Biogeosciences Discuss., 8, C3988–C4001, 2011
www.biogeosciences-discuss.net/8/C3988/2011/

Biogeosciences Discussions

Interactive comment on "North American CO_2 exchange: intercomparison of modeled estimates with results from a fine-scale atmospheric inversion" *by* S. M. Gourdji et al.

S. M. Gourdji et al.

sgourdji@umich.edu

Received and published: 29 October 2011

We thank the reviewers for their constructive comments and suggestions, which will help to improve the quality of the manuscript.

Comment: "This study presents the results of a geostatistical inverse modeling approach to constrain the carbon budget of North America at high spatial resolution (1x1 deg) for the year 2004. Findings are presented at temporal scales ranging from monthly to annual, and at spatial scales ranging from continental to regional (selected biomes/climate zones). The new results are put into perspective using several existing inverse modeling studies and bottom-up biosphere process models as references, discussing differences both qualitatively and quantitatively. Finally, the authors point out the importance of accurate data for incoming boundary conditions for CO2 in regional inverse modeling studies by comparing 2 different data sources.

Overall, this manuscript is well written and well structured, with methods and results described in a plausible way. Assuming the main objective of this paper is to demonstrate the potential of high-resolution geostatistical inversions at the regional scale, the authors generally succeed in demonstrating the potential of their method, and how it compares to existing carbon budget estimates for the same model domain. However, they place too much confidence in parts of their results, and some of their interpretations on model-model intercomparison are not supported by the findings. Also, they fail to explain what causes the large differences between results using different boundary conditions for CO2. Moreover, the text needs to be shortened significantly to improve readability. Finally, posterior uncertainty estimates are required, which I believe should be doable with the given modeling framework.

Summarizing, I believe that this manuscript provides novel findings that should be of interest to the readers of BioGeoSciences. The overall quality of the study is good, with just minor flaws (listed below) that should be straightforward for the authors to improve. My recommendation is therefore to accept this manuscript for publication after minor revisions.

MAJOR COMMENTS

First of all, this manuscript needs to be streamlined to improve readability. I understand that the authors prefer to describe ALL their methods in detail, and present ALL their results, but this carries the risk of burying the new and important material with less important stuff. I therefore recommend shortening this paper by at least 25%, including all sections. I added some suggestions for sections to reduce/cut further below."

Thank you for the helpful suggestions in terms of streamlining the paper. We will especially try to reduce the length of the methods, referencing previous work and putting

BGD

8, C3988–C4001, 2011

Interactive Comment

less central details in the supplementary material. Also, the suggestions regarding the most relevant results to present are also helpful.

Comment: “Second, as long as the authors do not use any datasets to evaluate the quality of their results, this is ‘only’ a model-model comparison, and all interpretations and conclusions should take this fact into account. Since there is no reference data available to validate spatial carbon budgets at the relevant scale, this approach is totally acceptable, and overall I acknowledge the efforts to demonstrate how their results compare against numerous (model) references. However, this approach implies certain limitations in the result interpretation, since you cannot truly evaluate which model is right and which is wrong in case their results differ. There are many passages in the text where the authors use such qualitative evaluations, and most of these statements need to be toned down since they are not supported by the data. See details below.”

We agree that this is only a model-model comparison, which is unfortunate given the lack of relevant validation data at large regional scales. Therefore, most of our efforts have gone into evaluating and improving our method and data sources to give us some confidence in our results relative to that of other studies. However, we acknowledge that even with the most careful efforts, the GIM results may still have unknown flaws and biases, and therefore any conclusions from this study are subject to revision in future work. We will carefully revise the discussion to reflect this.

Comment: “Third, one thing that puzzled me after finishing to read the manuscript is the rather unbalanced treatment of the boundary conditions issue. In great length, the authors discuss the role of prior settings, spatial and temporal flux patterns and how they relate to the reference models, only to state at the end that the impact of all these factors is dwarfed by the choice of the incoming boundary conditions for atmospheric CO2. They test 2 different options, with significantly different results, but do neither explain what causes these huge differences, nor do they give any indication which one may be the better choice. Consequentially, if the objective of this study was to add a new (geostatistical) number to the existing modeling results to constrain the North

BGD
8, C3988–C4001, 2011

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

American carbon budget, the current manuscript version fails to deliver an answer. Since this issue is obviously so important, the authors need to discuss it in more detail, and they need to provide a guideline which version they favor and why."

