Biogeosciences Discuss., 6, C3415–C3422, 2009
www.biogeosciences-discuss.net/6/C3415/2009/

Biogeosciences Discussions
Interactive comment on "A regional high-resolution carbon flux inversion of North America for 2004" by A. E. Schuh et al.
Anonymous Referee #1
Received and published: 4 December 2009

This paper describes the design and application of a new system to estimate the CO2 balance of the terrestrial biosphere of North America from observations of atmospheric CO2 mole fractions. The system combines many state-of-the-art components including the SIB biosphere model, the RAMS meso-scale atmospheric model, a Lagrangian particle model, and a Kalman filter optimization scheme. Novel in this approach is the separate estimation of ecosystem respiration and photosynthesis for each grid box of the domain. The authors carefully assess the sensitivity of their result to some of the typical inversion related assumptions. This leads to an interesting new way to assess a component of spread around the mean flux results using a jack-knifing approach with the observational dataset. The amount of work done by these authors is impressive and this paper is worth attention from the wider community of readers with an interest in surface CO2 balances. I recommend publication in Biogeosciences, but only after C3415
the following points are addressed thoroughly.

An important point of attention is that from the material presented, I found it hard to judge how well this system really works and what robust features of the North American carbon balance it can detect. This is unfortunate because, especially in the opening paragraphs, previous studies are described with important shortcomings while this paper is touted to overcome them. As an example, the abstract mentions the limit of optimizing with a coarse biome-dependent covariance structure compared to the high resolution biome independent approach taken here. But the paper does not indicate if there really are extra degrees of freedom given the large spatial correlations imposed,

With only eight towers, this is certainly a valid point, especially given the large correlation length scale imposed. However, this is a constraint of the available data in 2004 and not really of the "system". With additional observations, the covariance structure that is imposed here for regularization could be relaxed. This is somewhat true for the global system except that the state space starts becoming very large w/ increasing resolution (globally) and the need to keep many fluxes in the "lag space" due to the large domain that is employed and thus the computational burden should increase more significantly for global problems.

how these extra degrees of freedom help the final solution,

The solution does not constrain flux corrections to be the "same" across ecoregions. The east and west coast forests are given as an example. Another example could be the agricultural regions of the South Central region which have much more variable inter-annual weather (precip patterns for example) than the agricultural regions the northern Midwest.

and if respiration and photosynthesis bias factors are independently retrieved in the posterior solution.

It is assumed that respiration and photosynthesis will NOT be constrained nearly as well as net ecosystem exchange due to the atmosphere seeing the overall CO2 signal much more clearly than the components. Nevertheless, under certain conditions and in more constrained locations, these signals may be teased apart. This was the motivation for the plots that were displayed from the ARM site in Figure 11. The Bayesian paradigm gives us the ability to accommodate this possibility while not destroying the ability to see the more powerful combined signal (NEE).

Are the mentioned "subtle differences" between east and west coast forested regions now resolved?

Difficult to say with forests, per se, because the only real constraint is still in the N.E. United States. We have data in the N.W. United States starting in 2007 and data in the S.E. United States starting in 2008 but neither for 2004. However, one example that I would point out is in the C4/wheat belt of OK/TX where a negative (into ground) adjustment in NEE signal in inferred in the summer time while a positive (into atm.) adjustment in NEE signal is inferred in croplands farther to the north in Iowa. Both areas would likely exist in an "agricultural" biome but the inversion corrects each differently.

I ask you to devote a new section to such questions, in which also flux results are post-aggregated to carbon relevant sub-continental areas (such as biomes). These can then be compared to inventories, other models such as carbontracker, and to the biome-specific SOCCR results.

An additional two paragraphs was added to the "Comparison to CarbonTracker fluxes" section.

Also, I would like to see more information on the obtained match to the observations to assess the realism of the inversion: are CO2 mole fraction residuals indeed Gaussian and is the balance between assumed uncertainty (fluxes+data) and attained skill good (chi-squared)?

