Biogeosciences Discuss., 12, C6944–C6955, 2015
www.biogeosciences-discuss.net/12/C6944/2015/

Biogeosciences
Discussions
Open Access

Interactive comment on "The ability of atmospheric data to resolve discrepancies in wetland methane estimates over North America" *by* S. M. Miller et al.

S. M. Miller et al.

scot.m.miller@gmail.com

Received and published: 23 October 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.

1 Overall comments:

- **Concerning the language, the authors frequently make use of question sentences, often of rhetorical nature, which I personally find quite annoying within the context of a journal article. So I highly recommend rewriting these passages. Particularly these question sentences are often unnecessarily repeated throughout the paper.**

  We have revised this language accordingly.

- **The structure is clear and concise, though at some sections slightly unbalanced – in some cases (e.g. Section 2.1) you skip over many essential details and refer to existing manuscripts, while in others (e.g. 2.2) you provide many details where you could also have used citations instead.**

  We have added more detail to Section 2.1 on WRF-STILT and on the WETCHIMP model comparison project. We have edited Section 2.2 to rely more heavily on the existing literature.

- **The authors claim in the conclusions section (p.9357, ll.3ff) that bottom-up and top-down modelers should do a better job in joining forces to arrive at more solid estimates of methane emission budgets, a statement that I fully support (even though I don't think it belongs into a conclusion section in the way it is presented here ..). At the same time, they work out various systematic differences between bottom-up and top-down products within the context of this study, and attribute all the 'blame' to the bottom-up models, without even starting to discuss shortcomings in the inverse modeling approach which also lead to (well known) large uncertainties. Stated a bit provocative, it sounds like the authors' intention is to tell bottom-up modelers that they need to do a much better job, and better ask the top-down crowd how to do things right**

The reviewer makes a reasonable point. Both bottom-up models and top-down, inverse models have respective strengths and weaknesses/uncertainties; neither provides the final word on greenhouse gas fluxes. We certainly do not want to 'blame' anyone or any previous research effort, and we have revised the manuscript wording where possible to make this point clearer. Both top-down and bottom-up modelers often express a desire to meld efforts in a way that would leverage the respective strengths of each approach. This goal is easier said than done. As the reviewer points out, this statement in the conclusions section is likely too cursory or too provocative, and we have removed it. In the revised manuscript, we not only suggest future improvements to bottom-up models but also to top-down estimates.

- **Moreover, I was disappointed to find that the authors don't really make an effort to explain where such differences might stem from.**

  In the revised manuscript, we have augmented this discussion of 'why' or 'how' in Sections 4.1 – 4.3. Atmospheric data can often provide useful information on the magnitude, location, or timing of fluxes, but it is usually much more difficult to infer how or why these fluxes occurred. In several instances, we can hypothesize why model-data differences occur. For example, existing bottom-up flux estimates exhibit different spatial distributions over North America, and many of those differences appear to stem from the underlying wetland distribution. In the paper, we discuss this difference in context of atmospheric data; bottom-up models that are most consistent with the atmospheric data use wetland distributions that are based, at least in part, on land cover mapping. We also discuss discrepancies in seasonality and why these discrepancies may occur (Section 4.3).

- **However, a comprehensive interpretation of the observed differences as presented herein needs to include an extensive section that discusses the uncertainties and potential biases that stem from the atmospheric inversion part of the comparison.**

We have added a section to the supplement that highlights the largest sources of uncertainties in the top-down analysis conducted in the paper.

The analysis in this paper is based upon inversion frameworks developed in Miller et al. (2013), Miller et al. (2014a), and Miller et al. (2014b). Those papers discuss, in detail, the uncertainties and potential biases that stem from inverse modeling. For example, Miller et al. (2013) explore uncertainties in the estimated methane boundary condition, uncertainties in the estimated covariance matrix parameters (the parameters that define $\mathbf{\Psi}$), uncertainties due to atmospheric transport in WRF-STILT, uncertainties due to geological CH_4 sources, and uncertainties in the attribution of CH_4 to individual sources. Miller et al. (2014a) discuss uncertainties due to the nested meteorology domains in WRF-STILT, uncertainties in the methane boundary condition, uncertainties due to the sparsity of the CH_4 observation network, uncertainties in atmospheric transport estimated by WRF-STILT, and uncertainties in the covariance matrix parameters. Finally, Miller et al. (2014b) discuss uncertainties in the flux estimate due to the assumptions made by the statistical modeling framework. In addition to these inverse modeling papers, Nehrkorn et al. (2010) and Hegarty et al. (2013) also discuss atmospheric transport uncertainties in the WRF-STILT modeling framework.

