

Response to Reviewer 2

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 by 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.

We have changed the manuscript to focus on present day ozone and OLR and their correlation with ozone RF. For reasons detailed in the response to Reviewer 1, we believe these changes will address the bulk of your concerns. The correlation doesn't require directly merging observed OLR and ozone RF.

My recommendation would be to leave the idea of “constraining radiative forcings” aside and to compare observational and model parameters that are 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.

The application of Eq. 2 effectively does this, particularly if you were to consider the product of $H(\ln q^{\text{obs}})$ and $H(\ln q^{\text{m}})$ separately. The difference between these are shown in the Figures.

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.

This comparison was effectively done in Section 5.2 and Figure 2 of the original manuscript since GISS made the same calculation of Eq. 2 but using their own RTM and physical state. The differences between the two were substantial. However, per the request of Reviewer 1, we have removed that Section since fully explaining the results is outside the scope of the manuscript.

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.

While there are no reliable measurements of preindustrial ozone, there are at least aspects of preindustrial ozone that should be similar to present day, namely the natural background (accounting for the change in methane). However, to do so requires attributing the differences between the models and the data to natural versus anthropogenic. The LWRE combines both. The attribution question is one that we intend to pursue in a future analysis. We now compare to the Stevenson results simply on the basis of how models represent TES OLR and their correlation with ozone RF.

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.

A direct comparison of OLR between models and TES given the present day conditions is an important consideration. For example, if a chemistry-climate model systematically overestimates water vapor then they'll predict a more opaque atmosphere and consequently a decreased OLR sensitivity to ozone. We are working on new OLR products to do this but are not available at the time of this manuscript.

[With hindcast, I notice that Reviewer 1 denies the *instantaneous* character of 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.]

We agree that the *instantaneous* character of TES OLR is perfectly well defined concept. How we calculate spectrally integrated OLR is described in Section 3 and in detail in Worden *et al* (2011).

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”.

Sections 5.2 and 5.4 have been removed. We focus on the difference between TES and ACCMIP ozone and the consequences of those differences on observed OLR. The basis of an observational constraint is based up a correlation with that OLR and ACCMIP ozone RF.

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 methodology section in question has been removed. While it is true that assumptions are necessary, which were described in detail, we don’t think they were inconsistencies. But at this point, that question is moot.

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).

We l.23 doesn’t say anything about the spatial distribution other than the net irradiance has to be defined for some reference surface. But, that sentence is not necessary so we’ve removed it.

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).

We’ve incorporated your point as follows:

“In order for the ozone forcing to be taken as a reasonable proxy for the expected temperature response, RF is specifically defined as the change in net irradiance at the tropopause after stratospheric temperatures have relaxed to radiative-dynamical equilibrium but with the surface and atmospheric state held fixed.”

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.

The Stevenson et al, 2006 used the chemical tropopause for the radiative forcing calculations. Stevenson et al, 2012 does compare the differences between tropopause definitions, including the chemical tropopause. See, for example, Table 3 in Stevenson et al, 2012..

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.

We agree with this point. We take that to mean that for the present day, models need to predict both the right ozone distribution and the reference physical state to get the forcing correct. The “TES observed” radiative forcing includes both the ozone and the current physical state. It’s quite possible that the differences in Section 5.2 are due to different physical state between GISS and TES. If that’s true, then getting the reference (present) state consistent with observations is an important consideration for radiative forcing calculations. **However, as noted before, Section 5.2 has been removed.**

I do not think that RF_{cm} 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).

We believe you are referring to Eq. 8 in this case. For the purposes of this methodology, it was assumed that the RTM schemes and the ambient parameters are the same so that RF^c_m only reflects differences in ozone. However, it's true that if those quantities are not consistent, then there will be an additional model bias for present day $F(q^m)$ even if q^m is consistent with observations.

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.

We agree that the treatment of SW forcing was weak. We could have calculated a SW forcing based upon TES ozone and then make some assumptions about the SW albedo from clouds, etc. However, TES's strength is in the longwave. This is one reason why we've refocused only on the OLR.

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).

We now state: "We use TES observations to evaluate the ozone and ozone OLR in the chemistry-climate models that participated in ACCMIP." where ambient conditions for the OLR are referenced to the observed climate. From that standpoint, model evaluation is perfectly reasonable.

B) Minor remarks

- p. 23606, l. 11: Sentence needs rewording to avoid the double "with"

Changed the sentence to read:

In addition, changes in climate can affect the chemistry of ozone, including increasing the rate of OH production due to higher water vapor concentrations with warming temperatures, that will have impacts for methane and other gases.

