Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-786-RC1, 2019

Atmospheric Chemistry and Physics
Open Access
Discussions

Interactive comment on "Characterizing model errors in chemical transport modelling of methane: Using GOSAT XCH_4 data with weak constraint four-dimensional variational data assimilation" *by* Ilya Stanevich et al.

Anonymous Referee #3

Received and published: 12 December 2019

The authors aim at using the information contained in GOSAT XCH4 data to correct biases in in CH4 concentrations simulated by GEOS-Chem and improve the comparison with independent data such as TCCON XCH4, NOAA surface CH4 from flasks and HIPPO data.

General comments

The effort at assessing the errors in the (chemistry-)transport model before going into flux inversions is necessary and the issue remains too rarely treated adequately in inversion papers. It is therefore a very good idea to present a methodology to tackle this issue. The explanations of the methodology and most of the interpretations of the results, regarding mainly biases in the transport, are clear and interesting. Nevertheless, parts of the study seem convoluted and misleading:

1. what is the purpose of the "comparison" to the so-called "strong constraint" (SC) inversion? SC is said to consider that the model is perfect but this is not true since the covariance matrix R can contain errors in the model as well as errors in the measurement, representativity errors, etc. Moreover, it is always possible to invert transport variables/parameters along with the emission fluxes and boundary conditions - which may require assimilating more data, as is suggested but in the case of WC and only for glint measurements. Finally, the comparison between two data assimilating systems would be "fair" and useful only if the same information is provided to both. Here, it is not very clear what the configuration of the SC inversion is but it seems that it includes a lot less information than WC: for example, WC uses data to estimate the noise on GOSAT XCH4 (p. 9) but it is not stated how this knowledge is used in SC. I would suggest to simply drop this "comparison" between SC and WC: the inconsistency between assessing errors in the model (WC) and inverting fluxes without taking these errors into account (as seems to be the case in SC) is too strong.
2. what is the final aim of the characterization of the model errors? Is it to improve the model? In this case, the study lacks suggestions on how to do this (change convection schemes? use only high resolutions?). Is it to gain insights on how to invert fluxes? In this case, two questions are to be answered, which are not discussed as such in the paper:

(a) are the model errors too large compared to the errors due to the emission fluxes to allow meaningful inversions of the fluxes?

(b) if inversions are to be run with the model as it is, how can the errors in the model be taken into account so that the optimized fluxes are actually meaningful (e.g recommendations on the building of the R matrix for a so-called SC set-up)?

3. what does the proposed methodology bring to the assessment of the model errors? In the discussion of the results of this study, it is not clear what it brings compared to previous studies: the errors assessed in this paper seem to be already well-known through other methodologies.

Specific comments

General

Beware of the use of 4D-Var: it should be used only in cases when the problem is actually 4 dimensional. This is the case in meteorology, where the problem is on the initial conditions. Usually, it is not the case for flux inversions, which are problems on the boundary conditions: the relation between the 3D maps of fluxes through time is not taken into account in the model (only the relation of concentration fields through time is). It would be best to use only "variational" and drop the 4D in these cases of flux inversions.

Abstract

- p.1, l.6-7: "capable of differentiating the vertical distribution of model errors": what does differentiating mean here?

- p.2, l.2: "indicating the presence of resolution-dependent model errors": do you mean errors directly linked to the parameterizations in the physical core of the model or representativity errors also?
- p.2, l.4: "However, a major limitation of this approach is the need to better characterize the specified model error covariance in the assimilation scheme": this is true for all data assimilating systems in our domain. What does this study bring to this issue?

Section 1 Introduction

- p.2, l.21-22: "the impact of biases in chemistry and transport are often neglected": this is not exactly true, a lot of studies try and deal with biases in various ways e.g debiasing previous to the inversion itself, specifying adapted R matrices, etc. Please look deeper into the available studies.
- p.2, l.32: " In contrast to the traditional "strong constraint" (SC) 4D-Var method, the WC scheme does not assume that the model is perfect": same remark as above, the so-called SC method does not imply that the model is assumed to be perfect, it implies that all errors are taken into account in the R (and B) matrix so that what is not is "perfect".
- p.3, l.16-17: "Highly accurate aircraft CH_4 profile measurements would be an ideal source of information, but they are limited in space and time.": what about aircores?

Section 2 Data and Methods

- p.4, l.11-12: "due to the dependence of wetland emissions on the meteorological fields": do you mean that wetland emissions are actually recomputed from the references provided above with the regridded GEOS meteorological fields? Or computed on-line in GEOS-Chem?