The uncertainties that were estimated/developed in those existing studies are used throughout the current manuscript. For example, these uncertainties form the basis of the $\mathbf{R}$ and $\mathbf{Q}$ covariance matrices which are an integral part of the model selection analysis (see Eq. 1 and Fig. 3). In addition, the inversion estimates in Fig. 5 of the revised manuscript are shown in context of the confidence interval estimated from the inversion. This confidence interval accounts for limited data coverage, transport model errors, the finite resolution of the inverse model, and other error sources (as discussed in Miller et al. (2014a)).

2 Detailed comments:

- **P.9345f, Section 2.1: As mentioned above, I think this is extremely short. Even though the details might be given somewhere else, the reader needs more information to understand what approaches you used in the context of this study.**

  We have expanded this section to describe the data, atmospheric model, and WETCHIMP methane models in greater detail.

- **p.9346, l.16f: You never explain and/or discuss how the low temporal resolution (monthly) of emission fluxes is actually coupled to your mixing ratio observations, which probably have a temporal resolution of 1-3hrs (details also not given in the text)? Do you assume flat temporal trends in emission rates over the course of one month, then a step change to the rates of the next month? If so, you should add a sensitivity study how this low temporal resolution in the bottom-up products affect your inversion results. Didn't you have access to bottom-up products with a higher temporal resolution?**

  We have clarified this point in the manuscript. We are limited by the temporal resolution of the bottom-up products from the WETCHIMP study; those methane flux simulations have a monthly temporal resolution. With that said, observations during the first ten days of each month have footprints that extend into the previous month. As a result, the model estimate at any given site is often based upon wetland flux estimates from multiple time periods. In this way, the WRF-STILT model estimate is not a step change from one month to the next.

- **p.9346f.: Section 2.1 is almost of equal length compared to 2.1, even though also here you could refer many of the details to e.g. the Gourdji et al. reference. These are just minor details, but they make the paper appear unbalanced in parts.**

  We have expanded Section 2.1 to describe the data, atmospheric model, and WETCHIMP flux products in greater detail. Reviewer #2 asked a number of questions about the model selection framework, so we have also expanded Section 2.2 to better explain the statistical approach for a non-technical audience.

- **p.9347ff.: the strategy of the synthetic modeling setup needs to be rewritten in some parts. Some details are only given in the last paragraph, which are required earlier to understand the concept. For example, you only mention in the past paragraph that the 16 combinations of regions/seasons are optimised separately. Also, one thing that is not clear to me: in 1000 repeats different combinations of turning the 16 options for regions/seasons are randomly created. If each region/season gets an individual model fit through the BIC approach, why do you need the repeats?**

  The reviewer makes an astute point here. We have re-arranged the information in Section 2.3 as the reviewer suggested.

  The 1000 repeats are needed due to the random or stochastic nature of the synthetic data simulations. We add random noise to the synthetic data to simulate the effect of real-world modeling and measurement errors. We do not know the exact value of these modeling or measurement errors. Instead, we have an estimate of the properties of these errors (i.e., their variances and covariances), and we can simulate a plausible set of errors using these estimated properties and a random number generator. The results of the model selection can vary slightly, depending on the particular random numbers that we draw. Hence, we repeat the synthetic data experiments over and over again (1000 times in total) and average the results across all 1000 repeats. This procedure ensures that the model selection results are not the output of a single random number draw. We have clarified this setup in the revised manuscript.

- **p.9349, ll.25ff: Your assigned scaling factors for EDGAR emissions should**

be discussed in more detail as sources of uncertainty in the simulated mixing ratio time series! What about the influence of boundary layer height, which is certainly shallower in winter, and might thus exaggerate the influence of ground sources on mixing ratio changes in the atmosphere.

We do not scale the EDGAR emissions inventory in the revised manuscript. Many atmospheric CH_4 observation sites near wetlands are also located far from large anthropogenic emissions. As a result, any effort to scale the EDGAR inventory at these sites could be error-prone. Instead, we present the inventory as is.

Miller et al. (2013) and Miller et al. (2014a) explore in detail the possible influence of boundary layer height. Miller et al. (2013), for example, found no significant seasonality in their anthropogenic US CH_4 emissions estimate (Fig. S8 in that paper). Seasonal bias in the estimated boundary layer height could manifest as erroneous seasonality in the emissions estimate. The absence of seasonality in estimated US emissions suggests an absence of bias in estimated boundary layer heights.

