Dear Professor Davidson,

Thank you and the three respected reviewers for your time in reviewing our manuscript “Satellite methane observations constrain latitudinal distribution of wetland emissions and their climate sensitivity”. We greatly appreciate all the comments and suggestions toward improving our manuscript. We have addressed all the concerns as best as we could; please find the detailed responses below in blue fonts.

Please let us know if you need any further information or clarifications.

Yours truly,
Shuang Ma, Ph.D.
Postdoctoral Scholar, Jet Propulsion Laboratory at Caltech
Telephone: +1 8183542113
Email: shuang.ma@jpl.nasa.gov

--

Dear Dr. Ma:

Thank you for submitting "Satellite methane observations constrain latitudinal distribution of wetland emissions and their climate sensitivity" [Paper #2021AV000408R] to AGU Advances. I have received 3 reviews of your manuscript, which are included below and/or attached. As you can see, the reviews indicate that major revisions are needed before we can consider proceeding with your paper. I am therefore returning the paper to you so that you can make the necessary changes.

The three reviewers agree that the manuscript is generally well written, although they found numerous minor errors indicating some lack of attention to detail in editing by the coauthors. Reviewers #1 and #2 emphasize the potential importance of the methodological innovation of using the averaging kernel matrix for rigorous analysis of the dependence of the fluxes from the inversion on the prior information used in the inversion. All three reviewers note that the overall scientific findings are generally in line with recent developments in understanding of latitudinal patterns and controls of methane fluxes from wetlands, and so the real advance here is the innovation afforded by the Bayesian approach to the inversion modeling regarding averaging of the kernel matrix. Reviewers #1 and #2 like this approach and think it is worthy of publication, although they also list several caveats and potential limitations and shortcomings that need to be acknowledged and addressed in a revision. All three reviewers believe that there are some overstatements of results, including inferences made by the title. I recommend that the essence of the novel methodological approach, such as the statement on lines 117-119, as well as a succinct statement about why this innovation is important, should be incorporated into the abstract, one of the three key points, and perhaps the title. [e.g., an alternative title might be developed along the lines of: "A novel inversion model innovation to improve application of satellite methane observations to elucidate latitudinal distribution of wetland emissions and their climate sensitivity"].

Thanks for this nice suggestion! We strongly agree that we should emphasize our innovation in methodology. We added the novelty of our approach to key point #1: ‘We use a novel Bayesian approach to account for cross-correlations and spatial uncertainties in top-down flux estimates’. We also changed the title to fit our new findings in this revision (addressed in the next section). We struggled between ‘A Bayesian based inversion model innovation for quantifying the latitudinal distribution and temperature sensitivity of wetland emissions with satellite concentration data’ and ‘Satellite constraints on the latitudinal distribution and temperature sensitivity of wetland methane emissions’, and chose the latter because it’s shorter and easier to follow. The revised abstract also points out the innovation in methods: ‘Our comparison between bottom-up models and satellite-based fluxes innovatively accounts for cross-correlations and spatial uncertainties typically found in top-down inverse estimates, such that only the
information from satellite observations and the atmospheric transport model is kept as a constraint.’

The reviewers’ concerns about caveats of this approach should be fully addressed, and the emphasis on new insights into global patterns should be toned down to match what is really new and what is confirmatory of conclusions from other recent studies. None of us can tell for sure at this point whether the methodological innovation will truly be an advance that significantly influences further inversion model applications, but reviewers #1 and #2 appear sufficiently impressed that I believe it is appropriate for this journal to present it once properly scoped in a revision. Inversions using a combination of top-down satellite data and bottom-up constraints as priors are likely to become more common and important as the former become more available, so advances in the methodology in using these resources may be very important.

We completely agree! Our previous results emphasize the important of wetland extent map in determining model performance, which is not a new finding. The high-performance and low-performance models do not have significant differences in correlation of CH4 with climate/wetland extent. We agree we have not shown enough results to emphasize our new insights into global pattern. However, inspired by the reviewers’ comments, we dug into the data and found more interesting results, as shown in the new figure 4, and explained below:

Figure 4. Marginal RMSE performance (see section 2.4) of the factorial WetCHARTs model ensemble, with respect to individual temperature sensitivity of CH4:C emissions (Q10 CH4:C), global wetland emission scaling factors, and wetland extent datasets. The black dots represent medians, boxes represent the 25th- 75th percentile ranges, and lines represent 5th- 95th percentile ranges.

We find that 1) models using a Q10 equals 1 have significantly higher performance than the ones using larger Q10 values (top row in Figure 4); 2) models with the smaller global scalar perform better than those with larger scalars (bottom row in Figure 4); 3) models
use the GLOBCOVER derived wetland extent maps outperform those use the GLWD map.

These imply that smaller seasonal fluctuations in response to temperature, a smaller global total emission, and the satellite-based wetland extent agrees better with atmospheric CH$_4$ measurements, showing constraints on the parameters and drivers of terrestrial biosphere models.

We also find regardless of wetland maps, LPX-Bern and JULES always stay in the high-performance group, implying more realistic model structures or better choice of parameter values, which awaits to be further investigated.

Altogether we justify our conclusions on reducing uncertainties in the terrestrial biosphere model simulations by constraining model structures, parameters and drivers.

**In terms of placing constraints on terrestrial biosphere models predictions, we added a discussion in section 3.4 to clarify our statements:**

‘The large uncertainties in predicting wetland CH$_4$ emission come from 1) climate and wetland extent predictions (drivers) (Andresen et al., 2020; Zhang et al., 2017); 2) process controls (model structures); and 3) climatic sensitivities of CH$_4$ productions (parameter values) (Poulter et al., 2017; Riley et al., 2011). Our approach enables the evaluation of terrestrial biosphere models with satellite-informed estimates and places constraints for land model predictions by telling which wetland extent maps, biogeochemical processes, and parameters values used in the models are more realistic, as discussed in section 3.2 (Cox et al., 2013; Wenzel et al., 2016).’

The AGU editorial staff is concerned about the lack of availability of the WetCHART v 1.3.1 dataset. It is not consistent with our policy to make such data only available upon request to the authors. Is there a way to address this problem?

Thank you! We attached a DOI for the WetCHARTs v 1.3.1 data from ORNL DAAC repository. We are in progress of publishing WetCHARTs v 1.3.1 data on ORNL DAAC. The data has been under QA for two weeks ([track here](#)). The whole process may take a few more weeks to be available for the public. We could instead get a DOI from Zenodo public repository if the DAAC repository takes too long to pass before needed.

I found that the "future directions" section added at the end of the manuscript distracts from the nicely written summary that precedes it. I would recommend moving the future directions paragraph to the last paragraph in the discussion section and just before the summary section.

Thanks! We moved the future directions as suggested and shortened the summary section. Now the summary section really stands out.

Please submit a revised manuscript that addresses the reviews and any editorial comments by May 17, 2021.

Sorry for taking extra weeks to finish the revision.

In your revision, please follow our Checklist and use our Templates for the main file and any supplements. Please provide the following:

1. A response to reviewer file that lists each major comment and describes how the manuscript has/has not been modified in response to those comments.
2. A copy of the manuscript with the changes noted (e.g., highlighted, "track changes," italics or bold changes).

3. The final revised manuscript with changes incorporated and separate final figure files (figure parts should be combined into a single file), which will be used for publication if the manuscript is accepted. If final figures are already uploaded, they can be easily copied over to the next revision version.

4. If any, supporting information text, figures, captions, and small tables in single PDF file using AGU's template. Large data tables and multimedia should be uploaded separately.

AGU requires that data needed to understand and build upon the published research be available in public repositories following best practices. This includes an explicit statement in the Acknowledgments section on where users can access or find the data for this paper. Citations to archived data should be included in your reference list. All references, including those cited in supporting information, should be included in the main reference list and cited in-text. All listed references must be available to the general reader by the time of acceptance. AGU requires the corresponding author, and encourages all authors, to register for an ORCID.

Please check and verify authorship, and that all authors are included, have approved the revisions, and agreed to be listed in the order given. Authorship is final with publication. Responsibilities of the corresponding author are given here.

When you are ready to submit your revision, please login to your account (https://advances-submit.agu.org/cgi-bin/main.plex) and click "Revise 2021AV000408R."

I look forward to receiving your revised manuscript. If you have any questions, or need additional time to complete your revisions, please contact us at advances@agu.org.

Yours sincerely,

Eric A. Davidson
Editor
AGU Advances

---------IMPORTANT INFORMATION------------------------
Additional information on text preparation, formatting, acceptable file formats, supporting information, graphics preparation, and AGU style, is here.

Sharing your work is an important part of the research process, and AGU leverages and shares published research to promote the broader importance of Earth and space science. Learn how you can promote your paper, including how your paper can be considered for
Reviewer #1 Evaluations:
Recommendation: Return to author for minor revisions
Significant: The paper has some unclear or incomplete reasoning but will likely be a significant contribution with revision and clarification.
Supported: Mostly yes, but some further information and/or data are needed.
Referencing: Yes
Quality: Yes, it is well-written, logically organized, and the figures and tables are appropriate.
Data: No
Accurate Key Points: No

Reviewer #1 (Formal Review for Authors (shown to authors)):

This paper presents a nice analytical trick to use fluxes optimized through atmospheric inversion to validate fluxes from bottom-up process models in a novel way (or at least novel to me). One of the weaknesses of doing this directly is that there can be considerable influence of the prior still in the posterior fluxes, especially when the measurement constraint is lacking over certain times or regions. By applying the averaging kernel of the model to the process model fluxes (somewhat like calculating the smoothing error in the observation operator when comparing satellite column to modeled profiles), the authors have corrected for the influence of the prior. This is indeed a clever trick, with several potential future applications, but I am concerned that the authors understate some of the shortcomings of such an analysis, and overstate the information gained and the potential applications in the context of ESM projections.