- p.4, l.21: "the period of 1 February 2010 to 31 May 2010": this is a very short period of time, which does not allow for one full seasonal cycle for mid-latitudes. Why not work on a full year?
- p.4, l.22-25: "5.5 years ... initial condition for the analysis period": this seems a bit convoluted. Why not run a spin-up of about 9 years (life time of CH_4) OR assimilate data to obtain initial conditions?
- p.5, l.22-25: "The use of the alternative CO_2 fields did not change any of the findings about model errors in our study. There may still be unidentified biases in both retrieval products. However, the fact that both CO_2 fields were obtained using different methodologies gives us confidence in our results." Were both fields used for all the following results? But only one set of results is presented and there is no comment about a sensitivity test. Also, why not use the default field? How was it not satisfying?
- p.5, l.30-31: "However, we expect vertical structure to emerge from atmospheric transport patterns": what about the OH field patterns (e.g its vertical structure)?
- p.6, l.4-6: "Such precision could be enough in many regions of the world to improve knowledge about CH_4 a priori surface emissions. However, the presence of potential model errors significantly undermines this assumption." This is not clear: which regions? Improve how? By decreasing the uncertainties by how much? What would be the required ratio between the model errors and the precision and the expected improvement of knowledge?
- p.6, l.19-23: why use only flasks and not continuous measurements?
- p.7, l.2: "Retrievals are bias corrected based on comparisons with calibrated aircraft and AirCore profiles." Which aircraft profiles? Be sure not to use them in the assimilation if the validation data is to be kept totally independent. Why not use also aircores (others than the ones used by TCCON) as

validation data?

- p.7, l.27: "the adjoint forcing commonly used in 4D-Var": please clarify what "adjoint forcing" means.
- p.7, l.21-22: "This is the assumption that is employed in standard 4D-Var, which is also referred to as "strong constraint" 4D-Var because the model trajectory is used as a strong constraint in the optimization." Same remark as in the General comments about the use of "4D-Var": the concentrations are linked through time but not the fluxes.
- p.7, l.29-30: "The 4D-Var problem to estimate surface emissions is transformed into a 3D sources and sinks estimation problem": same remark as in the General comments: flux inversions are generally not actual 4D-Var and the cost function used is the same as what is shown after, but for the Q term. Does anything prevents Q from being included in R or in B? Please clarify the mathematical and technical differences between SC and WC.
- p.9, l.9-14: how long does it take to run for practical cases?
- p.9, l.18-23: "For each WC inversion ... about 10 ppb": this is a good idea but the validation data are not independent anymore since information contained in them is actually used in the inversions. Please clarify how this issue is dealt with when comparing to the data for validation.
- p.10, l.3: "Therefore, we did not attempt to characterize global pattern of model errors on shorter time scales": even with a period of three days for GOSAT data, information is available at shorter time scales e.g. in the GOSAT data and in the meteorology. Would this not make it possible to characterize patterns of model errors at shorter time scales?
- p.10, l.4-6: "Little is known about the a priori structure of the model errors": this is not what appears from discussions and references cited afterwards: at least some elements such as the role of the horizontal resolution or the

tropo-strato gradient are known. Is this information not usable in the inversion framework?

- p.10, l. 9-end and p.11, l.1-2: the issues described here are the same as when building the R and B matrices in the so-called SC case. Therefore, the exploration of the nature of errors in the modelled CH_4 uses the strong assumption of a diagonal Q. What are then the advantages of this methodology compared to the usual R and B matrices in the so-called SC, with various set-ups for R for example?
- p. 11, l.9-11: "Therefore, we considered a uniform structure for Q to be a satisfactory assumption for this initial assessment of model errors in the context of the WC 4D-Var analysis": is this statement justified by "expert-knowledge"? Is it not possible to design sensitivity tests to assess the impact of this strong assumption?
- p.12, l.9: "we believe that the OSSEs should reveal the best performance of the WC method": it is a bit dangerous to show the performances of a methodology only in best-case scenarios since the application to realistic cases may show the tool to be very limited.
- p.12, l.15: "for the real world applications, we expect less extreme model errors": would the method proposed here be able to characterize the errors if they are smaller? See also comment above.
- p.12, l.19-21: "Here, we intend to investigate the performance of the measurements and the assimilation method when no information is given about the sources and magnitude of model errors": to which configuration does this sentence refer to? Is it really useful to assess the performances in a case that is almost never implemented in actual inversions (even though the taking into account of model errors is never thorough)?
- p.12, l.21-22: "We also conducted SC 4D-Var assimilation experiment for comparisons with the WC approach in the OSSE with biased surface emis-

sions": see General comments. Without more details, it is not possible to understand the differences between WC and SC. If SC contains less information than WC, the comparison is not really meaningful and interesting.

Section 3 Results

- p.13, l.26 - p.14, l.4: how do the sensitivities in the model compare to the errors in the actual GOSAT data?
- p.14, l.7-8: "when using the SC method, we implicitly supply the assimilation with knowledge about the source of the bias": how?
- p.14, l.11-12: "Due to weak vertical sensitivity of the pseudo-data, it is difficult for the WC 4D-Var method to mitigate strong localized vertical bias." The low sensitivity is a problem for data assimilation in general. What is particular to WC here?
- p.14, l.13-14: "Instead, it compensates for the bias by applying relatively weak CH_4 state adjustment of the opposite sign in the column of the atmosphere above, particularly in the stratosphere": what are the consequences of the creation of such a dipole?
- p.14, l.15-20: the additional information explicitly provided to WC could also be used in SC in the covariance matrices R and B.
- p.14, l.33-34: "GOSAT retrievals possess sensitivity to biases in vertical transport and can distinguish...": this is not a property of the GOSAT data as such but of the whole data assimilation framework. Maybe "GOSAT retrievals contain information on the vertical transport, which can be used in our set-up to distinguish..." would be clearer.
- p.15, l.11-12: "we do not expect chemical biases to be as strongly localized as the biases associated with emissions and vertical transport": why?

