

Atmos. Chem. Phys. Discuss., referee comment RC2
<https://doi.org/10.5194/acp-2022-410-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on acp-2022-410

Anonymous Referee #2

Referee comment on "Characterization of errors in satellite-based HCHO/NO₂ tropospheric column ratios with respect to chemistry, column-to-PBL translation, spatial representation, and retrieval uncertainties" by Amir H. Souri et al., Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2022-410-RC2>, 2022

Souri et al present a detailed study highlighting four major shortcomings associated with FNRs and their ability to categorize ozone sensitivity. The sections about column-to-PBL translation, spatial representation error, and retrieval error are all well-written. The manuscript as a whole has understandable writing style and clear, well-made figures. However, I do have a few major concerns, mostly surrounding the modeling section of this work. I recommend that the manuscript be sent to the authors for major revisions.

VOC inputs for the box model: The modeled radical environment can be incredibly sensitive to changes in VOC inputs, especially in polluted urban areas. This manuscript is lacking detail about how VOC inputs were created, leaving readers to assume the authors used a simplistic approach that excludes many potentially important VOCs. As written, the authors' treatment of VOC inputs does not rise to the level established in previous modeling studies performed for the same field campaigns, leaving this reviewer wondering if the modeling presented in this study can represent the ambient radical environment.

The field campaigns modeled in this study have unique VOC measurement suites which require unique data engineering strategies to generate realistic VOC inputs. DISCOVER-AQ was served only by a quadrupole PTRMS, and features a very limited set of VOCs. The authors do not give adequate detail about how they generated VOC inputs based on these data. For example, previous studies (i.e. Schroeder et al 2017) generated speciated VOC box model inputs for DISCOVER-AQ using a fusion of VOC data from concurrent airborne campaigns (DISCOVER-AQ+SEAC4RS+FRAPPE). This enabled somewhat realistic estimation of VOCs that were not measured by the PTRMS during DISCOVER-AQ.

During KORUS-AQ, the whole air sampler was flown concurrently with a PTRMS, giving a richer suite of speciated VOCs. However, these two instruments had wildly different sampling cadences and integration times, with WAS measurements being incapable of resolving fine-structure details in pollutant gradients. As a result, previous studies (i.e. Schroeder et al 2020) fused the two datasets together to generate a pseudo-high-resolution set of VOC inputs for their box modeling work with KORUS-AQ.

As it is currently written, I have serious concerns about the VOCs used as model inputs, and thus have lower confidence in the results presented here. Can you show that the simplistic VOC inputs used in this study do not yield significantly different results from the two Schroeder papers?

Perhaps a more pointed observation: the box model inputs and outputs from the two Schroeder papers are publicly available online. What does this study gain by running its own model simulation – with questionable VOC representation – instead of using the freely-available Schroeder/Crawford data which has already been heavily vetted and used in multiple studies?

Model Setup: I have a few concerns with model setup:

Why use an arbitrary model run time of 5 days? Ideally, the model should be run indefinitely until it converges on a solution for key species, but I understand the desire to set a lower limit for the sake of computation. Do your outputs change if you use 4 days? Or 6 days? Or 20 days? Can you include a sensitivity analysis to back up your work – that is, show that your arbitrary choice of 5 days does not impact results?

If I understand this correctly, you calculate a unique dilution factor for each field campaign, deriving it empirically to yield the best agreement between measured and modeled HCHO. What is the physical basis for why one field campaign would have different dilution rates than another? Without further explanation, this feels like an arbitrary “correction factor” to game the model for better agreement with observations – which does nothing to tell you how well the model represents the underlying chemistry. Can you explain?

Model Validation: As written, the model validation section does not give me confidence in the model’s ability to represent the ambient radical environment (especially given the simplistic treatment of VOCs).

If one of the model parameters (dilution rate) is based on empirical model/measurement agreement, then comparing simulated values to observations is cyclical. If I am understanding this correctly, then Section 3.1 is incredibly problematic. Based on the description given in line 220, campaign-average simulated HCHO is not allowed be >5% off from campaign-average observed HCHO – its part of the model setup with an empirically derived dilution factor. How can you evaluate the model’s representation of the chemical environment with such a setup? For example, in line 254 you state that HCHO had a mean bias of less than 5% - which is meaningless because you’ve coded the model to do exactly that.

In practice, your dilution factor acts as a quasi-constraint on HCHO, which greatly influences calculated radical budgets. This eliminates your ability to truly test whether the model is capable of representing the radical budget from first principles. Furthermore, this does not allow you to test if your simplistic treatment of VOCs is adequate. You mention that the bias in PAN changes if you ignore dilution, but PAN can have a large impact on modeled radical and NO₂ concentrations. What happens to other test-species if you ignore dilution?

I don't buy the idea that this model, as currently setup, has proven itself as sufficient for representing ozone chemistry.

In short, you have questionable VOC inputs and a questionable model setup which prevents you from truly testing model performance. I'd suggest re-running the model with a fixed dilution factor based on reasonable physics. This will enable "true" unconstrained model runs, providing a testbed for evaluating model performance (and your VOC inputs). Or, you could run another, established model in parallel on a subset of data and compare the two. Or, you could use the freely-available model inputs/outputs from published studies from the same field campaigns, rather than re-invent the wheel.

Purpose of the paper: Finally, I would challenge the authors to include paragraphs in the Introduction and Summary sections describing the motivation for doing this work. Why are incremental improvements in our understanding of FNRs and ozone chemistry necessary? Martin et al first published their paper about satellite FNRs more than twenty years ago – yet, to the best of my knowledge, no regulator or policymaker has ever used satellite FNRs in their ozone planning strategies. Clearly FNRs were first developed as a potential tool for policymakers to fine-tune ozone mitigation strategies, but if policymakers have shown no interest in using these tools, why continue refining them? Is there a pathway for FNRs to be used by anyone outside of academia? What does the author think is preventing policymakers from using this tool - or is this simply a tool for academics?