Reviewer #2

Comments [2-1]: "Global methane budget and trend, 2010–2017: complementarity of inverse analyses using in situ (GLOBALVIEWplus CH4 ObsPack) and satellite (GOSAT) observations" presents long-term global inversions based on different available observation datasets. The authors present an inversion system based on the analytical solution of the Bayesian Gaussian problem which allow to better understand the weight of each piece in the system. The authors analyze the outputs thoroughly and use relevant comprehensive metrics to assess the usefulness of each type of observations.

The manuscript is well written, well structured and of significant importance for the community to be published in ACP after some weaknesses are properly addressed. Main problems are detailed in dedicated sections below and technical revisions are listed in Sect. 5. Overall, the manuscript is of high quality but falls short of properly exploiting the full potential of the system presented here. Sensitivity tests and additional inversions should be added to the manuscript (without computing additional response functions) to prove fully relevant to the community and to stand out of more regular inversion papers. It can be done with relatively little efforts considering all the material and the quality of the background work done to reach the present submitted manuscript.

Response [2-1]: We thank the reviewer for the positive and valuable comments. All of them have been implemented in the revised manuscript. In particular, we have performed a number of additional inversions to test the sensitivity of our results to the choices in cost-function construction (e.g. usage of observations, error assumption of the observations and state). Please see our itemized responses below.

Comments [2-2]: 1 Bias correction: p.7 l.191: Bias correction is mentioned. This is a critical point. It may have a huge impact on the inversions. Putting it under the carpet in one line is a little bit short. Please add details on this aspect and possibly some quantification of the impact of such a bias correction. Is the bias correction put in the constant c in eq. (2)? Or is it use on-line in the computation of GEOS-Chem? Or posterior to it? What is the impact on the response functions? If it is the constant c, please include (at least in supplement) your results with/without/with another bias correction to really see how sensitive your results are to that aspect.

Response [2-2]: Thanks for pointing it out. The bias correction is done off-line before the inversion. We have added the text briefly describing the procedures for bias correction, and a Figure S1 to show the influence of bias correction. We now state in Section 2.3 <u>"GEOS-Chem has excessive methane in the high-latitudes stratosphere, a flaw common to many models (Patra et al., 2011) especially at coarse model resolution. Following Zhang et al. (2020), we compute correction factors to GEOS-Chem stratospheric methane subcolumns as a function of season and equivalent latitude to match the measurements from the solar occultation ACE-FTS v3.6 instrument (Waymark et al., 2014; Koo et al., 2017). As shown in Zhang et al. (2020), the correction can be up to 10% at high latitudes during winter and spring. We apply the correction factors before the inversion to avoid wrongly attributing this model transport bias to methane emissions and loss. Figure S1 shows that the systematic differences in the posterior scaling factors of non-wetland emissions with vs. without bias correction are more prominent at the northern high latitudes, as also shown in Stanevich et al. (2020), but the global total emissions only differ by 1%. "</u>

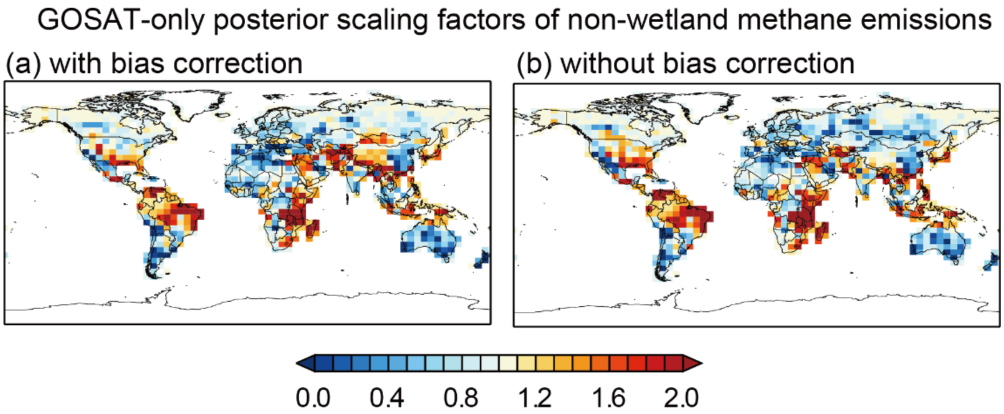

Figure S1. Posterior scaling factors of non-wetland methane emissions from GOSAT-only inversion (a) with GOSAT stratospheric bias corrections and (b) without GOSAT stratospheric bias corrections.