Firstly, the shortcomings: This clever trick does nothing to alleviate potential problems with model transport. If, for example, the interhemispheric mixing of the model is too fast/slow, this will directly over-/underestimate northern hemisphere emissions. Similar problems occur for errors in vertical mixing, tropopause height, and stratospheric gradient (the latter two of which are of particular concern when interpreting XCH4). This does not mean that the approach is invalid, but some discussion of these caveats is sorely needed.

Thank you. This is important indeed. We added a discussion on these caveats at the end of the introduction:

‘While our estimates account for the spatial resolution and error associated with the inversion of observations to fluxes, we do not explicitly account for errors in model transport and chemistry, such as interhemispheric mixing time, vertical mixing, tropopause height, and stratospheric gradient. These errors are important and can be accounted for if the corresponding posterior covariance is provided, possibly through multi-model comparisons (e.g., Schuh et al. 2010) or by integrating measurements with different sensitivities to the profile of CH4 (e.g., Jiang et al. 2013, 2017). We thus advocate studies that either characterize or mitigate these errors to offer better constraints
We note that the methodology presented in this study can only apply to atmospheric CH₄ inversions that calculate prior and posterior covariances (Turner et al., 2015). The methodology can also be approximated with other inversion methods (e.g., 4D-Var methods) if these can generate posterior covariance estimates (Liu et al., 2014; Chevallier et al., 2010).

Furthermore, GOSAT (unfortunately?) does not only measure methane emitted from wetlands. While that would simplify matters considerably, the interhemispheric distribution of methane fluxes and concentrations is also largely dependent on anthropogenic emissions, which have a distinct interhemispheric gradient with higher values in the northern hemisphere, and not inconsiderable uncertainties. This aspect is barely touched upon, and needs to be clearly explained and discussed, also in the context of the suggested application to carbon dioxide.

Thank you! Although GOSAT includes non-wetland methane emission, the posterior covariance for the wetlands automatically includes the cross-state error from the other terms and therefore they are accounted for in the error analysis (Worden, 2004). We added the above clarifications in our method (second paragraph in section 2.1). In addition, Fig. 6 of Zhang et al. (2021) shows that the error correlation between wetland & anthropogenic emissions are overall small on the regional scales.

Finally, I think the authors may be taking the potential applications in terms of ESM projections a bit far. This is fundamentally a data-driven constraint on process models, or perhaps better, on datasets upon which process models are dependent. Because that is one of the main findings, isn't it? That the same model can belong to either the high or low performing ensemble depending on which wetland extent map is used. (This is not really a new finding, but rather a new way to get at this information.) It is not clear how this sort of benchmark based on atmospheric measurements should be used to analyze ESM projections. There is no reason to think that the correlations that have been deduced for this 3- or 8-year period will hold for projections for the year 2100. Specifically, in the text:

Thank you for the comments. We completely agree!

As the editor and reviewers pointed out: Our previous results emphasize the important of wetland extent map in determining model performance, which is not a new finding. The high-performance and low-performance models do not have significant differences in correlation of CH₄ with climate/wetland extent. We agree that based on these conclusions, it is not clear how atmospheric estimates can be used to analyze ESM projections. And we did not conclude any new gains in terms of science.

However, inspired by the reviewers’ comments, we dug into the data and found more interesting results, as shown in the new figure 4, and explained below:
Figure 4. Marginal RMSE performance (see section 2.4) of the factorial WetCHARTs model ensemble, with respect to individual temperature sensitivity of CH4:C emissions (Q10 CH4:C), global wetland emission scaling factors, and wetland extent datasets. The black dots represent medians, boxes represent the 25th-75th percentile ranges, and lines represent 5th-95th percentile ranges.

We find that 1) models using a Q10 equals 1 have significantly higher performance than the ones using larger Q10 values (top row in Figure 4); 2) models with the smaller global scalar perform better than those with larger scalars (bottom row in Figure 4); 3) models use the GLOBCOVER derived wetland extent maps outperform those use the GLWD map.

These imply that smaller seasonal fluctuations in response to temperature, a smaller global total emission, and the satellite-based wetland extent agrees better with atmospheric CH4 measurements, showing emergent constraints on both model parameters and drivers data.

We also find regardless of wetland maps, LPX-Bern and JULES always stay in the high-performance group, implying more realistic model structures or better choice of parameter values, which awaits to be further investigated.

Altogether we justify our conclusions on reducing uncertainties in the terrestrial biosphere model simulations by constraining model structures, parameters and drivers.

In terms of placing constraints on ESM predictions, we added a discussion in section 3.4 to clarify our statements:

‘The large uncertainties in predicting wetland CH4 emission come from 1) climate and wetland extent predictions (drivers) (Andresen et al., 2020; Zhang et al., 2017); 2) process controls (model structures); and 3) climatic sensitivities of CH4 productions (parameter values) (Poulter et al., 2017; Riley et al., 2011). Our approach enables the evaluation of terrestrial biosphere models with satellite-informed estimates and places
constraints for land model predictions by telling which wetland extent maps, biogeochemical processes, and parameters values used in the models are more realistic, as discussed in section 3.2 (Cox et al., 2013; Wenzel et al., 2016).

To help illustrate the logic, we also corrected the misuse of ‘projections’ instead of ‘simulations’ at a few places. Please see our answers to specific questions below and also the revised text in the manuscript.

L292-292: I'm not sure I understand what you're trying to argue with the following sentence: "Our Bayesian methodology can be used to evaluate the probability of Earth System Model (ESM) projections based on the zonal profile of wetland emissions." This implies that only with this Bayesian methodology can you show that future wetland extent is going to be correlated with future wetland emissions... With no future data to see what the atmospheric methane distribution will be, it is hard to see what additional constraint would be brought here. Note that also the LP models have a high correlation of emissions with wetland extent - this Bayesian methodology was really more useful at sorting out which wetland extent/model combinations produce realistic fluxes, constrained by atmospheric methane measurements.

We strongly agree with the reviewer and thanks for pointing it out. We deleted this sentence and please find our new findings and corresponding implications in the above section.

L302-303: "can" seems a bit strong here, perhaps "might" or "could"? Also, I'm not clear on what is meant by using your analysis "as emergent constraints on ESM CH₄ predictions". Perhaps it would be more reasonable to suggest that your analysis could be used to test emergent constraints (in simulations from past years) that have been deduced from ESM CH₄ predictions? There is a logical flaw in arguing that a data-constrained result can be used to test the plausibility of climate projections, when there are no longer data to constrain the problem. This is essentially extrapolation into a regime that you haven't constrained with your model. Using this approach to test the reliability of process models (or perhaps more importantly, the robustness of the wetland extent simulated in ESMs?) based on current observations is reasonable, looking into the future, less so.

We agree! Thanks for pointing this out. The logical flow was ambiguous. We now added the following sentences in the discussion section 3.4 and clarify how present-day climate sensitivity can be used to constrain future wetland CH₄ emission scenarios:

"The large uncertainties in predicting wetland CH₄ emission come from 1) climate and wetland extent predictions (drivers) (Andresen et al., 2020; Zhang et al., 2017); 2) process controls (model structures); and 3) climatic sensitivities of CH₄ productions (parameter values) (Poulter et al., 2017; Riley et al., 2011). Our approach enables the evaluation of terrestrial biosphere models with satellite-informed estimates and places constraints for land model predictions by telling which wetland extent maps, biogeochemical processes, and parameters values used in the models are more realistic, as discussed in section 3.2 (Cox et al., 2013; Wenzel et al., 2016)."

The novelty of the approach is however significant, and worthy of publication. Whether this is best in a general-interest journal such as Advances, or rather in a more specialist journal is fundamentally an editorial decision. The methodology is really what is interesting and unique here, the scientific findings are actually not all that ground-
breaking, and the estimated latitudinal distribution and magnitude of wetland emissions is generally consistent with the findings of e.g. top-down estimates from the most recent GCP methane report, among others, as the authors point out in the discussion. The claim in the title, that this approach has allowed them to determine the climate sensitivity of wetland emissions also seems a bit overblown: Figure 4 shows rather that the climate sensitivity of the highest-performance and lowest-performance models are nearly indistinguishable. What is the knowledge gain here, really? Much of the difference between the upper third and the lower third is related to a better wetland extent map rather than an improved understanding of biogeochemical processes or climate sensitivity.

Thank you for the comments. We completely agree! Same as what we replied earlier: We agree that we did not conclude any new gains in terms of science in the previous discussions. However, inspired by the reviewers’ comments, we dug into the data and found more interesting results, as shown in the new figure 4:

We find that 1) models using a Q_{10} equals 1 have significantly higher performance than the ones using larger Q_{10} values (top row in Figure 4); 2) models with the smaller global scalar perform better than those with larger scalars (bottom row in Figure 4); 3) models use the GLOBCOVER derived wetland extent maps outperform those use the GLWD map.

Based on this, I would recommend this study for publication, but would suggest first walking back some potentially unrealistic scientific claims and including some caveats related to the uncertainties in the inverted fluxes.

Regarding data availability: Sometimes it is not entirely clear to me which data should be made available. The authors provide access to the GCP wetland simulations, WetCHARTs and CARDOMOM, but not e.g. the posterior wetland fluxes or the averaging kernel of the transport model. The AGU data policy states that "Data used to generate, or be displayed in, figures, graphs, plots, videos, animations, or tables in a paper" need to be made available. That has not been done, but many papers are published without making such data available.

Thank you! The dataset for the 2010–2012 posterior wetland fluxes and the averaging kernel of the transport model is available at an open-access repository (https://doi.org/10.5281/zenodo.4052518 ; Zhang et al., 2021). We attached a DOI for the WetCHARTs v 1.3.1 data from ORNL DAAC repository. We are in progress of publishing WetCHARTs v 1.3.1 data on ORNL DAAC. The data has been under QA for two weeks (track here). The whole process may take a few more weeks to be available for the public. We could instead get a DOI from Zenodo public repository if the DAAC repository takes too long to pass before needed.

