Author Responses to the Reviews Received from acp-2019-1152: "Assessment of pre-industrial to present-day anthropogenic climate forcing in UKESM1" by F. M. O'Connor

Dear Editor,

Thank you for agreeing to handle the above manuscript.

On behalf of my co-authors, I would also like to thank the two anonymous reviewers for their positive and constructive reviews of the submitted manuscript.

During the discussion phase, we discovered a bug in the experimental set of UKESM1 in the pre-industrial atmosphere-only timeslice experiment that acts as the control experiment (*piClim-control*) for all of the other perturbation experiments (*piClim-X*) discussed in the paper. Essentially, in the atmosphere-only configuration of UKESM1, we prescribe the land surface characteristics using output for vegetation fractions, leaf area index (LAI), and canopy heights from the coupled model. We found that the seasonal cycle in the pre-industrial climatology for LAI was 6 months out of phase, affecting surface albedo and hence, the top-of-atmosphere radiative fluxes.

As a result, we have re-run the simulations and re-evaluated all of the present-day effective radiative forcings discussed in the manuscript. All of the original data has been retracted from the Earth System Grid Federation (ESGF) and the updated data and an erratum have now been published: https://errata.es-doc.org/static/view.html?uid=5e70c479-9b19-ca91-a9ba-d11febe15377

As a result of the bug, and as agreed by email with the topical editor, we will make two sets of modifications to the revised manuscript from that submitted. The first set of changes will be to update the manuscript to take account of the LAI bugfix – these changes will be shown in the marked up manuscript in blue, e.g., the methane ERF is ~~old value~~ new value.

The second set of modifications will address the reviewers' comments. Below, we iterate the reviewers' comment in black with grey shading and our response to their comment will be *black italics*. Any new text will be shown in the marked up manuscript in red, and any deleted text will be like this: ~~deleted text~~.

I thank you again for your comments.

Regards,

Fiona O'Connor (on behalf of all co-authors)

First Set of Modifications:

As indicated above, the LAI bugfix meant that all simulations were re-run, and all the ERFs were re-evaluated. As a result, figures, some section text, and some tables all required updating. The following tables were updated:

Table 3 (Page 14-16 of the marked-up manuscript)

Table 4 (Page 21 of the marked-up manuscript)

Table 5 (Page 41 of the marked-up manuscript)

The following figures were replaced:

Figure 2 (Page 17 of the marked-up manuscript)

Figure 3 (Page 20 of the marked-up manuscript)

Figure 4 (Page 24-25 of the marked-up manuscript)

Figure 5 no longer on piClim-N2O but on piClim-HC (Page 27 of the marked-up manuscript)

Figure 6 (Page 30 of the marked-up manuscript)

Figure 7 (Page 32-33 of the marked-up manuscript)

Figure 8 (Page 34 of the marked-up manuscript)

Figure 9 (Page 36 of the marked-up manuscript)

Figure 10 (Page 39 of the marked-up manuscript)

Figure 11 (Page 43 of the marked-up manuscript)

Figure 12 (Page 48 of the marked-up manuscript)

Figure 13 (Page 50 of the marked-up manuscript)

Figure 14 – now removed

Figure 15 – now re-labelled as Figure 14 (Page 52 of the marked-up manuscript)

The following sections were updated:

Abstract - Pages 1-2 of the marked-up manuscript

Section 4.1 (Overview) – Pages 13-21 of the marked-up manuscript

Section 4.2.1 (CO2) – Pages 21-22 of the marked-up manuscript

Section 4.2.2 (N2O) – Pages 22-26 of the marked-up manuscript

Section 4.2.3 (now re-named as Ozone Depleting Substances) – Pages 26-27 of the marked-up manuscript

Section 4.2.4 (Methane) – Pages 28-31 of the marked-up manuscript

Section 4.2.5 (GHG) – Pages 28-31 of the marked-up manuscript

Section 4.3 (Aerosols) – Pages 31-35 of the marked-up manuscript

Section 4.4 (Ozone precursors) – Pages 35-45 of the marked-up manuscript

Section 4.5.1 (NTCF) – Pages 45-50 of the marked-up manuscript

Section 4.5.2 (Land use) – Pages 51-53 of the marked-up manuscript

Section 4.5.3 (Total anthropogenic) – Pages 53-54 of the marked-up manuscript

Section 5 (Conclusions) – Pages 54-57 of the marked-up manuscript

Response to Reviewer #1:

This paper is an important documentation of effective radiative forcing in a leading climate model, and should be published following modifications.

For reasons that I understand and appreciate, the paper has a rushed and inconsistent feel to it and is poorly written in some places. I have a large number of comments, some of them rather important for clarification of what has been done, and what the results indicate. Hence, I have recommended the paper needs major modifications before I believe it will be suitable for publication.

Response: Thank you for your positive comments and for your detailed constructive feedback on the submitted manuscript. We aim to address your detailed comments/suggestions in our responses below and in the changes applied to the manuscript.