For the most part, the residuals are symmetric and do not appear to deviate substantially from normality. There is a slight but pronounced positive skew to

the residuals indicating that when the residuals deviate most strongly, then come from instances where the observed CO2 is greater than the modeled CO2. This will be incorporated as a comment in the paper as well as one add'l figure as discussed below. The weekly chi-square innovation statistics are generally in the 0.5 range from Jan through May in the inversion to around 1 in the summer and after, with more variability in the statistic in the summer time. The low value, primarily in the winter, would seem to indicate some heterogeneity in the model-date residuals, seasonally, and that are assumed "observation" errors in the winter might be too large. In essence, this might lead to the Bayes model being weighted a bit too heavily towards the prior which might imply that the winter time NEE adjustment (towards a sink because of reduction of the source CO2) might be too weak, which seems to be corroborated by the magnitude of the negative winter-time biases in the northern stations in (new) "Biases" table given in paper. This also will be commented on.

Are seasonal biases in simulated mole fractions left after optimization of the bias factors?

There are some biases left in the optimization which will be shown in a new figure and table, partioning winter and summer residuals by tower via empirical pdfs and a table of bias terms.

Does the separate optimization of GPP and R match the full diurnal cycle of observed CO2 mole fractions and observed CO2 flux, or only the daytime when observations were introduced?

Nocturnal CO2 was explored but deemed to variable too provide meaningful data for assimilation (or direct comparison) at the current time. There is certainly ongoing work in this area by several people using u-star filtered results as well as those explicitly investigating nocturnal buildup of CO2 (D. Werth, Savannah River) via tracer release experiments. A comparison of R and GPP at the ARM site (Figure 11) provided very nice results similar to what is being asked for, but this is difficult at most eddy-covariance sites due to the small spatial-scale variability in landcover that can often drives the flux tower observations.

This is a second new analysis/paragraph that I ask the authors to write. Since the paper is quite lengthy already, space for these extra analyses should be created by removing some redundant sections (as indicated below) and by taking another critical look at the sometimes very long, interjected, sentences used. In addition to being hard to read, they also gave me the impression of being lectured sometimes on topics that could be assumed familiar to the readers. I realize this is a matter of personal style though and I do not mean to offend by suggesting this, as the paper is generally very well written. In addition to the points above, which I think are critical before consideration for publication in Biogeosciences, I have listed detailed questions and suggestions below.

Abstract:
line 6: As far as I can tell there is only one inversion based on biome-similar large regions (carbontracker), unless you count the CCDAS from Rayner et al., (2005). If you agree, either make this statement explicitly, or reword the sentence to be more general.

I don't quite agree. It does seem true that most current inversions use grid based approaches, however many inversions, such as those in the Transcom suite, certainly used very large regions whose choices were inspired by biomes and geographical distance. I added Transcom to the abstract in addition to CarbonTracker.

line 8: The example of east and west coast forests is an extreme example of when the biome-approach breaks down and seems specifically chosen to make readers doubt the logic behind such a choice. The authors know of course that a similar extreme example can be chosen for other approaches (such as isotropic distance correlations).

This is certainly true. I've reworded the abstract to state "an extreme example of....". A key component of the paper is to pursue the grid based approach so that a discussion of the shortcomings of the grid based inversion method might not be warranted in the abstract. Nevertheless, the point is well made. Extreme examples for the grid based approach are large decorrelation length scale inversions w/ isotropic correlations sitting over boundaries between ecoregions where one might a priori think the ecosystems would respond differently to external forcings. Additionally, I have made changes to the text on page 10199 where the reviewer had comments about the "degrees of freedom" of both approaches.

Since I feel that this paper presents an alternative to previous approaches and not clearly (unless more evidence is given) an improvement, I would ask the authors to reconsider this line of reasoning in the abstract, and perhaps add a more balanced description of the differences between the assumptions in this paper vs earlier work in the main text.

I will go through the paper and try to reword the text to provide better balance.

line 15: "...provide insights...", please be specific and quantitative. What insights are you referring to? List them, including a number describing the effect.

Agreed, removed. Sort of meaningless self-gratulation.

line 29: "... interesting hypotheses..." again be specific: What do you hypothesize after this study? What can the reader expect to learn?

