

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

**Anonymous Referee #2**

Received and published: 12 August 2015

Review of ‘The ability of atmospheric data to resolve discrepancies in wetland methane estimates over North America’ by S. M. Miller et al. May 2015

Summary

This manuscript presents an evaluation of a set of North American methane flux estimates from the global comparison project named WETCHIMP (Wetland and Wetland CH<sub>4</sub> Inter-comparison of Models Project). This study used the Bayesian Information Criterion (BIC) statistical analysis in an attempt to use atmospheric CH<sub>4</sub> data ‘to evaluate the magnitude, seasonal cycle, and spatial distribution of each WETCHIMP methane estimate. Of the seven available models, which have a magnitude, seasonal

C4229

cycle, or spatial distribution that is most consistent with the available data?’ (page 9345, lines 6-9)

It is useful to evaluate models and perform model-observation comparisons to help understand the performance of models. Using a different approach such as the BIC analysis here could be helpful in the understanding of the model results and performance. However, the pieces of analyses presented in this manuscript (sections 3, 4.1, 4.2, 4.3) seem disjointed and possibly conflicting. Thus the key question posed by the authors above was not answered in the manuscript, leading to the conclusion that: based on the results from this study, either the analysis is not working properly and/or atmospheric data has little ability to resolve discrepancies in wetland methane estimates using these types of analyses.

In general, a new method of statistical analysis (BIC for CH<sub>4</sub> wetland flux model selection) needs to be tested fully to understand the abilities and limitations to give confidence to the results. Examine the uncertainties for each model components, such as sensitivity of the methodology to transport errors, flux errors, background concentration estimates, other assumptions/approximations and their interactions, as well as the stability or robustness of the analysis. If a thorough analysis is not done, at least a full discussion of the potential problems is needed to put the preliminary results in the proper perspective.

This study raises many questions (in the following sections) that need to be addressed and the analyses modified. Therefore this manuscript would require major revisions before being accepted for publication.

General comments

Section 3

In Section 3: 4 of the 7 models were selected from WETCHIMP for the synthetic data experiment to evaluate the BIC method. Presumably the authors expected (without

C4230

sufficient reasons) the 3 remaining models to be 'unacceptable', including LPJ-Bern. Yet BIC results in section 4.1 indicate LPJ-Bern is selected most often when using real observations, bringing into question the value and correctness of this BIC method (in its current form). This suggests the rejection of LPJ-Bern could be a problematic assumption in section 3 and/or the BIC methodology is not working. Possibly other assumptions and approximations (such as uncertainty estimates, model transport errors, etc.) are incorrect, the whole new method needs to be evaluated much more thoroughly as noted above. In general, much stronger justifications are needed for model selection. The BIC analysis should be done for all WETCHIMP models in the synthetic data experiment. If the range of flux model variations is too large, then what is the range of applicability for the BIC method?

The authors stated 'By contrast, the observation network is largely insensitive to spatial variability in wetland fluxes across the US; in most instances, the model selection framework favors a spatially-constant model over a wetland model for the two US regions.' (page 9352, lines 9-12). Was 'a spatially-constant model' included in the work and the BIC model 'favors' or selected it over the other models? The comparison to 'a spatially-constant model' in general needs to be explained and documented more clearly. If the BIC method actually selects the 'spatially-constant model' (a clearly wrong model), maybe this is an indication of the lack of ability of the method.

#### Section 4

The 3 subsections, 4.1, 4.2, 4.3 seem to be unrelated and somewhat conflict with each other. Sections 4.2 and 4.3 did not make use of the information from 4.1 that LPJ-Bern and SDGVM performed best in the BIC analysis. In fact, the best model from 4.1, LPJ-Bern performed poorly in the flux magnitude and seasonal maximum comparison in section 4.2. There is no explanation on this discrepancy. This could be an indication of the problem in the BIC analysis in selecting a poorly performing flux model.

Section 4.3 is a comparison of all 7 flux models to the results of another inversion

C4231

model, Miller et al. 2014. Since inversion model results are highly uncertain, this simple model to model type comparison has little useful information unless the authors can show independently that the Miller et al. 2014 results are good and can serve as a benchmark for comparison. At the minimum, the WETCHIMP fluxes should be compared to a variety of inversion model results to see the uncertainty possible for the inversion model results. The authors should note in the manuscript that section 4.3 is only a comparison of models, which is different than comparison to 'available data' (page 9345, lines 6-9, see above).

#### Specific comments

Page 9345, line 1: 'biogeochemical models leveraged all available data', what is this vague 'all available data' referring to?

Page 9346: "background" concentration – the methane concentration of air entering the North American regional domain' Do all STILT particles always leave the North American regional domain during the model simulations? Provide more information on the estimation of the background concentration and the whole model settings to enable other scientists to check and/or compare results.

Page 9346, lines 21-24: repeating lines 23-27 on page 9344.

Page 9347, line 7: change 'first term in Eq. (1)' to 'first two terms in Eq. (1)'.

Page 9348, line 29: 'We also include a spatial constant or intercept term in X'. What is the physical or numerical significance of the 'spatial constant or intercept term'. How do the results compared with or without the 'spatial constant or intercept term'?

Page 9353, line 26: change 'GEIMS' to 'GIEMS'.

Page 9354, lines 16-20: 'The estimated contribution of anthropogenic emissions from EDGAR v4.2FT2010 is added to this background (in red). Note that the estimated scaling factors for EDGAR (Sect. 2.4) are  $1.7 \pm 0.3$  at Chibougamau,  $5.6 \pm 0.5$  at East Trout Lake,  $2.4 \pm 0.3$  at Fraserdale, and  $2.5 \pm 0.3$  at Park Falls.' Explain the meaning

C4232

of the different scaling factors and how realistic are they (up to 5.6x)? What are the spatial regions these scaling factors are applied to?

Page 9364: left figure contains wrong information 'Observation site (Fig. 4)'. Units in the right figure conflicts with caption.

Page 9368: Label each curve in Fig. 5 as in Fig. 4.

Supplement

Page 2, lines 56-57: 'In this equation,  $z_{\text{synthetic}}$  ( $n \times 1$ ) represents the synthetic observations generated for an observation site'. What is the number of sites ( $k$ ) and does  $n$  vary for each of the  $k$  sites?

Page 2, lines 71-72: 'The other WETCHIMP models, in contrast, predict much higher fluxes (Fig. 4)'. Fig. 4 only showed concentrations or mole fractions, change 'much higher fluxes' to 'much higher concentrations'.

Page 2, 80-82: 'The magnitude and spatial/temporal structure of these errors were estimated in Miller et al. (2013) for the US and Miller et al. (2014b) for Canada.' The referenced works were for different prior fluxes. Authors need to show why new error estimates are not needed.

Page 3: Correct the difference in units in the Figure and caption.

---

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