Reference:

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.

Zhang, Y., Jacob, D. J., Lu, X., Maasakkers, J. D., Scarpelli, T. R., Sheng, J.-X., Shen, L., Qu, Z., Sulprizio, M. P., Chang, J., Bloom, A. A., Ma, S., Worden, J., Parker, R. J., and Boesch, H.: Attribution of the accelerating increase in atmospheric methane during 2010–2018 by inverse analysis of GOSAT observations, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-964, in review, 2020.

Comments [2-3]: 2 Non-linearity of GEOS-Chem and OH chemistry. This is a little bit harsh to neglect it straight away. Could you run forward runs with your different posterior states and compare with what you get with the matrices Kx to have an idea of how negligible it is?

Response [2-3]: The GEOS-Chem methane simulation used prescribed monthly 3-D fields of global tropospheric OH concentrations taken from a GEOS-Chem simulation with full chemistry. With this regard the optimization of methane emissions is strictly linear. The only non-linearity emerges regarding the optimization of OH, because the sensitivity of the methane concentration to changes in OH concentrations depends on the methane concentration through first-order loss, but the variability of methane concentration is sufficiently small so that this non-linearity is negligible. We have tested that the $\mathbf{K}\hat{x}$ and posterior simulation of y has a small mean difference of 2±3 ppbv. We now state in Section 2.4 <u>"The optimization of methane emission and its trends is strictly linear by design because we use prescribed monthly 3-D OH fields as described in Section 2.2. There is some non-linearity regarding the optimization of OH, because the sensitivity of the methane concentration to changes in OH concentrations depends on the methane concentration through first-order loss,</u>

but we assume that the variability of methane concentration is sufficiently small that this non-linearity is negligible (we verify this assumption below)…. Comparison of the resulting Jacobian matrix to GEOS-Chem as *F*(*x*) – *Kx* - *c* shows a negligible residual difference of 2±3 ppb, verifying the assumption of linearity."

Comments [2-4]: 3 Regularization term: The authors use a regularization term to correct for ill-specified observation errors. However, their estimation is based on approximate matrices. Why not using the rigorous Chi-square criterion? such as in Desroziers et Ivanov (2001, https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.49712757417)

Response [2-4]: Thanks for pointing it out. We have made the revision to estimate the optimal value of the regularization parameter in the context of the Chi-square distribution. We have also tested the impact of using different regularization parameters on the global methane budget as discussed in [Response #2-5].

We now state in Section 2.4 "… For a given state vector element *i*, the expected value of $(x_i - x_{Ai})^2$ is the prior error variance $\sigma_{Ai}{}^2$. For an *n*-dimensional state vector with a diagonal prior error covariance matrix, the state component J_A of the cost function is the sum of *n* random normal elements

$$J_A(x) = (x - x_A)^T S_A{}^{-1}(x - x_A) = \sum_n \frac{(x_i - x_{Ai})^2}{\sigma_{Ai}{}^2} \quad (6),$$

and its pdf is given by the Chi-square distribution with *n* degrees of freedom (*n*=3378 in this case), with an expected value of *n* and a standard deviation of $\sqrt{2n}$. One can apply the same reasoning to the observation component J_O of the posterior cost function,

$$J_O(x) = (y - Kx)^T S_O{}^{-1}(y - Kx) = \sum_m \frac{(y_i - Kx_i)^2}{\sigma_{Oi}{}^2} \quad (7),$$

whose pdf follows a chi-square distribution with *m* degrees of freedom. However, this component is less sensitive to the choice of γ because of the large random error component for individual observations.