The estimated sink that we find in the South Central United States is what I'm drawing inference to. Although, the data is limited and the area is on the edge of the inversion domain and thus sensitive to boundary inflow, the strength and robustness of the sink demands some explanation and further investigation. Is the inflow not constrained enough? Could the existence of an uncaptured MidWest carbon sink impact the estimates? Or is there a large sink in an area … where for all intents and purposes, there should be a large sink? Interesting hypothesis for future work.

I've reworded the last sentence in the abstract to read:

"Additionally, the correlation of an estimated sink of carbon in the South Central United States with regional anomalously high precipitation in an area of managed forest and agricultural lands provides interesting hypotheses for future work."

p10198, line 15: I believe this statement is not true. Very few current inversions use coarseness of the estimated fluxes to save computational power. Rather, it is an explicit way to regularize the problem given a small number of observations.

Yes, if each inversion "cell" is treated independently in space than I could imagine that this is needed as a regularization constraint. This has been changed.

p10199, 1st paragraph: The discussion of grid box numbers versus biome regions is incomplete without some statement of degrees of freedom. For instance, NWP models use millions of unknowns and operate at 20x20 km global grids, but the number of degrees of freedom in the problem is much smaller because of the detailed covariance structure. Similarly, an inversion using a grid of 200km with an isotropic distance correlation of 1000km has 18 degrees of freedom. Not 540 (=30x18). This should be mentioned here already, and later in the paper (methods) explicitly calculated for this problem. I already mentioned this in the first review round, but I don't think carbontracker has only 17 ecosystems represented. The Peters et al (2007) paper gives the number 25. This suggests that your system (2x18 degrees of freedom, minus the propagated part of the covariances) is not so different from other inversions.

First, I certainly agree with the "nature" of the comment and there does appear to be 17 for North America (29 Olsen reduced to 19 for CarbonTracker w/ Water, "Non-optimized Areas" accounting for zero percent land area so 17), additionally w/ four of those accounting for less than 0.5 % of the land area each. So, effectively, around 13 or 14 at maximum. Of course, covariance between regions could significantly reduce these numbers as well. (http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/documentation_assim.html#ct_doc).

Nevertheless, it is important to note that Schuh et al. 2009 showed how biases can occur when using a relatively coarse fixed set of regions within an atmospheric inversion as opposed to a finer set of regions, even when assuming spatial-scale patterns of carbon flux errors on the order of 500 kilometers and greater. The argument was similar in nature to Kaminski et al. 2001 earlier arguments.

p10199, line 11: What is simple about your filter, and how does that allow you to work with all portions (what are these?) of your inversion. Be explicit.

I am dealing with the full covariance stucture and not a monte carlo estimate of the covariance structure as in most EnKF, also I am not using an localization scheme. So, in this sense, the filter is somewhat simple, no worries about how good my monte carlo estimate of the covariance matrix is performing or whether I am choosing the right "source regions" to impact the tower observions (localization). Just a simple grid over the domain w/ boundary conditions and treating all flux grid cells explicitly.

p10200, line 18: "estimate true fluxes of...", you mean 'measure' I presume?

Changed to "measure".

p10202, line 14:"global biosphere-transport model", you mean a coupled terrestrial biosphere and atmospheric transport model?

The two global inflow conditions were provided by PCTM-SiB and TM5-CASA, both "decoupled" models of the biosphere and transport.

p10203, 2nd paragraph: To my opinion, this section is long and unnecessary and could be removed to create space for the requested analyses.

Agreed, removed.

p10204, line 1: "... which provides an estimate": You mean to say that the carbotracker optimized CO2 concentration field includes the effect of sources and sinks outside your inversion domain. It now reads as though carbotracker gives you source/sink estimates.

I reworded to clarify.

p10204, line 6: "half-hourly", bring to start of sentence, it now reads as though someone sent you observations every half an hour.

Reworded.

p10204, line 12: "low quality modeled measurements" please rephrase to state more clearly what you mean.

Reworded to clarify : "As a consequence, robust afternoon snapshot observations, at 12, 2, 4, and 6p.m. LT, are used in order to avoid inversion model sensitivity to poor atmospheric transport modeling of extremely stable and stratified nocturnal atmospheric conditions near the ground".

p10204, line 14:"diurnally influenced CO2" please rephrase more exactly.