Upon reading the other reviewers comments: I agree fully with Reviewer #2, who managed to explain his or her concerns much more succinctly than I did. I'm not sure Reviewer #3 fully grasped the trick of projecting the fluxes through the averaging kernel of the inversion: this is a tricky idea for people who aren't very familiar with inverse modelling and/or satellite retrievals. However also this reviewer shared the concern that
the novelty of the results are overstated, as one of the main results is that having a better wetland extent map leads to better wetland fluxes - which is hardly novel.

I also had many, many minor typographic comments. I have a lot of patience with language errors when none of the co-authors are native English speakers, but in cases like this I wonder if all of the co-authors have even read the whole paper before submission. This seems to be a more common occurrence with journals that do not have an open review process - the authors know it'll get cleaned up before publication anyhow. This places an undue burden on reviewers, and makes me rethink reviewing for AGU journals going forward.

We apologize for the grammatical mistakes, and really appreciate your help offering suggestions on corrections. We have a few of our coauthors, and another native English speaker outside of the coauthor list to review the language.

Minor typographical concerns:

L3: affiliation of Zhang should be 2,3 I suppose (instead of 23)
    Thanks! Sorry for the typo!
L28: account -> accounts
    Thanks! Fixed.
L30-31: This is a bit awkward (especially for the plain-language summary):
    "hydrological controls on tropical wetland emissions are the dominant regulator of global wetland emission contributions to the atmospheric CH4 budget" Perhaps just:
    "hydrological controls are the dominant regulator of tropical wetland emissions" would be enough?
    Thanks! Corrected as suggested.
L40: Would recommend removing "corresponding variable" here.
    Thank you. Corrected.
L47: a global wetland CH4 emission -> global wetland CH4 emissions
    Thank you. Corrected.
L48: place -> places
    Yes! Fixed.
L60-61: The very first sentence of the abstract of Kirschke et al. (2013) is: "Methane is an important greenhouse gas, responsible for about 20% of the warming induced by long-lived greenhouse gases since pre-industrial times." Also the IPCC AR5 gives different values: "Emissions of CH4 alone have caused an RF of 0.97 W/m^2", but with a "concentration-based estimate of 0.48 W/m^2" (which I guess would be direct?), compared to 2.83 W/m^2 due to changes in the concentration of all well-mixed greenhouses gases. Not sure where the one-quarter number comes from, but it's not the references given. 
    We are sorry about using an inaccurate number here. Changes to 20% as reported in the referenced papers.
L63: budget -> budgets
    Yes! Fixed.
L72: emissions -> emissions; also maybe change "the key" to "key"? It's not the only uncertainty...
Thank you. Corrected.

L77 (and elsewhere): "to inverse" is not a verb. One could use "inverted fluxes", or, better, "fluxes resulting from atmospheric inversion".
Thank you. Sorry about the misuse of the word. The suggested changes are great. We’ve fixed it here and at line 37.
L89: estimate -> estimates
Yes! Fixed.
L96: ingredients -> the ingredients; steps -> the steps (I would also suggest a different word than "ingredients", perhaps "inputs", or "datasets", or similar? "ingredients" sounds like cooking.
Thanks. We replaced with ‘datasets’ as suggested.
L117: of a priori -> of the a priori
Thank you. Corrected.
L122: emission -> emissions
Thank you.
L123: as L117
Thank you.
L125 (and throughout): Perhaps it would be more appropriate to refer to 42 different wetland model setups or 42 different simulations of wetland methane fluxes, rather than "42 global wetland biogeochemical models". Changing the wetland extent map alone does not really make a new model... Also, the models themselves are not compared to the inversion results, but rather their simulated emissions.
Yes! Fixed.
L136: I'm not sure "year-difference" is completely clear. Perhaps: "the effect of using different time periods"?
Thanks a lot. It sounds much better.
L162-163: I guess just agricultural wetland emissions (rice) or agricultural emissions resulting from intentionally inundated fields are meant? This should be clarified.
Thanks! Corrected in the text: ‘Methane emissions from rice paddies are excluded in both WetCHARTs and GCP simulations.’
L171: factor -> fact
Thank you. Corrected.
L171: I think it would be clearer with: this cross-correlated information -> information cross-correlated
That’s right! Thanks! Fixed.
L176: is -> is the (twice)
Yes! Fixed.
L177: is -> is the
Thank you! Corrected!
L178: emission -> emissions
Thanks! Fixed.
L183: Perhaps should "therefore" be "then"? "Therefore" suggests causality related to what came before, as in, the projected fluxes are compared to the posterior fluxes to rank the models because the inversion-model mismatch is enlarged during the information-rich
period. But that's not really it I think...
You're right. There’s not a causality between two sentences. It’s now corrected as suggested.
L188: use posterior -> uses the posterior
Thanks! Fixed.
L192: is -> is the (twice)
Yes! Fixed.
L193: I don't think you need to include this definition again in parentheses...
Yes! Thanks for the correction.
L196: models -> model
Thanks! Fixed.
L197: emission -> emissions
Yes! Fixed.
L205: from -> from the
Thank you! Corrected!
L208: emission -> emissions; agrees -> agree
Yes! Fixed.
L214: from -> from the
Thanks!
L218: emission -> emissions; is -> are
Thanks! Fixed.
L221: Is it really "much less than", when the ranges overlap so much?
Yes! Thanks for the correction.
L222: emission -> emissions; is -> are; more than -> more than those of
Thanks! Corrected as suggested.
L226: "two-fold from the LP model" -> "a factor of two by the LP ensemble"
Thanks! Fixed.
L226: emission amplitude -> emissions
Thank you! Corrected!
L233: emission -> emissions
Thanks! Corrected as suggested.
L262: as a -> to be the; emission -> emissions
Yes! Thanks for the correction.
L265: emission -> emissions
Thank you! Corrected!
L267: Perhaps "steps" should be "factors"?
Yes! Thanks for the correction.
L273: indicate -> indicate that; emission -> emissions
Thank you! Corrected!
L282: of -> for
Thanks! Corrected as suggested.
L290-291: It's a bit circular to mention good agreement with the inversion results you used to determine which models performed well, isn't it?
Yes. The number from inversion results has been removed.
L297: are -> have been
Yes! Thanks.
L300-301: "we can't assert or refute" Perhaps better: "we can neither assert nor refute"
Yes! Thanks.
L301: would just use "is" (as you can't take a stand on either side)
Yes! Thanks.
L307: and -> and the
Yes! Thanks for the correction.
L311: spatiotemporal need not be capitalized
Yes! Fixed.
L313: remove "the"
Thank you! Corrected!
L323: and -> and the
Thank you! Corrected!
L327: "filters out prior estimate information from retrievals" Here it is key to specify that you mean the prior estimate information from flux retrievals (or estimated fluxes), otherwise it could lead to confusion with satellite retrievals (i.e. the statistically similar inverse problem of converting L1 observed spectra to L2 XCH₄). Also true for the use of "retrievals" in L328.
Thanks! Clarified in the text as suggested: ‘Specifically, our method filters out prior estimate information from flux retrievals and accounts for uncertainties and cross-correlations of estimated fluxes between regions.’
L334: include -> includes
Yes! Fixed.
L338: present -> present for
Yes! Thanks for the correction.
L339: emission -> emissions
Yes! Fixed.
L364: Not all the references appear in alphabetical order?
Yes! Thanks for the correction. The references now follow the alphabetical order.

Supplement:

The caption of S4 does not seem to match the figure. Where are the red- and white-dashed areas? Is it 2020 or 2010? Perhaps they should both say white-dashed areas and this map is only showing the diagonal and off-diagonal elements for the boreal North America region? While having one map might be relevant in terms of defining the regions, if you want to show the correlations between regions I would suggest two colour-coded matrices, one for January and one for July, with the different regions on the x and y axes. This would provide more information.

Sorry for the mistakes in the figure legend. We meant to show the uncertainty of North America emissions and its correlations with other regions. We revised the figure to clarify the numbers and make the figures easier to read by color-coding the cross-correlation numbers (in blue to red gradients):
Figure S4. Global map of diagonal and cross-terms in the averaging kernel matrix ($A$) in Jan and July of 2010 from atmospheric inversion, taking boreal North America region for example. Numbers in the boreal North America region (in grey) are $A$ diagonal values in Jan/July, representing the sensitivity of the inversion results to the satellite observation. The numbers outside of the boreal North America area are the cross-terms in Jan/July and indicate the sensitivity of other regions out of boreal North America that affect the inversion results for boreal North America CH$_4$ emission, the larger numbers indicate larger impact of cross-correlations.

L65: prior -> priori
Yes! Fixed.

Table S3: Why change the description from Table S2 (high/medium/low vs. top 1/3/medium 1/3/bottom 1/3)? These should be harmonized.
Yes! Thanks. The descriptions in the two tables are now consistent.

Reviewer #2 Evaluations:
Recommendation: Return to author for major revisions
Significant: There are major errors or gaps in the paper but it could still become significant with major changes, revisions, and/or additional data.
Supported: Mostly yes, but some further information and/or data are needed.
Referencing: Yes
Quality: The organization of the manuscript and presentation of the data and results need some improvement.
Data: Yes
Accurate Key Points: Yes

Reviewer #2 (Formal Review for Authors (shown to authors)):

General comments

This manuscript presents an interesting method for comparing independent flux estimates from e.g. land surface models to fluxes retrieved from atmospheric inversions. This
method rigorously takes into account the dependence of the fluxes from the inversion on
the prior information used in the inversion through the use of the averaging kernel matrix.
This method could be broadly applicable to other types of fluxes and to other species.
One limitation, however, is that the averaging kernel is not calculable in a lot of
inversions (e.g. in those using 4D-var algorithms), which is more often the case for
inversions using satellite data and at the global scale.

