Review on Wittig et al., egusphere-2023-2308

General comments

This study examines whether current observation network is capable for detecting future potential changes in CH_4 emissions in the Arctic. Arctic is important as vast amount of carbon is stored and could be released as Arctic warming proceeds, leading to positive climate feedback and enhance global warming. The authors use FLEXPART to generate synthetic observations (using current knowledge of fluxes and meteorological data) and examine whether those data can be used to detect scenario emission changes by an analytical inverse model. Inverse modelling has been widely used to quantify current and history of greenhouse gas budgets, but this study attempts to implement it also for studying future changes. This is a novelty.

The study challenges important questions in climate change, but I have few doubts and questions regarding their choice of methods. Particularly,

- The authors study the Arctic CH_4 emission changes in 35 years – this is rather short considering the processes of e.g. permafrost thaw, and in comparison to other scenario studies (which are often up to 2100). Because of the relatively short study period, the CH_4 emissions are needed to be increased unrealistically fast (20 % yr^{-1}), as authors point out as well, and therefore, the credibility of the results are weak. The choice of length and the increasing rate of emissions need to be justified, and at least add implications for more realistic changes.

- The above point leads to a conceptual question about “methane bomb”. In Introduction (P2 L7 – P2 L16), you use this term for both gradual and sudden methane release from the Arctic. As I understood, this study is about the gradual and continuous changes, and this needs to be clarified (Abstract, Introduction and Method).

- If I understood correct, you have generated synthetic data based on present/past prior emission information and meteorological data, which are used to constrain the future scenario fluxes. This would mean that observations would try to adjust emissions to current emission level. So the “detection limit” is when the observations cannot anymore constrain the fluxes to current emission level (+ uncertainty limit), i.e. the limit where observations cannot “see”. Is this correct, and what you aim to do? I would assume that it would be more meaningful if you generate synthetic mixing ratio data based on future emission scenarios, and constrain some prior fluxes with that data. With this, you could see if we can detect emission changes even if there are “missing information” in prior fluxes.

- The authors have examined the current and “extended” observation network, but due to the effect of the Russian war, substantial number of surface stations lack of data at current. How likely that we can still detect future changes in CH_4 emissions in Eurasia? How long of data lack is critical? I think these are very important questions. You may not need to rerun all simulations without those stations, but adding a few could bring really valuable information about future Arctic CH_4 study.

- Following the previous point, you have completely missed about the role of satellite data. I understand that it is challenging to do satellite inversion with Lagrangian models, but I would at least like to see some discussion about it. What if we have had “surface” data at satellite retrieval points?

Specific comments

P1 L13-15 Please add references to support your argument. I agree that CH_4 emissions from wetlands and other freshwater systems are probably a dominant source, but how large are the other natural sources?

P2 L3-4: Could you add information about how large are the anthropogenic CH_4 emissions in the Arctic in comparison to wetland emissions?

At end of Introduction: Please make it clear how many years of future scenarios you study.

P3 L1: Did you optimize the fluxes grid-wise or region-wise (121 sub-regions)?

P6 L7: Could you clarify by "only recently"? What is the year limit you have chosen?

P6 L8: "measurement of CH_4 columns" is originally not measuring mixing ratios, but to be used in inversion, you will probably only use the mixing ratio data. Also, satellite data also provide CH_4 column information, but those locations are probably not of satellites. This phrase should be clarified better.

Section 3.4:

- Could you possibly change the title to "Generating synthetic CH_4 mixing ratio data"?
- What is the temporal resolution of your generated data? 3-hourly?
- Initial concentrations means concentrations in each year (2008–2019)?

P8 L15-18: Please specify a bit more in detail how you have come to 506 different set-ups. It is unclear from the figures/tables as well as from the text. What are the different set-ups, did you change only emission scenarios (as the sentence is is in that section), or did you also use different synthetic data? Did you use different trends, or is all inversions have same trends as presented in Table 2? It is also unclear why there are two similar figures (Figure 2 and 4). Could you possibly combine them?

P9 L8-9: Why "only this region should be updated by the inversion"? Is East Eurasia strictly uncorrelated with other regions? Did you strictly set it so that observations are only constraining this region? If not, it is not surprising that other regions are also affected.

P9 L10: Is it so that the posterior emissions are much lower than the truth because the observations are generated using present-day emissions?

P10 L5-7: Is it really so that the "increase in the simulated scenarios is underestimated"? I wonder how strong are the regional correlations. Also, do you trust the "truth" or posterior estimates? You need to re-think how to put your arguments.

P11 L2: By "combine", do you mean that you only show the results of the region where you modified the trends, i.e. the effect of other regions are not presented? Please make it clearer.

P11 L 20-23: I am not sure what you wish to say. The applied trend is unrealistic, and you hoped that the inversion would detect the changes much earlier? Or you think that you should have applied a bit more realistic trend? What you mean by "more illustrative" – more, compared to what?

P12 L10-11: Is it really true that there is no influence about observations? What if there is a station over there? I would also guess that the observations in surrounding regions could affect the results.

P12 L13-15: This is interesting, but could be also due to the fact that many of the extended stations are often close to the currently available stations. Also, the emission magnitude near the station is important to consider – if we add stations where emissions are small, the effect could be minor.

P13 L1-2: Is there anything you could do to attribute those discrepancies to fluxes by changing some setups/uncertainties? Despite the minor effect on your results, do you still think those sites are important and could bring information about changes in trends in northern Europe or surrounding regions?

Technical corrections
Please use same terms for generated mixing ratio data (modelled, generated, synthetic, etc..)

Please check the spaces between units, and follow the journal role.

P1 L10 Remove "temperature"

P5 L10 Section Inversion framework
Please add section number

P11 L10: annual posterior emissions in year j and region r emis aj,r

P11 L12: Please move the *j* and *r* ranges on the right hand side of the equation, i.e.

$emiss^a - emiss^b < e, \quad j \in [2021, 2055], r \in [1, 121]$

You could put "is not fulfilled" in L10. Please also do so in Eq. 5 and 6.

P11 L16: "the threshold year is generally higher"
Do you perhaps mean "the year is generally **later**"?

P11 L25: "terms of detection limits, an increase of a few, up to 10 Tgy -1 , is necessary for statistically reliable detection."
Could you add e.g. in brackets how much they are in percentage?

P15 L18: "TROPOMI CH4" → CH_4 with subscript.

Figure captions: Use (a), (b) instead of "left" "right".

Figure 3 caption: I feel it would be more appropriate to say e.g. "Location of the sites where synthetic mixing ratio data are generated from", as you do not use actual observations at all.

Figure 5 y-axis: Are those units really correct? For example, in the bottom panel, 100 Tg/month of CH_4 from Arctic in 2020 does not sound at all realistic (even if it was annual emission). Y-axis label and caption does not have same units.

Figure 8:

- Please use more informative label in the color bars.
- The unit in color bar is [%], and color scales ranges between -10^3 to 10^3, i.e. 1000% change in emissions. Is this correct?
- Caption for (b): "Difference between the…" → "Absolute differences between.."?