“This review was prepared as part of graduate program course work at Wageningen University, and has been produced under supervision of Prof Wouter Peters. The review has been posted because of its good quality, and likely usefulness to the authors and editor. This review was not solicited by the journal.”

The authors are grateful for the comments made on the manuscript and have discussed with Wouter Peters the concerns raised by the contributor. Wouter has agreed the major suggested changes are not required. Below, in brief, we address these concerns and their relevance to this manuscript. Whilst several misunderstandings are addressed below several comments are constructive and will be included in a revised version of the manuscript with the response to the official reviews.

REVIEW

title: Quantification of methane emissions from hotspots and during COVID-19 using a global atmospheric inversion
author: McNorton et al.
DOI: https://doi.org/10.5194/acp-2021-1056

OVERVIEW

With this paper the researchers want to validate a method to quantity methane emissions. They used inverse modelling of in-situ and TROPOMI satellite observations to quantify methane hotspot emissions, to subsequently compare them with results of existing case studies. A forward model is used to translate previous methane emissions into atmospheric methane concentrations. Then they applied a 4D-Var Inverse model to detect methane emission hotspots based on methane concentration forecasts from IFS.

Overall, the paper is very comprehensive and the research is conducted well. In particular, the forward model seems to work correctly as there is no (large) overestimation of emissions compared to previous results. According to Cheewaphongphan et al. (2019), overestimation of inverse modelling (top-down) methods can be largely explained by errors in emission estimations of bottom-up approaches, like the forward model in this study. The comparison of their own results with the existing case studies is very elaborate and documented in a structured way. The separate page of figures for every part of the results makes it convenient to interpret and to compare between different case studies.

Nonetheless, I have my doubts about the novelty of this research. There are many examples of other papers using the same method to model methane emissions, see my major arguments below. On the bright side, what is very interesting about this paper specifically is that it compares emissions for a regular year with a year of COVID-19 pandemic slowdown. The researchers expected that the effect of the global slowdown can be compared with mitigation strategies to decrease greenhouse gas emissions, but during the COVID-19 slowdown period methane emissions continued to rise.

In my opinion, a sufficient explanation of this unexpected result is missing. Without this explanation, it looks like the aim of the research – which they described in the introduction – has not been achieved. Throughout the paper, it is more common for assertions to be made without sufficient or unclear explanations. Therefore, some added analysis and detailed explanations are needed.

If the proposed changes and additional explanations are adequately incorporated into a revisited version, l would recommend this paper to be published in the journal of Atmospheric Chemistry and Physics.

As mentioned above, Wouter Peters has confirmed to the authors that the major changes suggested are either not relevant to improving the manuscript or are beyond the scope of the work. Therefore, whilst we have updated the manuscript based on several minor comments the major comments will not be addressed for the reasons given below.

MAJOR ARGUMENTS

1. First of all, in this paper a 24-hour data assimilation window is used in the Inverse method. Another paper about implementation of four-dimensional data assimilation stated that "4D-Var using a 6- or 12-hour window performed better than 3D-Var over a 2-week assimilation period, whereas 4D-Var using a 24-hour window did not" (Rabier et al., 2000). This means using a 24-hour window is no improvement over already existing methods, so there should be a strong reason why the paper under review isn't able to use this shorter window. Barré et al. (2021) also performed 4D-Var, but here they did use a 12-hour window. From the paper under review it is not clear why they did not use this shorter window.

Using a larger 4D assimilation window means that data is combined over a larger time span. Apparently, a 24-hour window is too large and thus not specific for one moment in time anymore. By averaging over a too long period, the time component is eliminated which makes the method move more towards a 3D method. But for a good functioning 3D method, this 24-hour data assimilation period is in turn way to short.

Consequently, all results of the paper could be affected by the choice of window length, since more averaging possibly means less accurate detection of methane emissions peaks. The researchers must have a strong reason why to use a 24-hour assimilation window, because it is unlikely to yield the most reliable results. Therefore, I argue that it is necessary to perform all the model simulations again using a shorter assimilation window, unless the researchers can provide a very solid explanation on why that is not possible.

