Comment on gmd-2021-106

Anonymous Referee #1

Referee comment on "Variational inverse modeling within the Community Inversion Framework v1.1 to assimilate $\delta^{13}C(CH_4)$ and CH_4: a case study with model LMDz-SACS" by Joël Thanwerdas et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-106-RC1, 2021

General comments

This manuscript describes a new 3D variational inverse modelling framework for estimating fluxes of CH4 by source type, and includes the assimilation of d13C observations, that is, the ratio of the 13C to 12C isotopomers of CH4. The framework is based on the atmospheric chemistry transport model LMDz-SACS and the Community Inversion Framework (CIF). Although d13C observations have been assimilated in inversions of CH4 before, the new inverse modelling framework is based on variational data assimilation and allows for simultaneous optimization of CH4 sources and source signatures, which represents a new development. The new framework should contribute towards a better understanding of the global CH4 budget. The manuscript is generally well-written and presented and the level of methodological description is adequate. However, there are a few points that need clarification (see specific comments). Additionally, information about data and code availability is lacking. Under "Data availability" the authors give http://community-inversion.eu as the reference for the code, however, this is just a general website about the CIF. This website indicates the Git git.nilu.no/VERIFY/CIF but this is just for the generic version of CIF and not that pertaining to this paper. Also, under this section, details about where to access the observational and prior flux data should be given.

Specific comments

P1L2: suggest changing this to: "…indicating relative changes in the sources and sinks" as it is evident from the fact that the mixing ratios have been increasing that there must be a change in the sources and/or sinks and not just a variation but a change in one relative to the other.

P2L13: I think this sentence is potentially confusing and could be better formulated. What I think the authors mean is that without regularization the inverse problem is ill-conditioned (or ill-posed) giving no unique solution, hence the need for regularization e.g. by providing prior information. Also it is unclear to me what is meant by "no continuity with the data" - could the authors please explain this.

P2L21: Variational methods, such as the Lanczos version of the conjugate gradient algorithm provide the posterior error covariance matrix with little additional computational cost.

P2L33-34: I would suggest the authors give ranges for the various source categories to reflect how variable values within each category can be.

P3L2: I think the authors should precise that they are not consistent with the d13C observations and the prescribed d13C ratios.

P3L12: Thompson et al. 2018 used a variational method to optimize CH4 emissions and the OH sink with the AGAGE 12-box model. Perhaps the authors mean never in a variational inversion framework with a full 3D atmospheric transport model?

P4L15: All Bayesian methods require the inverses of R and B.

P4L17: I think you should specify the assumption, i.e. that the observation errors are uncorrelated.

EQ6-7: I'm confused about the value MTOT, is this the molar mass of CH4 in source FiTOT, if so then MTOT depends on the d13C ratio of CH4 in FiTOT.

P7L13: For the category "fossil fuels" could the authors please specify if this is only fugitive emissions or also combustion emissions, and if the source signature is considered the same for fugitive and combustive emissions?

P9L2: I think it would be good to include the references for the source signatures in the main part of the manuscript and not just in the supplement. Also, there is no reference given for the livestock category nor an explanation why this category had a time varying source signature and what the dependence on time was.

P10L7: Do the authors mean that the model, LMDZ-SACS cannot reproduce the high temporal frequency of CH4 or d13C or both? If it is d13C, weekly observations are not high frequency. Also do the authors have an idea why the temporal variability could not be reproduced? I think this needs to be better explained. Also why assimilating the curve fitted data was chosen as the solution rather than e.g. increasing the observation uncertainty, filtering or averaging the observations?

Fig. 3a) I think here "cost" (or "value of cost function") is meant and not "cost function" and it would help to specify that the x-axis is "iterations".

Section 3.1: I think somewhere the results of the adjoint tests should be presented since a new version of the model was developed, including its adjoint.

P14L16-19: Could the decreasing values of d13C in REF be also due to an underestimation of the atmospheric sink since reactions with OH and Cl enrich d13C?

P18L20: It would be helpful if it would be stated again that this is for NOISO and REF increments.

P19L6: It is interesting that in order to correct for the prior decreasing trend in d13C, the inversion increases the source signatures of all sources, this means that the increases in the d13C rich sources, such as biofuel/biomass burning, are not sufficient to correct this trend. In T3 and T4 these emissions increased significantly, since there was not the degrees of freedom to adjust the source signatures. The question is, what is more accurate, higher source signatures or high d13C rich sources? Also, this result depends of course on having the correct atmospheric sink. Although these questions cannot be answered in this paper, I think they warrant more discussion as these are key sources of uncertainty. Also, I think the statement "All source signatures are shifted upwards by the inversions" needs to be qualified, that is, there are the exceptions of T3 and T4 (which had very small prior uncertainties for the sources signatures) and the "natural" source.

P19L9: I think by "total fractionation effect" the authors mean the kinetic isotope effect of atmospheric oxidation, if so, I suggest changing this to be clearer about what is meant. Also, I think it would be interesting to include a test using alternative OH fields to see how strongly the results are affected by the OH sink estimate.

P19L14:18: Presumably this describes the results of the REF scenario, but it would be clearer to specify this.

Technical comments

P3L2: constrain -> constraint

P3L3: have -> has

P3L6: regrowth -> renewed growth

P4L16: This phrase is not grammatically correct, please change to "allowing for the inverse to be calculated easily"

P5L18: multi-constrain -> multi-constraint

P12L3-L16: This would be easier to follow if the list items (i.e. the different inversion tests) would be numbered.

P19L6: source signature -> source signatures

P21L16: relatively -> relative