Review on Thanwerdas et al., gmd-2021-106

General comments

The paper presents development of an atmospheric inversion model of methane (CH_4) by assimilating both CH_4 and $\delta^{13}C(CH_4)$ observations in order to optimizing global and regional CH_4 fluxes. The method allow to optimize several source sectors simultaneously, which is an advantage over CH_4 only assimilation methods that is more suitable for optimizing total budgets. As CH_4 is an important greenhouse gas, which has high mitigation potential, it is urgent to understand current budgets of both anthropogenic and natural sources. This work is highly relevant and valuable for increasing understanding regional source-specific emissions. Furthermore, the method is developed for Community Inversion Framework (CIF), which is flexible in various inverse modelling methods, such as transport models and optimization method. Therefore, this development will be beneficial for all CIF users.

The manuscript is generally well written and presented. I recommend the manuscript to be published, but would like to point out a few comments which could increase the value of the paper.

<u>Presentation of novelty</u>

As authors mention very briefly, this is a first attempt to carry out variational inversion assimilating $\delta^{13}C(CH_4)$ observations. Please mention this also in the abstract, and add slightly more details of the development in the Introduction, e.g. development of adjoint and implementation in CIF. From the Introduction, I was also not sure if such modelling has been done with LMDz previously, i.e. how well LMDz have been simulating $\delta^{13}C(CH_4)$?

<u>Categorization of the simulations</u>

I was not completely convinced about those S and T groups. Are they really needed? Did you categorize them based on results or really expected T groups to have higher variation before you started simulations? T1 is not only about changing isotope signature values and their uncertainty, but also degree of freedom (dof) in the optimization (I guess you optimize 10 flux categories?).

<u>Discussion on results</u>

Although this is a technical paper is not meant to evaluate the flux estimate nor $\delta^{13}C(CH_4)$ values obtained from the simulations, I would like to see briefly how your estimates are compared to previous studies. Or even simply mentioning in the Conclusion how you would do further analysis, including e.g. availability of evaluation data.

<u>Discussion on uncertainty estimates</u>

I understand that it is costly to calculate the full uncertainty from all simulations. However, you anyway present uncertainty in P12 L12. How was it calculated? From the cost function, you can speculate how dof and inclusion of additional data would affect the posterior uncertainty. Please comment on it in Section 3.5.

Specific comments

<u>Method</u>

Distribution of state vectors: did you assume all to be normal/Gaussian?

How did you derive the aggregated signature values?

What is the temporal and spatial resolution of prior fluxes?

What is the temporal resolution of the optimization?

Can you provide range of observation uncertainty (diagonals of **R**) for each stations, maybe by adding information in Table S3 and S4, and briefly mention ranges in the main text? This will help understanding the results on cost function and RMSE differences better.

Curve fitting data:

- Was there any specific reason why you decided to use smoothed data?
- After curve fitting, what is the temporal resolution of the data you assimilated? Did you generate same amount of $\delta^{13}C(CH_4)$ data in REF and S2?

Offsets in initial condition: How much offset did you need to add/subtract?

Results

P13 L9: "Consequently, the system is preferentially adjusting $\delta^{13}C(CH_4)$ over CH_4 values to reduce the cost function."

- Can you speculate why? Is it because observation uncertainty (diagonals of **R**) is relatively smaller in $\delta^{13}C(CH_4)$ than CH_4? The cost function show that the observational constraint in CH_4 is larger (probably main reason is amount of data?).
- For S2, I wonder why contributions of $\delta^{13}C(CH_4)$ and CH_4 are similar to S3. Did you assimilate same amount of $\delta^{13}C(CH_4)$ data in REF and S2?

P14 L30-34:

This could also be due to prescribed observation/transport model uncertainty.

Emission increments: Emission changes are large in regions with high emissions. Please mention.

Conclusion

Please expand how much work would be needed for switching transport models and optimization methods in CIF for the $\delta^{13}C(CH_4)$ data assimilation. Can we use e.g. initial mixing ratios, do we need to run spin-up and build adjoint if transport model is changed? How about changes in optimization methods? Can we use same state vectors and covariance structures?

Figures

Figure 3: Please add label of x-axis

Figure 4: Please add legend of posterior results from REF simulation, and perhaps use different color than green, as it's not S-group simulation? Please also add results from NOISO.
Figure 4 caption: I guess the figure is global **monthly** mean?

Figure 5: Prior CH_4 is same for all simulations, and $\delta^{13}C(CH_4)$ for some. Please consider minimizing.

Figure S2: Please add label and unit of y-axis. Caption is slightly unclear – what do you mean by "inferred with REF"?

Technical comments

P14 L22: The S-group provides a better match **to δ13C(CH4) observations** than...

P15 L4-5: AMY is not in South-East Asia.