The manuscript of Rabier et al. (2000) focuses on 4D-Var for meteorological applications and not for greenhouse gases. The non-linear nature of meteorological processes demands a short-window length, whereas greenhouse gas inversions are usually performed using offline transport and using much longer window-lengths (monthly to yearly). In this study the window-length is relatively short compared to most global GHG inversions. A limitation of the study is in fact the window length being relatively short and not being too long. As discussed in the manuscript future developments aim to extend this window length without interfering with the short-window required for meteorology. As shown in the results, the short-window length limits the number of days where inversions can be performed for specific case studies resulting from sparse observations. A further reduction in the window-length would decrease the observational constraints further and therefore be unsuitable.

2. My second issue is that I doubt about the novelty of this work. The paper of Barré et al. (2021) claims to be novel with a method that is highly similar to the method from the paper under review: it also used observations collected by TROPOMI and IFS methane forecasts produced by CAMS. The use of 4D-Var systems is not new, since Rabier et al. (2000) already described operational implementation of 4D-Var assimilation. A quick search yielded more papers that used a 4D-Var System to model methane emissions (e.g. Meirink et al., 2008; Van Peet et al., 2021; Yu et al., 2021).

However, new about this paper is the comparison of methane emissions between the COVID-19 and pre-COVID-19 situation. Most of the paper, though, is about testing the performance of their method and not the aforementioned comparison. If this comparison is supposed to be the aspect that distinguishes this paper from others, details and explanations are lacking in the conclusion (see also my next argument).

To overcome this issue of novelty I provide two options:

(1) If the novelty of the paper is about differences between the method of the paper under review and the method of other papers using 4D-Var systems, this has to be proved by explicitly pointing out these differences (at least with papers: Barré et al., 2021; Meirink et al., 2008; Van Peet et al., 2021; Yu et al., 2021). Now, it is not clear that there are obvious differences in methodology.

(2) If the novelty of the paper is about the comparison between the COVID-19 and pre-COVID-19 situation, the conclusion section needs to be rewritten. The presented results are not enough supported by argumentation. The results show methane emissions in 2020 to be higher than expected, but a solid explanation is missing. There has to be either a physical explanation or limitations of the method plays a role.

Concluding: one of these two scenarios has to be the case and need to be fixed as explained above. If not, then I strongly doubt about the novelty of this paper.

4D-Var inversions are indeed well studied and the methodology available in the literature is extensive. Our manuscript does not claim to be the first study to perform 4D-Var inversions for CH_4. However, we believe it to be the first study to perform inversions at high-resolution on a global scale using online transport and TROPOMI data. In particular, an online transport configuration allows better representation of the transport error since uncertainties in the initial meteorological fields are accounted for. Further novelty comes from the cases and time-period of study as mentioned by the contributor. In reference to Barré et al. (2021), it is a very different study in that they perform forecast experiments and look at observation-model departure statistics to detect potentially anomalous emissions. Their study is not a 4D-Var inversion and the only major similarity is the use of the IFS system. In the manuscript the evolution of 4D-Var inversions of CH_4 is discussed and several previous studies are mentioned which have informed the developments presented.

3. Thirdly, I am concerned about the lack of attention paid to the atmospheric sink of methane. The paper of Maasakkers et al. (2019) used a method that I think can be seen as the precursor of the method under review, but in addition it did include an OH depending changing sink for atmospheric methane. A changing sink means that for every assimilation step new atmospheric OH concentrations are used. The amount of methane oxidation is determined by the NOx – OH – CH4 reactions, which means that the available amount of OH for methane oxidation depends on NOx emissions (Stevenson et al., 2021). I think this atmospheric chemistry is very interesting during the COVID-19 period, because reduced NOx emissions induced by the global slowdown thus probably lead to an increase in methane concentrations. In contrast, the paper under review used a constant climatological OH sink and thus the strength of the atmospheric methane sink is not depending on atmospheric chemistry.