- 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).

The TES parameters and radiative fluxes are calculated such that they match the observed spectral radiances through the TES forward model. So, the forward model is a critical part of the calculation. The forward model itself has been extensively tested. Uncertainties are less than 0.1% for *spectral* radiances. Integrated band errors are much less than that. Most of the uncertainties here are driven by spectroscopy. We agree the doing a more rigorous comparison of the radiation modules with the TES radiative transfer code would be a useful activity but it outside of the scope of this manuscript. Stevenson et al, 2012 ACPD has done comparison with different radiance modules.

- p. 23609, l. 18: "... ozone profiles are biased high in the troposphere ($\approx 15\%$)..."; does this mean that the relative bias is almost uniform over the troposphere? Why, then, ...

Yes, the relative bias is approximately uniform throughout the free troposphere as can be seen in Fig 5-2 in the V004 validation report cite in the subsequent paragraph. The bias can be seen in other thermal IR spectrometers such as IASI. It's unclear why it's there though it could be a problem with spectroscopy over a mixture of errors with temperature.

- 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!

Yes, the bias there is the total atmospheric column including stratospheric ozone. TES retrieves an ozone profile from the surface to 0.01 hPa. We do not focus on the stratosphere as there are other more capable instruments, e.g., MLS, that measure that area. Stratospheric variability doesn't really drive OLR variability. To the extent that it does, TES captures it because TES measures the spectrally resolved IR radiances.

- 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?

That's a fair statement. Perhaps equally significant is that the NH bias has much less vertical structure than in V002.

- p.23611,l.8:AsIhavepointedoutinmymajorcomments,it isn'tobviousto 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.

Equation 2 is now $\Delta OLR^{\{j,m\}}_l$.

- 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/m² (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)”.

The confusing aspect of definition in Eq 3 is that change in OLR is with respect to $\ln q$. The logarithm of q has no units, so equation 3 is in units of W/m². The “fractional change” interpretation is derived from linear approximation of $\Delta \ln q \sim \Delta q/q$. We’ve changed the sentence to read:

“Note that the LWRE is a logarithmic change referenced to TES ozone and consequently it can not be used to calculate the change to a complete absence of ozone.”

Yes, the LWRE is a linearized about the atmospheric opacity from the reference ozone levels. The absence of ozone would lead to a more optical thin atmosphere and hence a difference sensitivity.

- 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.

LWRE and LIRK are two different quantities. The LWRE is derived from the LIRK as it is the integral of the LIRK (defined in Eq. 1 and again in Eq. 3) to the tropopause. Ozone is clearly another quantity: all three of which are shown in Fig. 1.

- p. 23612, l. 24: see major comment.

We have changed the text to read:

“In order for the ozone forcing to be taken as a reasonable proxy for the expected temperature response, RF is specifically defined as the change in net irradiance at the tropopause after stratospheric temperatures have relaxed to radiative-dynamical equilibrium but with the surface and atmospheric state held fixed

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 preindustrial 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)?

The reason “lw” was omitted in Eq 2 and throughout 4.2.1 is that the methodology is independent of TES. We wait until 4.2.2 to incorporate TES into the methodology. It’s at this point we included the approximation necessary for TES. However, this is moot since 4.2.1 and 4.2.2 have been removed.

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.

This is a definition includes both OLR and SW observations. The application to TES is in 4.2.2. It is radiative forcing if all of the difference between observed and predicted ozone can be attributed to ozone change, which is one extreme of the attribution problem—the other being that the difference can be completely attributed to natural, background conditions. Implicit in this assumption is that the present-day physical climate is well represented by the models.

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 $F(q_{\text{obs}})$ as being the “true” LWRE).

Eq. 8 defines the difference in net irradiance (LW+SW) between the observed ozone and modeled ozone. So, $F(q_{\text{obs}})$ is the observed irradiance, not the OLR. $F(q_{\text{obs}})$ not LWRE either because LWRE is a change in OLR whereas $F(q_{\text{obs}})$ is the absolute number. Eq. 8 posits that difference between predicted and observed ozone is attributed to a change in ozone not model bias.

For the updated manuscript, we examine delta OLR exclusively, which as you pointed out can be interpreted as “tropospheric ozone induced OLR bias due to model systematic errors in the simulated ozone field”

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.

