Responses to reviews of ‘Effective Radiative forcing from emissions of reactive gases and aerosols – a multimodel comparison’ by G. Thornhill et al

We would like to thank the two referees for their helpful comments and suggestions. Our responses to the reviewers comments are below – reviewer comments in black, our responses in red.

Anonymous Referee #1

This paper examines the effective radiative forcing (ERF) of reactive gases (e.g., CH4, NOx, VOC, O3, N2O, HC, NH3) and aerosols (e.g., BC, OC, SO2) to the climate system using multi-model output from the AerChemMIP experiments of the CMIP6 project. The contribution of each species to the total ERF is decomposed, and the differences of the calculated ERFs by various models are discussed. The paper is overall well written and easy to follow. I have some minor comments for the authors to consider before publication:

1. It is not clear how many ensemble members are used for each model. Can you please clarify this?
In most cases there was a single member – this will be clarified.

2. Fixed SST and sea ice are used in the ERF simulations. Is it the climatological SST of the 1850s?
These were the fixed SSTs from the Pre-Industrial run – mentioned in L. 174

3. The description of Eq. 5 is a bit vague. You may want to add “ERFaci” to line 240 after “The effect of the aerosol on cloud radiative forcing”.
This will be clarified with an improved description.

4. “A_trop” is used in text while “RA Trop” is used in Fig. 1. It’s better to be consistent.
This will be corrected to keep the reference to the quantity consistent

5. Line 373, “Fig. 6”! “Fig. 5”. And can you explain a bit more about how the total AOD is used to calculate the sum of the scaled ERFs?
We will correct the Fig. number reference and add more detail on how the scaling is calculated.

6.Table 6, second row, “Nox”! “NOx”
We will correct the typo

7. Line 465, please remove the brackets around “O’Connorl F. M. 2019”
Reference will be fixed.

8. Error bars are used in Fig. 9-10 to quantify the uncertainties due to interannual variability of model diagnostics. Is it possible to apply similar approach (error bars) to other figures where data are available?
We will add error bars where appropriate and we have the data.

9. Line 507, “RFs”! “ERFs”
This will be corrected to ERFs

10. Line 525, “+/-“! “_”, please also add explanation about the numbers follow the

sign, for instance, is it a standard deviation of multi-model output?
This will be clarified.

11. Lines 541-542, the overall aerosol ERFari from AR5 (-1.5 _ 0.4 Wm-2) is much larger than values reported here (-0.16 _ 0.03 Wm-2). Can you add some discussion about the differences?
Yes, discussion will be added here on the differences between the AR5 ERFari and that reported here (a difference in definitions, I think)

12. Lines 546-550, redundant
Redundant lines will be removed.

Anonymous Referee #2

Thornhill et al. "Effective Radiative forcing from emissions 1 of reactive gases and aerosols – a multimodel comparison"
This paper is a very important contribution to climate and atmospheric chemistry studies. It is critical for the current IPCC AR6 assessment. My apologies to the authors for the delay in my review – surprisingly, these isolation times do not make it easier to review. I am rushing to get this out and so there may be typos in this review.
AerChemMIP is a very important project that is trying to make sense of a complex, nonlinear system of chemistry and clouds. The design has some serious flaws, and we all knew that as it was being developed since there were obvious limitations in the number and complexity of experiments. First, there is the nonlinearity, which comes out clearly in these results: Exp(A) plus Exp(B) does not equal Exp(A+B). Further, and because of the nonlinear nature, the choice of reference atmosphere (1850) will produce quite different results than another (2014). Personally I would have much preferred to use the 2014 atmosphere as the reference atmosphere (at least we can compare the models to measurements) and then subtract the emission or concentration changes from 1850 to 2014. So, we live with AerChemMIP and use it. The analysis here is very good, but needs to develop more of an "assessment" view when reporting final numbers. They should reflect some (subjective) adjustment of values (for non-linearity) or bias (from 1850 atmosphere). Simply reporting the model average for Exp(A) and then Exp(B) will lead anyone who uses this paper to assume that the combined effect is the sum. That would not be good for either policy or attribution work. If the sum is the best answer (I would think that so), then the contributions of components may need to be expressed in % of total rather than in W/m2. I do not recommend a substantial rewrite, but rather a self-assessment by the author team of how to use these results.
We will make some assessment of the limitations and constraints of the experimental design, and also clarify the nonlinearity in the discussion of the results; we could add contributions in percent of total (where appropriate), if it improves the clarity.

The abstract starts off being very careful and clear, but then I get lost between emissions and composition change. For example, SO_2 is clearly an emission-based ERF, while CH_4 is an abundance-based ERF. VOCs & NO_x are obviously emissions (having direct RF) but O_3 is not emitted. The HC (L40: halocarbons? as opposed to NMHC?) are negative presumably from the ozone depletion, but this is an odd way to list the CFCs. Methane is singled out at the end of the abstract as increasing ozone, but ozone is listed separately. Is the methane based only on the change in CH_4 concentration? Further, I suspect that the methane includes stratospheric water vapor which is indirect as is ozone, so why the separate listings. This is just inconsistent and you

really need to present the framework and rules for partitioning and assigning ERF. Basically, AerChemMIP is a complicated set of overlapping experiments and thus the abstract with simple results is very difficult to write correctly. Try to keep it clear and simple.
We will clarify the abstract and add comments on the frameworks used, and where results are emissions vs. concentration based.