Thanks!
We agree with the reviewer that this method is only applicable to atmospheric CH₄
inversions with explicit prior and posterior covariances. We now introduced the
following text in the first paragraph of section 4:

“We note that the methodology presented in this study can only be applicable to
atmospheric CH₄ inversions with explicit prior and posterior covariances (Turner et al.,
2015). The methodology can also be approximated with other inversion methods (e.g.
4D-Var methods) if these are able to generate posterior covariance estimates (Liu and
Bowman, 2016)”

Although the method is interesting, I think the authors overstate its potential. For
example, it is not clear how this method could be used for validating the predicted fluxes
from Earth System Models, which is one claim made. Also, there is no mention of
systematic errors in the inversion and how these may influence the results.
1) Response to ‘validating the predicted fluxes from ESMs’:

As the editor and reviewers all pointed out: Our previous results emphasize the important
of wetland extent map in determining model performance, which is not a new finding.
The high-performance and low-performance models do not have significant differences
in correlation of CH₄ with climate/wetland extent. We agree that based on these
conclusions, it is not clear how atmospheric estimates can be used to analyze ESM
projections.

However, inspired by the reviewers’ comments, we dug into the data and found
more interesting results, as shown in the new figure 4, and explained below:
Figure 4. Marginal RMSE performance (see section 2.4) of the factorial WetCHARTs model ensemble, with respect to individual temperature sensitivity of CH$_4$C emissions (Q10 CH$_4$C), global wetland emission scaling factors, and wetland extent datasets. The black dots represent medians, boxes represent the 25th-75th percentile ranges, and lines represent 5th-95th percentile ranges.

We find that 1) models using a Q$_{10}$ equals 1 have significantly higher performance than the ones using larger Q$_{10}$ values (top row in Figure 4); 2) models with the smaller global scalar perform better than those with larger scalars (bottom row in Figure 4); 3) models use the GLOBCOVER derived wetland extent maps outperform those use the GLWD map.

These imply that smaller seasonal fluctuations in response to temperature, a smaller global total emission, and the satellite-based wetland extent agrees better with atmospheric CH$_4$ measurements, showing emergent constraints on both model parameters and drivers data.

We also find that regardless of wetland maps, LPX-Bern and JULES always stay in the high-performance group, implying more realistic model structures or better choice of parameter values, which awaits to be further investigated.

Altogether we justify our conclusions on reducing uncertainties in the terrestrial biosphere model simulations by constraining model structures, parameters and drivers data.

In terms of placing emergent constraints on ESM predictions, we added a discussion in section 3.4 to better illustrate our idea, and better inform readers of the existing issues in wetland CH$_4$ predictions:

The large uncertainties in predicting wetland CH$_4$ emission come from 1) climate and wetland extent predictions (drivers) (Andresen et al., 2020; Zhang et al., 2017); 2) process controls (model structures); and 3) climatic sensitivities of CH$_4$ productions (parameter values) (Poulter et al., 2017; Riley et al., 2011). Our approach enables the
evaluation of terrestrial biosphere models with satellite-informed estimates and places constraints for land model predictions by telling which wetland extent maps, biogeochemical processes, and parameters values used in the models are more realistic, as discussed in section 3.2 (Cox et al., 2013; Wenzel et al., 2016).

2) Response to systematic errors in the inversion and how these may influence the results.

We added a discussion on the caveats at the end of the introduction:

‘While our estimates account for the spatial resolution and error associated with the inversion of observations to fluxes, we do not explicitly account for error in model transport and chemistry, such as interhemispheric mixing time, vertical mixing, tropopause height, and stratospheric gradient. These errors are important and can be accounted if the corresponding posterior covariance is provided. We thus advocate studies that either characterize or mitigate these errors to offer better constraints for bottom-up estimates.’

Another concern is that the authors are sometimes not very precise in what they state, both in the methodology and in the results (see specific comments).
Thank you. We deeply appreciate your detailed comments and suggestions to improve this manuscript. We’ve addressed all the comments as listed below.

Lastly, I am concerned about the particular inversion that was chosen for this study as it uses a linear method to solve a non-linear problem (see specific comments). I think the impact of this on the fluxes from the inversion needs to be quantified, and thus the impact on the results of this study.

We agree with the reviewer that it is important to point this issue out in the text. As the reviewer said, the problem is non-linear because OH is optimized alongside with emissions. We justify that quasi-linearity can be assumed in this case, because prior methane concentrations are sufficiently close to the observations (relative difference <5%) so that the non-linear effect is small. Moreover, optimization of methane emissions is mainly informed by regional-scale patterns in GOSAT observations. Therefore, the inferred seasonality of wetland emissions should be insensitive to the non-linear effect of OH optimization, which, if any, is manifested mainly at the global scale. Below we present two analyses that support our claim. We add a supplementary section (Text S2) to include this discussion.

Given these concerns, I recommend major revisions.

Specific comments

P3L37: Please change "inversed from atmospheric CH4 concentration" to "from atmospheric inversions of CH4 concentrations" since it is not the concentrations that are "inversed" but rather the relationship between the fluxes and atmospheric concentrations. Thank you for pointing this out. We’ve corrected this sentence based on the comments from Reviewers #1 and #2.
P3L39-40: I do not agree that "the challenge of using top-down estimates to refine bottom-up models" is "mainly because of variable a priori emissions and corresponding variable spatial resolution" as the authors say. I think the challenges of the uncertainties in the satellite retrievals and the model representation errors, which translate into flux errors from the inversions are equally relevant.

Yes! We strongly agree with the reviewer’s statement. We rephrased the sentence here and elsewhere: ‘Despite the extensive coverage of satellite CH4 concentration observations, challenges remain with using top-down estimates to test and refine bottom-up models, mainly because of the uncertainties in the satellite retrievals, the model representation errors, variable a priori emissions, and spatial resolution.’

P3L44-44: By accounting for cross-correlations and uncertainties in the inversion estimates does not mean that only the information from the satellite observations is kept as a constraint. The fact that an atmospheric inversion is used means that the information provided is about the fluxes (that is having solved for atmospheric transport) so it is not "only" information from the satellite observations.

That’s right! We rephrased it as ‘Our comparison between bottom-up models and satellite-based fluxes explicitly accounts for cross-correlations typically found in top-down inverse estimates, such that only the information from satellite observations and the atmospheric transport model is kept as a constraint.’

P4L68: Please specify what is meant by "future climate feedback", isn't the climate feedback simply a result of the sensitivity of the model to meteorological forcing etc. or the "biogeochemical processes" as the authors write?

Sorry for the confusion. We corrected it as ‘The largest uncertainties in bottom-up modeling approaches stem from wetland inundation extent (driver data), biogeochemical process description (model structure), and responses to climatic variability (parameter uncertainty)’

P4L69: Please specify what is meant by "sampling" in this context.

Thank you! We specify in the text: ‘For top-down inversions, the uncertainties are due to concentration data retrieval, atmospheric chemistry, model errors related to atmospheric transport, and bottom-up information in the distribution of emissions’

P4L70: Please also add "atmospheric chemistry" to this list of top-down sources of error as the OH sink is a large source of uncertainty in the atmospheric chemistry transport models.

Thank you! It’s there! ‘For top-down inversions, the uncertainties are due to concentration data retrieval, atmospheric chemistry, model errors related to atmospheric transport, and bottom-up information in the distribution of emissions’

P4L74: Please change to: "Atmospheric inversions using space-borne observations provide a constraint on the global scale spatial-temporal variability of wetland CH4 emissions"
Thank you! It reads much better!
P5L79: Remove "explicitly" since the inversions do not estimate this at all.
Right! Thanks. Removed!
P5L81-81: Please change "limitations in the observing system" to "limited observational constraints"
Thank you! Corrected as suggested.
P5L82: This sentence is ambiguous, on one hand when the observational constraint is relatively weak, the inversion solution depends strongly on the prior estimates - I think this is what is meant here. When this is the case, it follows that "the interpretation of emissions processes" (based on fluxes from inversions) will also depend on the prior estimates.
Thank you. The logic is clearer following your suggestions:

‘When the observational constraint is relatively weak, the inversion solution depends strongly on the prior estimates. Much of the interpretation of emissions processes then depends upon the a priori, or prior estimates that helps constrain the solution.’
P6L102: Remove "emission" before "fluxes" since the inversion estimates "fluxes" and "emission" is redundant here. And again on L105 change "emissions" to "fluxes".
Thanks! Fixed!
L107: I think it's important to state that the inversion of Zhang et al. 2020 estimates CH4 emissions from wetland and non-wetland emissions separately. I also read Zhang et al. and I'm a bit concerned about this inversion since they attempt to optimize OH as well, which makes the problem non-linear (since the amount of CH4 lost due to the OH reaction depends on the CH4 concentration) but their optimization method assumes linearity. I'm surprised the reviewers of Zhang et al. didn't comment on this. I see that many of the same authors on Zhang et al. are also authors on this manuscript. The question is, what is the impact of using a method for the inversion that assumes linearity when the problem you are trying to solve is clearly non-linear? I think the error in this assumption needs to be quantified, and if it is large then another inversion result should be used in this study.

1) Response to 'state that the inversion of Zhang et al. 2020 estimates CH4 emissions from wetland and non-wetland emissions separately'

We agree we should state clearly that Zhang et al 2021 inversion method optimizes wetland & non-wetland separately. We added in section 2.1 (Methods):

The posterior wetland CH4 emission in our analysis is from a global flux inversion using GOSAT observations from 2010-2018 (Zhang et al., 2021). In addition to wetland emissions (resolved at 14 subcontinental regions and for individual months), the inversion by Zhang et al. (2021) also simultaneously optimizes anthropogenic CH4 emissions (2010-2018 mean and trend on a 4o×5 o grid) and tropospheric OH concentration (the main sink of CH4, hemispheric averages estimated annually) (Zhang et al., 2021).

