

Biogeosciences Discuss., 6, C3722–C3726, 2009
www.biogeosciences-discuss.net/6/C3722/2009/
© Author(s) 2009. This work is distributed under
the Creative Commons Attribute 3.0 License.

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 #2

Received and published: 29 December 2009

This paper presents results from an inverse modeling study aimed at quantifying the North American carbon balance for 2004. Several aspects of the paper are very appealing, including the use of multiple sets of boundary conditions, fossil fuel inventories, and sets of a priori covariance parameters. The authors also have a very thorough section interpreting the results of their inversion. There are, however, some fundamental aspects of both the approach and the results that are problematic, and in some cases not discussed in sufficient detail. As a result, it is difficult to objectively interpret the quality of the presented results, and therefore it is difficult to gauge the scope of the revisions that will be required.

GENERAL COMMENTS

First, the timescale on which the bias parameters are allowed to vary is unclear. Based on p. 10, the bias parameters change on a weekly basis. If the parameters are indeed allowed to change during each 7 day period, is there any temporal correlation assumed between weeks?

There is no explicit temporal correlation. The smooth evolution of the bias parameters are due simply to what is inherent in the data as well as the weight given to the posterior mean from the previous week, which is propagated forward as the prior mean for the current week.

If so, it is not described. If not, that would not be consistent with the authors' statement regarding the fact that these parameters are assumed to vary slowly in time. I would also suspect that this yields quite noisy bias parameters unless some reasonable temporal regularization is included, especially in the early parts of the year.

I am familiar w/ this effect. GPP is essentially unconstrained in the winter time, nevertheless, this does not seem to cause large problems in this inversion. The effect appears over a short window of time in the winter and this inversion has the winter time divided into two pieces, Jan/Feb of 2004 and Nov/Dec of 2004, potentially lessening the effect. This might be a more significant effect on larger time frame inversions. Additionally, there may be subtle constraints imposed by

the corrections to respiration and transferred to GPP via larger scale correlation constraints, although this is not known.

Contrary to the paragraph above, in other portions of the paper the authors seem to use the estimates of bias parameters from a previous week as a “prior” for bias parameters for the subsequent week. This leads to my second major methodological concern. The authors correctly state that if a posteriori estimates from one week are used as a priori estimates for the following week, then two problems arise. First, by definition, through the sequential Bayesian updating procedure, the uncertainties appear to decrease as time evolves. This would be reasonable if in fact the bias parameters were considered constant throughout time, but leads to unreasonably low uncertainty estimates otherwise. Second, the estimates of the bias parameters converge to constant values, because the incremental impact of new observations on the best estimates is low. The authors claim to solve both problem in equation 7, by adding a “grand” prior term to pull the bias estimates back their original priors (with high uncertainties). However, in an objective function such as the one in equation 7, the principle of Bayesian updating is still such that even if one adds a second set of “priors” with high uncertainty σ^2_{grand} , the a posteriori uncertainties will still by definition be lower than that in each of the individual terms in the objective function, and therefore be lower than σ^2_0 , which is the uncertainty that the authors had described as getting unreasonably low. In fact, by adding a third term (i.e. a third piece of information) to the objective function, the new a posteriori uncertainties will now become even lower.

I'm not sure of the logic here. The step-wise posterior variance of the beta parameters is given in Eq 9. The first term never changes and the third term, the likelihood portion, is a function of the current data. The only term that can perpetually decrease towards zero is the second term. Essentially, this provides the “grand” covariance as a lower bound to the posterior variance although you will likely have some non-zero contribution from the likelihood portion as well at each time step. Some schemes ignore the the σ_0 term (CarbonTracker) and others have traditionally ignored σ_{grand} and “inflated” the posterior variance from step to step (many EnKF schemes). This is essentially an average of the two approaches which I think is acceptable for my current work. The negative aspects of it are that the variance estimates start “high” and eventually “plateau” where both of the other methods attempt to provide larger amounts of variance at each filter step. This is likely a result of the other methods being employed over much larger periods of time. For the current paper, I think the best way to “visualize” the filter is as having a long “burn in” period. There is always “some” freedom for the filter but it is certainly more tightly constrained at the end of the year than at the start. This likely would need some modifications to run over many consecutive years.

The values of the best estimates themselves will be pulled slightly back to the a priori values (only slightly because these “grand” priors are given high uncertainties), but there is no guarantee that these priors are any better than the a

posteriori estimates from the previous step. This approach does not seem to make sense from a statistical perspective, or from a scientific one.

The incorporation of the “grand” prior has two main effects. The first is that the grand prior keeps the estimates of bias in respiration and GPP, which can be highly correlated at times, from diverging from zero in a correlated way. The second is that it allows “some” degree of covariance structure to be propagated through the filter. Filters such as CarbonTracker typically use a “fresh” restart and reset the covariance matrix at each filter step. My approach was a compromise, essentially an average of (1) a complete reset of the covariance matrix to the prior and (2) a full blown full propagation of the covariance matrix at each filter step.

