Atmos. Chem. Phys. Discuss., 14, C12015–C12027, 2015
www.atmos-chem-phys-discuss.net/14/C12015/2015/

Atmospheric Chemistry and Physics
Discussions

Interactive comment on "The potential for regional-scale bias in top-down CO_2 flux estimates due to atmospheric transport errors" *by* S. M. Miller et al.

S. M. Miller et al.

scot.m.miller@gmail.com

Received and published: 7 February 2015

We would like to thank the reviewer for suggestions and comments on the manuscript. The reviewer's detailed suggestions have been very helpful in improving the manuscript. Below, we have included the reviewers comments (in bold) along with our reply and the associated changes/updates to the manuscript.

- **My main criticism if focused on case study #1 and the framing of this experiment as a way to quantify the bias detection limit of carbotracker (CT). But CT does not work on a month by month basis lie your lambda scal-**

ing factoor, it does not consider signals site-by-site like here but instead a whole network, and it does not scale flux signals locally at each site like in your FSSR but over a large spatial area that is also seen by other sites. [...] Framing this experiment as a way to determine the balance between large-scale flux influences and transport errors is in that sense more appropriate, and I think describes better what was actually done.

  The reviewer brings up a great suggestion here, and we have re-framed case study #1 accordingly. As the reviewer points out, the goal of this investigation is not to re-estimate the uncertainty bounds on CarbonTracker. Rather, our goal is to understand the magnitude of these transport uncertainties relative to the fluxes. To that end, this case study provides ones means to relate these two entities (the transport uncertainties and fluxes) in the absence of an explicit model adjoint. We no longer frame the case study as a means to quantify the bias detection limit of CarbonTracker. Rather, as the reviewer suggests (below), we have re-framed the case study to examine the following question: how does the magnitude of the transport uncertainties compare against the afternoon, atmospheric CO_2 signal from regional surface fluxes? We have modified the manuscript text and figures accordingly.

- **Specifically, your comparison of transport noise (SSR) and flux biases (FSSR) is done in squared residual space which only measures the magnitude of a signal, but does not account for its sign. A bias in fluxes would typically manifest itself as a consistent over- or underestimate of the true concentrations observed and even if these are small (say 0.5 ppm) compared to the more random transport uncertainties (say 3 ppm), their consistency in sign over longer periods of time would make them detectable. In fact, in a Bayesian inversion the system would try to overcome this small bias as by design it strives for zero mean residuals even in the presence of large observation error covariances.**

The reviewer raises a very good point here, and we have clarified this point in the revised manuscript. As the reviewer explains, a Bayesian inversion will optimize the fluxes to minimize or remove any biases between the model and the observations. If a transport model is completely unbiased relative to the actual atmosphere, then the CO_2 budget estimated by an inversion should also be unbiased. (This statement assumes that other components of the inversion, including the observations, are unbiased.)

In contrast, the inversion may estimate an erroneous or biased budget if the atmospheric transport model is biased. For example, imagine a hypothetical transport model that consistently overestimates vertical mixing. One could construct a Bayesian inversion to optimize CO_2 fluxes using that transport estimate. The inversion will optimize the fluxes to minimize any model-measurement bias. However, the resulting flux estimate is unlikely to be correct; the inversion would erroneously increase the magnitude of the fluxes to compensate for errors in vertical mixing. The model would appear to match the CO_2 measurements, but the estimated fluxes would nonetheless be biased relative to the true fluxes. In this case, the bias in the fluxes would be undetectable with respect to the atmospheric observations. Stephens et al. (2007) adeptly discuss this topic in the context of atmospheric inverse modeling.

We have re-designed case study one in the manuscript to make this comparison more direct. Among other changes, the revised case study no longer uses squared residuals. We hope the revision makes this point about biases more transparent.

- **To overcome this criticism, I would suggest one of two approaches:**

  (1) is to try and change the metric so that it includes more sites at once and includes also spatial covariances between residuals. The new metric then also needs to account in some way for the sign of the residuals.

  (2) Is to write the question of this case study differently and to say that you'll try to estimate to what variation in flux magnitude the meteorological uncertainty corresponds for each site given a realistic surface flux from CT. This also means that most of the use of the word "bias" gets replaced by "flux signal".

  The reviewer makes two good suggestions for revising the manuscript. We have re-framed the case study according to the reviewer's second suggestion. In addition, we have also included multi-site comparisons in the revised manuscript, as per the reviewer's first suggestion. To this end, we have revised sections 2.4, 3.3, and Fig. 4.

