

The manuscript presents estimates of methane emissions, focused on Brazil and the Amazon Basin, using a top-down inversion approach, which is validated against independent measurements, and compared to a bottom up model. The manuscript is thorough and detailed, which will undoubtedly be informative for future studies. The conclusions are supported by the findings of the study, which are suitably appraised. As such, I hope to see its eventual publication in ACP. There is however, a major error in the methodology – as it is stated in the paper – which may or may not have a large impact on the results. This must be addressed before the manuscript can be considered acceptable for publication. In addition, below are a number of suggestions for revisions to improve the manuscript, followed by technical comments.

We thank the reviewer for his comments, which have significantly helped to improve the paper and clarify our results and message. We hope that we have addressed these concerns appropriately. We note this reviewer has withdrawn his comment regarding the major error in the methodology (see below and RC2). Our point-by-point response is given below, highlighted as blue text.

- 1) Line 150-151: “For accurate comparison between the retrieved XCH4 and those simulated by the model, the GOSAT averaging kernels were averaged similarly to the XCH4 and applied to the model vertical profiles.” This approach is not mathematically sound and therefore, as an absolute minimum, it must be checked that it has a negligible effect. Otherwise it could lead, quite likely, to an underestimation of the modelled XCH4 and thus overestimate emissions. The reason for the error here is as follows. If we simplify the maths to just two variables, and let  $X$  be the concentration at each of the  $i$  the model levels, and  $A$  be the diagonal of the averaging kernel, and  $Y$  be the XCH4, then we can say that  $Y = \sum_i a_i x_i$ . If we take the mean of each observation  $j$  of the total  $n$  observations in a grid cell then we want  $Y_{grid} = \frac{1}{n} \sum_j Y_j$ , which is equal to  $\frac{1}{n} \sum_j \sum_i a_{ij} x_{ij}$ . As there is only one  $X$  per grid cell in the model, we can simplify this to  $\frac{1}{n} \sum_j \sum_i a_{ij} x_i^{grid}$ . From the text, it states that instead the average of the averaging kernel was applied to the modelled profile, so  $Y_{grid} = \sum_i \frac{1}{n} (\sum_j a_{ij}) x_i^{grid} = \frac{1}{n} \sum_i x_i^{grid} \sum_j a_{ij}$ , which is not equivalent. Please let me know if this needs more clarification.

Although the reviewer has since retracted this comment (see below and RC2), we have clarified our reasoning for using this methodology. We have added text within the manuscript at line 247 as follows:

“Using a single model profile in each grid cell and model time step allows the use of averaging kernels that have been averaged in this way without introducing a bias, due to the distributive property of matrix multiplication.”

We thank the reviewer for bringing up this comment and for being open and responsive in further discussions regarding this method, leading to improvements to the text.

- 2) Line 197: Although each has completed 40 iterations, how do you check that the optimisation routine has converged within 40 iterations?

We have clarified this in the text as follows:

“The inversions were carried out for each year separately and each completed 40 minimisation iterations. For each year’s inversion, 40 iterations were enough for the cost function and its gradient norm to be judged to have converged (less than 1% variation through 5 consecutive iterations).”

- 3) The paper is thorough and detailed, although – as a matter of opinion – it is in places quite arduous to read. I suggest making use of supplementary material and moving some of the analysis here. For example, at line 533: This section should be moved to the supplement, and referred to in line 531. The application of the B-U model as a whole and its discussion would be better placed in the supplement to improve the conciseness of the manuscript. This could also apply to the validation against the site and aircraft data – refer to the outcomes in the main text but the details can be moved to a supplement. This should get across the key points of the paper, i.e. the emissions and their sources, better to the reader.

We have generally attempted to shorten the main text wherever possible but have also moved some sections to the supplement as suggested here. The section on the bottom-up model has moved to Appendix B, whilst the details of the inversion at the higher resolution has moved to Appendix C. We have left the analysis of the independent validation in the main text, however.

- 4) Code and data availability: It would be much more beneficial to the community if the results of this work (posterior emissions estimates in space and time) were publicly available. E.g. placing the spatial maps for the mean posterior emissions for Brazil and the posterior emissions estimates for Brazil each year in a public repository in e.g. netcdf format. I see that this has also been suggested by the topical editor, but I would like to reiterate its importance.