2) Response to ‘using linear method to resolve non-liner problems’

We agree with the reviewer this is important to point out in the text. As the reviewer said, the problem is non-linear because OH is optimized alongside with emissions. We justify that quasi-linearity can be assumed in this case, because prior methane concentrations are
sufficiently close to the observations (relative difference <5%) so that the non-linear effect is small. Moreover, optimization of methane emissions is mainly informed by regional-scale patterns in GOSAT observations. Therefore, the inferred seasonality of wetland emissions should be insensitive to the non-linear effect of OH optimization, which, if any, is manifested mainly at the global scale. Below we present two analyses that support our claim. We add a supplementary section (Test S2) to include this discussion.

P7L117: The dependence of the inverse solution on the a priori information can be expressed as a weighting between the "true" fluxes and the prior fluxes: $x_{\text{pos}} = Ax_{\text{true}} + (I - A)x_{\text{pri}}$ where A is the averaging kernel and $x_{\text{true}}$ are the "true" fluxes if these would be known. Thus the impact of the prior is not just "spatial uncertainties and cross-correlations" but rather all components that determine A, i.e. the observation operator, the prior and observation error covariance matrices, as well as the number of observations. Thank you! We revised as:

"These inversion diagnostics allow for an explicit representation of the impact of the prior estimates (including spatial uncertainties and cross-correlations) for the inversion-based estimates of wetland CH4 emissions (Rodgers, 2000)."

P8L149: I am not sure if it is generally true that all GCP models are land components of ESMs. It would be more correct to say that these are all land ecosystem models. Thanks! Changed!

P9L180: The approach here does not remove the impact of the prior information but allows the modelled estimate ($x_{\text{model}}$) to be "viewed" in a consistent way with the inversion accounting for the prior information. This would be the more correct way to state this.
Thank you! Corrected! This is indeed a more agreeable description with our approach. Fig. S1: It would be helpful to state what model is used here Thank you. We added the model information in the figure legend: ‘GCP v1 CLM4.5 modeled monthly CH4 emission’

L194-196: I think the list of biogeochemical models with the wRMSE should be given somewhere as it's relevant to see if there are similarities between the models that ranked as HP/LP.
Thank you. We had it listed in Table S2. A reference to this table here would be helpful as the reviewer suggested. So we switched the order of Table S1 and Table S2 since the wRMSE table is now reported ahead.

L201: I'm not exactly sure what is meant by this sentence. Do the authors mean that for a regional estimate they calculated a weighted average of each variable where they weighted by the amount of CH4 emission? This should be stated more clearly.
Yes. This is what we intended to say. We corrected this statement as suggested:

‘For a regional estimate we calculate a weighted average of each variable depending on the amount of CH4 emission.’

L284: Please add "emission" after CH4 in this sentence to be clear.
Thanks! Added.

L291: I think the reference Yuzhong et al. 2021 should be Zhang et al. 2020.
Yes! Thanks! We deleted this reference in this revision in response to reviewer#1’s
suggestion, that we don’t need to compare our results to the same data we use to evaluate models.
L294: Actually both HP and LP have equal sensitivity to precipitation and wetland extent. Also, surely the result that wetland extent strongly determines wetland emissions is trivial.
Yes! Thanks for pointing this out. We revised to state that ‘both HP and LP’ modeled CH₄ fluxes have high correlation with precipitation and wetland extent.
P16L302: Please specify that this is for tropical wetlands. In contrast to the tropics, the results show that non-tropical (and global) wetland emissions are strongly dependent on temperature.
Yes! Thanks for pointing this out. ‘we can neither assert nor refute the possibility that on decadal-centennial timescales temperature is the primary control on tropical wetland emissions’
Fig. 4. It is a bit difficult to distinguish which bars belong to which variable and it would help to either draw horizontal lines between the bars for the different variables or to increase the gap between them.
Thank you for the suggestion. We modified the figure 4 to easily distinguish bars to the variables.
P16L303: I think the claim that this analysis can be used as a constraint on ESM CH₄ predictions needs to be better justified. Using inversion flux estimates to validate biogeochemical results is sound, but in order to validate ESM predicted CH₄ emissions one first needs to validate the climate predictions. Also, in a future climate, it is unclear how changes in vegetation cover will affect the emissions, and in high latitudes predictions of hydrology are highly uncertain with some models predicting an increase in wetland extent and others a decrease (e.g. Andresen et al., The Cryosphere, 2019).
Thank you for the suggestions. Our reasoning here is ambiguous. We rewrite this paragraph to better illustrate our idea, and better inform readers of the existing issues in wetland CH₄ predictions:
‘The large uncertainties in predicting wetland CH₄ emission come from 1) climate and wetland extent predictions (drivers) (Andresen et al., 2020; Zhang et al., 2017); 2) process controls (model structures); and 3) climatic sensitivities of CH₄ productions (parameter values) (Poulter et al., 2017; Riley et al., 2011). Our approach enables the evaluation of terrestrial biosphere models with satellite-informed estimates and places constraints for land model predictions by telling which wetland extent maps, biogeochemical processes, and parameters values used in the models are more realistic, as discussed in section 3.2 (Cox et al., 2013; Wenzel et al., 2016).’
P17L332-P18L347: It is not clear to me how this section "Future Directions" links to the analyses and method described in this paper.
Thanks! This section talks both about remaining uncertainties and future directions. So we revised the title of this section as ‘Remaining uncertainties and future directions’ and moved this section up after the implications section, as suggested by the reviewer and the editor.

Technical corrections
This article addresses the emissions of CH₄ from wetlands. The study uses satellite-
derived emissions of wetland CH$_4$ to analyze a selection of process-based models and isolate those that provide the best fit to these satellite-derived emissions. This selection of refined models are analyzed to examine the correlation of wetland CH$_4$ emissions to the main underlying drivers of wetland extent, temperature and carbon availability. The manuscript is generally well-written and to the point. However, I have some questions regarding what the implications and novelty of this work are, since the satellite constraint on latitudinal distribution is already provided in the work of Zhang et al (2021), and the analysis of the sensitivity of CH$_4$ to the underlying physical drivers adds little new in its current format.

Thank you so much for these insights. In response, we listed our replies to your comments below.

Major comments:
This study is largely based on the work of Zhang et al (2021, hereafter Z21) which estimated global wetland CH$_4$ emissions to be 145 Tg/yr between 2010 and 2018. The selection of high performing process-based models (HP) are then calibrated to fit the results of Z21. The finding that the median emissions from the HP models are 148 Tg/yr is therefore hardly a surprise, demonstrating a calibration to the results of Z21. The regional emissions distribution of HP models shown in Fig. S2 is also seen to be very similar to that of Z21 (Figure 11). Again, given the HP models are selected based on their fit to Z21, this is perhaps to be expected. Therefore, it seems an overstatement to say in the abstract (also the plain language summary and key points) that the study places new constraint on the latitudinal distribution of wetland emissions. Surely this constraint is provided in Z21, and the HP models selected here are merely those that demonstrate the greatest consistency with that constraint?

Thanks for the comment. We agree that we are missing certain context in this statement, we revised it as:

‘Our study places new constraint on the latitudinal distribution of wetland emissions for process-based biogeochemical models.’

Top-down inversion results can be difficult to compare directly to bottom-up models because of its inherent uncertainty from model transport and chemistry, a coarser spatial resolution and absence of process-level emission controls. The main contribution of our innovative Bayesian methodology is that our analysis resolves the uncertainties from the prior estimates so that the atmospheric methane estimates can be used for evaluating land model simulations. We added a Table S4 to show the huge impact on the ranking of terrestrial biosphere model simulations before and after accounting for cross-correlations and special uncertainties.

The novelty of our study is using the averaging kernel matrix for rigorous analysis of the dependence of the fluxes from the inversion on the prior information used in the inversion.

The novelty of the work seems to boil down to an analysis of the underlying controls on what determines whether a model is in the HP or LP group. Currently this is limited to an analysis of correlation coefficients between CH$_4$ emissions and the underlying controls of
wetland extent/precipitation, temperature and carbon availability. This analysis shows no significant difference between HP and LP models. The major factor in what determines an HP vs LP model would seem to be the total global CH4 flux magnitude, which largely seems to be a result of the wetland extent map used. This is a useful result, but beyond this any further discussion is lacking. For example, to the uninitiated, what is the difference between the GLWD and non-GLWD extent maps? The underlying fluxes per unit area are presumably the same between the same models in GCPv1 and GCPv2. So what has really been learnt, other than the major control on the fit to the results of Z21 being the wetland extent map used?

Thank you for pointing this out. We completely agree!

We now added a description to show the differences between the GLWD and non-GLWD extent maps in the manuscript section 2.2:

‘GLOBCOVER is a global land-cover map using the fine resolution (300m) mode data from MERIS sensor on-board ENVISAT satellite. GLWD is an inventory-based product that synthesize a number of best available sources for lakes and wetlands on a global scale.’

We agree with the reviewer that our previous conclusions need to be extended. We also did more analysis with the data and will reply to the rest of this comment along with the next one.

The title and abstract indicate that the work provides constraint on the climate sensitivity of wetlands. I found this term a little confusing given its normal usage in climate sciences of the change in temperature given a change in radiative forcing. Perhaps, "sensitivity to climate forcings" might be a more appropriate term. Regardless, I did not really see where in the article this quantification was provided, e.g. the calculated change in CH4 emissions given a change in temperature or precipitation. The correlation analysis shows there is a relationship between all models (not just HP) and these climate drivers, but little on the form of that relationship. It seems to be a leap to claim that the work provides new constraint on this response of CH4 emissions to the underlying drivers. Narrowing down models into HP and LP seems to add little in this regard, precisely because the main difference appears to be the wetland extent map used, rather than the underlying process model dynamics.

Thanks for the comments! We agree that according to our results the high-performance and low-performance models do not have significant differences in correlation of CH4 with climate/wetland extent.

Inspired by the reviewers’ comments, we dug into the data and found more interesting results, as shown in the new figure 4, and explained below:
Figure 4. Marginal RMSE performance (see section 2.4) of the factorial WetCHARTs model ensemble, with respect to individual temperature sensitivity of CH4:C emissions (Q10 CH4:C), global wetland emission scaling factors, and wetland extent datasets. The black dots represent medians, boxes represent the 25th- 75th percentile ranges, and lines represent 5th- 95th percentile ranges.