Corrected to , "For most days, data at this tower consistently showed high CO2 concentrations in the 12p.m. LT records that were more consistent with typical morning CO2 than with well-mixed afternoon CO2"

p10204, line 15-19: This explanation is very unsatisfying. What "kind of systematic late venting" do you mean and what causes it? Is there some special boundary layer dynamics going on at this tower? If you do not know the cause, then simply state that you are using a shifted time inter val without fur ther reference to this unknown phenomenon.

You are correct, some unique boundary layer dynamics at the tower. I think it is important that the reader realizes that this is the likely cause and not artifacts in the data. Modeled CO2 and observed CO2 showed enough similarities to justify authentic unique dynamics at the tower.

Reworded to:
"One exception is the WKWT tower in Moody, TX. For most days, data at this tower consistently showed high CO2 concentrations in the 12p.m. LT records that were more consistent with typical morning CO2 than with well-mixed afternoon CO2. For this tower, mixed boundary layer conditions appeared to be better represented by snapshot observations shifted by 2 hours: 2, 4, 6, and 8p.m.LT.

p10205, top: Did you assume four independent observations each day at each site? Then mention that there is likely significant auto-correlation in these obser vations, that they do not inform independently on your biases, and that the 5.5 ppm should be scaled by a factor of sqrt(4) to compare to other typical inversion setups.

Good point. I replaced the sentence: "It should be noted that while it is possible to run inversions with artificially low prescribed observation errors, this will generally manifest itself in a~need to over tighten the a~priori covariance structure." with "We note that there certainly is expected to be autocorrelation in the errors within a daily time frame so that the "effective" number of independent observations is likely less than 4 each day. The end result is that the observational error term over multiple observations is probably estimated as being somewhat lower than reality. For example, a mean of 4 afternoon

observations has an estimated 2.75 ppm error, based on Gaussian 5.5 ppm independent errors for each observation. In reality, the error of the mean observation is probably larger due to likely temporal correlation in the observation errors."

p10205, line 18: "inversion" please replace by "biosphere model" or "SIB3".

Changed.

p10205, Section 2.4: The paper would be accessible by a larger community of data assimilation experts if the recommendations for notation of Ide et al (1999) were followed.
Please consider this.

I will review this paper. My current attempt was to make it accessible to a statistics audience and use a simple Bayesian updating paradigm as one might encounter in a Introductory Bayesian Statistics course. Certainly it would be useful to cast the methods in the future in both a basic state-space model formulation (as more familiar w/ classical statisticians) as well as a more detailed algorithmic (and notationally similar) fashion as seems to be the choice in the assimilation community. This is certainly a lesson learned and your advice will be followed in the future.

p10206, eq 4: Can you please state these assumptions in words as well.

I've added "... the Bayesian statistical assumptions are a Gaussian distribution on the "measurement" errors as well as a Gaussian distribution on the a priori distribution of <beta vector>, i.e. : "

p10208, eq 10: This would be a good place to discuss the degrees of freedom in your system based on the spatially explicit correlations in your grand prior.

I have included some comments on degrees of freedom at this point in the text. I have used metrics from D. Zupanski's chapter in "Data Assimilation for Atmospheric, Oceanic, and Hydrologic Applications" to estimate "E" dimension and degrees of freedom. For a two week inversion cycle (1000km décorr length), the degrees of freedom are typically between 2 and 8 giving a "measure" of the amount of independent information in the regression. With 1000 km decorrelation length scale and only 8 towers (6 of these are "somewhat" co-located as pairs as well), this is not surprising. The E-dimension is typically between 15 and 30 and gives a measure based more on the complexity of the underlying modeled atmospheric dynamics which drives the regression, as opposed to the quality of the underlying observations. Comments have been placed in the text at the area indicated by the reviewer.

p10208, line 15: I do not understand your explanation of the parameter alpha. It seems like an extra control (in addition to H0) on the correlation, but I couldn't find further reference to it in the paper.

I've changed to read: "The parameter alpha 0 controls what percentage of the covariance can be attributed to spatial covariance, as opposed to spatially independent errors, often termed "nugget" variance. While the "nugget" parameter is an important parameter if one is fitting a rigorous statistical spatial model to the errors, for regularization purposes alpha_0 is often set to zero which is what we will do for the remainder of the paper." This is probably important to include since it is a standard component of any spatial statistical model and adds clarification to spatial statisticians.

p10209, 1st par : This first section is too much lecturing for my liking, please consider deleting it as all the arguments can be found elsewhere again.