Major Comments:

There is a lack of clarity about some of the experiments which led to significant confusion. Some of it was dissipated by the text later in the paper but, for some, "mystery" persisted. Although I suspect these labels follow the various MIP protocols, they are unhelpful to a reader less immersed in the world of MIPs and clearer information, early in the paper, of what is and what is not included in various experiments is necessary. While Table 3 goes some way to doing this, it leaves various issues unclear; it, or an accompanying table, needs to be much more specific about what is, and what isn't included in different integrations

Response: Thank you for this comment. Unfortunately, we had no control over the experiment names used in the manuscript – these were pre-defined by the protocols for the Radiative Forcing Model Intercomparison Project (RFMIP; Pincus et al., 2016) and the Aerosol and Chemistry Model Intercomparison Project (AerChemMIP; Collins et al., 2017). However, we appreciate that the experiment names and/or the experimental set up for the different simulations may not have been clear enough for those not directly involved in the MIPs. In order to address this, we have done the following:

1. *In Table 1, we added an additional column to distinguish between those halocarbons that are ozone depleting substances (ODS) and those that are not. It clarifys that the piClim-HC simulation only perturbs ODSs and makes the difference in the present-day perturbations applied to the halocarbons between the experiments piClim-HC and piClim-GHG clearer.*

   See Table1; pages 8-9 in the marked-up document

2. *We also re-named Section 4.2.3 to further emphasise that piClim-HC is only considering ozone-depleting halocarbons as follows:*

   4.2.3 Ozone (O_3) Depleting Substances (ODSs)

   See Page 26 in the marked-up document

3. *In Table 1, we expanded the column titles to indicate that for greenhouse gases including methane, concentrations are prescribed and that for the aerosol, aerosol precursors and ozone precursors, emissions are used. Table 1 has a separate column already for methane and by expanding the column title names, we label the tropospheric ozone precursors as "O_3*

(VOC, CO, NOx) precursor emissions". We have also moved the aerosol column so that methane and the other ozone precursors are now side by side, but considered separately. This separation between methane and the other ozone precursors was inherited from AerChemMIP and so was retained here for consistency. In addition, instead of referring to non-VOC or non-NOx, we have made them more explicit, e.g., "non-VOC" is now replaced with "NOx" and "non-NOx" is replaced with "VOC & CO". And we have also added text to Section 3 to clarify the experimental setup of piClim-VOC, piClim-NOx, and piClim-O3 as follows:

"In all the experiments in Table 1, the emissions of primary aerosol and aerosol precursors (BC, Organic carbon (OC), and sulphur dioxide (SO2)) and O3 precursors (VOC, CO, and NOx) excluding CH4 for 1850 and/or 2014 were taken from Hoesly et al. (2018) and van Marle et al. (2017). In the case of the NOx emissions perturbation experiment (piClim-NOx), both surface and aircraft emissions were changed from PI to PD levels. In piClim-VOC, both VOC and CO emissions were changed from PI to PD levels while the experiment piClim-O3 perturbs emissions of VOC, CO, and NOx only, with the CH4 concentration remaining at PI levels. "

See Section 3: Lines 274-279 in the marked-up document and Table 1: pages 8-9 in the marked-up document

And we changed the title of Section 4.4 to "4.4 Ozone (O_3) precursor (VOC, CO, and NOx) gases"

See page 35 in the marked-up document

4. *There was one experiment missing from Table 1: piClim-4xCO2phys. This has now been added – See Table 1: pages 8-9 of marked-up document.*

5. *The way the greenhouse gas concentration in UKESM1 is prescribed differs for the different greenhouse gases. As a result, we added more detail on how these experiments were set up as follows:*

"In all cases, the GHG concentrations for 1850 and/or 2014 were taken from Hoesly et al. (2018), van Marle et al. (2017), and Meinhausen et al. (2017). However, the recommended concentrations for the different GHGs in UKESM1 are implemented differently. In the case of CO2, the prescribed concentration is uniform in mass mixing ratio throughout the model domain. For CH4 and nitrous oxide (N2O), the recommended concentrations are treated as lower boundary conditions (LBCs); their 3D distributions are modelled interactively by the UKCA chemistry scheme (Archibald et al., 2020) and coupled to radiation. For ozone depleting (OD) halocarbons in piClim-HC, their concentrations are prescribed separately (and consistently) in UKCA and in the radiation scheme. For the UKCA chemistry scheme, LBCs are prescribed for trichlorofluoromethane (CFC11), dichlorodifluoromethane (CFC12), and methyl bromide (CH3Br), all of which include contributions from other chlorine- and bromine-containing source gases not explicitly treated in UKCA. This approach ensures that the correct stratospheric chlorine and bromine loadings are used for the PD period. Further details on the species included in the CFC11, CFC12, and CH3Br LBCs can be found in Archibald et al. (2020). For the radiation scheme, the radiative effects of ODSs are handled by prescribing the mass mixing ratio of a lumped species (CFC12-eq) uniformly throughout the atmosphere, consistent with the UKCA LBCs, Finally, the piClim-GHG experiment collectively

perturbs all the GHGs from PI to PD levels, including non-OD halocarbons. For these latter GHGs, a uniform mass mixing ratio of a lumped species (HFC134a-eq), provided by Meinhausen et al. (2017), is prescribed in the radiation scheme."

See Section 3 (page 10; Lines 258-272) in the marked-up manuscript.

1. L33 starts the confusion about what NTCF is and whether it includes (following AR5) methane and the shorter-lived halocarbons. So a clear definition early in the paper is needed (or as indicated below, NTCF is not used at all, as it is incorrectly used as a shorthand for aerosol and ozone precursors).

Response: Although near-term climate forcers (NTCFs) include methane and short-lived halocarbons (e.g., Myhre et al., 2013), in the context of the AerChemMIP experimental protocol (Collins et al., 2017), the experiment piClim-NTCF only alters emissions of aerosol and aerosol precursors (SO2, BC, OC) and the ozone precursors (VOC, CO, and NOx). It does not include perturbations to methane and/or other short-lived climate forcers. To make this clear, we have modified the abstract as follows:

*"*Together, aerosol and O_3 precursors (called near-term climate forcers (NTCFs) in the context of AerChemMIP) exert an ERF of -1.03 ± 0.04 W m^{-2}, mainly due to changes in the cloud radiative effect (CRE). " – See lines 38-40 of the marked-up document

And a sentence was added to the section on the Experimental Set up (Section 3) to make this clearer as follows:

"Finally, although near-term climate forcers (NTCFs) include CH4 and short-lived halocarbons (e.g., Myhre et al., 2013), in the context of the AerChemMIP protocol (Collins et al., 2017), the experiment piClim-NTCF does not perturb concentrations of CH4 or other short-lived GHGs. It only changes emissions of aerosol and aerosol precursors (SO2, BC, OC) and O3 precursors (VOC, CO, NOx) from PI to PD levels; it is also referred to as piClim-aerO3 in the RFMIP protocol (Table 1)."