One outcome of the study was that while prior assumptions, transport, flux resolution, etc. are important in recovering fluxes at sub-continental scales, the boundary condition issue trumps all other setup concerns at the continental annual scale. A simple calculation using the mass of the atmosphere and the fractional area of North America can help one to see how small differences in the CO2 concentration of incoming air to the continent can lead to large differences in the annual continental budget. Given a residence time of about 10 days for air passing over the continent, it would take a flux of $\sim$1.5 to 2 PgC/yr over North America to bring the CO2 concentrations from 0 to 0.5 ppmv globally over the course of a year. If one considers mixing only within the lower 2.5 km of the atmosphere, then it would require a flux of $\sim$0.5 PgC/yr. These calculations are consistent with our findings of a 0.8 PgC difference in the North American continental budget for an average offset in the boundary conditions of $\sim$0.5 ppmv.

In terms of which of the two sets of boundary conditions may be a better choice, there are known seasonal biases in the CarbonTracker CO2 fields that have been corrected in the empirical, or GlobalView, boundary conditions. Using the GlobalView boundary conditions also returned a net annual continental budget more closely in line with estimates from an inventory approach, as mentioned on page 6807, lines 19-21. However, without more substantial validation of this empirical dataset in the peer-reviewed literature, it is difficult to make claims in this paper as to the relative quality of the results using the two different boundary conditions. Rather, this manuscript aims to demonstrate how important it is to accurately quantify these boundary conditions for continental carbon budgeting in future work.

Comment: "Finally, it is hard to put the quantitative results into perspective without posterior uncertainty estimates. These uncertainty ranges need to be provided to allow the reader to evaluate the quality of the findings."

BGD

8, C3988–C4001, 2011

As mentioned in the response to reviewer #2, we will include uncertainties in the final manuscript for flux estimates at spatially-aggregated scales (as shown in Figures 4 and 6). In order to recover realistic confidence intervals, we will rerun the inversions incorporating a priori temporal covariance, which were shown to be needed for obtaining accurate confidence intervals at temporally-aggregated scales in Gourdji et al. (2010). Recent work improving the computational efficiency of such calculations has made it possible for us to do this for a full-year inversion.

Comment: “RECOMMENDATIONS TO REDUCE MANUSCRIPT LENGTH p. 6777 (abstract), ll. 1-10”

We would like to include some motivation for the study within the abstract, but agree that that these statements can be shortened.

Comment: “p. 6780 (introduction), ll. 1-7”

We will think about how to cut this paragraph, given that this theme of the source of model spread is discussed throughout the paper.

Comment: “p. 6782f (methods, GIM): Since all of this has been described in detail elsewhere, this would be a good place to shorten and refer to other publications for more details”

We agree, and will do this.

Comment: “p. 6786, ll. 15-23”

We feel that this information about the relative constraint on fluxes throughout the continent is relevant for the discussion of results, given that we trust our flux estimates most in the well-constrained areas. However, we will try to shorten this.

Comment: “p.6797, first paragraph in Section 3.2.1: This is a repetition of the figure caption, and should be deleted in any case. Also applies to the relevant sections describing figures 4-6.”

BGD

8, C3988–C4001, 2011

We agree, and these statements will be removed.

Comment: “p.6799, Section 3.2.2: Overall, this is much too long. Should be reduced by 50% in length, removing much of the detail on specific BU models and ecozones.”

We will work on shortening this section, as also suggested by other reviewers. For example, we will replace the comparison to specific biospheric models with a comparison to the biospheric model median, as in Figure 3, and consequently shorten the text comparing GIM results to individual biospheric models.

Comment: “p.6804ff, Section 3.3.1: I do not think this section is necessary at all. You already discussed differences in fine scale spatial patterns on the monthly basis, and there are no new insights here.”

We will remove this section, moving important details to the discussion of Figure 6 (Section 3.3.2).

Comment: “MINOR COMMENTS p. 6780f, last paragraph of introduction: this should be aligned better with the abstract. For example, you fail to mention the boundary condition CO2”

We will add a statement to the introduction presenting the use of different boundary condition datasets.

Comment: “p.6781, Section 2.1: You should clearly state at this point what data year you are going to work on”

We agree, and will add this.