Yes, a bit "lecture"-y but I think it is an important point to reiterate. I do not have the luxury of add'l independent data to check the inversion against and it is important for the reader to understand how these inversions certainly can be overfit.

p10209, line 12: The statement that regional inversions have been shown to be very sensitive to inflow contrasts with you earlier statement (p10199,l16) suggesting that you are the first to investigate this. Please fix this.

I've eliminated this sentence. Just a note, previous attempts (by us, unpublished) were considered on a much smaller domain (500KM by 500KM) in the interior of the continent. Boundary variations in CO2 were typically stronger than the net effect of the inversion domain which creates signal to noise types of problems.

p10209, line 20: Is the number of alternatives used in the jack-knifing procedure 100?
I read later on that the results were based on 45 inversions?? Please clarify.

Changed. Yes, 45. Originally this was intended to be 100 but it was accidently run w/ only 45 but w/ reasonable results.

p10210, line 19: Since the matrices are unitless and the sigmas carry the true values, please add units to the quantities here.

The betas are unitless, G is in "ppm", sigma_obs is in "ppm". Please re-review, I think this should be correct.

p10211, line 12: This is the only statement on posterior correlations and it suggests that the bias parameters were not at all independently retrieved. Hence my request to discuss this in much more detail.

First, I will assume you are talking about the correlation of the correction factors and not the correlation structure of the errors in the correction factors. If the overall GPP flux is very accurately captured by the a priori model than one might consider that the beta corrections for GPP and Respiration might appear much more distinct. However, if annual GPP (and hence Resp due to annual balance in SiB prior) is significantly higher in real life than the model, say 50%, than Resp will also be approximately 50% higher. We know that the annual NEE sum is on the order of a few percent of annual GPP or annual Resp. Therefore, there are really two corrections that need to be made, a large scale correlated beta correction due to the approximate balance of the annual carbon cycle and a small scale beta correction to capture to very fine imbalance that the net sink creates. I've attempted to overemphasize the strong a priori constraint of annual NEE = 0 in SiB, which is conjunction w/ a source/sink on the order of a few percent of GPP/Re almost guarantees that the correction factors for Re and GPP will be correlated over large scales.

p10211, last par : I could not figure out the relation of this section to that on page 10209. Are you describing the same procedure again or is this a different test? I'm sure this is due to my lack of understanding, but likely other readers might also not follow the story here.

You are correct, it is the same procedure. I thought I needed to mention it in the Methods but I didn't want to go into detail until the Results section because the reader would have to recall all the specifics as I was presenting the results.

I've added a ", that was first mentioned in Section 2.5," to the first sentence of that paragraph.

p10212, line 4: The suggestion that the posterior error is too small is not really relevant to me: this number is so strongly method dependent that a comparison to Gurney and Peters does not say much. What is much more important (coming back to my main points above) is that the posterior errors are in good balance with the obtained skill (CO_2 residuals, flux residuals). If a comparison to other inversions is presented it would perhaps be more informative to know the uncertainty reduction.

While I have agreed with most of your points, I would have to disagree with this. By providing many many more degrees of freedom than justified, it should be relatively easy to fit the CO_2 record nearly perfectly and not necessarily taint any diagnostics. The example of a very fine grid resolution w/ completely independent correction factors and relatively tight prior variance bounds comes to mind. It is my impression that grid based inversions can dangerously overpredict fluxes, i.e. provide posterior estimates of fluxes w/ too tight of posterior variance because of the general unconstrained nature of the inversion

problem. There is a balance that is usually sought by seeking out a "decorrelation length scale" or an "optimal" number of independent inversion regions. Although this usually manifests itself in strange results, dipoles and so forth, I feel this is an important part of the inversion problem to always be aware of. Additionally, I, as well as many others, do not have an explicit representation of the classical dynamical noise term of the state space model which tends to lead to some instability in posterior variance estimates (leads to resetting the covariance at every filter step, inflation schemes, etc). Again, I don't think this can be over stated as a general concern in most inversion papers.

p10212, line 11: I assume you summed the posterior covariances of your filter, not just the variances?