(See Lines 279-283 of revised manuscript)

2. L33: Even after reading the paper, I did not really understand what "tropospheric ozone precursors" means and whether the impact that these have on lower stratospheric ozone (e.g. following what was shown in the Sovde et al. papers) was included or excluded from the analysis. Was ozone change above the tropopause prevented in piClim-O3 runs (and the associated individual runs such as piClim-NOx)? If not, I would class it as quite misleading to call this "tropospheric ozone" and would suggest an alternate name (such as "ozone reactive precursor gases" to make it clearer). If it was prevented, then it needs to be made clear what the criteria were for preventing ozone change in the stratosphere. There is the parallel issue for pi-Clim-HC, and whether it allows tropospheric ozone change resulting from the ODSs. An additional confusion is whether CH4 is classed as an ozone precursor. Again, it took some time before this became clear (or more precisely, less unclear) and Table 3 didn't help in this regard.

Response: In response to previous comments, we have modified the manuscript to make clear that the perturbation experiments involving what were previously called "tropospheric ozone precursors" only consider perturbations to VOC, CO, and NOx emissions; methane is considered separately. As indicated above, these have been re-named as "O3 precursors" (Table 1; Pages 8-9 of marked-up document). And in response to previous comments, we added the following to Section 3:

"In all the experiments in Table 1, the emissions of primary aerosol and aerosol precursors (BC, Organic carbon (OC), and sulphur dioxide (SO2)) and O3 precursors (VOC, CO, and NOx) excluding CH4 for 1850 and/or 2014 were taken from Hoesly et al. (2018) and van Marle et al. (2017). In the case of the NOx emissions perturbation experiment (piClim-NOx), both surface and aircraft emissions were changed from PI to PD levels. In piClim-VOC, both VOC and CO emissions were changed from PI to PD levels while the experiment piClim-O3 perturbs emissions of VOC, CO, and NOx only, with the CH4 concentration remaining at PI levels. "

(See Section 3; Lines 274-279 in the marked-up document)

For the ERF estimates, there is no masking applied. Atmospheric composition and the atmospheric state are both allowed to fully adjust in the perturbation experiments such that the ERF encompasses all changes to the TOA radiative fluxes due to whole-atmosphere changes in composition (including ozone) and/or other rapid adjustments (e.g. temperature, humidity etc..) as well as land surface rapid adjustments. The following is added to Section 4:

"ERF = ΔF, where ΔF is in response to whole-atmosphere changes in composition and/or other rapid adjustments; no masking of the response is applied. For example, the ERF quantified from piClim-HC minus piClim-control includes the direct radiative effect from the increase in ODS concentrations, the indirect radiative effect of the whole-atmosphere O_3 response as well as other changes to the TOA radiative fluxes due to whole-atmosphere and land surface rapid adjustments."

(See Section 4; lines 324-328 of the marked-up manuscript)

It is only when we consider the radiative forcing by tropospheric ozone – in order to compare against Stevenson et al. (2013), the IPCC 5th assessment report (Myhre et al., 2013), and other estimates – that we mask off the stratosphere and only consider the tropospheric O3 response. This was done solely to make the comparisons with Stevenson et al. and other estimates as comparable as possible. However, we recognise that methane and the other ozone precursors may affect ozone in the lower stratosphere and likewise, that ozone depleting substances may affect tropospheric ozone – we estimate the offset to tropospheric ozone radiative forcing by ODSs in Section 4.4.2. We add a statement to Section 4.4.2 to make clear that masking has been applied in this case:

"While the ERF captures changes in radiative fluxes at the TOA due to whole-atmosphere responses, here, we mask off the stratosphere and focus solely on the tropospheric O_3 response."

(Section 4.2.2; Lines 794-795 of marked-up document).

We do not assess the RF due to stratospheric ozone changes or total ozone changes. Stratospheric ozone depletion is discussed in the Morgenstern et al. paper in the context of the multi-model assessment of the ERF from piClim-HC. There is a paper accepted for publication by R. Skeie, which is a comprehensive assessment of the historical evolution of total ozone forcing from the CMIP6 models. It includes a detailed discussion of the results from UKESM1. For this reason, we do not explicitly discuss stratospheric ozone RF or total ozone RF here but refer to the relevant papers.

3. L160: piClim-HC gave me even more trouble. I actually thought HC stood for "hydrocarbon" until quite late in the paper. I then found it meant "halocarbon" but it was being used interchangeably with ODS, which was strange until I learnt that non-ODS halocarbons were excluded, for reasons that are not clear to me given the name of the experiment. And it wasn't until later in the paper that it became clear that the impact of HCs on ozone was included in these experiments, but still whether this influence remains stratospheric is unclear. This needed more detail in Table 3 or a new table making clear what is what.

Response: As indicated above, the experiment name piClim-HC was inherited from AerChemMIP (Collins et al., 2017). However, we acknowledge the potential confusion that arose because this was not made clear. As a result, we have added further details to Table 1 and to the manuscript (as indicated above) to clarify what greenhouse gases are perturbed in both piClim-HC and piClim-GHG and how the perturbation to halocarbons differ between the two (See Table 1 on pages 8-9 of marked-up manuscript; Lines 256-265 of marked-up manuscript).

