Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., author comment AC1
https://doi.org/10.5194/gmd-2021-39-AC1, 2021

Author Response to RC1

Edmund Ryan and Oliver Wild

Author comment on "Calibrating a global atmospheric chemistry transport model using Gaussian process emulation and ground-level concentrations of ozone and carbon monoxide" by Edmund Ryan and Oliver Wild, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-39-AC1, 2021

Responses by author to interactive comments by RC1 on "Calibrating a global atmospheric chemistry transport model using Gaussian process emulation and ground-level concentrations of ozone and carbon monoxide" by Edmund Ryan and Oliver Wild

(1) Reviewer's comment: The authors use reanalysis data from Flemming et al. (2017). It would help the reader if the authors make a clearer reference to this dataset by referring to it as the "CAMS interim Reanalysis", considering that ECMWF has published various atmospheric composition reanalyses. Also, even though atmospheric ozone in this reanalysis is constrained by satellite observations, the constraints on surface ozone concentrations are typically less, and are more governed by the model assumptions. To the least, it would be worthwhile to provide insight in the quality of surface ozone and carbon monoxide, e.g. as documented in Huijnen et al. (2020), and/or Flemming et al. (2017), and to what extent these aspects may alter your analysis. Also along these lines, the reference to "measurement data" (pp 15, line 7) may be confusing in this context.

Author's response: Thank you for pointing out the potential confusion in naming. We have amended the text to refer to this dataset in the manner suggested. We have also added a sentence to note the strengths and weaknesses of this dataset referring to these papers to give a more complete picture of the reanalysis data. We have replaced the reference to "measurement data" with "surface concentration data" to avoid confusion.

(2) Reviewer's comment: Page 10, line 7. The authors write a little cryptically: "We included p as one of the parameters to estimate for the reanalysis data and found values in the range 0.16-0.19." Not being a specialist in mathematics, could you give some interpretation of this analysis? Why is it in this range, and not much smaller (or much larger)?

Author's response: Thank you for highlighting this. The 0.16-0.19 range was derived from estimating the parameters using the reanalysis, prior to carrying out the model calibration involving the synthetic datasets. However, we appreciate that the way this sentence is phrased is not entirely clear. In the revised manuscript we have rewritten this and added extra detail to improve clarity.

(3) Reviewer's comment: Page 11, line 8: "In contrast, methods based on neural networks can require thousands of training runs.": This statement reads a bit unfounded.

Could you please either add a reference (e.g. a study which actually uses a neural network approach to model atmospheric chemistry), or further clarification? If I understand correctly, one of the key aspects compared to published use of machine learning in atmospheric chemistry is that the full CTM is replaced by the Gaussian Process Emulator, rather than 'just' the chemistry solver code as I am aware of (Keller et al., 2019)

Author's response: Thank-you for this comment. The point we make here is that more traditional machine learning methods for mapping model inputs to model outputs typically have hundreds of parameters, and for this reason thousands of training points are required. Neural network methods have not been applied yet to full CTMs for this reason, although they have been used for specific components of atmospheric models, as the reviewer notes. A key benefit of Gaussian Process emulation is the greatly reduced number of training runs that permit the methods to be applied to more complex models. We have rephrased the sentence to make this clearer. We do not cite the Keller et al. paper here, as the point we are making is not specific to atmospheric models, and because they used a random forest approach, but we agree that it is a good example of application of machine learning methods in the discipline.

(4) Reviewer's comment: The results from the sensitivity analysis (Sec. 3.1) indicate that surface ozone deposition is the largest driver to explain biases in modeled surface ozone concentration. While I completely understand from mathematical perspective, and experience with old-fashioned tuning in a chemistry transport model, that by changing (modifying) the dry deposition velocity for ozone has indeed a direct impact on surface concentrations, I find this result also a little worrying, in light of what would be a reasonable range of the ozone deposition flux. To get a handle on this, would it be easy and useful to quantify the change in annual total surface ozone deposition as suggested by this optimization method? In fact, the authors also rightly discuss the issues with this sensitivity analysis when including the simultaneous optimization of CO, pp22, line 6-9. When reading this manuscript, I had found it helpful if this aspect was already alluded to in Sec. 3.1 . Indeed, I agree that the use of this synthetic modeling can be very useful, but different optimization factors obtained using (in this case) different combinations of control variables also stress the danger of a false impression of a physically well-constrained parameter.

Author's response: Our aim in this paper is to demonstrate the feasibility of calibrating a full CTM, and we have intentionally chosen to use a simplified system with only 8 scaling parameters to demonstrate the method. As there are many sources of uncertainty that we do not consider here, we do not expect the calibration to generate the same results that it might when including all sources. However, we still expect it to provide useful insight. Surface O_3 is a little high over land masses in the model, and thus in the absence of consideration of uncertainty in chemical processes, the calibration suggests that dry deposition fluxes should be about 40% greater than the a priori value. Deposition accounts for 930 Tg(O_3)/yr in the model, so this would constitute a sink of 1300 Tg/yr (neglecting feedbacks); while this is at the high end of recent estimates, it remains within the spread of published model results. However, in the absence of consideration of all sources of uncertainty, we choose to highlight the ability to constrain processes rather than the value of the scaling factors that arise, for the reasons that the reviewer identifies. For the simple system we consider, the parameters are constrained well for this parameter as shown by the small posterior uncertainty. A more complete study addressing uncertainty in a much wider range of processes is needed to generate a more robust assessment of the scaling factors needed for a particular process. In Section 3.1 we already acknowledge the limited range of processes considered ("of the eight considered here") but to address the reviewer's concerns we add a statement at the end of Section 2.8 that alludes to the sensitivity of the calibration to the system considered.

(5) Reviewer's comment: Page 18, line 19: "along with a reduction in associated

uncertainty": Could you be more explicit on this? I don't directly see such decrease in uncertainty.

Author's response: Thank you for spotting this error. The uncertainty is reduced compared to using CO alone, but is comparable to that using O_3 alone, and therefore this phrase isn't needed. We have now removed this part of the sentence.

(6) Reviewer's comment: Figure 10, panels c/f/i : In almost any of the test configurations the parameter estimate for the Boundary Layer Mixing is approaching the maximum range that is given. Can you provide further interpretation in this aspect? Is the given uncertainty range for boundary layer mixing sufficient? Now it is written on page 21, l. 11-13 that this process "may not be represented well in the model". Or could this be an artifact of other (missing) processes not considered in this sensitivity analysis, or that this process is just treated differently between datasets.

Author's response: The range of the Boundary Layer Mixing scaling parameter is already large, spanning a factor of 100. The fact that the parameter estimates lie very close to the bounds suggests that the process is not well represented in the model. However, the reviewer is right to point out that other processes may be important; we are considering a simplified system here with only 8 scaling parameters and the boundary layer mixing parameter may thus be acting as a surrogate for uncertainty in processes not considered here, most notably chemical processes. We already acknowledge this point in Section 4.4 of the paper. Our aim in the paper is to demonstrate how well the calibration method works in a simplified system, and more complete coverage of uncertain parameters would be needed to provide a more robust assessment of specific processes. However, we have altered the text here to acknowledge the influence of other processes.

(7) Reviewer's comment: pp 21, l 13: "Our results suggest that dry deposition and isoprene emissions are represented relatively well in the FRSGC/UCI CTM " : Good to add phrase "with respect to the independent reanalysis data"?

Author's response: We agree that this should be clarified and have rephrased this sentence in the manuscript as suggested.

(8) Reviewer's comment: pp 23, l9: "…while it *is* effective …"

Author's response: Thank you for spotting this. We have corrected this in the revised manuscript.