We find that 1) models using a Q10 equals 1 have significantly higher performance than the ones using larger Q10 values (top row in Figure 4); 2) models with the smaller global scalar perform better than those with larger scalars (bottom row in Figure 4); 3) models use the GLOBCOVER derived wetland extent maps outperform those use the GLWD map.

These imply that smaller seasonal fluctuations in response to temperature, a smaller global total emission, and the satellite-based wetland extent agrees better with atmospheric CH4 measurements, showing emergent constraints on both model parameters and drivers data.

We also find regardless of wetland maps, LPX-Bern and JULES always stay in the high-performance group, implying more realistic model structures or better choice of parameter values, which awaits to be further investigated.

Altogether we justify our conclusions on reducing uncertainties in the terrestrial biosphere model simulations by constraining model structures, parameters and drivers.

We also added discussion on how our results places emergent constraints on ESM predictions in section 3.4:

‘The large uncertainties in predicting wetland CH4 emission come from 1) climate and wetland extent predictions (drivers) (Andresen et al., 2020; Zhang et al., 2017); 2) process controls (model structures); and 3) climatic sensitivities of CH4 productions (parameter values) (Poulter et al., 2017; Riley et al., 2011). Our approach enables the evaluation of terrestrial biosphere models with satellite-informed estimates and places
constraints for land model predictions by telling which wetland extent maps, biogeochemical processes, and parameters values used in the models are more realistic, as discussed in section 3.2 (Cox et al., 2013; Wenzel et al., 2016).

Specific comments:
Figure 1: This figure is not very self-explanatory i.e in what way is WetCHARTs an ingredient for evaluating biogeochemical models? How is this different from the 42 biogeochemical models themselves? Either some more info in the workflow is needed or the caption needs expanding.

Thanks for the comments! We updated the figure 1 and the figure legend to make it easier to read:

Figure 1. The workflow for constraining biogeochemical models with top-down (atmospheric inversion) CH₄ flux estimates. Datasets for our Bayesian methodology include 1) the prior estimates (WetCHARTs CH₄ emissions); 2) the posterior estimates and inversion diagnostics that allow for an explicit correction for the impact of the prior data; and 3) 42 biogeochemical land model simulations.

P6, L111: To what extent are the wetland emission estimates of Z21 negatively correlated with the separate anthropogenic estimates? Figure 6 of Z21 suggests this might be an issue in certain regions of the world.

Thanks for pointing this out! The posterior covariance for the wetlands automatically includes the cross-state error from the other terms and therefore they are accounted for in the error analysis (Worden, 2004). We added the above clarifications in our method (second paragraph in section 2.1). In addition, Fig. 6 of Zhang et al. (2021) shows that the
error correlation between wetland & anthropogenic emissions are overall small on the regional scales.

We should state clearly that Zhang et al 2021 inversion method optimizes wetland & non-wetland separately. So we added in section 2.1 (Methods):

The posterior wetland CH₄ emission in our analysis is from a global flux inversion using GOSAT observations from 2010-2018 (Zhang et al., 2021). In addition to wetland emissions (resolved at 14 subcontinental regions and for individual months), the inversion by Zhang et al. (2021) also simultaneously optimizes anthropogenic CH₄ emissions (2010-2018 mean and trend on a 4°×5° grid) and tropospheric OH concentration (the main sink of CH₄, hemispheric averages estimated annually) (Zhang et al., 2021).

P7, L131: "Both GCP version are chosen because..." The meaning of this line was little unclear. Do you mean you use two different GCP versions because they use different wetland extent maps?

Yes! Thanks! We revised to be clearer:
‘Both GCP versions are chosen because they are driven by different wetland extent maps with different spatiotemporal variabilities.’

P9, L175 Eq1: What is $\hat{x}_{\text{model}}$?

Thanks. We moved the definition of $\hat{x}_{\text{model}}$ up in the first line of the paragraph:
‘where $\hat{x}_{\text{model}}$ is the projected model emissions,...’

L194: Are the models evaluated globally with respect to the w-RMSE metric or independently on a region by region basis?

Thanks! The models are evaluated based on their global performance. We add in the text:
‘The wRMSE metric represent the overall model performance across the globe (14 regions by 36 monthly emissions).’

L203 Results: To be clear when analyzing the models are you analyzing the raw model output, or the observation operator transformed model output? I assume it is the former given the generally low averaging kernel values presented in the supplement. Unless N. Boreal America is a particularly poorly constrained region? The point being that if A is close to zero then the results will just be largely representative of the prior. It might be worth creating a table of mean averaging kernels for each region to show the constraint that the satellite data actually have on each region.

Thanks for the advice. We added Figure S5 to show the time series of averaging kernels for each region.
Figure S5. Diagonals of averaging kernels for each region.

L. 206: 5th-95th percentiles? Do these ranges refer to the percentiles of monthly emission estimates or the 5th and 95th percentiles of the 14 model medians (if such a thing exists)?

Thanks for asking. The 5th-95th percentiles are calculated from 42 data points: 3 annual emissions (2010-2012) x number of high/low performance models (14 HP and 14 LP).

L218: Saunois et al (2020) suggest that tropical wetlands are one of the more highly uncertain aspects of the wetland CH4 budget. Yet these results suggest there is very little variation in tropical emissions between HP and LP models. Is that right, or is the model range large and it is just coincidence that the medians are similar between HP and LP?

Yes! As the reviewer said the mean tropical emissions of HP and LP are similar, but 1) the uncertainty range is quite different, and 2) the subregional distributions are different (lower emission from Amazon in HP models).

L220-221: Why are the Amazon emissions smaller in the HP models compared to LP?

Good question! Although our model performance ranking is based on the global average (the top performance model does not necessarily outperform others in any of the 14 regions), we do see a relatively strong observational signal from the averaging kernel, indicating good amount of information from satellite observations. A lower Amazon emissions in the HP models may be closer to the atmospheric methane estimates.

L230: What do you mean by "match the annual CH4 emissions in the Alaska region"? Do you mean consistent within the uncertainties?

Thanks. Yes! We clarified in the text:

'Our spaceborne (GOSAT) data-constrained HP models is consistent with the annual CH4 emission in the Alaska region (1.66±0.95 Tg CH4 May-Oct, 2012) as well as month-to-month seasonality with independent estimates from CARVE (aircraft-based inversion of...
CH$_4$ emissions in Alaska, $1.80\pm0.45$ Tg May-Oct 2012.

L230: Are the CARVE emissions representative of exclusively wetland sources or do they include influences from non-wetland sources?

CARVE emissions represent wetland source only.

Table 1: Wouldn't it be worth including the Z21 posterior fluxes as well since this is the quantity you are calibrating your model selection on, rather than the prior?

We appreciate raising this question: we thought about it in the beginning but we then think adding another column could be distracting since the posterior fluxes is disturbed with errors from a priori (cross-correlations and spatial uncertainties). We actually validated model performance by excluding these errors from the posterior fluxes.

P14. L255: Why are LPX-Bern and JULES the best fitting models? What is it about the model parameterizations of these models compared to other models? In addition, I assume this is a global comparison, but do these models perform best for the majority of regions or is the best-performing model very region dependent?

LPX-Bern and JULES always stay in the high-performance group, likely due to more realistic model structures or better choice of parameter values. This is an interesting topic and awaits to be further investigated, since the parameter values and model structures of GCP simulations are not available from the data set we acquired.

We find that HP models usually do better in multiple regions (if not all) than the LP models, e.g. the figure below shows the projected LPX-Bern has better agreement with posterior fluxes especially in the South Hemisphere South America, temperate North America, North Hemisphere Africa, South Hemisphere Africa, temperate east Asia, and boreal North America.

![Figure 4](image-url)

Figure for responses. Anomaly of posterior CH$_4$ emission against projected GCP v1 LPX-Bern (HP) and GCP v1 TRIPLEX-GHG (LP) simulations.

Figure 4: Can you swap the order of HP and LP in the key to be the same as in the rest of the figure? It is just a bit confusing otherwise.
Thanks! We swapped the order of HP and LP in the key, and also added lines to help identify the groups. It’s now easier to read.

P. 16, L299: "Our analysis reveals how wetlands across the globe respond to variations in temperature, precipitation, inundation, and carbon availability." Does it? I cannot see any new quantification of how wetland CH₄ responds to changes in these variations. Would it not be more accurate to say the analysis narrows the range of a given set of wetland models that fit with a satellite-derived estimate of CH₄ emissions. But the response to variations in climate variables are qualitatively the same regardless of whether the model is in the HP or LP group. The main difference seems to be the use of a GLWD vs non-GLWD wetland extent map. So it is unclear whether narrowing the models down has really added anything here.

Thanks for the question. We agree!

Our previous results emphasize the importance of wetland extent maps in determining model performance, which is not a new finding. The high-performance and low-performance models do not have significant differences in correlation of CH₄ with climate/wetland extent.

This statement still holds despite no obvious differences found between HP and LP models. We revised in our manuscript that ‘from both HP and LP models we find high correlation of CH₄ emission with precipitation and wetland extent in the tropics and high correlation of CH₄ emission with temperature and carbon availability in the high-latitudes.’

Reference:

Andresen, C. G. et al. (2020) ‘Soil moisture and hydrology projections of the permafrost region—a model intercomparison’, Cryosphere, 14(2), pp. 445–459. doi: 10.5194/tc-14-445-2020.

Cox, P. M. et al. (2013) ‘Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability’, Nature. Nature Publishing Group, 494(7437), pp. 341–344. doi: 10.1038/nature11882.

Liu, J. and Bowman, K. (2016) ‘A method for independent validation of surface fluxes from atmospheric inversion: Application to CO₂’, Geophysical Research Letters, 43(7), pp. 3502–3508. doi: 10.1002/2016GL067828.