- **I find the discussion section a bit too short, and would like to see some more connections to other studies in this field. For example, some reflection could be added on the LETKF methods used by these authors in the past, and about the possible gain of co-simulating CO2 and transport errors. Also, there is room for some reflection on the covariations of CO2 surface fluxes, and those that shape the weather conditions (water and energy and momentum fluxes). What would the next step with this type of system look like when surface fluxes also become a function of the weather variables?**

  We have lengthened the discussion section to include these points, as suggested by the reviewer. For example, in section 3.4 of the revised manuscript, we discuss the possible gain of co-simulating CO_2 surface fluxes and transport errors. That approach could provide a more complete picture of how meteorological uncertainties affect CO_2 fluxes from the origin of the fluxes to the locations where we actually measure atmospheric CO_2. For example, Lin et al. (2011) explored how uncertainties in flux model drivers affected fluxes estimated for Canadian boreal forests. They found that uncertainties in downward shortwave radiation contributed to the largest uncertainties in the simulated fluxes. Similarly, Law et al. (2002) and Gourdji et al. (2012), among many others, have shown that both air temperature and specific humidity are drivers of CO_2 fluxes. These meteorological variables (e.g., downward shortwave radiation, temperature, and specific humidity) correlate with the persistent atmospheric transport uncertainties discussed in section 3.4. A future study could connect these uncertainties (in transport and flux estimation) to gain an even broader picture of how meteorological uncertainties affect CO_2 flux modeling and ultimately top-down CO_2 flux estimates.

- **Furthermore, these findings can nicely be connected to the error budgets presented in Pino et al., (2011) and in Williams et al (2011). Both take a look at the driving forces behind variations in CO2 in the PBL, one from a local and one from a larger perspective.**

  We have added references to both papers in the revised manuscript. Pino et al. (2012) argue that estimated morning PBL heights play a critical role to modeled CO_2 concentrations during midday. They examined transport errors at diurnal scales but point out that the role of different boundary layer processes could change when examined over longer time scales. Our analysis examines transport errors at both the diurnal and monthly time scales and can extend the arguments presented by Pino et al. (2012) to these longer time scales.

  Williams et al. (2011) argue that previous meteorological model discrepancies are usually due to overestimated vertical mixing. According to the authors, "However, the simple inverse proportionality between errors in vertical gradients and mixing only works when there are no systematic errors in the surface flux, horizontal advective transport, or non-linear vertical advective transport (i.e., synoptic-scale eddies)." In our analysis, we place these individual error sources, like those invested by Williams et al. (2011), in the context of other transport processes or uncertainties at sub-daily to monthly time scales.

- **Note that I remain a bit puzzled on the implementation of the SSR vs FSSR metric in equations 4 and 5 + the explanation in the supplement and would like to see some clarification.**

  We have simplified the approach in this section of the manuscript to make the methods easier to follow and more transparent.

- **p.23684: I could not find where the range is actually applied instead of the SDV**

  We have removed the phrase "or alternately the range" from the manuscript. In the revised manuscript, we primarily refer to the 95% confidence interval throughout the manuscript.

- **p.23684: What is the temporal resolution of these fluxes?**

  We reformatted CarbonTracker fluxes to a 6-hourly resolution. This resolution is identical to the CAM model time step. We use this 6-hourly resolution for all model simulations presented in the manuscript. Figure 2, by contrast displays monthly-averaged CT fluxes. The primary objective of this figure is to illustrate the spatial and seasonal distribution of the fluxes. We do not use these monthly-averages in the actual model runs or analysis. We have clarified this point throughout the manuscript.

- **p.23684: So this means that the feedbacks of meteorological errors on carbon exchange are not accounted for? In other words, different weather does lead to different water exchange, but not other carbon fluxes. Okay, I got it.**

  The reviewer is correct here. We have added a sentence to section 2.3 clarifying this point.

- **p.23685: Larger than most means more than 32 if k=64 members?**

We agree with the reviewer that this text is ambiguous as written. We have re-framed this section of the methods accordingly.

- **So this suggests that for p to get to 0.05, there must be 64*0.05 = 3.2 elements in A (eq 5). And when there are four or more SSRs in the set that are larger than FSSR then you have proven the null-hypothesis that bias in fluxes is indistinguishable above transport uncertainties. This seems quite strict to me.**

  Oh wait, I think there might simply be a typo here and you actually meant 0.5 instead of 0.05? Sorry, I spotted this kind of late because 0.05 is such a typical p-value in statistics...

  We have simplified the approach in this section of the methods and no longer use a hypothesis test or associated p-values. In the revised manuscript, we estimate confidence intervals in modeled atmospheric CO_2 and compare those uncertainties against the surface flux signal. We no longer test an explicit hypothesis.

- **Can you elaborate in the main text how this temporal covariance is accounted for. I am sure the Supplement gives info but I'd rather like to understand it here.**

  The reviewer makes a great suggestion here, and we have elaborated on this point in section 2.2 of the manuscript.

