Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2020-871-AC2, 2020

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Atmospheric methane source and sink sensitivity analysis using Gaussian process emulation" *by* Angharad C. Stell et al.

Angharad C. Stell et al.

a.stell@bristol.ac.uk

Received and published: 27 November 2020

We would like to thank the reviewer for their helpful comments. In this document, we reply to each comment, providing extra detail and outlining how we have updated the manuscript.

My only major suggestion is that perhaps the observational dataset, currently just used to show that the CTM simulations encompass realistic values, could be incorporated into evaluation of the emulator simulations. For instance, looking at Fig. 4, the CTM appears to underestimate the observed global mean δ^{13}C-CH$_4$

Printer-friendly version

Discussion paper

value considerably (panel b). Could the emulator simulations be used to posit the drivers of the CTM underestimate? I understand that it would be unreasonable to meaningfully look at millions of simulations one-by-one, but perhaps the optimal values of the largest drivers of global mean δ^{13}C-CH_4 (from Fig. 7b) could be identified? I.e., which combinations of inputs are needed to close in on the observed global mean δ^{13}C-CH_4? This could be done for all the observed metrics shown in Fig. 4, if sorting through the emulator simulations to find observation-matching values is feasible.

This is a very similar point to the major comment by Reviewer 1, and one to which we gave a great deal of consideration before submitting the manuscript. As we wrote in our response to Reviewer 1:

We decided against going down this route because we felt that the most effective way to combine model sensitivities (in this case derived using Gaussian process emulation) with observations is through a full Bayesian inverse analysis. This will require some additional methodological development (to effectively make use of the Gaussian process) and much more involved consideration of model and prior uncertainties. We felt that adding this material would make the paper long, less readable, and may take focus away from the emulation method and the sensitivity analysis, which we feel are novel and important in their own right. Therefore, we hope the reviewers will agree with our suggestion that a full inverse analysis would best be presented in a follow-up paper, which is currently in preparation.

To answer the more specific element of the reviewer's comment regarding the principal cause of disagreement with the observations, we note that the cause of the global mean δ^{13}C-CH_4 offset can be considered qualitatively using the sensitivity analysis itself. The parameters that are responsible for the largest proportion of the output variance are the δ^{13}C-CH_4 source signature of agriculture, the magnitude of the Cl loss, and the magnitude of the freshwater source. It is these parameters that the output is

most sensitive to that are most likely to be adjusted to reach the optimal solution, for example, in an inversion.

The reviewer also raises an interesting suggestion that the ensemble could be examined to find the subset that best agrees with the data. We have indeed tried such approach, e.g. "history matching" as referenced in line 443 (now line 476 in the revised paper), for example, by attempting to find some subset of the parameter space that is consistent, within some uncertainty, of the observations. However, we found that, given the high dimensionality, even with an efficient emulator, it was extremely expensive to derive a statistically meaningful ensemble from a purely random exploration of the space. Therefore, as we note in our response to Reviewer 1, we feel that the most promising approach will be a Bayesian method, which can be explored more thoroughly in a follow-up paper.

While Gaussian process emulation has not been used for study of the methane budget specifically (as far as I am aware), it was recently used to evaluate the CH4 lifetime due to loss by OH. Please see and cite Wild et al., Global sensitivity analysis of chemistry–climate model budgets of tropospheric ozone and OH: exploring model diversity, https://doi.org/10.5194/acp-20-4047-2020

We thank the reviewer for reminding of us of this important paper that should have been cited. We have added this reference to the revised version of the paper.

L93: The authors allude here to the Gaussian process inputs being maintained in their original spatial resolution. Does this mean all inputs are 2-D fields at 12x11.25 degrees resolution? Or are some 3-D? An explicit statement of exactly what is being fed into the Gaussian process emulators would be helpful, particularly regarding the inputs' dimensionality.

In terms of the input to MOZART, all input fields are interpolated to the model resolution

of 12.00° N × 11.25° W, with emissions (and the soil loss) being 2D, and the other losses being 3D.