Perhaps even more critically, it appears (lines 10-13 p. 10206) that only 7 days of data are used to constrain each set of bias parameters. This is an unreasonably short window, and much shorter than any time lag that has been shown to be appropriate in the application of Kalman filtering approaches in the past (e.g. Peters et al. 2005 JGR, Bruhwiler et al. 2005 ACP, Peters et al. 2007 PNAS etc.). Also, in a classical Kalman filter, the state space remains unchanged, and one only steps through the observations. In this way, the same state space is sequentially updated, while stepping through the observations. In fixed-lag filters such as the ones references above, each element of the state space is instead updated a fixed number of times, until the newest observations are no longer informative of past variability. If the authors are instead solving for a new weekly set of biases with each set of observations, then they are essentially solving a series of “batch” inversions, with the very unreasonable assumption that each week of observations only contains information about that same week’s bias parameters. This is not a Kalman filter, and also is not justified from a scientific perspective.

Like I have indicated in the paper at the start of section 2.4, this is really a “modified” Kalman filter. I have presented the algorithm from an algorithmic standpoint but using statistical notation. Most commonly, these equations are presented from a control theory or numerical analysis framework. In terms of a state-space formulation (and using the variable notation in my article), a Kalman Filter can be represented as a set of two equations:

$$(1) \beta(t+1) = \Phi * \beta(t) + V(t+1)$$

$$(2) y(t+1) = G * \beta(t+1) + W(t+1)$$

In our case, Φ is the identity matrix and $V(t+1)$ is identically zero meaning we do not explicitly have a dynamical model of the change in the betas. Furthermore, our equation (1) changes slightly to include a prior distribution that works as a regularization agent and in the process ensures a lower bound on the posterior variance estimates for the betas.

$$(1) \beta(t+1) = (\beta(t) + w \cdot \beta(0)) / (1+w) + V(t+1)$$

$$(2) y(t+1) = G \cdot \beta(t+1) + W(t+1)$$

$$V(t+1) = 0, W(t+1) \sim N(0, \text{SIGMA}(\text{obs})), \beta(0) \sim N(0, \text{SIGMA}(\text{grand}))$$

As you have mentioned, the betas are “fixed but unknown” state space parameters, w is a scalar controlling the strength/value of the prior information, y are the observations, etc.

Furthermore, there is no real lag process employed here. There are two important points to be made here about that.

(1). The global inversion problem is different from the regional inversion problem. The time lag window generally must be larger in a global inversion problem. For example, carbon fluxes originating from the early summer in Siberia may affect Europe several weeks to months later. Furthermore, the domain is cyclical in the global problem, fluxes from Siberia may affect Siberia several months later. In the regional inversion model, this is avoided, or at least transferred to the constraining global inversion model which provides boundaries. At 10 meters per second, the wind will likely cross the land area of our domain in less than 7-10 days. We employ back trajectories for 10 days or until the particle crosses out of the boundary. So, the window of time to consider flux sensitivity of an observing site to a flux source is very similar to the length of time for which a bias is estimated, hence the simple lag 1 approach.

(2) Recall that we are estimating GPP and Respiration instead of NEE. First, I obviously must caveat this discussion with the fact that the data certainly should not constrain the component pieces, GPP and Resp, as well as NEE. However, having said that, I believe the conceptual model of a smoothly varying bias in respiration and GPP is better than the conceptual model of smoothly varying bias in NEE. With an exception to GPP in the winter time which can be difficult to estimate, the ratios of modeled GPP to observed GPP (and Resp) appear better constrained than the ratios of modeled NEE to observed NEE. I currently have a paper in prep concerning this very topic and from preliminary analysis, it seems to be supported quite well from the data and I believe it would be supported, at least qualitatively, by the ecological “forward-modeling” biosphere community as well.

If I have misinterpreted the numerical implementation of the approach, then this needs to be clarified in the manuscript. Furthermore, I find it somewhat frustrating that the parameters that are actually being estimated through the inversion are the bias parameters, and yet the authors never present these results. Instead, they jump directly to time series and maps of GPP, ER, and NEE. Although I understand that these are the quantities that we are ultimately interested in, one is left with the

suspicion that perhaps the estimated biases were not reasonable, and that this is why they were not presented. These are, however, critical in order to evaluate the effectiveness of the approach. Do these bias parameters vary in a reasonable way both spatially and temporally? How does the apparent uncertainty of these parameters evolve over the course of the year? Are their absolute values reasonable within the context of what the authors understand about the underlying biospheric model? etc.

This is a very good point. The explicit bias maps are very useful for diagnosing the inversion, although, I would not agree that they are critical to always show. In fact, these estimates are rarely showed in papers using “real” data but more often in “pseudo-data” papers like Schuh et al. 2009 (JGR). The important point to make is that the inversion allows the flexibility to incorporate GPP and Resp while at the same time, not requiring it. In other words, the worse case scenario is that you are allowing a meaningless parameterization of NEE in the inversion while the best case scenario is that you capture individual estimates of the component fluxes of NEE. Practical “reality” is somewhere in the middle. The data likely constrains Resp and GPP on certain scales in space and time while not on others. However, there should be no damage done to NEE estimates at places where the individual pieces can not be identified. Large time-scale patterns of reductions of Resp and GPP are not necessary in order to produce the small net carbon sink we know exists over the domain in 2004. However, this seems to match well to emerging data which shows that SiB3 seems to have one of the stronger GPP estimates among the enzyme kinetic models, which themselves have significantly higher GPP estimates than the Light Use Efficiency models like VPRM. Furthermore, concurrent work from the North American Carbon Program synthesis project clearly shows biases in driving meteorology which could clearly drive the SiB3 prior GPP/Resp estimates to be too high. Than, finally, the independent flux tower results shown for the ARM site support the fact that the model is able to separate these components out for at least one site which appears to be well constrained by CO2 data.