Unexpected are the results that suggest that the methane emissions are higher in the COVID-19 situation than the period before. In the reasoning of the paper, they referred to two different sources. Stevenson et al. (2021) shows that the results of these higher than expected methane emissions can be explained by the use of a changing OH depending sink in the atmosphere, because this over time decreasing sink is strong enough to explain the increase in methane emissions. The paper of Weber et al. (2020) suggest that the effect of this OH sink is too small to explain the excess in methane emissions.

The paper under review does not decide which of these two contrasting result is the most likely. Independ of the possible effect of an OH based changing atmospheric sink, the researchers did not manage to fully explain the higher than expected methane emissions. Because, according to the aim of the research, this is the most important conclusion, I think there is a need to share their opinion about the strength of an OH depending atmospheric sink. If needed, they have to perform additional analysis to quantify the effect of a changing sink. If they doubt this effect to be large enough to explain the higher than expected methane emissions, they definitely have to come up with other possible reasons. Without a plausible explanation, the main result of the research is not sufficiently supported.

We believe the concern raised here is partly valid, although it is already addressed and the challenges of including OH are discussed in the manuscript. Including varying OH either requires online chemistry which would increase the computational cost of the system significantly, making it no longer viable to run for extensive periods. Alternatively, an offline OH field could be used for 2019 and 2020 based on full chemistry forward simulations; however, the uncertainty of the derived OH fields is considerable (see contrasting results from Stevenson et al. 2021 and Weber et al. 2020). As discussed in the manuscript we believe the influence of OH variability during the pandemic may have influence CH_4 lifetime and we include reference to a recent study using the IFS to investigate these changes. Including, and even remarking on, such changes is beyond the scope of this study; however, it will be investigated in the future. Several explanations for the trend in emissions are indeed mentioned in the manuscript, however remarking on human activities in detail would be unwise given the focus of the study is to investigate where and when emissions change and less so on socio-economic drivers of the changes.

MINOR ARGUMENTS

1. The reduced anthropogenic activities because of the COVID-19 pandemic, gave the researchers the possibility to look into the effects of potential climate mitigation strategies to decrease greenhouse gas emissions (Diffenbaugh et al., 2020). The paper and the source referred to does not explain why the researchers believe this to be a legitimate analogy. It seems a reasonable comparison, but in the conclusion it appeared not to be true: the methane emissions increased during the COVID-19 period, whereas the purpose of climate mitigation strategies is to reduce these emissions. In the paper, no further reflection is made on this comparison. I recommend explanation for this, as it is an interesting question why this analogy not seems to work.

As discussed above the socio-economic intricacies of the influence of COVID-19 are not the aim of this manuscript. The relation of COVID-19 to climate mitigation strategies focuses on a step change in emissions, these could either be a decrease or increase, either way it serves as a good testbed to explore shifts in emissions.

2. In the method section is stated that “prior emissions errors are assumed to be independent between 24-hour inversion cycles”. This assumption is said to be made because not enough is known
about temporal error correlations. If this assumption is not valid and errors are dependent, there may be biases in the results, especially when the period coved by the model is extended. This extension of the modelled period is exactly wat is done with the comparison between the pre COVID-
19 and during COVID-19 period. It would be good to indicate how bad this assumption is expected to
affect the results by applying a range of error correlations and compare these results to the original
results.

With no available estimate of these correlations, various assumptions would be speculative. This also ties into the point below.

3. In the method section is stated that “posterior errors in methane emissions and 3D state are not propagated forward across data assimilation cycles”. This shortcoming is said to be a technical
limitation of the system and will be addressed in subsequent versions. This limitation can cause a
bias in the results, but the researchers did not indicate how large they expect this effect to be. I recommend to perform some additional model runs with distorted initial posterior emissions and 3D
states, to determine the effect on the final results compared to the original results.

Such analysis of error correlations is not possible with the system as described owing to limitations in the uncertainty calculations, this is mentioned throughout the manuscript. This is an area of future development. It is unclear what is meant by “distorted initial posterior emissions and 3D states”. If this is perturbing the prior with information from the posterior of the previous window then as discussed throughout the manuscript this requires substantial modifications to the system. As discussed with Wouter Peters, further simulations should not be conducted as it is beyond the scope of this study.