Poulter, B. et al. (2017) ‘Global wetland contribution to 2000–2012 atmospheric methane growth rate dynamics’, Environmental Research Letters, p. 094013. doi: 10.1088/1748-9326/aa8391.

Riley, W. J. et al. (2011) ‘Barriers to predicting changes in global terrestrial methane fluxes: analyses using CLM4Me, a methane biogeochemistry model integrated in CESM’, Biogeosciences, 8(7), pp. 1925–1953.

Rodgers, C. D. (2000) Inverse Methods for Atmospheric Sounding, Series on Atmospheric, Oceanic and Planetary Physics. WORLD SCIENTIFIC. doi:
doi:10.1142/3171.

Turner, A. J. et al. (2015) ‘Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data’, Atmospheric Chemistry and Physics, 15(12), pp. 7049–7069. doi: 10.5194/acp-15-7049-2015.

Wenzel, S. et al. (2016) ‘Projected land photosynthesis constrained by changes in the seasonal cycle of atmospheric CO2’, Nature. Nature Publishing Group, 538(7626), pp. 499–501. doi: 10.1038/nature19772.

Worden, J. (2004) ‘Predicted errors of tropospheric emission spectrometer nadir retrievals from spectral window selection’, Journal of Geophysical Research, 109(D9), pp. 1–12. doi: 10.1029/2004jd004522.

Zhang, Z. et al. (2017) ‘Emerging role of wetland methane emissions in driving 21st century climate change’, Proceedings of the National Academy of Sciences of the United States of America, 114(36), pp. 9647–9652. doi: 10.1073/pnas.1618765114.

Authors’ Response to Peer Review Comments on Second Revision of Manuscript (2021AV000408RR)

Dear Professor Davidson,

Thank you and the three reviewers for reviewing our manuscript “Satellite constraints on the latitudinal distribution and temperature sensitivity of wetland methane emissions”. We are very appreciative of the time and effort that you and the reviewers have spent in improving our manuscript.

We have carefully gone through reviewer’s remaining comments and have responded to their concerns and suggestions, which we believe have greatly improved our paper. Please find our detailed responses below in blue fonts.

Please let us know if you need any further information or clarifications.

Yours truly,

Shuang Ma, Ph.D.
Postdoctoral Scholar
Jet Propulsion Laboratory at Caltech
Telephone: +1 8183542113
Email: shuang.ma@jpl.nasa.gov

Dear Dr. Ma:

Thank you for submitting your manuscript entitled "Satellite constraints on the latitudinal
distribution and temperature sensitivity of wetland methane emissions" [Paper #2021AV000408RR] to AGU Advances. I have now received 1 reviews of your manuscript, which are included below and/or attached.

The authors have adequately responded to most of the reviewers' comments. However, one reviewer has a several additional comments and questions that should be considered. None seems like a show stopper.

Based on the review comments, your manuscript may be suitable for publication after minor revisions.

The feedback provided in the reviewer assessments of your manuscript is important and should be taken into account as you complete your revision. I encourage you to submit a suitably revised version of your manuscript by July 6, 2021.

Upon submission, we will need to receive the following:

1. A response to reviewer file that lists each of the comments and describes how the manuscript has/has not been modified in response to those comments.

2. A copy of the manuscript with the changes noted (e.g., highlighted, "track changes," italics or bold changes). Please upload the article with tracked/highlighted changes as a response to reviewer file.

3. A copy of the revised manuscript with the changes incorporated which will be used for publication if the manuscript is accepted.

4. In addition to addressing the remaining important technical issues raised by reviewers, please also ensure that AGU data policy is addressed in the Acknowledgements section and that the key points report what is learned from the study.

5. All files in publication-ready formats.
   ***Publication-ready formats for article files are limited to Word and LaTeX (Excel is also acceptable for tables only). Figure files must be individually uploaded as .eps, .tif, .jpg, or .pdf files and all parts of the same figure need to be combined in one file.

6. AGU has officially joined with many other publishers in a [commitment](https://eos.org/agu-news/agu-opens-its-journals-to-author-identifiers) to include the ORCID (Open Researcher and Contributor ID) for authors of all papers published starting in 2016. Funding agencies are also asking for ORCID's.

   Including the ORCID as part of published author information in papers will better enable linking of content and accurate discovery across individuals, similar to the way DOIs have enabled reference linking across journals. Given a specific scientist's permission, AGU can also add published papers to his or her ORCID record. See our statement [https://eos.org/agu-news/ag...](https://eos.org/agu-news/agu-opens-its-journals-to-author-identifiers). We can also provide credit to you through ORCID when you serve as a reviewer.
If you have not already created an ORCID or linked it to your GEMS record, please do so as soon as possible. This will need to be completed for us to accept your paper. You can both create and link an ORCID from your GEMS record.

***New supporting information guidelines***
AGU now requires that supporting information be included in one file, except where limited by file type or size. Please see Supporting Information Guidelines in Author Resources (https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Supporting-Info-Requirements).

When you are ready to submit your revision, please login to your account (https://advances-submit.agu.org/cgi-bin/main.plex), and click "Revise 2021AV000408RR."

I look forward to receiving your revised manuscript. If you have any questions, please contact the editor's assistant at advances@agu.org.

Sincerely,

Eric A. Davidson
Editor
AGU Advances

-----------IMPORTANT INFORMATION------------------------
Additional information on text preparation, formatting, acceptable file formats, supporting information, graphics preparation, and AGU style, is here.

Sharing your work is an important part of the research process, and AGU leverages and shares published research to promote the broader importance of Earth and space science. Learn how you can promote your paper, including how your paper can be considered for additional publicity or for the issue cover if it is accepted.

-------------------------------------------------------------------
Reviewer #2 Evaluations:
Recommendation: Return to author for minor revisions
Significant: The paper has some unclear or incomplete reasoning but will likely be a significant contribution with revision and clarification.
Supported: Yes
Referencing: Yes
Quality: Yes, it is well-written, logically organized, and the figures and tables are appropriate.
Data: Yes
Accurate Key Points: Yes

Reviewer #2 (Formal Review for Authors (shown to authors)): 
General comments

The authors have gone to some length to address the concerns raised in the first review and this version of the manuscript is much improved. However, I have a few remaining, mostly minor, concerns.

Text discussing the caveats of the method, which was said to have been added at the end of the introduction appears to be missing.

We apologize for this mistake - thanks for pointing out! Instead of placing this discussion in the introduction, we put it in section 4.0 Remaining uncertainties, future directions, and summary:

‘While our estimates account for the spatial resolution and error associated with the inversion of observations to fluxes, we do not explicitly account for errors in model transport and chemistry, such as interhemispheric mixing time, vertical mixing, tropopause height, and stratospheric gradient. These errors are important and can be accounted for if the corresponding posterior covariance is provided, possibly through multi-model comparisons (e.g., Schuh et al. 2010) or by integrating measurements with different sensitivities to the profile of CH₄ (e.g., Jiang et al. 2013, 2017). We thus advocate studies that either characterize or mitigate these errors to offer better constraints for bottom-up estimates. We note that the methodology presented in this study can only apply to atmospheric CH₄ inversions that calculate prior and posterior covariances (Turner et al., 2015). The methodology can also be approximated with other inversion methods (e.g., 4D-Var methods) if these can generate posterior covariance estimates (Liu et al., 2014; Chevallier et al., 2010; Chevallier et al., 2007; Yi et al., 2020)”

About the general applicability of the method to inversions using e.g. 4D-Var methods - the authors write that the method can be used for any inversion for which the prior and posterior error covariance matrices are available. However, the method requires the Averaging Kernel, which depends on the posterior and prior error covariance matrices (or an alternative formulation using the prior and observation error covariance matrices) but also the transport matrix. Could the authors please explain if they have derived another formulation of the Averaging Kernel which does not require the transport matrix? If not, then please state that the availability of the transport matrix is a requirement of this method, as is the prior and posterior error covariance matrices.

Thanks for this suggestion!

We realized that we did not show how we calculated the averaging kernel (\( A \)) in our manuscript, and that there are other equations for \( A \) that require a transfer matrix (e.g., Yi et al., 2020). Thus, we added to section 2.3: ‘The calculation of \( A \) requires the prior and posterior covariance matrix, as described by Zhang et al. (2021) and Rodgers (2000).’

We explain in more detail here in response to the reviewer’s comment:
We use the same equations to calculate the averaging kernel (A) as Zhang et al. (2021):

\[ A = I - \hat{S}S_a^{-1}. \]  

\[ \hat{S} = \left( \gamma K^T S_o^{-1} K + S_a^{-1} \right)^{-1}. \]

(E1) (E2)

where \( I \) is the identity matrix, \( \hat{S} \) is the posterior covariance matrix, \( S_a \) is the prior covariance matrix, \( \gamma \) is the regularization parameter, \( K \) is the transfer matrix, \( S_o \) is the observation error covariance matrix that includes contributions from the instrument error and the forward model error.

In this calculation, the transfer matrix is reflected through posterior covariance, but the averaging kernel does not directly require a 4D-Var system to compute the transfer matrix, which is hard to generate.

There is another formula for calculating \( K \) as in section 2.2.3 from Yi et al (2020), which would require the computation of the entire transfer matrix, as mentioned by the reviewer:

\[ A = GK \]  

\[ G = S_a K^T (KS_a K^T + S_o)^{-1} \]

(E3) (E4)

where \( G \) is the gain matrix that describes the sensitivity of fluxes to observations. However, according to Rodgers (2000, Chapter 2, equations 2.78 and 2.79), the equations E1 and E3 are equivalent. So, the essential terms needed for our Bayesian methodology is the prior and posterior covariance matrix.

