Review of Kuai et al, "Attribution of Chemistry-Climate Model Initiative (CCMI) ozone radiative flux bias from satellites"

In this paper, Kuai et al. quantify and attempt to interpret model biases in outgoing longwave radiation (OLR) in the wavelength band of strong absorption by tropospheric ozone, 9.6 microns. Using assimilated data, they first construct instantaneous radiative kernels (IRKs) to quantify the sensitivity of OLR at 9.6 μ to four driving variables: tropospheric ozone, water vapor, atmospheric temperature, and surface temperature. The IRKs are calculated at a range of altitudes and across the globe. They then apply the observation-based IRKs to the biases in the four variables in an ensemble of climate models; this step yields the contributions of these variables to biases in OLR at 9.6 μ in each models. The subsequent analysis shows the importance of accurate representations of tropospheric ozone and water vapor in calculating OLR in this wavelength region.

The topic of the paper is important. Understanding the contribution of tropospheric ozone to current and future climate change requires models that can accurately simulate OLR in the relevant wavelengths. The calculations in the paper are fairly straightforward and yield some interesting results – e.g., that the overestimate of tropical tropospheric water vapor present in many models leads to an underestimate of OLR in the ozone wavelength band. But the paper feels like an early draft: much of the analysis in the paper is shallow and the paper is not well written.

I recommend acceptance only after major revisions.

Major comments.

1. A key reason to improve OLR at 9.6 μ is to improve estimates of the radiative forcing from the change in tropospheric ozone from the preindustrial era to the present-day or from the present-day to the future. The manuscript leaves the reader wondering how important are the biases uncovered in the manuscript. Do the biases have much of an effect on ozone forcing estimates over time?

2. In discussing the sources of biases in the four driving variables, the manuscript provides some detail about how modelers have struggled to improve estimates of tropospheric ozone (e.g., Oman et al., 2011 and 2013; Stenke et al., 2013; Revell et al., 2015). But nothing is said about model biases in the other three variables. Are these biases well-known or new to the community? If known, what are the potential reasons for these biases?

3. The paper focuses on OLR under clear-sky conditions, which is fine. But cloud cover varies significantly among models. Could variation in clouds also affect model estimates of ozone radiative flux – i.e., OLR at 9.6 μ? The text should include a discussion of how the focus on clear-sky conditions affects the conclusions.

4. The paper is poorly written with many lapses in English, about 1-3 per paragraph throughout. Those co-authors for whom English is a first language should read the manuscript much more closely and rewrite as needed. As is, the manuscript gives the impression that the co-authors did not read it carefully.

Minor comments.

Page 2. Line 10. Abstract should make clear that the overestimate of atmospheric opacity in the models is due to too much tropospheric ozone and/or water vapor. Also, the large number of significant digits looks suspect.

Page 2. Line 22. The text should make clear whether these estimates of radiative forcing are instantaneous or adjusted.

Page 3. Line 18. Text should clarify why the western Pacific is particularly opaque.

Page 6. Line 38. Reader wonders whether surface characteristics matter to OLR calculations, or are these captures by T_s?

Page 9. Line 41. “The flux bias due to Ta is found to be negligible in all models, which indicates that the Ta is relatively accurate.” The statement should probably be qualified to say simply that the modeled Ta estimates provide reasonable radiative fluxes at 9.6 μ. The temperatures may not be so accurate for other purposes.

Page 10. Lines 20-25. The authors might consider reordering the figures so that the figures showing biases in each driving variable come right after the figures showing the biases in the fluxes attributable to that variable – e.g., Figure 9 right after Figure 5. That way readers can more easily see how the driving variables affect fluxes.

Page 11. Lines 20-25. Are these ozone biases, reported in ppb, weighted by air density? If not, perhaps they should be, given that the opacity of the atmosphere is affected by total ozone column, not average ozone mixing ratio in that column.

Page 12. Line 8. What updates did Oman 2011 and Oman 2013 implement in GEOSCCM to improve the model’s representation of ozone? Readers will want to know.

Page 12. Line 25. The text should provide a reference to substantiate the claim that MRI-EMS1r1 overestimates lightning NOx emissions and underestimates convective updrafts in the tropics.

Page 12. Line 35. This reader does not understand “fact” 3, which gives this explanation for the bias in the radiative flux: “the systematic bias throughout the tropical troposphere, when propagate into the TOA flux, causes an accumulated bias in the radiative effect.”

Page 13. Line 20. Why do the Ta biases in the tropics “shift between positive and negative vertically” in most models?

Page 14. Lines 6-35. The text should clarify what of value is learned in the analysis of how biases in 9.6 μ band affect biases in the entire OLR band.

Pages 14-16. The conclusion should emphasize that the paper focuses only on clear-sky OLR.

Figure 4. The caption could say that the black curves are the same as the colored curves in Figure 3.

Figure 5. What do the black curves represent?

Figure 13. Equations should be shown only for those correlations and slopes that are statistically significant.

Table 3. Are the mixing ratios for ozone and water weighted by air density in each layer?