1. Reviewer 2 Comments

This manuscript presents the results of a multi-model intercomparison of methane emissions from the West Siberia Lowlands. The West Siberia Lowlands are a good choice for this study – big and important, some good data (but not enough to know the answer), and important climate gradients, particularly non-permafrost to permafrost. The intercomparison includes inverse and forward models of varying complexity and emphasis, and thus represents a diversity of approaches. Overall, it represents the state-of-the-art in regional/global methane modeling, and should be of interest to readers of Biogeosciences.

The paper is very clearly written and the tables and figures are also clear (a few comments on the figures below). I recommend minor revisions before final publication.

GENERAL COMMENT

The concluding recommendations are not unexpected, but it is useful to have them spelled out and backed up by the analysis of multiple models of multiple types. It would be interesting to read any conclusions/recommendations you reached at this stage about model representation(s) of biogeochemistry?

SPECIFIC COMMENTS (numbering added by author)

1. p. 1915, l5-7. Why aggregated from 25-km to 0.5°? There is probably a good reason, which you should provide.
2. p. 1926, l5-7. Comparing soil moisture content between mineral and peat soils – what do you mean by 'content'? by mass or volume, or by degree of saturation? This needs a more careful explanation.
3. p. 1931, l3-4: this is true for UW-VIC (GEIMS) in the north only.
4. p. 1934, l1-3. This isn't clear, and as I try to interpret it, it doesn't seem like a general conclusion in keeping with points above.
5. p. 1934, l4-21. Would an interactive N cycle also be a longer-term influence? Did the N-cycle (stocks and/or fluxes) change substantially over the _10 year simulations for those models that included it?
6. p. 1934, l22-28. This paragraph may be more specific to a limited set of models than should be included in the paper.
7. p.1935, l5. 'larger' or 'large'?
8. p. 1937, l17-19. Well, really, from a climate change point of view, CH4 is well-mixed in the atmosphere and has a c.10-year lifetime, so to first order (which is where we are with this collection of models) long-term mean emissions is probably good enough. Not satisfactory, and not a goal, certainly, but not necessarily any worse than the other results at this point. Until we have more confidence in the models, this is probably still as good as any of them.
9. Refs missing – at least Walter et al. 2006; Pace et al. 2004 (I didn't do a thorough check, but you should).
10. Table 2. A footnote should define I, M, M+, and T.
11. Fig. 5. Interesting figure! I suggest moving I, T, M and gray symbols to upper right (above legend (and adding that to figure 12 upper right), and then either reduce area in upper left to 800 (all match), or reduce all areas to use more of the graph.

12. Fig. 5 & 8 & 12 (in particular). Increase font size in legends (there is space in upper right). As many model names are similar, it is difficult to tell them apart when the font is small.
13. Fig 12. Explain 'Tair-dominated' and 'Finund-dominated' and associated lines at 0.7 in caption, for the benefit of most of your 'readers'.

2. Author Response

General

Unfortunately, as indicated on page 1935, line 5, the scatter in model results arising from other differences (differences in how methane-contributing areas are delineated and differences in soil thermal physics) was so large that it prevented us from seeing clearly the effects of biogeochemical representations across all models. Those cases in which a single model was run with different biogeochemical configurations did illuminate some potential effects of biogeochemical representations (e.g., page 1934, lines 22-28). In response to your question on N cycle and C stocks (specific comment #5), we have added some information about the LPX-BERN simulations in this regard to the results and discussion. But we feel that point (e) in the abstract sums up our biogeochemical findings: they had relatively smaller effects than the large errors due to poor wetland area constraints and inaccurate soil thermal physics schemes (or, in the case of nitrogen limitation, the factor was only examined in a single model, preventing us from separating out artifacts of model implementation). To discriminate among biogeochemical schemes would require another model intercomparison focusing on models that use similar (accurate) wetland areas and soil thermal physics, to eliminate these sources of noise.