We also specify that the ERF derived from piClim-HC minus piClim-control is the result of the direct radiative effect of ozone-depleting halocarbons, the ozone depletion that they cause, and any other rapid adjustments in the following:

"ERF = ΔF, where ΔF is in response to whole-atmosphere changes in composition and/or other rapid adjustments; no masking of the response is applied. For example, the ERF quantified from piClim-HC minus piClim-control includes the direct radiative effect from the increase in ODS concentrations, the indirect radiative effect of the whole-atmosphere O_3 response as well as other changes to the TOA radiative fluxes due to whole-atmosphere and land surface rapid adjustments."

(See Lines 324-328 *of the marked-up document).*

4. L160: To find that piClim-VOC really meant VOC+CO was also confusing.

Response: As indicated previously, we do not have the option of changing this experiment name. However, as indicated previously, we updated Table 1 and added clarity on what was perturbed in piClim-VOC - See Section 3; Lines 278-279 and Table 1: pages 8-9 in the marked-up document

5. L160: Added clarity would be given if Table 3 could make clear the distinction between experiments that impose changes in emissions and those that apply changes in concentrations (or in lower boundary conditions). Even after reading the paper, I am not entirely clear on some of these (e.g. methane). For this table, HC should be spelt out (or replaced by ODS) and it should be explicit what gases are included in the Trop O3 precursors. The authors should also reconsider whether "aerO3" is a better descriptor than NTCF, or removed entirely.

Response: Thank you for this suggestion. We have modified the column titles in Table 1, as indicated previously. And we have altered the titles of the rows in Table 3 to make clear whether a change in concentration or emissions is responsible for the quoted ERFs – See Table 1 on pages 8-9 and Table 3 on pages 14-16 of marked-up manuscript.

6. An additional and quite prevalent issue is that many of the important references are to papers that are submitted (but unavailable) and in some cases only "in preparation" – some key results was in these papers and I felt a bit teased by the allusion to these, without being given either sufficient information or the underlying punchline. I am particularly concerned about the overlap between this paper and Andrews et al. (2019 - submitted) which, from the title of that paper includes forcings in UKESM1, and the Morgenstern et al (2019 - in preparation) paper that contains a discussion of a key result (HC-driven net forcing is strongly negative) when no details at all are presented here, beyond a cursory (4 line) section. This also leads to a very mixed feeling to the paper. Tropospheric ozone is handled in depth, while stratospheric ozone is dealt with in a cursory way, with a reference to the in preparation paper.

Response: We acknowledge that this was problematic and resulted from a high number of papers being submitted ahead of the submission deadline for papers to be considered in the 6th assessment report of the IPCC (December 2019). We were definitely reticent about what to include here on the results from piClim-HC because, at the time, the Morgenstern et al. manuscript was in review.

Nevertheless, in direct response to your concern about the overlap with the Andrews et al. (2019) paper, we can now confirm that Andrews et al. has been published – see https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019MS001866

Despite the title, the only forcings discussed in that paper are based on the UK's physical climate model, HadGEM3-GC3.1 from the RFMIP experiments. Our paper is quite distinct - it quantifies ERFs in UKESM1 and includes the RFMIP experiments as well as the more comprehensive set of experiments from AerChemMIP. As a result, there is a greater range of ERFs investigated in this manuscript and for those RFMIP experiments that are common between the two papers, there may be differences in ERFs due to the additional Earth System interactions included in UKESM1. As a result, we include comparisons with the estimates from Andrews et al., (2019), where appropriate.

In the case of Morgenstern et al., that paper has subsequently been re-scoped, includes additional models, and has been re-submitted and hence, still under review. However, the leaf area index bugfix outlined above has substantially reduced the magnitude of the present-day ERF from piClim-HC (from -0.33 in the submitted manuscript to -0.18 W/m2 in the revised manuscript). It is still an interesting result and we now include a short section on the ERF from piClim-HC for completeness and a new figure (Section 4.2.3, page 26-27 of the marked-up manuscript).

In relation to the balance between the treatment of stratospheric and tropospheric ozone forcing, this was in large part due to the Morgenstern et al. and Skeie et al. papers. Neither of these papers discuss tropospheric ozone RF in detail or attribute it to different factors. Hence, why we focussed more on tropospheric ozone RF in this paper. The impact of stratospheric ozone depletion (from piClim-HC) on the present-day forcing by tropospheric ozone was quantified in this study. And we now acknowledge that by masking off the stratosphere, we are not fully accounting for the RF resulting from methane and the other precursor gases (e.g., Sovde et al., 2011 & corrigendum):

See Section 4.4.2 - lines 815-818 and lines 823-825 of the marked-up manuscript

Of the other papers that were cited but were in review or in preparation at the time of submission, we can now confirm that the following are available:

Andrews et al., https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019MS001866, 2019.

Andrews and Forster, https://www.nature.com/articles/s41558-020-0696-1?proof=t, 2020.

Bellouin et al., https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019RG000660, 2020.

Mulcahy et al., https://www.geosci-model-dev-discuss.net/gmd-2019-357/.

Sellar et al., https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019MS001739, 2019.

Thornhill et al., https://www.atmos-chem-phys-discuss.net/acp-2019-1205/.

These citations have been updated – See Lines 1316-1318; 1320-1321; 1344-1348; 1559-1563; 1644-1650; 1731-1735 in the marked-up manuscript.

7. I am concerned that it is hard to understand what features in the geographical plots are signal and which result from unforced variability. This needs to be discussed in more detail, especially for forcings such as N2O that are inherently small. This is one clear disadvantage of the ERF and partly the reason for my comment at 104/127, below, about ERF being the metric of choice. I do not believe it is the metric of choice in all circumstances, especially for small forcings.

Response: We thank the reviewer for this comment. We have now adapted Figures 3-6 (Page 20, Page 24, Page 27, and Page 30) and Figures 8 (Page 34), 12 (Page 48), and 14 (Page 52), such that areas of the globe, where the unforced variability is outside of the 95 % confidence level, are masked out in white. Indeed, taking N2O as an example, the ERF over most of the globe is not statistically significant at the 95 % confidence interval, except for the longwave clear-sky (LWcs') component, which has a clear signal mainly over the oceans. This is made clear in the Figure captions and when referring to the contributions of the different components to the total ERF (See Lines 386-412 of revised manuscript).