Figure 4 shows the dependences of $J_A(\hat{x})$ and $J_O(\hat{x})$ on the choice of the regularization parameter γ, for the in situ and GOSAT observations. The in situ observations are sufficiently sparse that γ = 1 (no regularization) is expected. In the case of GOSAT, however, γ = 1 would yield $J_A(\hat{x}) = 6n \gg n + \sqrt{2n}$ which indicates overfitting, while γ = 0.1 yields $J_A(\hat{x}) \approx n$ which is the expected value and is used here…."

Comments [2-5]: 4 Computation cost and sensitivity tests. It is nowhere stated what is the computation cost of the system (computing response functions on the one hand, solving the matrix products on the other hand). Once the response functions are computed it is in principle quite straightforward to change parameters in the R/B matrices to see the impact.

I think the main strength of the system presented here comes from this very fact (otherwise, a variational inversion would give posterior fluxes at reduced cost, even if DOFS can be retrieved easily). This is a critical limitation of the present paper.

Different horizontal and temporal correlations should be tested in the prior matrix, as well as standard deviation of errors, to see the impact of such modifications, given that we never really know how good are our prior/obs errors.

More critically are observation errors. Even though the observation data set is very large, it should be possible to imagine a matrix that is diagonal only by block, allowing to consider correlations between GOSAT neighbour observations, while keeping it possible to compute the inverse easily. As stated by the authors, the inversions are not consistent with each others (Fig. 13). This comes probably from ill-specified error matrices, which the authors have the tools to inquire into.

Response [2-5]: Thank you for pointing it out.

1) We have added the following text in Section 2.4 (Analytical Inversion) to clarify the computation cost of the system <u>"A requirement of the analytical approach is that the Jacobian matrix be explicitly constructed, requiring *n* + 1 forward model runs. Building the Jacobian matrix for the 3378 state vectors in this 8-year period study requires about one million core hours (8 cores × 36 hours per simulation × 3378 simulations). However, this construction is readily done in parallel on high-performance computing clusters."</u>.

2) We have also conducted a number of additional inversions to examine the results with different error assumption and ingestion of observations. For the prior standard deviation of state vectors (non-wetland emission trends and OH), we test their different magnitude (decrease by 50%) but not their distributions (correlations) due to the lack of objective information on the later. For the observation error, the ability to test off-diagonal assumption is also limited by the calculation of $\boldsymbol{S_O}^{-1}$ which involves inverting a matrix with ~10^{12} elements. Therefore we test the unknown observation error correlations by changing the regularization parameter γ.

We have added a new Figure 13b, and now state in Section 2.4 <u>"We will make use of these advantages in comparing the ability of the in-situ-only, GOSAT-only, and GOSAT + in situ inversions, and to test how choices in cost-function construction affect our conclusions including changing the regularization parameter γ, changing the prior error estimates, and using different types of in-situ observations. Our analysis will focus on results from the base inversions with the default settings, but we will use results from the sensitivity inversions to address specific issues."</u>.

And in Section 3.5 we state <u>"We examine in Figure 13b the sensitivity of the global methane budget optimization to the choice of different regularization parameter γ (and therefore observation error $\boldsymbol{S_O}$) and prior error of methane emission trends and OH concentrations. We find that reducing γ or prior errors of trend and OH by 50% yields consistent estimates of anthropogenic emissions and OH concentrations as compared to the default inversion, with differences within 3%. Decreasing the weighting of observations in the inversion (i.e. assuming larger observation error) enlarges the posterior error and pushes the posterior estimates closer to the prior estimates. Assuming a lower prior error for OH concentration from 10% to 5% results in lower methane lifetime (closer to the prior) and higher emissions, and also reduces the error correlation between the optimization of methane emissions and OH, while assuming a lower prior error for non-wetland emission trends leads to an opposite effect. Our results are consistent with Maasakkers et al. (2019), which shows that different assumptions of error distribution and magnitude tin their analyses have relatively small results. We also find that having the shipboard and aircraft measurements in the in-situ-only inversion pushes the estimate to be more consistent with the GOSAT-only</u>

inversion (Fig.13b), implying that the shipboard and aircraft measurements by emphasizing the methane in the remote atmosphere play a similar role as satellite measurements in global methane budget optimization."