  Both spatial and temporal covariance are built into the transport errors estimated by CAM-LETKF. The CAM-LETKF system includes 64 different ensemble members. At the first time step, we launch 64 weather forecasts simultaneously, one for each ensemble member. At the end of the first 6-hour time step, we optimize these ensemble members collectively to match meteorological observations, and the spread of these ensemble members represents our posterior uncertainty in the meteorology. We then use these optimized ensemble members as initial conditions for the next time step and re-launch 64 simultaneous weather forecasts.

  Transport uncertainties within one ensemble member can easily persist over many time steps. For example, if the PBL height in one ensemble member is lower than the ensemble average at one time step, it will probably be lower than average at the next time step. In this way, transport uncertainties or errors can persist over many time steps.

- **p.23687: This suggests you indeed used fluxes including a diurnal cycle.**

  That statement is correct. We have updated the methods section to make this point clearer to the reader.

- **p.23688: I think this is an absolutely wonderful conclusion to draw, and hope it will get a prominent place in the abstract and conclusions**

  Thank you for the encouraging suggestion! We have modified case study #1 to focus more specifically on these conclusions. Furthermore, we have made these points more prominent in both the abstract and conclusion.

- **p.23688: I do not think this case study uses an appropriate question, as your test is not a correct metric to determine the minimum size of flux biases that are detectable through atmospheric CO2.**

  We agree with the reviewer here. We have re-framed case study #1 based upon the reviewer's suggestions above.

- **p.23688: This effect of measurement bias was explored by Masarie et al., (2011), please reference.**

  This is a good suggestion, and we have included this reference in the revised manuscript accordingly.

- **p.23689: What does the number 0.3 represent?**

  A correlation coefficient of $R^2 = 0.3$ does not represent any specific threshold. Rather, we simply wanted to show the meteorological variables that correlate best with the transport uncertainties (instead of including 60 different scatterplots). We have modified this section of the revised manuscript. Instead, we now show the two variables that correlate most closely over land regions and over the ocean (four total variables).

- **p.23689: Since this point is now mentioned a second time, a reference to Pino et al., (2012) is in place as he already showed such PBL-CO2 error relations.** We have included this reference in the revised manuscript.
- **p.23689: Again, your analysis is very nice but this conclusions is not correct. Since one of the authors is associated with the CT group at NOAA, perhaps a synthetic inversion could be done to prove this statement beyond my doubt?**

  We agree and have re-framed case study #1 accordingly.
- **p.23689: This second part is very nice. Can you speculate how this conclusion might change if the interactions between the meteorological variables and the CO2 fluxes themselves were included in a follow-up study?**

  The reviewer poses an interesting question: what would be the effect of including these meteorological uncertainties in the bottom-up or biogeochemical model that generates the CO_2 fluxes? The uncertainties in estimate CO_2 fluxes would likely increase. We have added a discussion on this point to section 3.4. Refer to the discussion earlier in this reply for more detail on this point.
- **p.23690: I find the discussion section a bit too short, and would like to see some more connections to other studies in this field.**

  We have expanded the discussion accordingly (see the discussion earlier in this reply for more detail).
- **p.23693: You could compare these to the posterior flux uncertainty in CT and show that they are at least as large indeed.**

As per the reviewer's suggestion, we have re-framed case study #1 to de-emphasize any direct comparison against the posterior uncertainties in CarbonTracker. As such, we would hesitate to make that comparison explicit here.

Furthermore, it might be difficult to make a direct comparison in this instance. CAM–LETKF estimates the variances and covariances due to transport errors. This information is often incorporated into one of the covariance matrices in a Bayesian inversion. This matrix is often termed the 'model-data mismatch matrix' or 'observational error covariance matrix'. This covariance matrix is then combined with the prior covariance matrix to compute the posterior uncertainty. Hence, this suggestion would require comparing somewhat different quantities. In other words, if we compared transport uncertainties against the posterior flux uncertainty we would be comparing two very different covariances matrices to one another.

- **p.23694: What do the letters below the x-axis indicate?**

  We have removed these letters from any analogous plots in the revised manuscript. The letters below the x-axis that figure indicated whether the CO_2 measurement sites were marine ("M"), short towers ("S"), or tall towers ("T").
- **p.23694: Why do we only see the land CV? Was the constant flux also only applied over land? This was not clear to me from the description yet.**

  In the revised manuscript, we discuss these results over both land and ocean regions (in section 3.4 and Figs. 6-7).
- **p.23695: The variables 1,2 and 4 look very similar as one would expect from meteorological principles. In the same way, 5 and 6 are closely related. What is perhaps more interesting is that (1) the PBL height which in the end is most directly related to the CO2 mixing ratios is not shaped the same as these primary drivers. This stresses the need for a meteorological model**

to calculate the (co)variances of transport errors rather than to just use some simple proxy. And (b) is that the CV of temperature and CO2 are very similar which is because they are shaped by the same large scale synoptic systems. This is also discussed in the Williams et al., (2011) paper, and the driving power behind the LETK methods shown by Kang, Kalnay, Liu, and Fung (co-authors here). Perhaps this is worth to mention in the discussion.