I am not showing "spatial" covariances here. I am summing up "marginal" cell specific variances for this plot which include the covariance between Resp and GPP terms for each cell. This is in contrast to the bounds given in Figure 4 from one particular inversion realization. In this case the full covariance structure is used because I'm plotting a domain-wide flux, which of course has to incorporate spatial correlation as well as Resp/GPP correlation.

p10212, line 23: This paragraph again repeats earlier statements and could be re-moved or shortened.

I've collapsed the first 3 sentences into 1:

Inflow of CO2 from the boundaries has typically been a large concern of regional models (Gerbig et al., 2003; Peylin et al., 2005) and should be investigated.

p10215, line 6: Did you also experiment with alpha0?

No. Refer to previous comments about alpha0.

p10215, section 3.4: I really enjoyed reading this section, and the analysis presented. The idea to use some performance weighted measure of multiple inverse realizations is interesting and could be mentioned in the abstract as novel.

I appreciate it. I would certainly agree that it was interesting to see these results. One pitfall I would be careful of when employing a performance weighted approach is the potential overfitting of the model leading to inaccurately weighted realizations. Nevertheless, I appreciate the complement and I do think it is an interesting direction to pursue.

p10216, line 26: A third explanation for the prominence of prior patterns in posterior flux fields is that the prior was quite good to begin with. After all, the size of the adjustments is as large as the flux itself in some locations suggesting that priors were not too tight, and data helped change fluxes. Whether your system was too rigid or not can of course be diagnosed from the filter statistics.

I've reworded and changed this last sentence a bit: "In light of somewhat limited data for both inversions, this result is not surprising and most corrections must conform implicitly to the a priori flux fields in some fashion, whether through the coarse biome regions or large decorrelation length scales .

p10217, section 3.6: The promised comparison to Ameriflux data now only entails a figure that shows improvements at ARM. What about the rest of the dozens of sites in Ameriflux? To pre-empt (other peoples') idea that you're cherry-picking I suggest to include a table or figure showing statistics for as many sites as possible in the network.
Without this, your promise of a comparison to Ameriflux level-4 data is overstated and should be revised.

Agreed. I've changed to read comparison to Southern Great Plains flux site. At the time this manuscript was written, there were only a couple level 4 data sets available. Most of those were sites that were somewhat complex and heterogenous, not ideal to compare large scale fluxes to. ARM is unique in that the area is somewhat large and homogenous, when compared to mixed forests/ag/grassland areas.

p10218, line 3: remove "approximately", your numbers are exact.

Removed.

p10218, line 5: the carbontracker website currently shows a number of -0.69 ±0.51 PgC/yr for 2004 while it encompasses a larger area than your domain. Did you match up the domains to get to 0.9PgC/yr or is this simply a wrong number quoted? Or has carbontracker revised its numbers down since the Peters et al (2007) paper and if so, why?

Carbontracker updates its numbers nearly annually now. I believe I was using an "old" number of about 0.89 PgC/yr possibly from 2007B. The website does not seem to contain old source/sink estimates for specific years. Moot point. I've changed it to read 0.69 PgC/yr now.

p10219, line 13: "given its late diurnal venting", what was first speculation has now become a given, but I still object to this without some more detailed explanation.

I have now changed to read: "CO2 observations at the top of the tower did not appear to be well mixed until well after 12PM L.T.. Additionally, the tower is located relatively closely to both the model boundary and the ocean and is in close proximity to fossil fuel sources of major metropolitan areas and oil refining facilities near Houston and Galveston."

p10219, line 18: The answer whether there indeed is a correlation to cause this behavior is in your covariance matrix, please investigate this.

While it seems a likely cause for missing what we believe is a sink in the northern Great Plains, there were no indications of this correlation in the covariance matrix for the annual NEE fluxes. There were very small amounts of larger positive correlation induced by the spatial smoothing of the prior as well as large amounts of very very weak negative large scale correlation indicating constraints at larger scales. Therefore, I have removed this sentence.

p10220, last paragraph: your discussion of crop harvest removal and its relation to the estimated NEE is in contrast with the Peters et al (2007) notion that a sink is seen but a source not. Please explicitly state this contrast.

