Review of "Experimental Protocol for Phase 1 of the APARC QUOCA (QUasibiennial oscillation and Ozone Chemistry interactions in the Atmosphere) Working Group" (revised)

by C. Orbe et al.

Recommendation: Accept after final (minor) revision

The authors have responded satisfactorily to most of my comments on the original paper. The revised version is suitable for publication after additional minor revision to address a few remaining specific comments listed below.

Specific Comments (line number):

(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.

(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.

(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.

(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".

(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).

(325) “Ming(2016))” → Ming (2016)

(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?

(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, F_z, which cannot be calculated accurately from F_z 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).