- **p.9350ff, Section 3, first part: I'm lacking a summarising conclusion/discussion here. To what extent does the ratio of natural to anthropogenic emissions influence the detectability of wetland fluxes? And to what extent is the network configuration responsible?**

  We have expanded the first synthetic data study to explore these questions in greater depth. For example, we explore how these results change if we set anthropogenic emissions to zero. Similarly, we explore how these results change if we set atmospheric transport errors to zero. These expanded results are summarized in Fig. 3 and Section 3.

- **p.9351f, Section 3, second part: since the patterns displayed in Figs. 3a and 3b are virtually the same, the question arises whether you can truly separate the 2 effects you are looking after. After all, it boils down to the**

same question: What is the ratio of natural and anthropogenic emissions in a certain region/season, and how well is the observation network designed to capture these signals. I therefore strongly recommend to explain better if the 2 steps of virtual experiments truly provide different answers!

Part one of the synthetic data experiments asks a question of magnitude and part two asks a question related to the spatial distribution fluxes. In particular, part one investigates whether the observation network can detect any kind of atmospheric pattern from wetlands. Part one asks a basic question about about the detectability of wetlands over patterns from anthropogenic sources or from model errors. Part two investigates whether it matters to the observation network where those fluxes are located. Part two is pre-requisite for the real data experiments in Section 4.1.

In many ways, it makes sense that the first and second synthetic data experiments produce similar results, but that result is not necessarily guaranteed. In regions with large wetland fluxes, those fluxes often display high spatial variance. The first synthetic data experiment often produces positive results when the wetland fluxes are large. The second synthetic data experiment often produces positive results when the fluxes display high spatial variance. Hence, the first and synthetic data experiments often produce similar results. Furthermore, we show in the revised manuscript that these results are not necessarily due to the ratio of natural and anthropogenic emissions (refer to Section 3 of the revised manuscript).

Fang et al. (2014) validated and tested the model selection framework used in the second synthetic experiment. That study used atmospheric observations and model selection to differentiate among spatial patterns in CO_2 flux estimates for North America. We have elaborated on this discussion in the revised manuscript.

- **p.9351ff, Section 4.1: I think it is a very important finding that plausible spatial patterns in CH4 emissions from bottom-up models are only based**

on land cover maps, not on the remotely sensed inundation maps. Here, you provide the only detailed recommendation to the bottom-up community how their model estimates can be improved! So this definitely deserves a more detailed discussion, and a more prominent place in the conclusions.

We have expanded this discussion in Section 4.1 and have featured this result more prominently in the conclusions.

- **p.9354ff, Section 4.2: these results basically indicate that none of the bottom-up models is useful for North American regional simulations ... the summertime emissions seem to be extremely overestimated, so that the resulting seasonal courses in modeled data are opposite of what the observations show. This isn't discussed at all ... ??? I think what definitely needs to be added here is an uncertainty estimate of the background data set, and the scaling factors of the EDGAR emissions. Given the substantial overestimates in summertime emissions by virtually all models, it's hard to imagine how these models could be re-calibrated to match the observations, given that the other modeling components are correct ...**

  We would hesitate to say that none of the bottom-up models is "useful" for North American regional simulations. We would argue that there is an opportunity to tune the seasonal and spatial patterns in these bottom-up estimates. Similarly, the revised manuscript also offers several suggestions for improving future top-down emissions studies.

  We do not scale the EDGAR emissions inventory in the revised manuscript. We have also added a discussion on boundary condition uncertainties to the supplement. This discussion mirrors the boundary condition uncertainty analysis in Miller et al. (2013) and Miller et al. (2014a).

- **p.9354, l.12f: you need to provide an explanation why you restricted your time series analysis to only a few sites, and why you chose those 4.**

  We chose those four sites because they are located near large wetlands and in regions where the synthetic data experiments produced positive results; the wetland methane signal is easier to distinguish at these sites relative to others. In the revised manuscript, we have added plots of all remaining US and Canadian sites in the supplement (Fig. S4).

- **p.9355f, Section 4.3: I think it's not enough to base the seasonality analysis on relative flux contributions from each month alone. Since most of the bottom-up models (as shown in Fig.4) have very high flux emissions rates in summertime, it may well be that the shoulder season fluxes are matching the inverse product quite well, while there are simply way off in the peak of the warm season.**

  We have added a second panel to Fig. 5. That panel shows the absolute budget in Tg CH_4 per month. The revised version of Fig. 5 therefore allows the viewer to compare both the relative and absolute seasonal cycle in each model.