L53- aerosols are chemically reactive. try again – aerosols and gases that …
This will be re-worded as “chemically-reactive gases”.

L56- climate feedbacks on natural emissions is a tough one for the non-CO2 species. I do not think AerChemMIP did anything on that.
The discussion of natural emissions will be removed. These are covered in Thornhill et al. 2020.

L68- "conditions" do you mean SST, I doubt you prescribed different chlorophyll or DMS?
We will clarify which conditions are referred to here.

L69- why aerosols here? It seems more like and ie than and eg, why not just say aerosols and gases.
It is simply meant as an exemplar of what is perturbed – hence the eg., but could be re-worded to include ‘aerosols or gases’

L71- again the contrast of 'aerosols and chemistry' is really a poor description of AerChemMIP. Aerosols are a chemically reactive species (most of them are, they were created by gas phase chemistry). This Intro really must have a better inclusive discussion of greenhouse gases (including ozone), of indirect greenhouse gases (CO) and of aerosol (primary and secondary). This is at odds with the content of AerChemMIP.
This section will be expanded to clarify the indirect roles of precursor species as recommended by the reviewer.

L90- again, every time I read this, it sounds odd and misleading: " aerosols and reactive chemistry" are not opposites.
This will be reworded as “aerosols and reactive gases”

L93- again, this phrasing sounds wrong: " anthropogenic and reactive species" is this two separate species or is it 'both'
This will be reworded as “aerosols and reactive gases”

L103- what does 'down to' mean? from 0.001 to 560 hPa?.
This will be clarified in terms of levels of the atmosphere..

L107- typo 'is includes'
This will be corrected.

L102-L165 This information should really be in a table somewhere, not in the text. Focus on the results. In fact, most of this is already in Table 1.
Agreed – although some text discussing how model differences may affect results is appropriate here, with references to Table 1 and Table S1 for detailed information.

L172-173- You should say 'emissions' with the NTCFs, also you should note that methane is an SLCF, which your statement seems to preclude. Also, I thought that SLCF was new preferred, but…

The use of NTCF was retained because the experiment was called piClim-NTCF, so this nomenclature was kept to aid in understanding which experiment was referred to. The comment in brackets L. 173 makes the point that this nomenclature has changed. Other references to the term NTCF can have SLCF added parenthetically to point out the change in nomenclature.

The wording will be clarified to describe methane correctly.

L174- There is a serious problem with the AerChemMIP as defined and we realized this at the time, but did not address it: viz. because of the large changes in atmospheric chemistry and oxidants between 1850 and 2014 (including the ozone depletion), it is not clear that the effect of today's NTCF emissions in today's atmosphere are anywhere close to those calculated here for the PI atmosphere. This needs to be addressed when compiling results and clearly adds to the uncertainty. The reason why it is dangerous is that it could totally misrepresent the magnitude of the response if we were to cut NTCF today.

The reviewer has a good point here. We will add a discussion of the issue likely differences between perturbing emissions in a pre-industrial and present day atmosphere.

L200- "ocean state" implies much more than SSTs – do you really mean that.

This will be changed to “SST”.

.

L235- I am confused here, it seems like the direct forcing would include, not ignore, the absorption and scatting by aerosols. Is this a typo.

Agreed, this is unclear as written - we will re-write this to make it clear that this is the difference between the aerosol-free flux and the total flux.

L239- this section is very confusing as written. I am sure it is simpler that it seems but some of the writing seems incorrect. In this line surely you mean the aerosol direct radiative effect, since "radiative effects due only to aerosols" would imply both ari and aci. The notation for ERFcs,af is inconsistent between eqn and text (comma or not). Eqns 4 & 5 sum to give not ERF, so where is this missing 'surface albedo' term (ERF-ERFcsaf) and does it matter? It must certainly be counted as an aerosol ERF. [OK, I see this in Table 3 later, is the cs,af just noise?]

Agreed – this section needs clarification in the writing and description, consistency in the subscripts will be corrected, and the surface albedo term eqn. will be added in to complete the definitions.

L254ff- This is odd, you said just above the SARF is calculated from ERF – A_trop, so of course this should give 0% difference. Are you just checking the math? Also, with SARF calculated as a residual term, and the ability to denote and sum all the A_trop being highly uncertain, SARF would not appear to be very certain. In fact the SARF term would depend on the models' ability to diagnose A_trop correctly.

I think we defined the SARF here as the (IRF+A_{temp_strat}). There is the ERF as calculated from the kernels using the sum of the adjustments and the IRF, which is compared to the direct calculation of the ERF in 3.1.1. I will clarify what is being compared here.