In a 4D-Var system, the posterior flux uncertainty is often approximated using a Monte Carlo approach (Chevallier et al., 2007, Liu et al., 2014). In a Monte Carlo approach, an ensemble of prior states and observations accounting for the prior and observation gaussian error statistics are generated, and the standard deviation of the ensemble posterior fluxes gives the estimates of posterior flux uncertainty. In this case, the impact of transport model has been implicitly taken into account. So, the averaging matrix could be approximated using the same method as the analytical solution, provided that the posterior covariance matrix is available.

We think the above paragraph will be of interest to the community, especially for people using the 4D-Var approach. So we adapted it and added to the manuscript, at the end of the first paragraph in section 4.0 Remaining uncertainties, future directions, and summary:

*In a 4D-Var system, the posterior flux uncertainty is often approximated using a Monte Carlo approach (Chevallier et al., 2007, Liu et al., 2014). In a Monte Carlo approach, an ensemble of prior states and observations accounting for the prior and observation gaussian error statistics are generated, and the standard deviation of the ensemble...
posterior fluxes gives the estimates of posterior flux uncertainty. In this way, the averaging matrix could be approximated using the same method as the analytical solution, provided that the posterior covariance matrix is available.

Specific comments

Keypoint 3: Please state what "25% less" is in reference to.

Thanks! Sorry for the ambiguity. The “25% less” is in reference to the low-performance models. We find it challenging to keep both central messages in Keypoint 3 within the 140 characters limit. In this revision we decided to keep only one central message here: ‘We find ~25% less global emissions from best-fit models and a lower-than-expected sensitivity of global wetland CH4 emissions to temperature.’

P2L27 (and P3L39): Suggest changing "CH4 budget" to "CH4 source" since the budget incorporates also atmospheric loss.

Right! Thank you.

P2L28 (and P3L47): I suggest removing "refine" since you are not refining the bottom-up models in this study, just comparing them with "satellite-informed" flux estimates.

Thanks! We agree that this sentence was ambiguous. We did not modify any land models, but we did bring up a ‘high performance ensemble’ in place of the ‘full ensemble’, showing reduced uncertainties.

So instead of saying we refined ‘models’, we now say: ‘…we refined bottom-up estimates of wetland emissions’, at both P2L28 and P3L47, as shown below:

‘To test and refine bottom-up estimates of wetland emissions using satellite-informed top-down CH4 fluxes, we present an approach that comprehensively accounts for prior flux assumptions, posterior cross-correlations, and spatial uncertainties in top-down flux estimates.’

‘Here, we use satellite-based top-down CH4 flux estimates (2010-2012) to test and refine 42 bottom-up estimates of wetland emissions that use a range of hypothesized wetland extents and process controls.’

P5L82: Please change this sentence since they way it is written it implies that all inversions only use space-borne observations, which is of course not the case.

Sorry for our mistake! The correction is below:
Atmospheric inversions use space-borne observations of atmospheric column CH₄ to provide a constraint on the global scale spatial and temporal variability of wetland CH₄ emissions.

P6L102: The method described in this study only accounts for random uncertainties in atmospheric inversions, and not systematic uncertainties, due for example to atmospheric transport errors. Therefore, please add "random" before "uncertainties".

Thanks! We added “random”!

P9L157: Could the authors please add a bit more of an explanation for "C respired as CH₄". Do the authors mean simply the amount of CH₄ produced by methanogens, and if so, how is temperature sensitivity of methanotrophy (microbial methane oxidation) accounted for or is this not included in this parameter? Also, for CO₂ respiration, do the authors mean autotrophic or heterotrophic respiration (or both)? I think in general a bit more explanation the parameter Q₁₀ CH₄:C is needed.

Thanks! We made changes accordingly. These explanations help the readers to understand the model:

‘Q₁₀ CH₄:C is the relative CH₄: C respiration for a 10 °C increase. A Q₁₀ CH₄:C value equals 1 thus indicates CH₄ respiration is as sensitive to the temperature as heterotrophic CO₂ respiration. In different WETCHARTs configurations, Q₁₀ CH₄:C equal 1, 2, and 3 to represent low, medium, and high temperature sensitivity of CH₄ emissions. Any value smaller than 1 is not considered since the temperature sensitivity of CH₄ emissions is higher than CO₂ respiration across ecosystem scales (Yvon-Durocher et al., 2014).’

Then at the end of the method description, we added:

‘This approach empirically provides first-order constraints on the role of carbon, water and temperature on the spatiotemporal variability of wetland CH₄ emissions, and indirectly accounts for the individual processes of production, oxidation, and transport pathways from belowground to the atmosphere. Variants of this formulation have been used within a range of wetland CH₄ emission models (e.g., Zhang et al., 2017; Bloom et al., 2012; Melton et al., 2013).’

P12L233: The authors write that larger delta_wRMSE values indicate a better model agreement with atmospheric estimates, w.r.t the model subset. However, I would think the opposite, i.e., smaller delta_wRMSE values would indicate better agreement, as values of wRMSE_MM smaller than wRMSE mean that the specific model process agrees better than the mean of the model subset. Could the authors please explain.

Thanks for asking! We apologize that our description was not clear. The reviewer is right that if the wRMSE of a model is smaller than the mean wRMSE, then this model outperforms the average. In fact, in our equations wRMSE_MM is the reference mean and not the wRMSE of the model. But as the reviewer asked, we realized it is more
intuitive to calculate the marginal performance of each model $m$ by subtracting the model wRMSE with the average wRMSE, so we reversed the order on the right-hand side in equation (3):

$$
\Delta \text{wRMSE}_{(m,p)} = \text{wRMSE}_{MM(m,p)} - \text{wRMSE}_{(m)}
$$

We revised the x axis values in figure 4 to reflect the changes in equations. We intentionally reverse the x axis so that models with higher performance distribute on the right-hand side. So it is more intuitive to understand the figure.

We also added description in the figure 4 legend to help digest the figure:

‘Negative values on the x axis indicate the wRMSE of a model is smaller than the reference mean, and thus the model has a better performance than the average.’

P15L278: Could the authors please explain what the "emission scaling factor" is.

Thanks! A description of the ‘emission scaling factor’ was in the second paragraph of section 2.2, we revised it to make it clear:

‘Global wetland emissions scaling factors are used to calibrate models with three prescribed global CH4 emission rates: low (124.5 Tg CH4 yr^-1), medium (166 Tg CH4 yr^-1), and high (207.5 Tg CH4 yr^-1) from 2000–2009 top-down wetland CH4 emission estimates (Saunois et al. 2016).’

‘The global wetland emissions in WetCHARTs models are calibrated with global emission scaling factors so that for each model the total emission in 2009-2010 amounts to 124.5 Tg CH4 yr^-1, 166 Tg CH4 yr^-1, and 207.5 Tg CH4 yr^-1 based on the uncertainty range of top-down wetland CH4 emission estimates (Saunois et al. 2016).’

P15L281 (and P16L315): The authors state: "atmospheric constraints indicate lower temperature sensitivity and [...] generally leads to better agreement with atmospheric CH4 measurements. However, the parameter compared is the Q10 CH4:C, i.e. the rate of C respired as CH4 versus CO2. So in fact the result is lower temperature sensitivity of CH4 relative to that of CO2. Please also see my comment concerning P9L157 about what Q10 CH4:C represents.

Thanks for your comment! We realized that the way in which we demonstrated ‘a lower temperature’ could be confusing. Our intention was to say, ‘a lower temperature sensitivity ($Q_{10}$ CH4:C equal 1, comparing to $Q_{10}$ CH4:C equal 2, and 3)’, not, ‘a lower temperature sensitivity ($Q_{10}$ of CH4, comparing to $Q_{10}$ of CO2)’.

We revised at both P15L281 and P16L315 to make it clear, highlighted in bold:
Relative to the median WetCHARTs configurations, atmospheric constraints indicate a low temperature sensitivity and a low mean global wetland emissions scaling factor, and satellite-based wetland extent (GLOBCOVER) generally leads to better agreement with atmospheric CH4 measurements.

Altogether, we find that low global total emissions, a low temperature sensitivity, and satellite-based wetland extent are the primary factors driving biogeochemical model performance against the global inverse flux estimates.

We revised P9L157 as:

Q10 CH4:C is the relative CH4: C respiration for a 10 °C increase. A Q10 CH4:C value equals 1 thus indicates CH4 respiration is as sensitive to the temperature as heterotrophic CO2 respiration. In different WETCHARTs configurations, Q10 CH4:C equal 1, 2, and 3 to represent low, medium, and high temperature sensitivity of CH4 emissions. Any value smaller than 1 is not considered since the temperature sensitivity of CH4 emissions is higher than CO2 respiration across ecosystem scales (Yvon-Durocher et al., 2014).

The method requires the Averaging Kernel, which depends on the posterior and prior error covariance matrices (or an alternative formulation using the prior and observation error covariance matrices) but also the transport matrix (please the general comment about this). Therefore, I think this should also be added as a constraint for this method.

This is a great point. Please see our reply to the second general comment above.

I suggest changing this to "relative uncertainties larger than that of wetland fluxes" since overall the freshwater body emissions are significantly smaller than that of wetlands, and thus also the absolute uncertainty.

Thank you so much. We agree and have made the change as suggested.

Is there a correlation between wetland extent (GLOBCOVER versus GLWD) and global wetland emission, for instance, does GLOBCOVER have a smaller extent and thus models using this have a smaller global wetland emission?

Thanks for the question! We revised the descriptions in section 2.2 to make it clear that the global WetCHARTs CH4 emissions are explicitly determined by the scaling factors:

The global wetland emissions in WetCHARTs models are calibrated with global emission scaling factors so that for each model the total emission in 2009-2010 amounts to 124.5 Tg CH4 yr-1, 166 Tg CH4 yr-1, and 207.5 Tg CH4 yr-1 based on the uncertainty range of top-down wetland CH4 emission estimates (Saunois et al. 2016).

The different wetland extent maps do however contribute to the overall wetland CH4 seasonalities and fractional distributions at 0.5 resolution, which together account for the GLOBCOVER and GLWD delta_wRMSE values in Figure 4.
