Atmos. Chem. Phys. Discuss., 11, C1946–C1950, 2011
www.atmos-chem-phys-discuss.net/11/C1946/2011/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "A comparison of different inverse carbon flux estimation approaches for application on a regional domain" *by* L. F. Tolk et al.

Anonymous Referee #2

Received and published: 18 April 2011

The paper "a comparison of different inverse carbon flux estimation approaches for application on a regional domain" compares performances from different inversions of regional CO2 fluxes over The Netherlands for a 15-day period in May-June 2008. The study explores the impact of different assumptions affecting the inverse flux estimates in a pseudo-data experimental framework.

The overall paper is well presented and the scientific questions are relevant for regional scale inverse studies. But several points are missing and conclusions go far beyond the actual results, requiring revisions before the manuscript is published.

- The approaches you use include very different assumptions which are key elements

of the system (error correlations, number of unknowns, ...). You explore here all these aspects at the same time, and without really identifying the true causes. For example, when you assume one correction factor per vegetation type (case 1), you assume that one factor can be optimized to improve the fluxes in each region. It includes large error correlations within the regions and few unknowns over the domain, leading to a reduced flexibility and a prescribed diurnal cycle. Then, you split this factor for respiration and GPP (case 2), increasing the degree of freedom, and changing the diurnal cycle definition. You have to clarify and organize your results. What comes from the diurnal cycle? What is the impact of the increase of unknowns? Could you define shorter periods of time in NEE to fit better the diurnal cycle? Is the choice of 15 days relevant? What is the aggregation error in time due to this assumption?

- In your conclusions, you recommend the use of case 3 and case 4 compared to the other solutions. But results in figure 4 show large differences between the different methods and no real improvement in either case. One can notice large biases for crops using two factors in case 3 (pixel based inversion with factors for respiration and GPP), and very little changes everywhere in case 4 and case 1. One can wonder if your results are limited by the vegetation type homogeneity (ecoregions) assumption, or by the lack of data in critical areas (over the crops), or by a lack of flexibility in case 1 and more corrections but no better results in the other cases. Does it mean that splitting in two factors is better? It is a change of assumptions but it is hard to conclude clearly looking at your results. Then you perform one case on a pixel based system (case 3), increasing your degree of freedom and changing the underlying assumptions (less flux error correlations or more distributed in space). One more time, this is a change in the assumptions more than a change of "approach". You have to clarify all these elements. Figure 2 shows that lots of noise is introduced in the results without any structure with the pixel based system. Is it due to your flux error correlations? Do you have enough members to capture relevant signals in the EnKF? You should perform a third inversion as case 1 but at the pixel scale to really explore what is the impact of the correlations, and what is the impact of the ecoregion assumption. Case 2 is not really better in

general. You could also test different correlation lengths or structures to identify their importance.

- Similarly, Figure 2 leads to several critical issues and questions. Basically, your case 1 shows no significant change, suggesting too few unknowns and no flexibility. Then case 2 generates very positive fluxes over grass, and very negative fluxes over crops, both far from the truth (indicated in fig 4 as well). Case 3 presents very noisy results, with almost no structure in space. It seems clear that flux error correlations are critical for the different cases. Are you sure this has anything to do with the approach? I would suggest that this is a direct consequence of the flux error correlation assumption, the temporal behaviour you assume (diurnal cycle), and the number of unknwons. Finally, the case 4 shows little changes compared to the prior and no real structure in the posterior compared to the truth. Figure 2 also shows clearly that this method is not far from the prior. Defining which "approach" is best is overstated, mainly based on a statistical criteria, the RMS from table 3, even if all cases give biased estimates for the main regions (crops being a large component in terms of annual NEE). Finally, in figure 5, it seems clear that for the RG 1.0 case, the nighttime fluxes are unrealistic, with a negative value almost constant. Why the other cases give more reasonable estimates? Is it only due to the perfect correlation between resp and GPP?

- You have to identify more clearly which assumptions are critical (flux error correlations, number of unknowns, flexibility of the system,...). What signals can be retrieved with your current observation network? Are the fluxes strong enough that your inversion can detect them? May-June is the very beginning of the growing season, with 5-10 umol.m-2.s-1 for crops. What is the impact of a 15-day time period? Is it too long and might lead to biased estimates? Many aspects of the system should be explored before you conclude about methods, results, and potentials of such a system.

And finally, your discussion section is out of the scope of the paper. How can you argue about longer term issues, correct carbon flux balance at regional scale or not, and aggregation errors, with your pseudo data test over 15 days? You also discuss

non linearity without showing consistent results with this statement.

This study is based on different potential approaches that one may use to perform a regional scale inversion. You have to seperate clearly what is the impact of each component on the final results. A more rigorous analysis is required to identify the importance of the diurnal cycle, the spatial distribution, the number of unknowns without splitting in different factors simultaneously, the importance of the flux strength (very weak during your period), the correlation structures in the flux errors (length scale value), and the prescription of the diurnal cycle.

Technical comments:

-introduction: your bibliography doesn't include any of the regional scale inverse studies already published. Some examples are Gerbig et al., 2003, Lauvaux et al., 2009, Schuh et al., 2009, Goekede et al, 2010, Gourdji et al., 2010)

3358-1: "recently" is inappropriate, or explain why the 3 studies you cite are different compared to past inversions.

13: missing "has"

3360-9: refer to appendix A

3361-12: your length scale seems much smaller compared to the different studies you cite. Can you explain why? Why don't you diagnose the flux error correlations from your prior-truth residuals?

3363-1: what crops are included in "crops1"?

3365-1-2: the prior error covariance can't be equal for all cases, only the overall error.

3.6: What is the purpose of this paragraph? You Want to show that early June is not very different in terms of fluxes compared to late May? This metric has no clear value as written here. You could run a second period (early June) of inversion and propagate the information from the first one to see if it helps. Otherwise, you just detect if the two

periods are similar or not.

-Discussion:

3373-21: "non-synthetic" = real

3374-12-14: only true if you include error correlation between periods, or propagate the optimized factors. But is it a good solution? Does it really impact your system? You can't answer this point with the actual analysis. This is an assumption.

3374-15-22: This paragraph is out of the scope of your paper. You suggest that a highly sophisticated system able to optimize model parameter by ingesting several types of data and information in addition of co2 concentrations will perform better. This is a very general comment, not related to this study.

3374-23-26: other papers have compared their results to other methods and found reasonable results using real data compared to bottom-up approaches (e.g Goekede et al., 2010). Could you discuss or explain the differences in your case? Several pseudo data experiment have found that one can retrieve reasonable patterns and signals (Gourdji et al., 2010). They tested different time windows for the fluxes, and found large differences between one value, 4 time windows per day, and every 3 hours. You conclude with only four cases that the method is not suitable as it is now. Why your conclusions are different from other studies?

3375-1-3: The two sentences are beyond your analysis. Your study is limited to 4 cases and 15 days of pseudo data. Do you think this is enough to conclude about the needs for future inversions?

-Conclusions: See above the general comments.

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 3355, 2011.