Comment: “p.6784, Section 2.3.1: Since the authors raise the issue themselves, they should try harder to explain why they base their study on this relatively data-poor year. Especially for hi-res inversions, more towers should make a big difference.”

We agree. 2004 was originally chosen as a year of interest due to the large availability of aircraft data. A follow-on study (Mueller et al., in prep.) will present results from 2008

BGD

8, C3988–C4001, 2011

Interactive Comment

using 35 towers across the continent, and assessing the additional information content of these new data-streams relative to the network in 2004.

Comment: “p.6786f, last paragraph of Section 2.3.3: this last paragraph is confusing - I got the message after reading it several times, but you may think about rewording to clarify what kind of aggregation and model comparison you are talking about.”

We will rework this paragraph for clarity, and potentially merge it with the preceding paragraph.

Comment: “p.6787, Section 2.4: not clear how temporal covariances in Q were treated in the end in this study. The specific paragraph needs clarification.”

We did not include a priori temporal covariances for the fluxes presented in the Discussions manuscript. However, for the revised and final manuscript, we plan to add temporal covariance into the Q matrix in order to provide realistic confidence intervals with our flux estimates. As mentioned in the discussions manuscript, this may affect flux quality in times of the year with rapid change, but overall, we believe that these assumptions should improve both flux estimates and uncertainties at spatially and temporally-aggregated scales. We will update this section to reflect these changes.

Comment: “p/6787, l.24 (and throughout the text): I do not think ‘biome’ is the right term for your selection of sub-regions. You are talking about a climate zone that is dominated by a certain plant functional type, which are in no way homogeneous biomes. An alternative suggestion could be ‘ecozone’, or something like that.”

We will change the terminology from biome to eco-region, as in Gourdji et al. (2010).

Comment: “p.6788, Section 2.5.1: I think this needs more information for readers who are not familiar with the ’model of the trend’ approach. You may want to give an example, e.g. how correlation with temperature may link your flux fields to seasonal cycles, etc.”

We were trying to reduce text on information that had been covered in previous studies

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

(e.g. Gourdji et al., 2008, 2010), although we can add a statement here referencing these papers, and further explaining how the covariates included in the model of the trend provide ancillary information about CO2 fluxes that is consistent with the atmospheric data constraint.

Comment: "p.6789, ll. 6ff: related to the previous comment, to allow a correct interpretation of your results, you need to provide more detail on how exactly your auxiliary variables are used to produce the final flux fields. In particular, you need to explain how you can set up an inversion with just this one driver that can obviously only explain parts of the fluxes, as compared to the more comprehensive use of a set of NARR variables below."

We will better explain how the trend component in GIM is used to explain a portion of the flux variability, with the remaining variability that is seen by the atmospheric measurements (but not explained by the flux covariates) represented through the spatio(temporally)-correlated flux residual from the trend. The Simple inversion presented here is conceptually similar to using a prior of zero for the biospheric fluxes in a Bayesian setup.

Comment: "p.6789, ll. 15ff: Add spatial resolution of the NARR dataset"

In the geostatistical approach, the auxiliary variables (including both the NARR and fossil fuel inventory datasets) that are included in the trend have to be defined at the resolution of the estimated fluxes (i.e. 1x1). We will make this clear in the manuscript.

Comment: "p.6790, Section 2.5.2: please provide spatial resolution for VULCAN and datasets used in the other regions"

Again, these are defined at 1x1 for the purposes of this study, and this will be further clarified in the manuscript.

Comment: "p.6791, Section 2.6: You should include some general statements here on the value of model-model intercomparisons, and the interpretation of differences in the results. For example, you should mention that, in principle, even good agreement with other models wouldn't mean that this model works well."

We agree, and will add a statement to this effect. However, we do believe that at spatially-aggregated scales in the well-constrained areas of the continent, the GIM results have been extensively validated with synthetic data testing, and should provide insight into the biospheric model spread. Without any confidence in our results relative to those of other models, it would be hard to draw any conclusions from this model inter-comparison at all, which would be unfortunate.

Comment: "p.6793, Section 3.1.1: These findings need some more clarification, parts of which may be covered by some more background info on the model of the trend concept. I do not fully agree that a beta of 1 is the desired outcome here in any case. What about some incidental spatiotemporal correlation between anthropogenic and biogenic flux fields? In that case, wouldn't the fossil fuel fluxes be scaled to cover parts of the biosphere signal?"