Specific

1. This was for consistency with model results. We have added a few words to that effect.
2. Thank you for catching this. We have replaced "reductions in soil moisture content" with "larger sensitivities of water table depth to evaporative loss".
3. We have qualified our statement with "in the North".
4. We agree; the use of poorly constrained model features can lead to poor performance in any application and is not unique to the modeling of high latitude wetland methane emissions. We have removed this point.
5. Nitrogen limitation had substantial effects on mean CH4 emissions and minor effects on carbon stocks in the LPX-BERN simulations. While the effects on mean CH4 emissions were large, we cannot separate out the effects of model implementation due to only LPX-BERN simulating this effect. The effects on carbon stocks and trends in CH4 emissions were small over the 12-year period, again calling attention to the need for longer study periods (although this topic need not be limited by the observational record). We have added a few sentences describing these effects to the results and discussion sections.
6. We would prefer to keep this paragraph. While the features discussed here only applied to a small number of models, they nonetheless gave us some idea of the sizes of uncertainties due to these features (small) relative to uncertainties due to other features such as soil thermal physics (large). The features discussed here are biogeochemical in nature, addressing the reviewer's general comment. In addition, we have incorporated our answer to the reviewer's previous comment (#5) into this paragraph.
7. Thanks for catching this; we have changed this to "large".

8. We have rephrased the final sentence of the paragraph to have a less critical tone towards the Bloom et al (2010) product.
9. In fact, Pace et al. (2004) was not missing. But yes, Walter et al. (2006) was missing, as well as Tarnocai et al. (2009), and we have added those references. There also was a typo in our citation of Berrittella and van Huissteden (2011), which we have fixed. Thank you for catching that.
10. We agree, and have added footnotes explaining these codes (and other aspects of the table). If the editor prefers, we can move this information into the table caption.
11. We agree, the symbol definitions are better in the upper right, next to the legend box. We have moved them there. We can't give the panels all the same x limits since the areas in the WSL panel (upper left) are the sum of the areas in the south and north (lower left and right, respectively). In addition, data points fall very near the x- and y-limits of the WSL panel, so we cannot shrink it without losing those points. However, we reduced the maximum x value in the south and north panels to 700 (from 800). In addition, we removed some of the intensity lines, and we labeled all panels with letters (a, b, c) and moved the labels to the upper left of each panel.
12. We agree, the legends were quite small in these figures. We have expanded them.
13. We are not sure that we understand this request. The caption of Figure 12 already contains the following text:

    > "F_{inund}-Dominated" and "T_{air}-Dominated" denote correlation thresholds above which inundated area or air temperature, respectively, explain more than 50% of the variance of CH_4 emissions.

    We think that this text addresses your question. Could you clarify your request? Perhaps you were referring to the symbol definitions for circles, triangles, squares? Just in case, we have also copied the text describing these from the caption of Figure 5 and pasted it here. However, this makes the caption rather lengthy – perhaps the editor can give us some guidance here?

3. Author's Changes in Manuscript

General

(page and line numbers refer to the Word document *with markup shown*)

To address the reviewer's questions about biogeochemical formulations here and in specific comment #5, we added the following text to page 19, line 19 – page 20, line 3:

> Nitrogen limitation influenced intensity in LPX-BERN, the one model that included it. Although we did not plot results from the two LPX-BERN configurations that lacked nitrogen-carbon interactions in Figure 5, we compare results from all four LPX-BERN configurations in Table 6. In LPX-BERN (N) and LPX-BERN (DYPTOP-N), the nitrogen limitation imposed by nitrogen-carbon interactions substantially reduced NPP, relative to LPX-BERN and LPX-BERN (DYPTOP), leading to a reduction of mean annual CH_4 emissions of approximately 20% over the entire WSL over the period 1993-2010. This reduction was slightly larger than the difference in emissions between

simulations using the Sheng2004 map to prescribe peatland area (LPX-BERN and LPX-BERN (N)) and simulations using the DYPTOP method to determine peatland extent dynamically (LPX-BERN (DYPTOP) and LPX-BERN (DYPTOP-N)). In addition, the reduction in emissions due to nitrogen limitation was concentrated in the northern half of the domain, in contrast to the reduction due to dynamic peatland extent, which was concentrated in the southern half of the domain. Nitrogen limitation also reduced trends in CH_4 emissions over the entire WSL over the period 1993-2010, through reductions in soil carbon accumulation rates. However, both these trends and their reductions were very small (< 0.5% per year in most cases) and statistically insignificant over the study period.

