Dr. Clara Orbe
NASA Goddard Institute for Space Studies Code 611 New York · NY 10025
clara.orbe@nasa.gov
clara.orbe@gmail.com

Dr. Luke Western
Editor,
Geoscientific Model Development

September 24, 2025
re: manuscript number: egusphere-2025-2761 Title: "Experimental Protocol for Phase 1 of the APARC QUOCA (QUasibiennial oscillation and Ozone Chemistry interactions in the Atmosphere) Working Group"

Dear Dr. Western,

We thank the referees for their second round of reviews of our manuscript. We appreciate the care with which the reviewers have examined the manuscript throughout this process and have considered all feedback thoughtfully. In particular, we have addressed the additional 8 points raised by Reviewer 1, as well as the request to change the formatting of the appendix figures by Reviewer 2. Our responses to the two reviewers are addressed herein and reflected in the revised version of the manuscript.

With this revision we provide two versions of the revised manuscript, one of which includes the corrections highlighted in red. The point-by-point responses to the referees' comments are also included. We hope that the manuscript is now acceptable for publication in GMD. I confirm that my coauthors concur with the submission of our manuscript in its revised form. The revised version of the manuscript has been resubmitted electronically.

Yours sincerely,

Dr. Clara Orbe

Response to Reviewer 1

Recommendation: Minor revision

(45) " inadequate vertical resolution": Note that there is a new paper that examines systematically the impact of vertical resolution in NCAR's CESM model: Simpson et al., J. Adv. Mod. Earth Sys., 17, e2025MS004957, https://doi.org/10.1029/ 2025MS004957.

➔ Thank you for pointing us to this reference. We now cite this paper. Indeed, we were already familiar with it and find it interesting that vertical resolution does improve the QBO in that model, although this improvement does not translate to improved teleconnections. That certainly is an interesting result that warrants further research.

(183) "transient experiments may be complicated by": The only thing that unavoidably complicates analysis of coupled-ocean integrations is slowly evolving internal variability (e.g., ENSO). On the other hand, there is no reason why "anomalous triggers" (volcanoes, wildfires) need to be included in such integrations. At the risk of sounding pedantic, the point of the original comment (which the revision does not address) was that the "anomalous triggers" have nothing to do with why one would choose to carry out time-slice instead of transient runs.

➔ We maintain that "anomalous triggers" like volcanoes and wildfires can generate anomalies in stratospheric ozone that can be nontrivial to remove, as would be needed to generate the ozone annual cycle forcing fields used to constrain the PD NINT experiments. We do not agree with the reviewer that only ENSO – and not these types of events – would need to be considered. While this type of analysis would certainly not be impossible, it is likely that modeling centers would use a broad set of techniques to remove these influences from resulting ozone trends, generating even more (undesired) spread in the forcings among the models. No changes to the manuscript.

(196) "the influence of NOx": This makes it sound like NOx is one among several factors that drive the ozone QBO above 20 hPa. However, NOx is the main catalytic loss mechanism for ozone between about 20 and 5 hPa (Brasseur and Solomon, 2005, Fig. 6.1), and several studies conclude that NOx mostly explains the ozone response in the upper stratosphere, e.g., Butchart et al. (JGR 2003), Anstey (Nature Rev. 2024), etc. I understand that the authors wish to have as many models as possible participate in the QUOCA exercise, but I wonder about the usefulness of any model that does not include a proper representation of NOx chemistry. However, since I am not too familiar with "simplified [ozone] mechanisms", I will defer to the authors in this matter.

➔ We agree wholeheartedly with the reviewer that NOx will be very important. Indeed, we are eagerly anticipating results from two GISS model experiment submissions, which will compare the ozone feedback on the QBO generated from simplified versus fully interactive ozone schemes that respectively ignore and include NOx chemistry. Quantifying this effect within a single model will help isolate the contribution of NOx to the ozone QBO.

(197) "nor at lower levels … column ozone aloft": It might be clearer to write something like "nor at lower levels, since the overlying ozone column can modulate the ultraviolet radiation that reaches the lower stratosphere and affect infrared transfer between layers".

➔ OK—we have included this suggested rephrasing. While this does significantly lengthen the sentence, we agree that it makes the point clearer. Please see the revised manuscript.

(205) "similar in magnitude … more standard approaches": That is good to know but note that my original comment asked about the motivation for using separate 30-year segments to define the ozone climatology for the PD-NINT runs. I expected that 30-year segments would have statistical properties similar to the 90-year climatology (which you now show to be the case); however, I wondered whether there were specific reasons why your methodology would be preferable to the conventional approach. The answer to this question appears to be that there is no compelling reason to use 30-year segments but also no downside. (If I have understood this correctly, this comment does not require any additional revision).

➔ There is nothing special, per se, about 30 years versus more years. We just wanted to use enough years so that the statistical properties of each resulting climatology were not substantially different. The reviewer seems to appreciate this point. No changes to the manuscript.

(325) "Ming(2016))" -> Ming (2016)

➔ Thanks for catching this typo, which has been fixed in the revised version of the manuscript.

(325) "as close to … as possible": Why as close as possible to the native pressure grid? I would think one wants the output on the actual native grid to ensure accurate calculation of vertical derivatives. Or perhaps you are thinking about grid-cell midpoints vs. interfaces?

➔ Sorry, our original phrasing is unintentionally (and unnecessarily) confusing – we have now removed "as possible". Please see the revised manuscript.

(327) "verify consistency with …Table B2": This is fine as long as such comparisons are limited to the TEM quantities included in Table B2. However, one quantity not included in that Table is the acceleration associated with the EP flux divergence due to resolved waves, div(F). Calculation of div(F) involves the vertical derivative of the vertical component of the EP flux, Fz, which cannot be calculated accurately from Fz values interpolated to standard levels (plev42). On the other hand, all TEM quantities, including div(F), can be calculated offline from the output (Table B3) on the native model vertical grid (plevTEM).

➔ *Correct* -- comparisons are limited to the TEM quantities included in Table B2. No changes to the manuscript.

Response to Reviewer 2

The revisions have largely improved the quality of the paper. I have no more questions. The only comment is that the figures in the appendix should be placed in portrait.

➔ Thanks for taking the time to review our revised manuscript. As requested, we have changed the figures in the appendix to portrait mode.