Review of “Observational constraints on ozone radiative forcing from the Atmospheric Chemistry Climate Model Intercomparison Project (ACCMIP)” by K. Bowman et al.

I cannot recommend this manuscript for publication in ACP, because I feel that it leaves a number conceptual inconsistencies between the measured and modelled quantities unresolved. At the very least, these inconsistencies and their treatment are not explained in a convincing way. As it stands, I cannot accept the “observed forcing” as a “constraint” for the “modelled forcing”, because in fact different entities are compared. Partly, this may be just a semantic problem and could be overcome be improved formulation (the current text confused me at several stages). However, I suspect that an observational constraint of model forcings can only be created, if the way of determining the model forcings is adjusted to what the TES radiative fluxes actually describe (in fact, the authors appear to think in a similar direction (p. 26319, l. 5)). Constraining the tropospheric ozone change itself may be a different animal, but this is obviously not the primary aim of this paper.
My recommendation would be to leave the idea of “constraining radiative forcings” aside and to compare observational and model parameters that a really comparable. Besides present-day tropospheric ozone that could mainly be the parameter called by the authors the total longwave radiative effect (LWRE, p. 23611). It could be evaluated by means of the TES kernel for each model ozone field, and compared with the TES LWRE purely derived from observations. Beyond, a consistent LWRE counterpart could be calculated by means of the model radiation schemes. In the latter process, ambient parameters (temperature, clouds, water vapour etc.) could be exchanged between observational ones from TES and those simulated by one (or more than one) from the models. Only in this way, I think, can you get rid of the problem that you do not know the observed pre-industrial ozone and, thus, cannot without arbitrary corrections calculate a radiative forcing counterpart from observations. In a last step you could then include into the comparison the actual radiative forcings from the models (Stevenson et al., 2012) and try to derive consistent conclusions. To this end, however, the models should provide the *instantaneous* longwave radiative forcing rather than the stratosphere adjusted forcing, in order to get rid of this inconsistency, too
[With hindcast, I notice that Reviewer 1 denies the *instantaneous* character TES “forcing” and LWRE. I’m not quite sure whether he is right, because it appears to depend on how measurements in the 9.6 µm band are translated into spectrally integrated OLR. Hence, a discussion of results from either kind of calculation method in the model world might help the interpretation.]

A) Main concerns

- The fundamental question is what the authors really wish to constrain: The ozone perturbations simulated by the models? The quality of the radiation schemes of the models? And do they intend to quantify the influence of other (ambient) parameters to the forcing, too? This would require a step-by-step evaluation, which the paper does not provide, and for which I have given a recommendation above. Or is the intention to ”constrain” the net effect of a combination of all these impact parameters? This, as sections 5.2 and 5.4 suggest, results in the necessity of weakly

defined corrections and unclear speculations that make the comparison worthless, at least in the sense of “constraining”.

- My main concern is that I perceive inconsistencies of the methodology framework (as pointed out in section 4) with the paper’s declared objective. Radiative forcings in general, and those of Stevenson et al. (2012) in particular, are defined through ozone changes relative to the pre-industrial state, when all other parameters have to remain fixed to a reference state in the radiative transfer calculation (reference may be either the pre-industrial or present-day parameter set). On top of this, the question arises how to account for stratospheric temperature adjustment when comparing simulated longwave radiative forcing and observed longwave radiative flux change (OLR).
- The description of radiative forcing in sub-section 4.1 seems largely correct, but it also contains some details that irritated me. In contrast to the authors’s notion (p. 23612, l. 23) the radiative forcing term is generally *not* applied for the spatial distribution of the radiative flux imbalance induced by a radiatively active tracer (although a respective use of the term in the context of the present paper is acceptable if properly introduced). Rather, radiative forcing is generally (and not mainly by IPCC!, l. 24) taken at the tropopause or at the top of the atmosphere, *because only in these cases* it is reasonable to assume that sign and magnitude of the forcing can be taken as a proxy for the expected temperature response. The stratosphere adjusted radiative forcing at the tropopause (p. 23613, l. 8) is considered as being best suited for this purpose, especially in case of ozone changes (Fels et al., 1980; Hansen et al., 1997). As the authors correctly state, this leads to a dependence of the radiative forcing value on the tropopause definition which can be and has been quantified (e.g., Forster et al., 1997). I add that using the chemical tropopause in the process, as done in the present paper, is rather unusual; I would feel more comfortable, if the quantitative consequences would have been tested for an example case.
- I am not sure that the authors have given adequate attention to that the TES radiative kernel is only valid for the ambient parameters forming the TES atmosphere. So even if the ensemble model mean or any individual model would provide a perfect simulation of the ozone change between present-day and pre-industrial times, and if it would involve a perfect radiative transfer model, the “TES observed” radiative forcing could still be different simply because the model(s) provide(s) a different reference background state. I cannot see how this possibility is accounted for in the GISS/TES comparison described in Section 5.2.
- I do not think that RF^{c}_{m} in Eq. 9 should be called a radiative forcing or a radiative forcing change. My impression is that it forms a flux difference (correction term?) that reflects the biases of the model present-day ozone fields, of the radiation scheme used for calculation the modelled fluxes, and perhaps also of differences in the ambient parameters (water vapour, temperature, clouds etc.) between the TES atmosphere and the modelled atmosphere(s).
- I feel that the role and the treatment of the shortwave radiative forcing component is insufficiently addressed when it comes to comparing net forcings. This is a further point (independent of those discussed above) that renders the observed forcing results questionable as a constraint for the simulated net radiative forcings.
- Figure 6 and some hints in the text (e.g., p. 23617, l. 3f.) suggest to me that the authors interpret their results in a way of model quality evaluation. This is certainly unjustified, both in view of conceptual inconsistencies that have been recognized and tried to be corrected, and those that (to my opinion) have not been accounted for (see above).