We also added a table (Table 6) summarizing these results from LPX-BERN.

Specific

1. Page 8, lines 21-24: these lines now read:

   "For both products, surface water area fractions (F_w) were aggregated from their native 25 km equal-area grids to a 0.5×0.5° geographic grid and from daily to monthly temporal resolution, for consistency with model results."

2. Page 19, lines 4-8: these lines now read:

   "While this allowed LPJ-Bern to make emissions estimates in the South, the much lower porosities of mineral soils resulted in larger sensitivities of water table depth to evaporative loss than those of peat soils. These drier soils led to net CH_4 oxidation in much of the South."

3. Page 24, lines 1-2: we inserted "in the North".
4. Page 26, lines 19-21: we removed these lines.
5. As mentioned in our response to the general comment, we have added a paragraph discussing the effects of nitrogen limitation in LPX-BERN, page 19, line 19 – page 20, line 3:

   "Nitrogen limitation influenced intensity in LPX-BERN, the one model that included it. Although we did not plot results from the two LPX-BERN configurations that lacked nitrogen-carbon interactions in Figure 5, we compare results from all four LPX-BERN configurations in Table 6. In LPX-BERN (N) and LPX-BERN (DYPTOP-N), the nitrogen limitation imposed by nitrogen-carbon interactions substantially reduced NPP, relative to LPX-BERN and LPX-BERN (DYPTOP), leading to a reduction of mean annual CH_4 emissions of approximately 20% over the entire WSL over the period 1993-2010. This reduction was slightly larger than the difference in emissions between simulations using the Sheng2004 map to prescribe peatland area (LPX-BERN and LPX-BERN (N)) and simulations using the DYPTOP method to determine peatland extent dynamically (LPX-BERN (DYPTOP) and LPX-BERN (DYPTOP-N)). In addition, the

reduction in emissions due to nitrogen limitation was concentrated in the northern half of the domain, in contrast to the reduction due to dynamic peatland extent, which was concentrated in the southern half of the domain. Nitrogen limitation also reduced trends in CH_4 emissions over the entire WSL over the period 1993-2010, through reductions in soil carbon accumulation rates. However, both these trends and their reductions were very small (< 0.5% per year in most cases) and statistically insignificant over the study period."

We also added a table (Table 6) summarizing these results from LPX-BERN.

We also added the following lines to the discussion section (page 27, lines 14-19):

"Similarly, nitrogen-carbon interaction had a substantial latitude-dependent effect on mean CH_4 emissions for LPX-BERN (Table 6). Again, the size of the effect could be model-dependent, and potential impacts on sensitivities to climate change might become more apparent over a longer analysis period."

6. There were no edits specifically related to this comment, but we edited this paragraph in response to comment #5, above.
7. Page 27, line 24: changed "larger" to "large".
8. Page 29, lines 28-31: the text now reads:

   "Thus, while Bloom2010 provided a useful estimate of long-term mean emissions, it was less helpful in constraining model responses to climate drivers."

9. Page 27, line 22: fixed spelling error in citation of Berrittella and van Huissteden (2011); page 38, lines 16-19: removed citation of Hauglestaine et al (2004); page 40, line 23: added doi for Liu et al. (2013); page 47, lines 1-3: inserted citation for Tarnocai et al. (2009); page 48, lines 11-14: inserted citation for Walter et al. (2006).
10. Pages 52-58, tables 2-3: We have added footnotes underneath the tables to explain the column headings (in addition to changing the wording of the column headings to be more consistent with our terminology).
11. Figure 5: updated the figure accordingly.
12. Figures 5, 8, and 12: updated these figures (primarily in the legends, but also in symbol codes and in replacing "Finund" with "Fw").
13. Page 73, lines 9-14: added the following text to the caption:

    "Circles denote models that used satellite surface water products alone (corresponding to code "S" in Table 2) to delineate wetlands. Triangles denote models that used topographic information, with or without surface water products (corresponding to code "T" in Table 2). Squares denote models that used wetland maps with or without topography or surface water products (corresponding to code "M" in Table 2)."