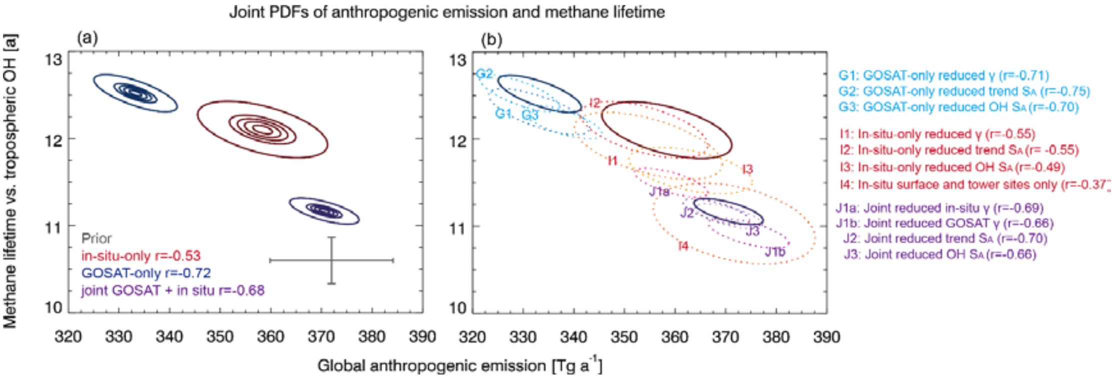

Figure 13. Joint probability density functions (PDFs) of global mean anthropogenic methane emission and methane lifetime against oxidation by tropospheric OH optimized by different inversions. Panel (a) shows the results from the prior and the three base inversions. The prior estimates are shown in grey with bars representing the prior error standard deviation. The thick contours show probabilities of 0.99 (outermost), 0.7, 0.5, 0.3, and 0.1 (innermost). The error correlation coefficients are given inset. Panel (b) shows the 0.99 probability contours from the three base inversions along with the same contours for ten additional sensitivity inversions using reduced values of the regularization parameter γ (0.05 instead of 0.1 for GOSAT, 0.5 instead of 1 for in situ); reduced errors for the methane emission trends on the 4°×5° grid (5% a^{-1} instead of 10% a^{-1}); reduced errors on annual hemispheric mean OH concentrations (5% instead of 10%); or surface and tower data only in the in-situ-only inversion.

Comments [2-6]: 5 Technical comments. p.4 l.89: aircraft measurements: those can be particularly challenging to ingest inversion systems as CTMs never really excel in representing the vertical distribution of CH_4 concentrations. Plus it is never clearly stated whether or not they are really used in the inversion or only in the posterior evaluation. Please discuss more about the aircraft measurements and justify better their use (is it only vertical profiles, very hard to assimilate? or transects, easier to use?)

Response [2-6]: Thank you for pointing it out.

1) The aircraft measurements are used in the inversions, as stated in the original text (L122-124) "We obtain in this manner 157054 observation data points for the inversion including 81119 from 103 surface sites, 27433 from 13 towers, 827 from 3 ship cruises, and 47675 from 29 aircraft campaigns.". We have added a Figure S2 to also address [Comment #1-4], which shows that the posterior model can well fit the aircraft methane measurements measuring the background (e.g. in the Southern Hemisphere), but indeed some discrepancies emerge in the northern mid-latitudes, reflecting the difficulty in modeling methane vertical distributions or optimizing emissions near source.