Related to this point, all the figures that involve maps have aspect ratios, color schemes, and other plotting choices that hinder the interpretation of the presented results. For example, whereas figures 1 and 2 are presented at the native resolution of the datasets used in the analysis with a high contrast color scheme, and figure 5 (which presents uncertainties) is still presented at the native resolution of the inversion, all figures showing inversion results (Figure 7, 8, 10) use a washed out color scheme with smooth contouring that make it difficult to interpret the spatial variability and overall magnitude of the estimated fluxes.

If the reviewer wishes, I can change the color scheme, however I only use a high contrast scheme when dealing w/ categorical data or a large range of arbitrary magnitude data. From an annual source/sink standpoint, it appears to me much clearer to use a simple red/blue scheme which emphasizes the sources/sinks and deemphasizes areas with estimated net balance of carbon

(NEE=0). This is emphasized even more in Figure 11 where it is hoped that the simple color scheme will allow the reader to quickly process the differences in flux estimates between prior/posterior and model. I would ask the reviewer to reconsider this request, but if need be, it is not difficult and I will convert the color scheme on Fig. 7 and Fig. 11.

Many figures also have very small fonts, inconsistent color schemes, etc., which further complicates the interpretation of results.

******I will work to correct this w/ the publication staff. This would seem to depend upon the size of the figures in the journal and be somewhat image specific.******

OTHER SPECIFIC COMMENTS

The use of a single uncertainty for all towers does not seem justified (p. 10204-5). At the very least, I would expect that the degree of mismatch will differ between tall and short towers.

Agreed, I did use a very simple “pooled” error estimate for this. Some experiments were performed w/ varying lengths of time employed, although it was not clear what was “ideal”. I have added additional diagnostics, winter-summer histograms and biases as an add'l figure which will hopefully make this transparent the reader.

I have also added the following sentence to the “observations” section: “This is a simple assumption and we certainly don't expect to the error to be completely homogeneous across towers although at what temporal scale the observation error should be estimated is still somewhat uncertain.”

Related to this issue, do the a posteriori fluxes from the inversion reproduce available observations to within the assumed mismatch error (figure 3)?

To some degree, yes. Again, a pooled estimate is just that, pooled, and is not as descriptive as a tower specific or better yet a season BY tower type approach. So, yes, there will be mismatches. I have included a comment on chi-square stats which indicate reasonable agreement in the summer and residuals in the winter that are LESS than those assumed.

Overall, how are the results of the inversion evaluated? Do the a posteriori bias estimates remain close to their prior values? How are other model assumptions evaluated? Are results equally reliable throughout the continent?

Most of the corrections are, of course, made over areas that are more heavily sampled by the 8 towers' footprints. Outside of this, most of the posterior bias estimates remain somewhat close to their prior values. For example, the SE U.S.

has little “correction” and higher variability (Figure 5) due to low sampling coverage as well as very strong a priori fluxes.

Have any OSSE studies been conducted to evaluate in an idealized setting whether the presented approach can accurately constrain biases in a setup analogous to the one presented here? In other words, how do the authors evaluate whether they can trust their estimates?

I believe a “pseudo-data” type paper is what is meant by OSSE study? In this case, yes.

Schuh, A. E., A. S. Denning, M. Uliasz, and K. D. Corbin (2009), Seeing the forest through the trees: Recovering large-scale carbon flux biases in the midst of small-scale variability, J. Geophys. Res., 114, G03007, doi:10.1029/2008JG000842.

The paper lacks references to several papers that have taken related approaches, either through the implementation of a Kalman filter approach to the solution of the inverse problem, the use of statistical methods for estimating error covariances, the use of correlation length scales to describe the covariance of fluxes or errors, specific choices for observations error variances, etc. The authors need to put their approach and their modeling choices more clearly within the context of other recent inversion papers, both to give credit to this earlier work, and to better justify the changes from previous approaches that they have chosen to implement.

There are at least 12 references to atmospheric inversion papers in the text. Additional ones that should likely be included because of the “regional” – centric view of this paper would likely be Carouge et al. 2008a,b and Lauvaux et al 2008. Although the argument for the assumed isotropic error structure we impose (regularization) is different than their general argument (accurately capturing the spatial error structure via the observations), I’ve added reference to Mueller et al. 2008 and Gourdji et al. 2008, both of which use prior fluxes to characterize the spatial error structure. I will add refs/descriptions from these to the text. S. Gourdji has a very nice pseudo-data paper out in discussion now (ACPD) which is very applicable but it is probably too early/late to reference/integrate it in this paper.