The reviewer makes a great point here. We have also added an analysis over ocean regions, and the errors here correlate most closely with zonal winds. This added analysis further supports the reviewer's comment above on the role of synoptic scale systems. Also, these variables cannot explain all of the uncertainties, and this result stresses the need for a meteorological model to calculate transport errors over the use of a single proxy for transport errors (like PBLH). We have added a discussion of these points to section 3.4 of the revised manuscript.

- **p.19: This 5% I guess corresponds the p=0.05 probability stated in the main text. That suggests this was not just a typo, and I remain confused on equations 5 and the use of this test.**

  We have removed the hypothesis test from case study #1 to make the analysis simpler and more straightforward. Concomitantly, we have removed most equations to streamline and simplify the revised text.

- **p.19: This is a nice illustration of the properties of the SSR, which I think correctly assumes transport errors to be normally distributed around a zero mean. But the problem I have is in the comparison to FSSR, which for a biased flux would not just be a residual around some mean, but an actual signal with a sign and a spatial pattern. See for instance the figures S9, S11, and S14 that both represent winter conditions. A shift of the fluxes by 10% upwards would lift both lines for the ensemble mean upwards by 0.5-2.0 ppm and reveal a systematic offset (if the model mean was a bit more**

  unbiased which it is not without data assimilation of the fluxes) at three locations.

  If the atmospheric transport errors were completely uncorrelated from one model time step to another, then it might be relatively easy to distinguish a bias in modeled concentrations caused by an erroneous flux estimate. However, the atmospheric transport errors estimated in this study are often correlated in both space and time. In other words, these errors are modeled as a multivariate normal distribution, and the covariances in this distribution can be large. As a result, transport errors could bias the model relative to the measurements over many time steps. In that case, it could be very difficult to distinguish the difference between sustained model-data differences due to the fluxes or due to transport errors. We have revised and re-framed case study #1 to better explain and more prominently feature the role of spatially and temporally correlated transport errors.

References

Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., Petron, G., Nehrkorn, T., Eluszkiewicz, J., Henderson, J., Wen, D., Lin, J., Fischer, M., Sweeney, C., and Michalak, A. M.: North American CO_2 exchange: inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, Biogeosciences, 9, 457–475, doi:10.5194/bg-9-457-2012, 2012.

Law, B., Falge, E., Gu, L., Baldocchi, D., Bakwin, P., Berbigier, P., Davis, K., Dolman, A., Falk, M., Fuentes, J., Goldstein, A., Granier, A., Grelle, A., Hollinger, D., Janssens, I., Jarvis, P., Jensen, N., Katul, G., Mahli, Y., Matteucci, G., Meyers, T., Monson, R., Munger, W., Oechel, W., Olson, R., Pilegaard, K., U, K. P., Thorgeirsson, H., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: Environmental controls over carbon dioxide and water vapor exchange of terrestrial vegetation, Agricultural and Forest Meteorology, 113, 97 – 120, doi:10.1016/S0168-1923(02)00104-1, 2002.

Lin, J. C., Pejam, M. R., Chan, E., Wofsy, S. C., Gottlieb, E. W., Margolis, H. A., and McCaughey,

J. H.: Attributing uncertainties in simulated biospheric carbon fluxes to different error sources, Global Biogeochemical Cycles, 25, doi:10.1029/2010GB003884, 2011.

Pino, D., Vilà-Guerau de Arellano, J., Peters, W., Schröter, J., van Heerwaarden, C. C., and Krol, M. C.: A conceptual framework to quantify the influence of convective boundary layer development on carbon dioxide mixing ratios, Atmospheric Chemistry and Physics, 12, 2969–2985, doi:10.5194/acp-12-2969-2012, 2012.

Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langenfelds, R. L., Steele, L. P., Francey, R. J., and Denning, A. S.: Weak Northern and Strong Tropical Land Carbon Uptake from Vertical Profiles of Atmospheric CO_2, Science, 316, 1732–1735, doi:10.1126/science.1137004, 2007.

Williams, I. N., Riley, W. J., Torn, M. S., Berry, J. A., and Biraud, S. C.: Using boundary layer equilibrium to reduce uncertainties in transport models and CO_2 flux inversions, Atmospheric Chemistry and Physics, 11, 9631–9641, doi:10.5194/acp-11-9631-2011, 2011.

Interactive comment on Atmos. Chem. Phys. Discuss., 14, 23681, 2014.