B) Minor remarks

1. p. 23606, l. 11: Sentence needs rewording to avoid the double “with”
2. p. 23609, l. 11: It is important to know whether and to which extent parts of TES results (parameters and radiative fluxes) depend on the TES forward model and will, thus, automatically deviate if another radiation module is used. Future versions of the paper should point out (or recall) what is known about this. As an alternative, the quality of the radiation module(s) of the ACCMIP model(s) could be subjected to an individual quality check (see my recommendations in the introducing part of this review).
3. p. 23609, l. 18: “… ozone profiles are biased high in the troposphere (≈15%)…”; does this mean that the relative bias is almost uniform over the troposphere? Why, then, …
4. p. 23609, l. 20: … can the bias of total (tropospheric?) ozone column be less (i.e. 10%)? Does the difference origin from stratospheric contributions? However, it was not mentioned that profiles in the stratosphere are included, too!
5. p. 23610, l. 7: Do I gather correctly from this that the V004 products are less biased than the V002 products described before in the NH, and much less biased than the V002 products in the tropics and SH?
6. p. 23611, l. 8: As I have pointed out in my major comments, it is not obvious to me why the left part of equation 2 should have the character of a radiative forcing according to usual definitions. Formally the superscript “c” should be defined here once again, as its being mentioned in the introduction does not suffice to understand what is meant here.
7. p. 23611, l 17: Again, I have problems with the terminology and/or the wording here: LWRE is obviously *not* a fractional change when its unit is W/m^2 (Figure 1; btw. does the column bar rather indicate a fractional change?). Consequently, the reasoning of this sentence (“Since the …”) remains fuzzy. I understand, however, that the value of LWRE must not be taken as the OLR increase due to the absence of any atmospheric ozone due to non-linearities in the concentration/radiative impact relation (saturation effects). The last sentence evidently gives the (inverse) value of LWRE as defined in l. 15; this would become clearer if this sentence would be shifted upwards, behind “… Worden et al. (2011)”.
8. p. 23611, l. 25: You seemingly have now shifted from LWRE to LIRK, please indicate. Further shifts to TES ozone occurs in p. 23612. l. 5, without being announced. This whole paragraph is written in an unnecessary confusing way.
9. p. 23612, l. 24: see major comment.
10. p. 23613, l. 19: It seems that beginning with this equation (and continuing in section 4.2) the subscript “lw” (Eq. 2) is omitted. Please indicate that, from here on, you nevertheless refer to the longwave flux, exclusively. Let me add (as I’m not completely sure from the description) the assumption of mine that in Eq. 4 the TES kernel is applied to the modelled present-day and pre-industrial ozone. So I agree that RF^m is actually a radiative forcing here, because all parameters are fixed through use of the kernel. Btw., why do you omit the “*i*” here (compared to Eq. 2)?
11. p. 23614, l. 9f.: I think this is *not* a radiative forcing according to the definition framework given in the introduction, because it is a longwave radiative flux (OLR) difference between two ozone fields actually coupled to a different ambient parameter sets.