These two paragraphs have been completely rewritten. I'm currently principally involved in the Mid-Continent Intensive project of the NACP which is studying the Midwest in particular. I don't want to "step on" future publications but I must comment in this paper, relative to what I now know from emerging research on that project.

p10221, line 10: "sources" should be "forests" or "ecosystems"?

Yes, changed to ecosystems.

p10222, line 1: I'm not sure I understand your low NEE resulting from a lack of constraints oh high GPP regions. The Intex data was not shown here but if it really holds the key to some problems in your estimated NEE it should be shown and discussed.

I will remove the following: "NEE estimates for the entire domain appear on the low side of estimates derived from global inversion models, which is understandable given the lack of constraint on some key regions of high annual GPP, and hence potentially high annual NEE. This was corroborated by a comparison to INTEX aircraft data which shows the existence of a deficit in GPP over the southeast which would, when all other things are considered equal, potentially inflate the domain-wide sink closer to levels estimated from global models such as CarbonTracker."

Although I did not include discussion of INTEX, I wanted to mention it simply because of the work it took to rerun new fluxes through the transport and compare to aircraft data. What I learned from it was that there were significant errors in the flux prior in that area of the United States in mid summer due to high plant stress in SiB. Nevertheless, we have no CO2 data available in that area of the domain to do much about it. Furthermore, it doesn't appear to lead to any unusually low/high annual NEE estimates, i.e. stays near the annual prior NEE values.

In conclusion, interesting, but not pertinent to the paper.

p10222, line 15: A 30

??

Figures:
Figure 1 is to me quite uninformative and could be removed in favor of some more substantial results as suggested by my first two points. The IGBP soil class map can then be referenced through a URL or in supplementary material.

Figure 2 could follow this recipe.

Agreed, removed.

Figure 3: I would like to see a histogram of these differences, preferably by season. Labels need to be added with mean differences and the prescribed model-data mismatch.

Figure 4: 100 inversions? Sorry for my confusion.

Again, sorry , no confusion, just mistake. Originally was supposed to be 100 inversions but final run only had 45. Changed to 45.

Figure 6: Can you add histograms of the differences, or seasonal/monthly means? The systematic differences peaking at 10 ppm are large and support my idea that the flux differences they induce are large, not small.

I've added a "BOLD" smoothed estimate via a local regression low pass filter w/ a 30-day span window for ease of viewing as well. I was assuming that the viewer would immediately see that the difference in the inflows is, to the first order, a difference in amplitude of the underlying prior seasonal cycles. I'm hoping the scale of variability over any particular window of time could be approximated by the viewer by eyeing the time series and this provides an additional source of information.

Figure 10: The adjustments in both systems do not look ecoregion shaped, or isotropic distance shaped liked the covariance matrices suggest. Can such patterns be seen in the beta-factors themselves, but they are wiped out by the flux patterns in RE and GPP?

Well, I would say if you gave somebody the ecoregion map used by CarbonTracker and they had some knowledge of inversion systems, one would be able to guess pretty quickly which results are based on an isotropic scheme and which on a fixed biome region approach. The results I present from SiB are

certainly more smooth in space. However, you have to remember that is the result of corrections being performed every week on a prior flux map that is also changing every week. Additionally, the isotropic portion is forced from the prior but there is also a lingering component that is carried from week to week as well as the actual data from the week that is currently being estimated. So the posterior covariance will not be exactly isotropic.

Although the prior will be evident in the CarbonTracker results, it is not as clearly evident as one might think. Envision a biome region split into two spatially distinct areas, one very strongly drawing down in June and the other in August. If a correction is made in June but not August then the associated "change" induced by the inversion over a 3 month period will likely only show the region that was active in June because the correction factor times the other biome area is nil because it is not active at that time.

So in some, sense , yes, they should certainly be evident in the beta factors but less so when convoluted w/ the prior fluxes.

Figure 11: Could you make these figures wider than tall instead of taller than wide?

Done.

Figure 12: This could follow figure 1 and 2 to make room for my requested additions.

Agreed, removed.