L260- With BC, the long-standing problem is that some models get far too much in the stratosphere and that would cause a very large SARF.

We will check the vertical profiles of BC in these models.

L285- in this figure and some others, please define carefully what the shorthand for the terms means.

We will make sure that terms in the figures are properly defined.

L296ff- This is a very good discussion of the aerosol components!

Thank you, nice to hear what is good about the paper!

L321- This would be a much better lead off to the aerosol section and analysis, begin with the big picture before the weeds.
Agreed – the overall values should come before the effort at analysing and breaking down the components (or 'the weeds'…)

L343 – just put this table into the figure, it is just a summary of the bars anyway.
I think it is useful to have the multi-model mean numbers and errors separated out so the figure is not cluttered, but the table could go in the Supplemental materials, and the multimodel mean bar added to the plot.

L354- Please, stop wasting space, Table 5 and Fig 4 has the same information. If you want to show AOD, then add it to the figure.
We think the numbers are useful to have separately, but the table could be moved to the Supplemental material.

L370ff & Fig 5- I do not understand the purpose of this AOD scaling, it really does little to help. The figure shows the key data: ERF from parts vs ERF from all. The ERF-parts consistently over accounts for the combined ERF. This is as expected wince the cloud effects are largest in a pi atmosphere with little background aerosols. So herein lies one of the fundamental problems with the AerChemMIP that must be acknowledged and accounted for. I am not sure that scaling by AOD is any more justified than just scaling the indirect to match – An interesting question is whether the ERF-ari sum of the parts equals the whole? Keep this fudge factor simple since there is no correct way to do this.
The idea was to assess the breakdown of the overall ERF by the constituent aerosol, to get a sense of the relative importance of the different aerosol species in the combined piClim-aer experiment, and to illuminate model differences (if any) in this breakdown. It is imperfect, but we thought a useful additional data point.

L393- Again, this title jars a bit, SO2 and NH3 are reactive gases.
This will be renamed to reactive greenhouse gases

L406- How big is the error in GHGas ERF if one ignores clouds? I would think large.
We will add a discussion of the expected size of the cloud masking effect

Fig 6 & Table 6, can easily combine and understand the Std Dev better.
As noted above (L 354), having the multimodel mean numbers separated into a table makes it easier for the reader to use them– the table can be moved to the supplemental material and a multimodel mean bar added to the plot, together with std error bars.

L419- Here is a case where some assessment is due as to how well these model simulations are accurate in the sense of including all the effects. As noted the N2O-O3 link is important and missing in some, and the other key link between N2O and tropospheric OH and methane (and maybe aerosols) is established but missing here.
We will make it clearer as to which processes are included in which models. For the relevant impacts on stratospheric/tropospheric ozone, and methane lifetime, only the models with appropriate chemical processes are considered.

L431-441- This is a very interesting and important discussion about the additivity of the components. I suspect that the CH4 result is similarly affected.
BTW, where is the effect of stratospheric H2o form CH4 noted or counted?

The effect of stratospheric H2O is included with the WMGHG. It is not diagnosed separately.

L480- Yes halocarbons, but also N2O, and N2O may be more important since it depletes ozone in the tropics (there are papers on this). I do not see what difference stratospheric Cl will make on the total lifetime – hopefully the Stevenson paper becomes referencable. Also, you need to be careful here since the methane feedback factors ff, apply to the PI atmosphere, and that is different from the present, particularly lower CO and NOx…. . The feedback factor used for GWPs etc, is the current one, not the PI one, so these results should NOT be used to change any previous assessments.
We will mention the change due to N2O as well. We will clarify that the f factors are starting from a pre-industrial atmosphere.

Fig 11 is really hard to understand or see clearly, it will need a cleanup.
This figure will be revised to only show the multi-model mean.

L525- Here is a very important statement and I am not sure that you have put together all the reasonable uncertainties or non-linear scaling. The individual components here must be adjusted to recognize that the total ERF (with all simultaneously) is much less than the sum of the components. Thus you cannot recommend the individual results without scaling and without increasing the uncertainty.
This discussion will be revised to explain the differences between the individual components and the total, and the implications for uncertainty.

L561- Same as above. The individual values will not sum correctly and so these do NOT reflect the ERF of NOx emissions as we progress from 1850 to 2014. Thus they should not be used as part of an assessment until they are more critically evaluated and put in context.
These are all very important AerChemMIP results, and the analysis here is highly valuable, but their use in attribution and related studies should reflect the bias and uncertainties in combining nonlinear parts that were calculated separately, and in basing these all on a pi-atmosphere.

This discussion will be revised to explain the differences between the sum of NOx and VOC and the total. The ERFs are typically defined starting from a pre-industrial atmosphere, there is no unique way to reflect the ERF of NOx emissions as we progress from 1850 to 2014 (subtracting components from a present atmosphere would overestimate the ERF).