Other Comments (but note that these are not exhaustive)

25: "altering the sign" – it would help readers to know what the sign has been altered from and to.

Response: This has changed due to the updated simulations following the LAI bugfix. However, we have been more explicit with the updated results:

"Ozone (O3) precursors gases consisting of volatile organic compounds (VOCs), carbon monoxide (CO), and nitrogen oxides (NOx) but excluding CH4, in addition to exert a positive radiative forcing due to increases in O_3. However, they also lead to oxidant changes, which in turn cause an indirect aerosol ERF. The net effect is that the ERF from PD-PI changes in NOx emissions is negligible at 0.03 ± 0.04 W m^{-2}, while the ERF from changes in VOC and CO emissions is 0.33 ± 0.04 W m^{-2}."

See Abstract Lines 34-38 of marked-up manuscript

49: I would say the change of concentrations resulting from changing emissions is the first part of the cause-effect change

Response: This has now been altered as follows:

"An important part of this cause-effect chain from activity to climate response, mediated through the atmosphere and the land surface, is quantifying changes to the Earth's radiation budget, often termed radiative forcing."

See line 59 of revised manuscript

65-94: The intro, with its focus on LLGHGs and tropospheric ozone, feels unbalanced. Why are other forcings not discussed here (or is this detail inappropriate here)? The split between tropospheric and stratospheric ozone forcing is to some extent artificial and a little dated (see e.g. Sovde et al.), because the mapping of the drivers of ozone change does not map exclusively on to ozone change in the troposphere and stratosphere, and that could be clearly acknowledged 104 and 127: "preferred" and "metric of choice" – by who?

Response: We agree that the introduction to the study was unbalanced. In response, we have added some background on aerosol forcing (Lines 108-120 of revised manuscript) as follows:

"Aerosol forcing involves a wide range of physical processes. These include (i) direct changes to the radiation budget through scattering and absorption of both SW and LW radiation (e.g. Haywood and Boucher, 2000), (ii) indirect impacts on the radiation budget by changing the microphysical properties of clouds (Twomey et al., 1977), and (iii) changes in the distribution of cloud cover or condensate that follow on from perturbations in cloud microphysics (Albrecht 1989) or radiative heating by aerosol (Hansen et al., 1997). Direct aerosol radiative forcing can be calculated using offline radiative transfer models in a similar manner to GHG and O3 forcing, whereas assessing impacts of aerosols on clouds requires simulations in atmospheric models. The 5th assessment report

(AR5) of the IPCC recommended the ERF framework as an effective way of assessing the overall aerosol forcing as it enables the more complex cloud impacts to be evaluated as part of the climate's rapid adjustments (Myhre et al., 2013). To simplify terminology, AR5 also made a clear distinction between components of the forcing driven by Aerosol-Radiation-Interactions (i.e. the direct or instantaneous forcing) and Aerosol-Cloud-Interactions (that include all indirect or semi-direct cloud-related forcings). Despite wide-ranging and on-going research, the role of aerosols remains the leading source of uncertainty in PI to PD climate forcing, due to the difficulty in constraining the sensitivity of clouds to changing microphysical processes (Bellouin et al., 2019)"

And we have added a section on land-use RF estimates (See lines 122-131 of revised manuscript) as follows:

"In the case of land use, RF estimates have been made using a single GCM simulation with a double-call to the radiation scheme (e.g. Betts et al., 2007) or by comparing paired simulations that include rapid adjustments (e.g. Andrews et al., 2016). However, the choice of RF calculation is not the major source of differences in RF estimates. Similar to O3, uncertainty in PI land-cover is a major source of uncertainty in land-use RF (e.g. de Noblet-Ducoudré et al., 2012). Historically, deforestation has been the dominant type of land-use change, and this causes a positive RF due to increased CO2 emissions and a negative RF due to increased surface albedo. Here, we include the effects of land-use CO2 emissions in the CO2 RF estimates and the land-use RF is due to biophysical changes, predominately albedo. Deforestation has a much larger effect on albedo in snowy regions and model biases in snow cover also contribute to uncertainty in land-use RF (Pitman et al., 2011). Land-use RF estimates also vary due to different time periods being considered (Myhre et al., 2013a), because, unlike many other forcings, there was substantial land-use change before the industrial revolution."

In response to the comment on ozone and the strat/trop split, we agree with the reviewer's comment. In the manuscript, we acknowledge it in the Introduction and add the Søvde et al. (2011) reference:

"Other uncertainties in Stevenson et al. (2013) arise from neglecting the change in O3 in the lower stratosphere attributable to changes in O3 precursors and the contribution from stratospheric O3 depletion on the modelled changes in tropospheric O3 (e.g., Søvde et al., 2011a; 2011b)."

See Lines 102-104 of revised manuscript

In relation to the ERF being the metric of choice, we have now moved the first statement (and removed the duplicate statement) and provided an additional reference to support it as follows:

"Although the uncertainty associated with an ERF tends to be larger than that of RF (or SARF), it is more representative of the climate response than the traditional RF (Hansen et al., 2005; Forster et al., 2016). As a result, it is now the preferred metric of choice for ranking the drivers of climate change (Boucher et al., 2013; Forster et al., 2016)."

See Lines 158-161 of revised manuscript.

117: Xia et al – although it would be tiresome to say "found in their model" on all occasions in this paper, I think it is important in cases, as it is a result of a single model study

Response: Added (Line 154 of marked-up manuscript)

119-121: This needs clarification. For RF it is a requirement that tropopause and TOA forcings are identical (not "nearly identical") following SARF, so in a sense the TOA/trop distinction doesn't matter. However, where it does matter is that RF requires the specification of the tropopause, which

is always to some extent arbitrary. whereas ERF does not – the model just does its own thing. This is one advantage of ERF.