- **p.9356, ll.10ff: It's a plausible explanation that air temperatures are significantly decoupled from the conditions in the soil (where CH4 is produced) for fall, but not for spring ... even if you can show through NARR that soils start thawing in April, this isn't possible without air temperatures that are appropriately high ...**

  Another possible explanation is that the temperature threshold for CH_4 production may be too high in some existing models. Most models predict relatively small fluxes when soils are cold but still above freezing. And most predict dramatically larger fluxes in the summer when both soil and air temperatures are at their peak. Our analysis suggests that the shape of the seasonal cycle may be broader than that predicted by many models; the relative difference between cool-season and warm-season fluxes may not be as great as predicted by many bottom-up estimates. This conclusion is supported by flux measurements taken across the

North Slope of Alaska by Donatella Zona (University of Sheffield). Her paper is currently under review. We have added to this discussion in Section 4.3 of the revised manuscript.

- **p.9357, ll.3ff: As mentioned already above, this hasn't been discussed earlier, and I don't think this is the proper place to start with this kind of agenda. I agree with the general statement, but if you want to place it in a publication you need to be more constructive. Your results show that there are obviously still large discrepancies between the methane signal that is simulated by WETCHIMP models, and the methane signal as seen from the atmospheric observations. Still, you don't offer any conclusions how information from atmospheric methods might be used to improve the bottom-up models ...**

  The reviewer makes a good suggestion here. We have removed this statement from the conclusions.

- **p.9357, 2nd paragraph (ll.8ff): I think this part of the conclusions needs more details. You just list your basic findings, without even attempting to interpret where these differences come from. Also, you seem to assume that any atmospheric inverse modeling product (or the approach to link tower observations to surface fluxes through atmospheric inverse modeling) can be regarded as the 'truth', and all discrepancies with bottom-up products can be attributed to shortcomings in the latter.**

  We have revised the wording of the article to make the top-down, atmospheric analysis sound less absolute (relative to bottom-up modeling). We have also included more discussion in Sections 4.1-4.3 and in the Conclusions, discussion that emphasizes plausible reasons for any discrepancies between the top-down analysis in the paper and existing bottom-up estimates of wetland fluxes.

References

Fang, Y., Michalak, A. M., Shiga, Y. P., and Yadav, V.: Using atmospheric observations to evaluate the spatiotemporal variability of CO_2 fluxes simulated by terrestrial biospheric models, Biogeosciences, 11, 6985–6997, doi:10.5194/bg-11-6985-2014, 2014.

Hegarty, J., Draxler, R. R., Stein, A. F., Brioude, J., Mountain, M., Eluszkiewicz, J., Nehrkorn, T., Ngan, F., and Andrews, A.: Evaluation of Lagrangian particle dispersion models with measurements from controlled tracer releases, J. Appl. Meteorol. Clim., 52, 2623–2637, 2013.

Miller, S. M., Wofsy, S. C., Michalak, A. M., Kort, E. A., Andrews, A. E., Biraud, S. C., Dlugokencky, E. J., Eluszkiewicz, J., Fischer, M. L., Janssens-Maenhout, G., Miller, B. R., Miller, J. B., Montzka, S. A., Nehrkorn, T., and Sweeney, C.: Anthropogenic emissions of methane in the United States, P. Natl. Acad. Sci. USA, 110, 20 018–20 022, doi:10.1073/pnas.1314392110, 2013.

Miller, S. M., Worthy, D. E. J., Michalak, A. M., Wofsy, S. C., Kort, E. A., Havice, T. C., Andrews, A. E., Dlugokencky, E. J., Kaplan, J. O., Levi, P. J., Tian, H., and Zhang, B.: Observational constraints on the distribution, seasonality, and environmental predictors of North American boreal methane emissions, Global Biogeochem. Cy., 28, 146–160, doi:10.1002/2013GB004580, 2014a.

Miller, S. M., Michalak, A. M., and Levi, P. J.: Atmospheric inverse modeling with known physical bounds: an example from trace gas emissions, Geoscientific Model Development, 7, 303–315, doi:10.5194/gmd-7-303-2014, 2014b.

Nehrkorn, T., Eluszkiewicz, J., Wofsy, S. C., Lin, J. C., Gerbig, C., Longo, M., and Freitas, S.: Coupled Weather Research and Forecasting-Stochastic Time-Inverted Lagrangian Transport (WRF-STILT) model, Meteorol. Atmos. Phys., 107, 51–64, doi:10.1007/s00703-010-0068-x, 2010.

Interactive comment on Biogeosciences Discuss., 12, 9341, 2015.