- p.15, l.30-32: "The perfect observing system would completely remove the initial condition bias at the start of the assimilation period (on February 1). However, what is shown on February 1 is just an 8% reduction in the bias in each of the eight regions, relative to the a priori, with the rest of the bias propagated onto the assimilation period". What are the consequences of this?
- p.16, l.3-4: " This suggests that additional vertical correlation between forcing terms in the stratosphere would be beneficial to accelerate convergence in the stratosphere." Why not test the sensitivity to such correlations?
- p.16, l.11: "residual high latitude bias, which resembles noise or bias in the GOSAT observations." Is this due to the period of interest being winter in the Northern hemisphere?
- p.16, l.19-20: "the SC assimilation leaves significantly larger residual biases." What is the impact of these on the optimized fluxes and on the uncertainty reduction?
- p.17, section 3.2.1: the TCCON and NOAA data are not actually independent from the WC (see above). How is this dealt with? If they are actually independent from the SC, the comparison between SC and WC against the fit to these data is not meaningful.
- p.18, l.11: "similar errors": similar to what? The sentence is not very clear to me.
- p.18, l.9-13: these sentences seem a bit convoluted. Is the idea to state that a station such as IZO sees the upper troposphere and is best compared to the model's upper troposphere?
- p. 18, l.27: what about the impact on GOSAT retrievals of the still long nights around the North pole during the period of interest?

Section 4 Discussion of Model Biases

- p.19, l.22-23: "a stratospheric bias introduced in the system through the initial conditions ": is there no other way than the ICs to introduce a bias in the stratosphere?
- p.19, l.29-30: "it shows that the SC assimilation attempts to correct the positive high-latitude stratospheric CH_4 bias at the expense of surface emissions": this seems to be a direct consequence of the set-up of the SC inversion. More details on this inversion are required to discuss its results.
- p.21, l.6-14: can you deduce some recommendations for improvements in the model? If not, how can this knowledge be used for setting up inversions for fluxes?
- p.21, l.15-23: same questions as above: recommendations for improvements in the model? Information for flux inversion?
- p.22, l.25-26: what could be concluded? Recommendations for improvements in the model? Information for flux inversion?
- p.23, l.13-14: "Finally, the results strongly suggest that the WC assimilation and the GOSAT observations have the potential to diagnose transport errors at both model resolutions" What would prevent the assimilation of data to bring information on errors at any resolution (if the data are relevant)?

Section 5 Conclusions

- p.24, l.1-2: " However, characterizing these correlations will be challenging": is it possible at least to design sensitivity tests to explore various possibilities for these error correlations?
- p.24, l.5-6: "Initial comparisons suggested that GEOS-Chem was affected by biases not solely related to discrepancies in surface emissions." This is

not very informative as it is well known and flux inversion would be relatively straightforward otherwise.

- p.24, l.13-14: "Meanwhile, the results showed that running the a priori model at 2x2.5resolution produced better agreement with TCCON observations than the a posteriori fields from the SC 4D-Var surface emission optimization at 4x5." What can be deduced from this result?
- p.24, l.31-32: "glint measurements": a lot of other data could be assimilated, not only satellite data related to the concentrations of a given species.
- p.25, l.2: "if the model were assumed to be perfect, as is the case in SC 4D-Var": see General comments: this is not so simple.
- p.25, l.7: "Potentially, any CTM may be improved if the signal from the surface emissions can be separated from other model errors." This is only relevant if the objective is to optimize fluxes. A given CTM and set up may work very well with errors compensating one another for other objectives (e.g forecasting). Be more specific on what improvements could be made, how and to which purposes.
- p.25, l.14: "regional-scale analysis at higher spatial resolution": how in practice? The link to what it is used for in this study is not plain for me.

Technical corrections

- p.10, l.33: "the errors in the model CH_4 simulation" -> the errors in the modelled CH_4 concentrations?
- p.11, l.19: "The sensitivity of the GOSAT observations to the modelled state" -> The sensitivity of the equivalent of GOSAT observations to the modelled state
- p.15, l.13: "in the Fig.8" -> remove "the"

- p.18, l.10: "inland" -> island?
- p.18, l16: homogenize including/excluding Sodankyla.
- p.22, l.9: "lofted" -> lifted

Tables and Figures

- Tables 1, 2: please use consistent names for the columns in the legend and table.
- Fig. 4: the black boxes are not so easy to see: maybe use a very different color (try pink?)
- Fig. 15: do not repeat in the legend what appears in the graph itself so that is it easier to read and may be kept on one page.
- Fig. 16: reduction (third column) may be easier to evaluate in % and absolute value.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-786, 2019.