Response: Thank you for this comment. We have now added more clarity following the suggestion as follows:

"The SARF and ERF differ in where the change in radiative fluxes is diagnosed. The SARF, diagnosed at the tropopause, requires the tropopause to be defined but the ERF has the advantage of being diagnosed at the TOA, with no need for a tropopause. While the SARF at the tropopause and TOA are, by definition, identical, this is not the case for the ERF; climate forcings can lead to adjustments in stratospheric circulation and therefore changes to dynamical heating above the tropopause. Changes in dynamical heating are thus balanced by radiative divergence across the stratosphere, explaining the change in net fluxes between the tropopause and TOA."

See Lines 167-172 of marked-up manuscript

154: "addressed the strong negative" – the nuance here (and in other places in the paper) is that it was too strong negative (i.e. unrealistic). Is that the intention?

Response: We have altered the statement to state explicitly that the aerosol ERF in the GA7.0 configuration of HadGEM3 was -2.75 W m^{-2}, leading to an unrealistic negative total anthropogenic forcing for the 20th century. The improvements implemented in Mulcahy et al. reduced the magnitude of the aerosol ERF to -1.45 W m^{-2} (Lines 209-213 of marked-up manuscript)

268: "from energy budget constraints" – I have no idea what this means and hence whether the Andrews and Forster (2019 - submitted) paper is in some ways a superior assessment to that given here. Should we (or Andrews and Forster) be concerned that the UKESM forcing falls outside their stated range? This is another problem with the heavy reliance on unpublished papers.

Response: The Andrews and Forster study has now been published and the citation updated (Lines 1320-1321 of marked-up document) – see https://www.nature.com/articles/s41558-020-0696-1?proof=t. We have now added more detail on the approach used by the Andrews and Forster study to constrain the total anthropogenic ERF at the present day (Lines 370-372 of marked-up document). In addition, the updated simulations with the corrected seasonal cycle in the prescribed leaf area index result in the total anthropogenic ERF from UKESM1 changing from 1.61 to 1.76 W m^{-2}; this corrected estimate now falls just within the range of estimates from Andrews and Forster.

282: "some evidence" – I did not know what this meant. It either is, or it isn't? Or is this meant to imply that the derived forcing is no larger than the unforced variability in this region, and hence is not a robust signal? There is no indication of statistical significance in any of the plots, or discussion of it and this should be rectified in some way. See Major Comment 7 282: This paragraph is the first place where it is made clear that the HC calculations include the ozone forcing from ODSs (and indeed that HC=ODS) and leaves the reader unclear whether this is the case for (e.g.) methane and CO2.

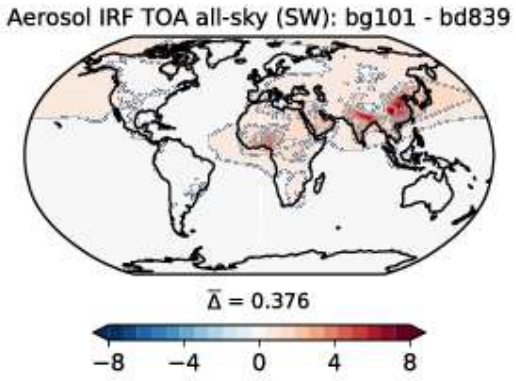
Response: We have removed any reference to "some evidence" and made the statement more explicit that the GHG ERF is strongly positive everywhere except for the southern hemisphere high latitudes, where it is negative as a result of the contribution from the piClim-HC experiment. As outlined in previous responses, the plot of the top-level forcings (Figure 3: Page 20 of marked-up document) masks out those areas where the signal is not statistically different from zero at the 95 % confidence interval. It means that the negative ERF in the southern high latitudes from the piClim-HC and piClim-GHG experiments is a robust signal and not due to variability.

287: The GHG forcing is rather inhomogeneous. Is this statement that the aerosol forcing is more so based on an objective measure of inhomogeneity or a subjective assessment?

Response: Thank you for this comment. We have since carried out a chi-squared test for homogeneity and found that the global distributions of the ERFs from the top-level experiments (piClim-GHG, piClim-aer, piClim-O3, piClim-LU, and piClim-anthro) are all inhomogeneous. However, it does find that the aerosol ERF distribution plot is more heterogeneous than that for GHGs. We also quantify the percentage of the global area giving an ERF statistically significant from zero and outside the variability at the 95 % confidence level. (Lines 386-412 of marked-up manuscript)

288-290: Is this conjecture or the result of model analysis?

Response: This is a statement based on our model analysis. Over bright surfaces, the aerosol IRF can become positive if the aerosol mixture is sufficiently absorbing (e.g. Haywood and Boucher, 2000) and this happens in our model over parts of the Sahara, Arabia and Southern Asia as shown in Fig 3c. To ascertain this, we looked at IRF and ERF from the individual aerosol components (BC, OC, sulphate) as well as the total aerosol ERF. The BC IRF was particularly high in those regions as shown in the figure below.



292: Here it is unclear whether ozone change from methane is in the GHG or the trop O3 plot

Response: We have now added in a number of places in the document that methane is considered in the piClim-CH4 experiment, while the other ozone precursor gases (VOC, CO, and NOx) are considered in the piClim-VOC (VOC and CO), piClim-NOx, and piClim-O3 (VOC, CO, and NOx) experiments. We add explicit reference to VOC, CO, and NOx here to directly address this (Lines 402-405 of marked-up manuscript).

296-297: Is the land use forcing over the ocean real (resulting from downstream rapid adjustments) or just indicating unforced variability in the model?