In terms of the input to the Gaussian process, the inputs are scaling factors of these fields, i.e. the input to the Gaussian process is 28 numbers (one for each parameter), but this will respond as if the 12.00° N × 11.25° W field had been scaled.

This has been clarified in the revised paper by adding to line 129 (now line 131):

"The model input fields are 2D for sources and the soil sink, and 3D for the remaining sinks."

Additionally, the following has been added to line 201 (now line 218):

"In this work, the input parameters are the 28 scaling factors in Table 2, and the outputs are the MOZART hemispheric average mole fraction and δ^{13}C-CH_4 values."

Table 2: For the Trend values in the final column, the units are given as "%". Since trends are usually expressed as a rate, I would recommend noting the time period (I believe 2000-2012, based on my interpretation of the text) in the Table header information.

We agree, the revised version of the paper has both the time period (1996-2012) and the units as % yr^{-1} in Table 2.

L195: I would hesitate to say that the loss of CH4 by OH is linear; the abundance/loss of CH4 has a feedback on the abundance of OH (see, e.g., Holmes et al., JAMES, https://doi.org/10.1002/2017MS001196). This would likely only influence results regarding large perturbations to CH4, so may not be relevant here, but it should probably be noted.

We agree that the loss of CH4 by OH is non-linear, this is discussed in the paragraph starting line 322 (now line 341):

"The multiple linear regression accuracy can be improved by considering the non-linearity of the mole fraction with respect to the OH loss. By using a log-transformed OH parameter to estimate the mole fraction, the RMSE becomes 11 ppb (the complete residual distribution is shown in Fig. 6). Multiple linear regression using a log-transformed OH parameter still has a significantly larger RMSE than the Gaussian process, implying that the remaining small non-linearities and parameter interactions are important for predicting the output value. This finding suggests that inverse modelling studies that have assumed linear and independent sensitivities between observations and source and sink parameters may have under-estimated their posterior uncertainties."

The Gaussian process does not assume linearity, and the mean function in line 195 (now line 210) could equally be set to zero and it would perform similarly well. We have added the following to the revised paper to clarify this:

"A linear mean function does not stop the Gaussian process from being able to model non-linear relationships."

L360: It would be interesting to assess the role of altered spatial distributions of OH, both in the horizontal (i.e., more NH OH as many global models simulate) and in the vertical (i.e., what if there's more OH in the free troposphere than anticipated by Spivakovsky et al.?). It is understandable if this is beyond the scope of the current study but would make a good future direction.

We do agree that this would indeed be interesting. However, as we have noted in our response to Reviewer 1, it was not included in our emulator design, because we made the decision early on to focus on uncertain magnitudes and trends in sources and sinks, rather than spatial distributions. We have acknowledged as much in line

279 (now line 297) and have added another line to the revised manuscript when discussing the sensitivity of the interhemispheric difference to the input parameters (line 360, now line 385): "However, had the uncertainty in the hemispheric distribution of OH been included in our analysis, it would likely have explained a larger proportion of this sensitivity."

L380: "are a serious" should be "is a serious"

We agree, this has been changed in the revised version of the paper.

L388: I would be interested to see a bit more discussion regarding the freshwater source of CH4. Some context regarding what is known about these emissions (that these are distinct from wetlands, what we know about the mechanism (bacteria?), that they are perhaps close in magnitude to wetlands emissions, etc.) would be helpful to the reader without them having to refer back to Saunois et al. This is potentially a very interesting finding, and some context could help raise awareness of this issue in the community.

The following sentence has been added after line 388 (now line 419) to address this:

"Freshwater bodies emit methane by bacteria breaking down organic matter in an anaerobic environment, as in wetlands, and the freshwater emissions are potentially of similar magnitude to wetlands, but more uncertain (as seen in Fig. 1)."

NB: both "fresh water" and "freshwater" are used in several locations; I suggest maintaining consistency.

This is intentional as "fresh water" is a noun whereas "freshwater" is an adjective.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-871, 2020.

Printer-friendly version

Discussion paper