The reviewer is correct that there is some correlation between the anthropogenic and biospheric fluxes. For example, when including both the fossil fuel inventory and NARR variables within the trend, we see non-neglible covariance between the betas on the fossil fuel inventory and biospheric variables such as snow cover and specific humidity, perhaps because warmer areas tend to be more populated in North America. We will try to reword the discussion here to emphasize the uncertainty associated with this result.

Comment: "p.6794, ll. 20ff: I have some problems with this part of the X interpretation. specific humidity should be a driver for GPP rather than for RH. The way you present the role of specific humidity makes it sound more like an artifact - if you've got precipitation at different timescales included, you should have the soil moisture well covered for RH. So adding a temperature measure should do a better job to complete the picture for RH."

BGD

8, C3988–C4001, 2011

The specific humidity variable has an 83% correlation with air temperature, and therefore, this variable is most likely helping to explain the temperature sensitivity of RH. It is not entirely clear why the air temperature variable itself was not selected by the variable selection routine. It is possible, for example, that specific humidity is not only helping to explain the RH signal, but also to modify the GPP signal as captured by the evapotranspiration variable. Given that the specific humidity and evapotranspiration variables are 41% correlated, the positive β on specific humidity may imply lower correlation of GPP with evapotranspiration at high humidity when the stomata are already open all the time. We will clarify this discussion in the revised manuscript.

Comment: "p.6795, ll. 22ff: The last section of this paragraph is confusing. Explain better!"

The statement made here is similar to those on pages 6788 and 6789, which also prompted the reviewer's concern. Without being repetitive, we will attempt to further clarify in the manuscript the role of the model of the trend in GIM.

Comment: "p.6797, ll. 5ff: I do not agree with this statement! In both cases, there could be the same type of systematic error that causes biases, so better agreement between classes of models doesn't necessarily mean that these models are more accurate."

Although we cannot rule out systematic errors affecting all inversions (or all biospheric models), we do believe that the atmospheric data constraint should be useful for evaluating process-based model assumptions at large regional scales. If the flux estimates from the inversions are similar despite very different setups, transport models, etc., this provides some hope that the inversions can give insight into the biospheric model spread. However, we will add a qualifying statement to this paragraph indicating that systematic errors in the inversions cannot entirely be ruled out, and that therefore all conclusions drawn from this inter-comparison must be tempered with this uncertainty.

Comment: "p.6798, ll. 7ff: You should mention seasonal biases in inverse modeling results here as well. For example, what about the role of boundary layer height dynamics

BGD

8, C3988–C4001, 2011

across seasonal timescales?"

There is ongoing research in co-author John Lin's research group to quantify boundary layer biases in WRF when compared to radiosonde measurements, but preliminary results show larger relative biases during the growing season than in the winter months, contrary to the results presented in the manuscript showing closer agreement between GIM and the biospheric model median during the growing season. However, given recent errors found in the radiosonde database, this work assessing boundary layer biases is subject to future revision.

Comment: "p.6799, ll.24ff: these findings may be artifacts - they depend a lot on the setup of the aggregation regions, and how they correspond to spatial patterns in the NARR variables."

We think that it is unlikely that this result is an artifact of the aggregation regions, especially at these large spatial scales. This result is consistent with results in Gourdji et al. (2008), as well as numerous sensitivity tests with both real and synthetic data, showing that the auxiliary variables have only minor impact on spatially-aggregated flux estimates in well-constrained regions of the domain.

Comment: "p.6801, ll. 12ff: I think this is an over-simplification of the whole matter - the results depend on the balance between Q and R matrix settings - if too much confidence is placed on Q, it will always bias the results, or rather influence it strongly."

The reviewer is correct that the variances used in Q determine adherence to the priors in synthesis Bayesian inversion studies. In the geostatistical approach, we optimize Q and R parameters simultaneously using the atmospheric data, and have thus far consistently found the components in the trend to have little influence on the final flux estimates at aggregated scales in well-constrained areas, when the covariance parameters are properly optimized, as they were here. However, we should not assume this to be true as well for other synthesis Bayesian inversion studies, whose flux estimates are presented for comparison here. We will rephrase the discussion here to reflect this complexity.