Response: The land use ERF over the ocean was predominantly due to unforced variability and is not a statistically significant signal at the 95 % confidence interval. These areas have now been masked out in white from the new Figure 14 (page 52 of the marked-up manuscript).

298: This negative forcing over the mid-lat continents is a striking result. Is this the first paper to indicate it? Also, presumably the Land Use forcing in these areas is mostly "historical". This paper understandably does not address the time dimension of the evolution of the forcing, but perhaps some added text could be added to indicate when this forcing would have been most active.

Response: The strong negative forcing over mid-latitude continents is a common result in GCM studies of historical land use change (e.g. de Noblet-Ducoudre et al., 2012). In the large regions with a strong negative ERF, forests have been replaced by agriculture, which has remained upto the present day, and the associated SW forcing has also persisted. Comparing transient coupled UKESM1 simulations with and without historical land use change, shows the albedo response in North America gradually grows from the 1850s to the 1960s, with a northward expansion starting after about 1900 (not shown). The albedo response in Eurasia gradually increases from 1900 until the 1980s.

The following text has been added at Lines 1089-1091 of the revised manuscript.

"The large regions of negative ERF in the northern mid-latitudes are consistent with previous model studies (e.g. de Noblet-Ducoudre et al., 2012); they are caused by agricultural expansion and the forcing is expected to have gradually increased from about 1850-1970 in North America and from 1900-1980 in Eurasia."

303: I wonder if adding an additional frame with the zonal-mean forcings of each component would be useful for bringing out the overall larger-scale structures of the forcing.

Response: Figure 3 (Page 20 of the marked-up manuscript) which plots total GHG, aerosol, ozone precursors, land use and total anthropogenic ERFs now includes zonal mean plots. This is a very informative addition and we thank for the reviewer for the suggestion.

303: It is striking that the geographical plots in each figure differ, presumably because different coauthors were responsible for them, with different conventions in the headers, and in the case of Figure 8, a different projection. Could these be made more consistent? Personally, I find the addition of global mean values on each frame, as in Figure 4, quite helpful.

Response: Thank you for this observation and your assumption that different co-authors produced the different figures was indeed correct! We have now made the global distribution plots as consistent as possible, e.g., masking out regions (with white) where the signal is not statistically different from zero at the 95 % confidence interval, adding in global mean values, using the same order of subplots, using the same projection, figure captions, etc..

See Figures 3 (Page 20), Figure 4 (Page 24), Figure 5 (Page 27), Figure 6 (Page 30), Figure 8 (Page 34), Figure 12 (Page 48), and Figure 14 (Page 52).

323: Here and elsewhere it should be made clear if the concentrations are surface values or mass-weighted means over the whole atmosphere (and whether a single global mean values is used everywhere in the model for each LLGHGs).

Response: This has previously been addressed in response to another comment - See Section 3 (page 10; Lines 258-272) in the marked-up manuscript.

328: 121 ppm. Presumably the expressions in Etminan et al. could be used to derive their forcing for the same CO2 change?

Response: Yes, indeed. The equivalent from Etminan et al. is 1.80 W m^{-2}. The revised ERF is, however, 1.89 W m^{-2}. The text has been updated accordingly (Lines 446-447 of the marked-up manuscript).

331: Some of the CO2 SW forcing likely comes from the SW bands of CO2, but this is implicitly discounted here. Or are these bands not included in the radiation code?

Response: Yes, the radiation scheme does include SW forcing from the CO2 bands in the near-IR. However, as you say, the text implicitly discounts this. We have updated the text to note that the direct effect is included but is small compared to the LW based on results from the radiation codes in RFMIP (Pincus et al., 2020) – Lines 453-448 in the marked-up manuscript. Using this assumption, we estimate the clear-sky component of the ERF in the SW is dominated by the rapid albedo response in our model.

352: These two estimates overlap within the uncertainties.

Response: Given the leaf area index bugfix and the new simulations, our updated ERFs are no longer overlapping. The new ERF is now substantially larger than the cited value – See line 480 in the marked-up manuscript.

358: "noise" – here and throughout, it is important to guide the reader as to which features are noise and which are robust. Could this be done via masking or stippling? For example, is the marked negative feature in clear-sky longwave in Figure 4c over Russia and its surroundings, robust?

Response: Figure 4 (and other figures) have been modified to show only regions where ERF components have significant signals at the 95 % confidence interval; all other regions are masked with white. With updated ERFs shown in the figure, the negative contribution to the ERF from the clear-sky longwave (LWcs') component over Russia and its surroundings are no longer significant (and now appear masked white).

Other global distribution plots also now only show regions where ERFs (and/or its components) have a signal statistically different from zero at the 95 % confidence interval

363-365 and Figure 5: 364 talks about "increased ozone" (in the UTLS) but no hint is given as to what drives this feature. Is it a "self-healing" effect of the overlying depletion? The caption to Fig 5 needs to make clear this is for the N2O experiment.

Response: The updated ERF due to N2O following the leaf area index bugfix now shows a marked difference from the original ERFs. We have now dropped figure 5 and the associated discussion. The reviewer may well be right about the self-healing but we have not assessed that. The finding is consistent with the literature (e.g., Morgenstern et al., ACP, 2018).

396-398: There is no hint of what is going on here, except a reference to a paper that is "in preparation". This is very frustrating and leads me to question whether the HC/ODS results should be shown here at all, if there is no accompanying analysis of the striking result. It became a bit unclear to me why, for example, the tropospheric ozone is discussed in such detail in this paper, but not stratospheric ozone. Also note that the assumption that HC=ODS is made here without elaboration, although later there is some hint at this.

Response: We appreciate the reviewer's comments and have now included a figure and some discussion of the ERF of ozone-depleting HCs. The forcing from stratospheric and/or total ozone forcing are discussed in two other papers: Morgenstern et al. (In review) and Skeie et al. (Accepted) manuscripts. As a result, it would not be appropriate to include here.