2) We have also added an additional inversion using only surface and tower observations in the inversion and compared the results with the In-situ-only inversion (which ingest all in situ observations) in Fig.S3 and Fig.13b. Comparison of Figure S3 to Figure 8a-b shows that

adding the aircraft and shipboard observations to the surface and tower observations increases the DOFS for constraining non-wetland methane emissions from 96 to 113 (18%), and reflects the upward correction in the South America which is consistent with the GOSAT-only inversion (Fig.8d). We also find in the Figure 13b that adding the aircraft and shipboard measurements pushes the inversed global methane and OH levels more consistent with the GOSAT-only inversion, however, it makes the inversion less effective in optimizing the global methane budget and OH. These results thus illustrate the ability of aircraft and shipboard measurements in the inversion.

We now state in Section 3.2 <u>"We find that the DOFS from the in-situ-only inversion observations are mostly (85%) from the surface and tower measurements (Fig.S3)."</u>

We also state in Section 3.5 "…<u>A sensitivity inversion using only the surface and tower measurements in the In-situ-only inversion yields *r*=-0.37 (Fig.13b). It indicates that in situ observations, in particular surface and tower measurements, are more effective than the satellite observations in constraining methane emissions independently from the sink by OH.</u>", and <u>"We also find that having the shipboard and aircraft measurements in the in-situ-only inversion pushes the estimate to be more consistent with the GOSAT-only inversion (Fig.13b), implying that the shipboard and aircraft measurements by emphasizing the methane in the remote atmosphere play a similar role as satellite measurements in global methane budget optimization."</u>

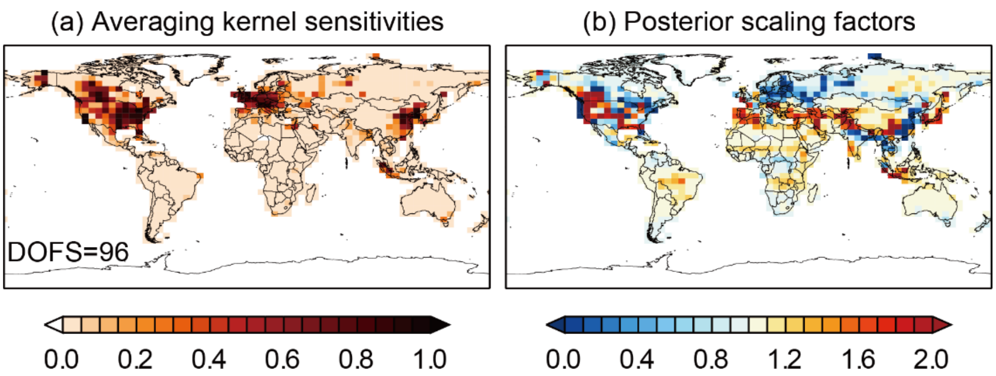

Figure S3. Same as Figure 8a and 8b but from a sensitivity inversion using only surface and tower methane observations.

Comments [2-7]: p.4 l.104: how exactly the linear trend are computed as response functions? same for OH? A start of explanation is given p.8, but additional information would be welcome.

Response [2-7]: We now state in the text to introduce the construction of response functions (Jacobian matrix K) in Section 2.4: <u>"We construct the Jacobian matrix K explicitly by conducting GEOS-Chem simulations with each element of the state vector perturbed separately. For the linear emission trend elements, this is done by perturbing the 2010-2017 emission trend in each grid cell from 0% (the best prior estimate) to 10% a^{-1}; for OH, this is done by perturbing yearly hemispheric OH fields by 20% without modifying the spatial or seasonal distribution."</u>

Comments [2-8]: p.7 l.163: What is the corresponding total error on the prior budget when using your prior distributed errors? Please represent it on Fig. 13

Response [2-8]: We have revised Fig.13 accordingly.