Comment: “p.6802, ll. 22ff: This is the kind of statement that isn’t supported by your results, since you don’t have reference data to prove which model is ‘right’. there’s no reason to believe that the inversion works better than any other model, as long as you prove it with data!”

We plan to remove this discussion comparing GIM results to individual biospheric models, although as suggested in the paper, given the care put into validating our results with synthetic data testing, we believe that our results can give insight (based on the atmospheric data constraint) into the spread of the biospheric models. While no statement can be conclusive as to which model is right, it is interesting to at least draw preliminary conclusions from the comparison, assuming that one model may be better than another based on the knowledge of the strengths and weaknesses of the model setup.

Comment: “p.6803, ll. 4ff: Same as above”

Again, we plan to cut this discussion of individual biospheric models, in comparison to the GIM results. We will replace it with a comparison to the biospheric model median at the monthly biome-scale.

Comment: “p.6803, ll. 10ff: some of these conclusions need to be put into perspective. #1 is largely dependent on the setup of Q, so in this form it isn’t generally valid. #4 should be toned down, noting that the inversions might be as wrong as the bottom-up models.”

We will think about ways to rephrase conclusions #1 and 4 in this paragraph, while still remaining faithful to our findings (in terms of #1), and having cautious confidence in our GIM results (#4).

Comment: “p.6803f, first paragraph of Section 3.3: maybe you should note here that the annual fluxes, though subject to large errors, are the one number that people want to see! So even though they may be uncertain, this is the one aspect that should deserve most attention for future improvements!"

We do state exactly this on page 6792, lines 16 to 21, although we can re-iterate it later on as well. This is an important point.

Comment: "p.6805, ll. 14ff: please clarify how exactly you handled the fossil fuel fluxes for the bottom-up models."

The biospheric models all estimated NEE directly, and therefore there was no need to separate the fossil fuel signal from this model output. In the annual total plots, the fossil fuel emissions from the inventory were simply added to the flux from the biospheric models.

Comment: "p.6806f, ll. 27ff: not sure if this discussion is helpful at all!"

This text presents a hypothesis to explain why the GIM inversions show weaker net sinks in this biome relative to the other inversions. We will reword to minimize speculative text here.

Comment: "p.6807, conclusions section: As mentioned above, many of your statements here should be toned down since your results do not support qualitative statements on specific model performances. Particularly the last paragraph on p.6809 is largely speculative!"

We will rephrase speculative statements, but we do stand by the general conclusions from the study. In terms of the last paragraph on p. 6809, the strong sources seen in the GIM results in March and October are supported by observations showing rising CO2 concentrations at many of the towers in these months. We will add a plot showing these CO2 measurement time series to the supplementary material to support the conclusions drawn in this paragraph.

Comment: "p.6810, ll. 3ff: I don't think you can use the observed differences between Simple and NARR to conclude that grid scale inversions are impossible to date. These differences should be obvious, given the database differences between both. The main problem is rather the availability of atmospheric observations."

We did not mean to imply that inversions were impossible at this scale, but instead that the atmospheric data constraint does not allow for conclusions to be made about net fluxes at the grid-scale. (We will also rephrase the statement in the text to clarify this.) The reviewer is correct that the limited availability of atmospheric observations is a strong reason for this. There are, however, other reasons as well, including, for example, limitations associated with transport models, boundary conditions, etc. In terms of the direct impact of including more CO2 observations in inversions, however, preliminary work using the 35-tower network in 2008 (i.e. Mueller et al., in prep.) shows that the differences in the continental budgets with the two sets of boundary conditions are just as pronounced as with the 2004 network. So, while the increased density of observational data can help to improve the quality of flux estimates at sub-continental scales, the boundary condition issue will still exist for continental budgeting.

Comment: "p.6829, Figure 3: The color scales should be uniform for all figures"

We experimented with this, but since the primary goal of this figure was to compare the spatial patterns across models (within a given season), we thought it best to use different color scales to better visualize differences in spatial patterns. Otherwise, in one or more of the seasons, it would be difficult to see differences across models.

Comment: "p.6831, Figure 5: explain which boundary condition is used for which row of figures!"

We will add an indication of which boundary condition is used to both the legend and the figure.

Interactive comment on Biogeosciences Discuss., 8, 6775, 2011.