This is correct and we have since uploaded our prior and posterior monthly CH<sub>4</sub> fluxes for South America in netcdf format to the data archive of the Centre for Environmental Data Analysis (CEDA, <https://www.ceda.ac.uk>). The data can now be cited as follows:

Wilson, C.; Chipperfield, M.; Gloor, M.; Parker, R.; Boesch, H. (2021): Posterior South American monthly mean surface flux of methane (2010-2018) produced using the INVICAT 4D-Var inverse model. Centre for Environmental Data Analysis, 17 February 2021. doi:10.5285/88224a922439441fa6644b4564dcd90c.

We have added this information to the manuscript.

#### **Technical comments:**

Throughout: Units for ACP should be expressed in exponential form, i.e. Tg yr<sup>-1</sup> and not Tg/yr.

We have adapted our terminology as suggested.

Throughout: Be consistent with  $T_g(\text{CH}_4)/\text{yr}$  and  $T_g/\text{yr}$  (e.g. line 81). I would recommend stating explicitly in its first occurrence that  $T_g/\text{yr}$  refers to  $T_g(\text{CH}_4)$  and from thereon just writing  $T_g/\text{yr}$ .

We have adapted our terminology as suggested.

Throughout: There is an error in the nomenclature used throughout the manuscript. The manuscript often refers to the prior/posterior mole fraction when referring to a single parametric estimate; the prior/posterior are distributions of values. A better usage is *a priori* and *a posteriori*, or more explicit, the prior/posterior mean.

This is a good point; the former nomenclature is a little careless. Instead of correcting every instance of this through the text, which we believe reduces readability, we have inserted the following text in Section 2.2.1:

“The inversion input is in the form of an *a priori* mean flux value for each grid cell along with an error covariance matrix for these values, and the output is made up from an *a posteriori* mean grid cell flux value and error covariance matrix. Mean *a priori* and *a posteriori* atmospheric mole fractions of  $\text{CH}_4$  are also produced. For ease, throughout the remainder of this text, we will refer to the mean values of the *a priori* and *a posteriori* fluxes as ‘prior fluxes’ and ‘posterior fluxes’, respectively. Similarly, the mean *a priori* and *a posteriori* mole fractions will be referred to as prior and posterior mole fractions.”

Line 74: Basin is capitalised (not consistent with earlier use of basin).

We have changed all uses of ‘basin’ to lower-case for consistency.

Line 75: This sentence is confusing: it reads as though fires contribute to the number of wetland sources. Consider splitting this sentence as e.g. “as well as a number of other wetland sources in S America, emissions from. . .also contribute to methane emissions.”

We have changed this as suggested.

Line 77: Overlap in which sense? The emissions processes? In space? In stochastic error?

We have changed the text to clarify that flux contributions that we now know to be through trees is not extra contribution to the total but has likely already been counted as part of the wetland flux in some inventories.

Line 79: Consider “variability” rather than “variance”, so as not to confuse with statistical variance.

Yes, this has been fixed.

Line 79: 'Earlier estimates' is inexact. Is this when the research was carried out, or the emissions from the year(s) in question? Specify the time period that you are discussing.

We have been more exact in our language here – “In studies published in the 2000s and early 2010s...”.

Line 142: ‘. . .and found that the two agreed within their respective errors’. This sentence is meaningless without a description of the probability content in which these dataset agree (e.g. the 1 s.d. uncertain regions for both datasets overlap).

We have clarified this, explaining that ‘biases between the satellite retrievals and the flask observations were not significantly different from zero’.

Eq(1): Why are the numbers in bold, as well as the brackets? Please remove.

Neither the numbers nor the brackets are in bold.

Line 170: Specify that it is the inset of Figure 1.

Done.

Line 171: ‘until 2014’ should not be in parenthesis.

Done.

Line 172: Until when? Present?

Yes, we have now clarified this.

Line 178: Space needed between 500 and m

Done.

Line 190: It is worth mentioning the species here, considering it is only 4 citations.

Done.

Line 194: Are these 5.6 degree square grids? Horizontal is vague.

Yes, we have clarified this.

Line 196: Use the latex command `\citep[ERA-I][{}]{dee_reference}` here.

We have altered this citation.

Line 215: They are given 250

We have corrected all instances of 'were given'.

Line 222: What was the functional family of the spatial correlation imposed? The most defensible choice here is a Matérn covariance structure (see Stein 2012, Interpolation of Spatial Data: Some Theory for Kriging), although it seems that this is not the case here.

