

Geosci. Model Dev. Discuss., referee comment RC1  
<https://doi.org/10.5194/gmd-2022-45-RC1>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on gmd-2022-45

Anonymous Referee #1

---

Referee comment on "Integrated Methane Inversion (IMI 1.0): a user-friendly, cloud-based facility for inferring high-resolution methane emissions from TROPOMI satellite observations" by Daniel J. Varon et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2022-45-RC1>, 2022

---

Varon et al describe a user-friendly interface for conducting inverse modelling of TROPOMI CH<sub>4</sub> data to estimate regional CH<sub>4</sub> emissions. The underlying approach is based off previously published work using the GEOS-Chem chemical transport model to calculate the relationship between emissions and atmospheric column mean concentrations and uses a well-documented analytical inversion approach. The novelty of this work is in providing a user-friendly interface so that inversions can be performed in the cloud without the requirement to download either model or observational data (which can be large) to a local computer, and without any expertise required in either running the GEOS-Chem model or Bayesian statistics. The Integrated Modelling Inversion offers the potential to be used by any interested party, after allowing for the costs of using Amazon Web Service. As such, it has the potential to be a useful resource for the scientific community and is within the scope of GMD. Of course, there is a danger that a lack of user expertise may result in spurious analysis of the outputs, but that bridge will just have to be crossed when it comes to it. Overall, the paper is well structured and the methods adequately described. An inspection of the config file confirms that it could be used with limited expertise.

### Specific comments:

P2. L32: "*...is a major focus of climate policy*"

Whose policy? Perhaps give an example. E.g. the COP26 global methane pledge.

P2. L42: *It exemplifies the emerging concept of "bringing compute to data"*.

I think it would help to explain this jargon. At first glance it just looks like whoever came up with it doesn't know the difference between a noun and a verb.

P2. L47: "*were limited by...sparse sampling (GOSAT)*"

GOSAT is still active (along with GOSAT2) so it sounds odd to refer to it in the past tense like this.

P2. L62 "*Analytical solution takes advantage of the linearity of the relationship between methane emissions and concentrations (Maasackers et al. 2021)*"

This seems a peculiar citation. I hardly think that a paper from last year was the first to exploit this linearity. Analytical methane inversions date back several decades.

P3. L69-70: "*and for stakeholders lacking technical expertise.*"

Who are these stakeholders? And isn't some technical expertise still required to interpret the results? If you don't understand how or why the results are being generated isn't there a danger that you might over-interpret the results without accounting for the many limitations of the satellite data, model and statistical analysis?

P4. L95, "*in a public S3 bucket*"

This is just jargon. What does it mean?

P4. L105: "*caused mainly by aliasing of surface albedo errors into the methane retrieval*"  
Are there other potential causes of regional biases beyond albedo?

P4. L110: "*and with SWIR albedo < 0.05*"

I assume this should say ">0.05"? Otherwise you'd be filtering out almost all data.

P4. L110-112: What percentage of "good quality" observations do the additional filters remove?

P4. L124: "*and distributed vertically following a GEOS-Chem simulation at 4x5*"

Why use the 4x5 simulation when 2x2.5 is available and might produce more accurate fields?

P5. L137: "*and the sinks are relatively smooth*"

Do you mean spatially smooth? And is this referring to the offline sink fields in the model?

P5. Fig 1: This probably doesn't matter for this example work but is there an option for the buffer regions to include oceans? If your domain were further east then presumably you would need to somehow account for oil and gas fields in the Gulf of Mexico within the buffer regions. Is that also covered by the 75% adjustable default?

P8. L206: *"The first step is to geo-locate the TROPOMI pixel..."*

Are the TROPOMI data geo-located only over the 0.25x0.3125 inversion domain of interest (i.e. the Permian basin in this work) or over the full computational domain or some other spatial domain?

If the first, is there a danger that important signals downwind of the region of interest are missed?

P8. L211-213: I'm not sure the figure is entirely clear. Do you prescribe the model surface pressure as the satellite surface pressure or the opposite way round? Does this run the risk of not conserving mass in the model column?

P8. L227: *"...except under polar vortex conditions (Zhang et al 2020)."*

This reference is surely wrong. Stanevich et al (2020) were the ones to show the nature of the GEOS-Chem latitudinal bias and errors were not entirely limited to polar vortex conditions. This leads back to the earlier point of why use the 4x5 simulation for boundary conditions rather than 2x2.5?

Stanevich, I., Jones, D. B. A., Strong, K., Parker, R. J., Boesch, H., Wunch, D., Notholt, J., Petri, C., Warneke, T., Sussmann, R., Schneider, M., Hase, F., Kivi, R., Deutscher, N. M., Velazco, V. A., Walker, K. A., and Deng, F.: Characterizing model errors in chemical transport modeling of methane: impact of model resolution in versions v9-02 of GEOS-Chem and v35j of its adjoint model, *Geosci. Model Dev.*, 13, 3839–3862, <https://doi.org/10.5194/gmd-13-3839-2020>, 2020.

P9. L244-245: Does the regularization parameter essentially limit the resolution of inferred emissions? Couldn't one just achieve a similar effect by using coarser resolution grid cells (i.e. coarser resolution perturbations), thus reducing runtime and AWS costs? The choice of the regularization parameter seems highly arbitrary, would the intended non-expert user understand this?

P11. L303-304: *"we find it gives a successful approximation"*

What counts as a successful approximation? It might be wise to show some evidence for this.

P12. L321 *"using a non-informative uniform prior estimate"*

But isn't it the specification of Gaussian PDFs that enables one to perform a quick analytical inversion? Prescribing an alternative PDF would surely require an alternative inversion approach.

P16. L401-404: Presumably if you have the Jacobian matrix already there's no need to perform a final forward GEOS-Chem simulation, since you can just take the product of x

by K to generate modelled observations.

P18, L458: "*a better gridded bottom-up inventory*"

How do you know it is better? What has it been validated against?

P 20 Section 4.3

How do costs scale with domain size or number of state vector elements?

P22. L534: "*option to use non-uniform prior and observation error covariance matrices*"

Does this also include considering the off-diagonal terms in the error covariance matrices?

P22. As it stands it appears there is only the option to run for a fixed time window with constant emissions assumed during this period (or at least a constant posterior scale factor applied to the temporally variant prior emission fields). Does the IMI have the capacity to run say a 1-year run resolving monthly emissions, or does each monthly run need to be conducted independently of others?