We were well aware of validity of this assumption, which is why we stated that this was a “key assumption”. The primary value here is to look at the “worst case” scenario given that differences between observed and modeled ozone are a combination of both bias and errors in changes in ozone as was discussed in the paragraph.

Nevertheless, we’ve switched focus to look exclusively at the change in OLR induced by the difference between observed and modeled ozone.

14.p. 23615, l. 11: As evident from my major comments, the paper manuscript fails to convince me that the claim formulated here holds.

We changed the text to read

“TES observations, which directly measure outgoing longwave radiation (OLR) in the 9.6 micron band (where ozone absorbs thermal infrared radiation) and have the spectral resolution to disentangle the geophysical quantities, e.g., temperature and clouds, driving OLR variability, can provide useful information about ozone radiative forcing”

Hopefully this will be more acceptable to the reviewer.

15.p. 23615, l. 17: There are two components contributing here: $iRFlw(toa)$ is different from $adjRFlw(trop)$, 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).

Yes, the Shindell and Stevenson papers look at the difference between TOA and tropopause as well as adjusted/instantaneous differences. Those differences were within the models themselves. The contribution did include both.

However, this section and Table 2 has been removed.

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.

We agree that the SW calculation is a weak component of the analysis as we did not use the TES ozone distributions for the SW calculation. In principle, one should use satellite observations of the SW to get this number. This part has been removed.

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)?

Yes, the GISS modeled ozone with TES ozone run through the GISS RTM. This was compared against the OLR change from applying the TES IRK to the GISS/TES differences. The ambient parameters are different for each. This may explain the difference. However, to fully explain it requires a more thorough investigation, which we'll pursue in the future. For now, we removed this section.

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.

We agree.

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?

It's justified to the extent that if the global mean TOA irradiance is zero, then there's no energy imbalance. Shindell and Faluvegi [2008] have shown that there's a regional climate response to regional forcing. However, they did not explore it in the context of balanced errors. That's worth further investigation.

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).

A direct test of OLR will be pursued in the future.

21.p. 23620, l. 7: As I expressed before, from several reasons the term radiative forcing appears out of place to me here.

This has been changed to delta OLR

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 must ask, are the modelled ozone fields not directly evaluated with the ozone sonde results, if these are regarded as the more credible observational basis?

The models are evaluated against ozone-sondes in the companion paper Young et al (2012) ACPD. While ozone sondes are more accurate than satellite measurements,

they do not have the coverage and temporal sampling afforded by satellites. The variability of the ozone sondes is quite high because they are driven by a lot of small scale, local processes. Consequently, they are not a terribly strong constraint on global model distributions.

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.

We changed the sentence to the following:

As quantified by the LIRK in Fig. 1, the thermal contrast between the temperature at which the ozone absorbs in the troposphere and the surface temperature along with the low humidity, which decreases atmospheric opacity, contributes to importance of the tropics and subtropics relative to the midlatitudes.

24.p. 23620, l. 25: typo “clomplementary”.

Done

25.p. 23620, l. 27: Wouldn’t it 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.

We changed Figure 1 so that the color scale for TES ozone saturates at 150 ppb. This should help show the domain of integration.

26.p. 23621, l. 26: “tropical” should read “tropics”.

Done

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

We restated to:

Northern mid-latitude high retrieval biases in the TES data will not strongly effect the ensemble OLR mean because of the lower radiative sensitivity. On the other hand, the ACCMIP SH sub-tropical OLR high bias is a robust result with respect to the relatively low TES retrieval bias there.

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.

This line has been removed.

30.p. 23623, l. 6: I find it awkward that RFmobs 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.

The shortwave was included by assuming that the SW component (taken as indicated in the text from Stevenson (2012)) of the modeled RF was correct. Consequently only the LW component is constrained. Eq. 6 includes both LW and SW as a theoretical formulation. Application to TES was done in 4.2.2. In any case, this part has been removed.

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.

Hopefully the revised text will appear less quixotic. There's no presumption that the relationship between modeled and observed OLR and that of radiative forcing is straightforward. On the other hand, there should definitely be implications for ozone radiative forcing if models do not predict present day ozone in radiatively important regions. The new results, which are based upon the correlation of model OLR bias with model RF, suggests that this is a fruitful direction. More sophisticated approaches are described in the conclusions.

32.p. 13641, caption Fig. 3: Please, reformulate to "... between the ACCMiP modelled ozone and the TES ozone ..." for the sake of clarity.

Done

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_{cm} , has been used for a quantity displaying vertical dependence.

Done

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).

We have endeavored to improve the clarity of these plots.