We think that we previously used slightly unclear language in this section, and we have now clarified. Where we previously referred to the normal distribution of the uncertainty, which could refer to the underlying assumptions of the inversion methodology and the quadratic cost function, we now state:

"Both NAT + AGR + BB and FF sectors had spatial correlations imposed between grid cells, based on Gaussian covariance functions with correlation length scales of 500km."

We hope that it is now clear that we have not imposed a covariance structure such as that suggested by the reviewer. Whilst a Matérn structure can be favourable when representing some natural correlation structures, many previous studies, similar to ours, have used Gaussian spatial error functions. Due to the short correlation length used in our study, relative to the coarse nature of our model grid, model grid cells in South America are usually correlated strongly only with cells immediately next to them. In our case using a Matérn structure would therefore change the imposed covariances only in a minor way. In the future at higher resolutions, we would consider other covariance structures such as that suggested.

Line 223: Again, probability content of this uncertainty needs defining.

See comment above.

Line 226: L-BFGS is a general minimisation routine, not a method to explicitly derive a covariance matrix in the context of uncertainty. This needs rewording.

Yes, we have reworded this as follows:

'...using output from the L-BFGS method that we employ to minimise the cost function...'

Line 228: The 'cost function' is introduced for the first time here, and needs expanding upon for readers unfamiliar with the method.

Agreed, we have added a brief description of the cost function at the end of the first paragraph of this subsection.

Line 229: This should be the other way round – the lack of off-diagonals would give smaller emissions uncertainty than expected when including off-diagonals i.e.  $var(\sum_i(x_i)) = \sum_i(var(x_i)) + \sum_{i \neq j} cov(x_i x_j)$ .

We have changed this.

Line 237: Again, use `\citep[MeMo][ref]`.

We have altered this citation.

Eq 2.: Brackets are bold, when they shouldn't be.

The brackets are not bold.

(Paragraph starting 274: This paragraph reads well and is explicit. Ideally much more of the paper should read like this.)

Okay, thank you. We have attempted to reproduce this style elsewhere.

Line 301: 'The posterior error-weighted mean residual model-satellite mismatch' needs defining. It's unclear what this is.

We have clarified this: "The posterior (prior) mean model-satellite mismatch, weighted by the observation uncertainty,"

Figure 2: Consider earlier comment about the use of the terms prior/posterior. Note that the red-blue colour bar is not colour blind friendly.

See previous comment concerning the prior/posterior terminology. We understand that a red-blue diverging colour palette is, in fact, colour blind friendly and is one of those suggested by the ColorBrewer 2.0 website, which is recommended by the journal's submission criteria for colour blind people.

Line 310: Is this the mean of the prior/posterior means?

Yes, see previous comment concerning the prior/posterior terminology.

Line 318: This should say the spatial distribution of the posterior mean.

Yes, see previous comment concerning the prior/posterior terminology.

Line 345: What correlation do you assume when you make the assumption that they are highly correlated between years?

By reporting the mean uncertainty, we implicitly assume a correlation of 1 between the years' uncertainties. We do this rather than choose a relatively arbitrary correlation value. We have stated this in the text.

Figure 5: These colours are very difficult to differentiate. I suggest revising the shadings/colours in the figures.

We have changed the shading in this figure.

Line 458: The curve fitting programme (note 'programme' if using British English) needs elaborating on. What is the programme? What curve does it fit?

We have added extra text into the Appendix regarding this algorithm. British English uses the spelling 'program' when referring to computer code. However, we have adapted to use the word 'procedure' instead.

Line 462: If talking about 'no significant trend', the (statistical) significance of this trend must be given.

We have now stated that  $p = 0.36$ .

Line 474: Does 'here' refer to Section 4? If so, this should be stated as such: "Section 4 also shows..."

We have changed this to 'Figure 8e also shows...'

Line 551: Use `\citep[e.g.][]{}{ref}`

We have altered this citation.

Line 585: An extra closing bracket is present.

This has been corrected.

Line 586: Consider 'substantial' instead of significant.

Done.

Line 598: Which other models besides Bloom et al.?

Changed to 'another model'.

Appendices Figures: Consider making these colour blind friendly.

We understand that a red-blue diverging colour palette is, in fact, colour blind friendly and is one of those suggested by the ColorBrewer 2.0 website, which is recommended by the journal's submission criteria for colour blind people.