Comments [2-9]: p.8 l.208-213: observation error: it is not clear what ensembles are taken. Do you separate each station? Some regions for GOSAT? etc.

Response [2-9]: We now state in Section 2.3: <u>"For in-situ observations, we derive ε_0 separately for the ensemble of background surface sites (Dlugokencky et al., 1994), non-background sites, tower sites, shipboard measurements, and aircraft measurements, while for GOSAT observations ε_0 is calculated for each 4°× 5° grid cell."</u>

Reference

Dlugokencky, E. J., Steele, L. P., Lang, P. M., and Masarie, K. A.: The growth rate and distribution of atmospheric methane, J. Geophys. Res., 99, 17021, http://doi.org/10.1029/94jd01245, 1994.

Comments [2-10]: p.9 l.284: not correct. The other way around. the analytical solution is the solution of the Bayesian Gaussian problem. The cost function is derived from the formulation of the Gaussian problem when the analytical solution cannot be computed explicitly. Actually, writing the cost function in Eq. (1) in a paper using analytical inversions is superfluous; the factor gamma can be introduced differently.

Response [2-10]: We have rephrased as "The analytical solution to the Bayesian optimization problem, as done here, has several advantages relative to the more commonly used variational (numerical) solution."

Comments [2-11]: p.11 l.376: This warning should also be repeated in the method section. Actually as response functions are computed for each pixels individually, why not duplicating the corresponding time series to separate sectors in the target vector? This would not add new response functions to compute and allow you to assess how good is the distribution in sectors. You could even imagine specifying different correlation lengths to different sectors.

Response [2-11]: We cannot separate sectors at the level of individual grid cells because they will all have the same response function. We can separate sectors for ensembles of grid cells and this is precisely what we do with the matrix *W*. We have added the following text in Section 2.4 <u>"We cannot separate individual sectors within a 4°×5° grid cell because they will all have the same response function (Jacobian column). However, we can aggregate results spatially and by sector…"</u>

Comments [2-12]: p.11 l.382: Is GEOS-Chem really suitable with very coarse resolution to constrain US emissions? the resolution is fine for background sites, but what about sites nearby emission hotspots. Representation errors will likely bias your results at such stations, making it very important to filter properly data prior to the inversion.

Response [2-12]: Thanks for pointing it out. We agree that representation errors will likely bias results at stations near source regions, and it is important to filter properly data prior to the inversion. As already mentioned in Section 2.1, we address this problem by <u>"For surface and tower measurements, we use only daytime (10-16 local time) observations and average them to the corresponding daytime mean values. We exclude outliers at individual sites that depart by more than three standard deviations from the mean."</u>. Still this might be insufficient

to properly interpret sites nearby emission hotspots. A high-resolution inversion (e.g. Turner et al., 2015; Sheng et al., 2018) would be preferable to better interpret the in-situ observations near emission hotspots and to understand the spatial pattern of US anthropogenic methane emissions.

Reference:

Sheng, J.-X., Jacob, D. J., Turner, A. J., Maasakkers, J. D., Sulprizio, M. P., Bloom, A. A., Andrews, A. E., and Wunch, D.: High-resolution inversion of methane emissions in the Southeast US using SEAC⁴RS aircraft observations of atmospheric methane: anthropogenic and wetland sources, Atmos. Chem. Phys., 18, 6483-6491, http://doi.org/10.5194/acp-18-6483-2018, 2018.

Turner, A. J., Jacob, D. J., Wecht, K. J., Maasakkers, J. D., Lundgren, E., Andrews, A. E., Biraud, S. C., Boesch, H., Bowman, K. W., Deutscher, N. M., Dubey, M. K., Griffith, D. W. T., Hase, F., Kuze, A., Notholt, J., Ohyama, H., Parker, R., Payne, V. H., Sussmann, R., Sweeney, C., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data, Atmos. Chem. Phys., 15, 7049-7069, http://doi.org/10.5194/acp-15-7049-2015, 2015.