MINOR ISSUES

Several of the suggested minor issue changes have been uptaken and will be included in the revised manuscript.

line 38 “The change in energy and fuel demand is estimated to have reduced oil and gas CH4 emissions by 10 % for 2020 when compared to 2019 (IEA, 2021)” The reference is a figure that does
not contain information about this estimated 10% reduction. Pleas correct reference.

line 61 / line 148 / line 202 Reference to Barré et al. is wrong, year of publication is 2021 instead
of 2020.

line 69 “For this paper, the focus on CH4 emissions allows to benefit from greater observability from remote-sensing (compared to CO2)" Why is this? Please explain.

line 121 "background errors for the meteorological variables at initial time are constructed based on a climatology, and therefore are not flow-dependent" Please Explain, unclear what you
mean.

Line 126 "… at the relatively high increment resolution used (~80km) CH4 sectors are rarely collocated." Is this true? I think in a quare of 80 times 80 kilometres often multiple sectors occur.

line 132 "Globally, constant wetland uncertainties were estimated at 58%, taken as the standard deviation from the WetCHARTs ensemble (Bloom et al., 2017)." From this paper it is not
evident where the researches get this 58% from. Please explain.

line 140 "Total grid cell uncertainties, used in the control vector, were calculated with the error propagation method." Is this a common method? Explanation or reference is lacking.

line 160 "…, we perform simulations from January, when slowdown restrictions were limited to China, to June for 2019 and 2020." Period January to June in 2019 is clear why, but which period in
2020 exactly and why?

line 171 "All subsequent experiments used the mapped prior uncertainty, typically ranging from 50-150%." What is the mapped prior uncertainty? Please explain.

line 192 "we distributed total posterior emissions into 6 sector specific categories; energy, agriculture, waste, other anthropogenic, wetlands and fires". What is included in sector 'other anthropogenic'? Please explain.

line 385 “Given the limitations of our system we have typically focused on anthropogenic emissions…”. How is the system limited? Why can that be fixed by focussing on only anthropogenic emissions? Please explain.

line 459 “Future developments will adapt the system for use to constrain CO_2 emissions based on a hybrid-ensemble system that will extend the assimilation window and utilise observations of co-emitted species” What is a hybrid-ensemble system? Why is extension of the assimilation window a good thing? Please explain.

figure 1 Part B would be very useful to explain in a bit more detail how the 4D-Var system works, but in the text there is no reference to this figure. I like to see some explanation in the text about this figure.

figures 2-10 It would be good to make the link between the text and the figures a bit more clear when reading the paper, for example make it very obvious that there is one page with figures for every part of the results. In addition, the section names/numbers can be included in the description of the figures.

figures 2-10 It is not clear why for some case studies a pi-chart and overview map is provided and for others not. An overview of the setting of the case study area and a graph of the relative contributions of all the sector emissions would probably be good to have for all the case studies.

SOURCES (NOT REFERENCED IN PAPER)

Meirink, J. F., Bergamaschi, P., and Krol, M. C. (2008). Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: method and comparison with synthesis inversion, Atmos. Chem. Phys., 8, 6341–6353, https://doi.org/10.5194/acp-8-6341-2008.

van Peet, J., Houweling, S., Marshall, J., Nunez Ramirez, T., and Segers, A. (2021). Inverse modelling of global methane emissions using TROPOMI, EGU General Assembly 2021, online, 19–30 Apr 2021, EGU21-14510, https://doi.org/10.5194/egusphere-egu21-14510.

Yu, X., Millet, D. B., and Henze, D. K. (2021). How well can inverse analyses of high-resolution satellite data resolve heterogeneous methane fluxes? Observing system simulation experiments with the GEOS-Chem adjoint model (v35), Geosci. Model Dev., 14, 7775–7793, https://doi.org/10.5194/gmd-14-7775-2021.