per se.

C8007

Since the vertical gradients, trends, and even seasonal cycles in concentration examined here depend critically on performing the flux inversion correctly, I would prefer not to have to worry that the results depend on errors in the EnKF or end effects from the short span examined here.

I wish the authors would discuss a bit more whether they feel the GEOS5 fields result in more accurate CO₂ simulations than the old GEOS4 fields, when run through GEOS-Chem. What is the reason for this, in terms of the winds and vertical mixing fields themselves? (Perhaps just examining the runs using the un-optimized fluxes might be best for this.) What do the mean flux results obtained here, especially with respect to the large uptake by the northern land biosphere, imply when interpreted in light of Stephens et al (2007)? Is GEOS-Chem a "bad model" in the context of that analysis because it gets such a large northern land uptake? How do the vertical profile results obtained here compare to those obtained in Stephens et al?

Some ideas for future work: run the a priori fluxes through GEOS-Chem at the 0.5x0.66 deg (latitude/longitude) native resolution of the GEOS5 met fields to reduce representation errors. (The optimization could still be carried out at a coarser resolution in that case, but the fine-resolution prior would hopefully give a better initial fit to the data.) Examine how well GEOS-Chem models global SF6 data: does the model give good inter-hemispheric gradients for SF6, or is the mixing too slow?

Detailed comments:

L22-24: could this bias in trend reflect a lag of earlier fluxes reaching the upper troposphere? How is spin-up done here?

P18031 I16: Was the Palmer 2006 run done with the balanced CASA as well, or did it have realistic inter-annual variability in the land biosphere? If it used CASA as well, then I don't think this global correction is good enough, since most of the uptake may be in the extratropical NH, and a global correction would not correct these well. A lot of the biases seen here may simply be due to not having a spatially-correct correction

C8008

of these initial errors.

P18034: Given that the number of unknowns is only $4 \times 12 \times 22 = 1056$ here, it seems like it would have been better to simply do a straightforward single matrix inversion, rather than using the ensemble Kalman filter. The EnKF has several disadvantages for this problem: 1. there are approximations involved in the ensemble implementation of the Kalman filter that introduce error and degrade the estimate beyond what the traditional Kalman filter would give; 2. the 8 month window assumed here introduces error by neglecting the effect of transport beyond 8 months; 3. a given month of flux is constrained by later measurements only out 8 months in the future, rather than out to the end of the span; and 4. the a posteriori covariance estimate is only an approximation (despite the use of a sample of ensemble members as large as the number of unknowns) of the true one. While the EnKF certainly has advantages for larger problems, for this case it seems like it would have made more sense to use the simpler batch inversion (as was done in TransCom). If the authors were to do the inversion with the batch approach at this time, it would have the added advantage of testing how accurate their EnKF method actually is, in this case (something that is sorely needed in the literature at the moment).

P18036, I1-4: Since you don't give a list of what these sites are, you should at least say how many sites there are in each category (white circles, red squares). The problem is that some sites might be aircraft vertical profiles, in which case we do not know how many levels you use; also, some sites are close to being on top of each other and are hard to discern from the figure. A list would be best, but failing that, some more details in the text should be given.

P18036, I 5: The whole discussion of relative weights and the selection criteria for including stations in the study is not at all clear: please give more details. What do you mean by a relative weight of 6.0, for example? What are the weights relative to? How is the 1 ppm uncertainty for transport factored into the discussion?

C8009

P18036 I20: the -4.4, -3.9, and -5.2 PgC/yr for land+ocean for 2004-2006 are a bit stronger than I get for those years, based on fossil fuel minus atmospheric increase. Is this because the land uptake must account for the biofuel and fire emissions, as well, for those years? If so, please provide us the global annual totals of biofuel and fire emissions that you assume for those years. If not, how do you account for the differences?

P18036, I24: Just pointing to the Stephens et al (2007) is not sufficient here: if you think that transport differences between G4 and G5 are at the root of the flux differences in 2005, you should give some evidence of this. What differences in transport, in particular, do you think explain these results?

P18037, I 9: Given the great interannual variability that is generally thought to occur, especially in the land biosphere, there is little meaning in comparing your results for 2004-2006 with the 1992-1996 Transcom values. It would be more useful to compare your results to the various estimates for 2004-2006 found in the literature. Or, failing that, to estimates for longer spans that will be less liable to the particular span considered.

P18037, I 15-16: "Our global annual G4 and G5 a posteriori estimates have much stronger sinks over northern continents . . .; we acknowledge that the most likely cause is due to changes in biospheric activity." I think that it is quite speculative to say that the most likely cause is an increase in biospheric activity. If you want to investigate that possibility, do the same analysis using data from 1992-1996 to see whether your results are more in line with Transcom or not. A more likely explanation is that using an 8-month window in the EnKF likely is too short to adequately constrain the tropical land regions: they will float to whatever value is convenient to satisfy the global constraint, allowing the northern land to go wherever it wants to go to satisfy the global extratropical N/S gradient.

P18039, I 8-10: why not avoid this circularity by only using those sites not used in the

C8010

flux inversion for the evaluation?

Figure 6: hard to see the differences between model and data on this plot. . . a difference plot would be better for this (we know the a priori is bad, given the use of balanced CASA – that should not be the point of this graph, and that is all that one can see from it).

18040 I 25: this inaccuracy in fitting the trend would improve, presumably, if a span longer than 3 years were to be used. As it stands now, there is not enough of a penalty in missing the trend to compensate for the benefit in other aspects of the cost function fit.

18041 I 1: why would the larger-than-expected land sink explain the inaccurate trend estimate? It is only the CHANGE in a sink over time that affects the trend, not the value of the sink itself.

18041 I 20-21: "Over southern middle latitudes the model has a smaller seasonal cycle and lower concentrations than observed by AIRS, suggesting errors in the fluxes and/or atmospheric transport." Given the variability in the AIRS data, it is easily possible that the difference could be due to AIRS problems, as well. This ought to be noted.

18041 I 22-25: Over northern tropical latitudes, the a posteriori model seasonal cycle is in good agreement with AIRS, but has an amplitude much smaller than the sparse GLOBALVIEW aircraft data that span 5–8 km. We did not observe the difference in seasonal cycle with the ground-based GLOBALVIEW data, suggesting that incorrect model vertical transport plays an important role in the discrepancy between the model and data." You have not displayed the GEOS-Chem model results sampled at the location of the GLOBALVIEW measurements here – is it possible that sub-sampling the model properly might improve the fit to the GLOBALVIEW data?

Section 6.2.3: You might want to just remind the reader here that the CONTRAIL data were not used in the flux inversion, so that they provide a good test of the vertical

C8011

transport in the model.

Editorial comments:

P18028 I13-14: you should reword this: the TransCom experiment was AN intercomparison project, not THE one.

P18031 I13: "inavailability"

P18031 I17: "correction" not "correct"

P18031 I9: "geographical distribution"? missing word. . .

P18035, I 6: instead of "includes", consider rewording this to "accounts for".

Interactive comment on Atmos. Chem. Phys. Discuss., 10, 18025, 2010.