Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC1
https://doi.org/10.5194/gmd-2021-339-RC1, 2021

Comment on gmd-2021-339

Anonymous Referee #1

Referee comment on "Order of magnitude wall time improvement of variational methane inversions by physical parallelization: a demonstration using TM5-4DVAR" by Sudhanshu Pandey et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-339-RC1, 2021

The authors present an implementation of a so-called physical parallelization for variational flux inversions (PPVI): from a previously described PPVI aimed at carbon dioxide (CO_2), they add developments to take into account the chemical reactivity of methane (CH_4).

General comments

The developments described in this paper are particularly relevant since long-term methane inversions are now run by several teams and the issue of the trends in methane emissions by various types of sources is still under study. Nevertheless, I think the presentation is too sloppy as it is: the work must be better introduced and described. technically, several notations are unclear in the mathematical description. Moreover, although I am not an native English speaker, I think the writing has to be improved.

The introduction to the paper is off the mark. It remains very general and not precise enough on variational inversion. Some examples: the state vector, in most inversions, not only consists in emissions but also includes initial conditions or boundary conditions for area-limited domains; the analytical approach is alluded to compared to the variational one but it is never explicitly stated that the analytical approach cannot be used for non-linear problems (which may be the case with reactive species); conversely, it is not stated that the variational approach does not provide full posterior uncertainties as a by-product of an inversion (either none are obtained, or truncated ones). I think the introduction does not target the right readers: people who may be interested in PPVI already know the whys and hows of analytical and variational inversion. It would be more useful to clearly state in which cases and why this implementation of PPVI is interesting e.g. for variational inversions of reactive species at scales at which chemistry is to be taken into account but the precision is not so important i.e. not for non-linear chemistry.

In Section 2 Physical parallelization for variational inversions, it must be made very clear which parts are the general or Chevallier (2013) developments and which are specific to this work and therefore, to methane. It should make it possible to understand whether the developments are also applicable to other species (e.g. CO). A discussion on the assumptions required to apply this PPVI and its limitations is necessary, either in this part or in the introduction or in the discussion.

In Section 3PPVI Performance test, not all the information required to understand (and reproduce) the simulations are available. The main information missing is how the posterior uncertainties are obtained: which approach is used? What are the assumptions? Even the simple approach of using Congrad as a minimizer and using the uncertainties obtained with a truncation requires to specify at least this truncation threshold and how it is expected to affect the resulting uncertainties estimates.

Specific comments

- Section 2 Physical parallelization for variational inversions:

- p.3 l.93 in Eq.2: it should be H^* and not H^T - or the assumptions which make H^T equal to H^* should be stated. It would also be safer to add a bracket: H^* [$R^{-1}(H(x^i)-y)$]
- p.3 l.96: same remark as above: H^* is the adjoint, if H^T is used, it means that the problem is linear, which must be stated explicitly from the beginning.
- p.4 l.116: why this conversion factor?
- p.4 l.119 in Eq. 5: the notation for H changes suddenly from italics i.e. an operator with no particular characteristics to bold i.e. a matrix (probably): see above for the issue about H and its various spin-offs being linear or not and adjust notations accordingly.
- p.5 l.134 in Eq 6 and seq.: the notation * for the adjoint appears here: please make this consistent with the beginning of the Section. Moreover, H is bold so probably a matrix i.e. for a linear problem so that * and transpose are the same: this is not clear at all for the reader.

- Section 3 PPVI performance test

- p.7 l.210-211: the difference between both posterior emissions must be compared to the difference with the prior to be said to be small - or not. Better still, the uncertainties on the three estimates must be taken into account for such a comparison.
- p.7 l.213-214: same remark as above for the regions: how much is the deviation from the prior compared to the 5% between the two inversions? What about the uncertainties on the emission estimates?
- p.7 l.214: the posterior uncertainty is alluded to here but nowhere is it stated how it is computed. Since the full posterior uncertainties are not a by-product of the variational inversion, the way they are computed must be described (truncated from Congrad?

ensemble? Monte-Carlo? other method?).

- p.7 l.215 seq.: I guess the correlation coefficient used here is simply the correlation of the time series. There are other characteristics of the inter-annual variability which could be interesting to look at e.g. are the uncertainties the same?
- p.8 l.233-234: what about the uncertainties? Without an explanation on how they are computed, the times given here are read as times for one inversion and may therefore be a lot smaller than what is actually required to get the full range of results (i.e. emission estimates + uncertainties).
- p.9 l.263 seq.: the specification of the OH fields is one of the main issues for methane inversions today, particularly as the vertical distribution of OH is crucial when using satellite data. A sink defined as simply as suggested here (even with an annual change) does not really solve the scientific issue. The optimization of the sink, as described in Section 4.3 is one of the possible ways forward.

- Figure 1

It would be useful to distinguish between the general (CO_2) PPVI and the elements which are particular to this work i.e. the CH_4. for example, the sink does not appear in this figure. Please also check the consistency of the notations (matrices, operators, vectors,...).

- Figures 5, 6 and 7

How are the uncertainties obtained? Does a 2-sigma interval make sense?

Technical corrections

Throughout the text, "a priori" and "a posteriori" are used: shouldn't it be "prior" and "posterior" instead?

There are many writing mistakes, such as sentences where words are missing (e.g. p.7 l.194: "the PPVI results are good agreement with the results from serial") or superfluous words remain: the text must be re-read carefully by the authors before being checked by a native speaker.