12. p. 23614, l. 16f. (Eq. 8): This, now, is *surely not* a radiative forcing (nor a radiative forcing change); it might rather be called an "tropospheric ozone induced OLR bias due to model systematic errors in the simulated ozone field" (if we accept $F(q_p^{obs})$ as being the "true" LWRE).
13. p. 23614, l. 20: "A key assumption …"; this sentence again touches the root of my criticism (see major comments), as I think it's essentially not correct. An unbiased (tropospheric) ozone change is only one component potentially inducing a difference of observed and modelled radiative forcing. All other components influencing ozone radiative forcing ought to be unbiased, too.
14. p. 23615, l. 11: As evident from my major comments, the paper manuscript fails to convince me that the claim formulated here holds.
15. p. 23615, l. 17: There are two components contributing here: ${}_{i}RF^{lw}$(toa) is different from ${}_{adj}RF^{lw}$(trop), because its adjusted vs. instantaneous, but also because its toa vs. tropopause. The latter holds, because only the net adjusted radiative forcing is the same at the toa and the tropopause; the lw and sw components are not. From a classical model inter-comparison (Shine et al., 1995) I notice substantially larger differences between longwave instantaneous (Fig. 3, ibidem) and adjusted (Fig. 7, ibidem) forcing of tropospheric ozone (25-30%) than indicated here. Please clarify whether your correction really includes both contributions, and what that means for your methodology and conclusions (especially when turning to the net forcing, Table 2).
16. p. 23615, l. 23: I assume that the TES and modelled ozone distributions have been used as input to a radiative transfer model to calculate the sw radiative forcing. Please indicate, which radiation model has been used and how large the differences were. Shine et al.'s (1995) results suggest that the model-induced uncertainty of tropospheric ozone RF is *not* negligible.
17. p. 23618, l. 1: What has been compared? The GISS model ozone with the TES ozone? Or the GISS radiative flux change and the TES radiative flux change caused from the TES retrieved ozone? Or both? And what about the ambient parameters (see major comment)?
18. p. 23618, l. 16: I think a better understanding is indeed required, otherwise the paper will always miss its objective according to the title.
19. p. 23619, l. 3: Is this really a justified statement? Just because part of the quite large difference at certain latitudes compensate in the global mean?
20. p. 23619, l. 5: Here, the authors themselves touch on what I feel necessary to be done, before the idea of establishing an "observed constraint" may be introduced (see major comments).
21. p. 23620, l. 7: As I expressed before, from several reasons the term radiative forcing appears out of place to me here.
22. p. 23620, l 1f: I am aware that the present paper does not claim to establish a constraint to tropospheric ozone but rather to its forcing. Nevertheless, I feel that a constraint to ozone forcing will be hard to provide from TES when even the basic field cannot exactly be constrained due to a potential bias in the observations themselves. Why, I musk ask, are the modelled ozone fields not directly evaluated with the ozone sonde results, if these are regarded as the more credible observational basis?
23. p. 23620, l. 13: "The strong thermal contrast in the tropics …"; You refer to the temperature difference between the absorber (ozone) and Earth's surface, don't you? Please, formulate more precisely. The CMAM model may suggest itself as an example to underpin what this sentence (and the next one) shall express.

24. p. 23620, l. 25: typo “clompl**e**mentary”.
25. p. 23620, l. 27: Wouldn’t is be useful to show the chemical tropopause for the various models in either Figure 3 and Figure 4 (or both)? Or, as a compromise, to enhance the “ENS” panel of Figure 3/4 to display the ensemble mean chemical tropopause? That would help the reader to recognize the domain used for vertical integration.
26. p. 23621, l. 26: “tropical” should read “tropics”.
27. p. 23622, l. 3: “NH high retrieval biases …”; I fail to understand what this sentence is meant to express. Why should the NH have “lower radiative sensitivity” than the SH?
28. p. 23622, l. 5: The meaning of this sentence is even less obvious to me than that of the preceding one.
29. p. 23622, l. 8: As indicated above, I disagree that the result of the tables can be interpreted in the way the authors do.
30. p. 23623, l. 6: I find it awkward that RF_m^{obs} is to indicate a net forcing (including the shortwave – how?) here, while in Eq 6 it’s only the lw component. Please, be more precise.
31. p. 23624, l. 21: I got rather lost over this concluding section, probably because I do not agree to what is stated here, viz., that the comparison between observed and modelled radiative forcing (and, thus, its constraining) is straightforward. Consequently, the perspectives developed from here on sound somewhat quixotic to me.
32. p. 13641, caption Fig. 3: Please, reformulate to “… between the ACCMiP modelled ozone and the TES ozone …” for the sake of clarity.
33. p. 23643, caption Fig. 5: Caption should included that this is the flux at the chemical tropopause (as in the text on p. 23620), because in Fig. 3 the same symbol, iRF^c_m, has been used for a quantity displaying vertical dependence.
34. Figures 3, 4, 5: In any future version of this paper these figures ought to be revised to improve their readability (size, numbers along the axes, numbers along the colour bar).

C) References (only if absent from the paper manuscript)

- Fels, S.B. et al., 1980, Stratospheric sensitivity to perturbations in ozone and carbon dioxide: radiative and dynamical response. J. Atmos. Sci., 37, 2265-2297.
- Forster, P.M.F. et al., 1997: On aspects of the concept of radiative forcing. Clim. Dyn., 13, 547-560.
- Hansen, J. et al., 1997: Radiative forcing and climate response. J. Geophys. Res., 102, 6831-6864.
- Shine, K.P. et al., 1995: Radiative forcing due to changes in ozone – a comparison of different codes. In Atmospheric Ozone as a climate gas (Eds. Wang, W.-C. and Isaksen, I.S.A., NATO ASI series, Springer, Berlin, 373-396.