414-415: Again, this allusion to a different submitted paper, with no hint of what is found, is frustrating

Response: The O'Connor et al. paper has received reviews but requires some additional analysis before publication. So, rather than expanding on their findings, we include some comparisons of the

methane ERF with the AerChemMIP multi-model ensemble, and cite the discussion paper of Thornhill et al., which is available at:

https://www.atmos-chem-phys-discuss.net/acp-2019-1205/

as follows:

"The UKESM1 ERF quantified here is at the upper end of estimates from the recent study of AerChemMIP multi-model ERFs by Thornhill et al. (2019). They found that the multi-model mean ERF was 0.70 W m^{-2}, with a standard deviation of 0.22 W m^{-2}. They attributed part of the inter-model spread to different complexities in the representation of interactive chemistry in the respective models i.e. some models only captured the direct radiative effect of CH_4 (e.g., NorESM2) while others (e.g., UKESM1) also included indirect contributions from CH_4-driven changes in O_3 and SWV. However, the contribution to the ERF from tropospheric adjustments differed in both magnitude and sign, with UKESM1 being the only model with a positive rapid adjustment. The relative contributions of the direct and indirect forcings to the total CH_4 ERF quantified here and the mechanism behind the positive rapid adjustment can be found in O'Connor et al. (2020)."

See lines 580-588 of revised manuscript

423: The frames in Figure 6 are in a different order to those in Figure 4, and there are a different number. Commonality of presentation would help the reader. Again, the caption to this Figure does not make it clear it is for methane. (All captions should be checked for this – e.g. Figure 12 suffers the same lack of specificity.)

Response: As indicated above, the figures have now been more consistent, both in terms of including global mean values, order of sub-plots, map projection, figure captions, etc. - See Figures 3 (Page 20, Figure 4 (Page 24), Figure 5 (Page 27), Figure 6 (Page 30), Figure 8 (Page 34), Figure 12 (Page 48), and Figure 14 (Page 52) in the marked-up manuscript

440: This is the first place the reader learns that the non-ODS halocarbons are not included in HC.

Response: This has now been made clearer both in Table 1 (Page 8-9), Table 3 (Page 14-15), and in the description of the experimental setup in Section 3 (Lines 253-266 of marked up manuscript). We have also re-named the section (Section 4.2.3 – Page 26) presenting the results of piClim-HC as "Ozone (O_3) Depleting Substances (ODSs)" to again make clear that the experiment piClim-HC only considers ozone-depleting halocarbons.

444: "warrant further investigation" – I agree. this difference is interesting and one factor not considered (spectral overlap) would likely push the difference in the wrong direction (i.e. make GHG even lower than the sum of individual components)

Response: As the reviewer points out, this was indeed potentially very interesting. We had quantified IRFs (using an offline radiative transfer code) and found that the sum of the individual IRFs was higher than the total GHG IRF, suggesting that spectral overlap and/or saturation effects were occurring but were in the opposite sense to what the ERFs quantified from the UKESM1 simulations were indicating. We chose not to include this at the time of writing, because we were considering doing a separate study on the non-linearity of the GHG ERF.

However, the apparent non-linearity has somewhat disappeared with the updated simulations, in which we have corrected the seasonal cycle in the prescribed leaf area index! We now find that the sum of the ERFs from piClim-CO2, piClim-CH4, piClim-N2O and piClim-HC is 2.93 ± 0.08 W m^{-2},

indicating that the ERFs add linearly and agree with that from piClim-GHG (2.92 ± 0.04 W m^{-2}). See the end of Section 4.2 (Lines 626-629 of the marked-up manuscript).

473: There is a minor ambiguity in using "sum", "all" and "total". As I understand the Total=All? This also applies to the text from 513-521.

Response: Yes, the reviewer is correct – we had used "total" and "all" interchangeably. We have removed all references to "total" to avoid any doubt. See lines 649-659 and the caption of Figure 7 (Page 33 of the marked-up manuscript)

484: "many other models" – but aren't versions of HadGEM included in the Stjern et al. 2017 paper?

Response: Well, not quite. Stjern et al. (2017) included HadGEM2 (the UK's CMIP5 model) and an early development version of HadGEM3 that still had the mass-based aerosol scheme (CLASSIC) from HadGEM2. UKESM1 now has a new and totally different aerosol scheme (GLOMAP-mode; Mann et al., 2010; Mulcahy et al., 2019) with quite a different representation for black carbon. As a result, the UKESM1 results presented here are quite separate from those included in the Stjern et al. study.

494-496: This sentence seems out of place, especially without any supporting reference.

Response: Agreed, this has now been removed (Lines 686-687 in the marked-up manuscript)

550: I had a lot of comments on Section 4.4 as it was particularly unclear in my view. One underlying issue, raised above, is whether the tropospheric precursor gases are allowed to change lower stratospheric ozone, and whether that change is included here. Without a clear statement, it is very hard to understand this section and place it in the context of prior work. Similarly, the lack of clarity on whether CH4 is included or not, as an emitted species, is unhelpful.

Response: We have addressed the lack of clarity with the experimental set up in our previous responses above (Section 3, Table 1, Table 3).

For the ERF estimates, no masking of the response is applied. However, when estimating the tropospheric ozone RF in order to compare with previous estimates, we mask off the stratosphere. We add the following statement to make this clear: "While the ERF captures changes in radiative fluxes at the TOA due to whole-atmosphere responses, here, we mask off the stratosphere and focus solely on the tropospheric ozone response" to Section 4.4.2 (See Lines 794-795 of the marked-up document).

559: This figure purports to show "tropospheric column" but the panels are labelled in ppbv. Is this the mean (mass weighted) ppbv over the depth of the tropophere? How is the troposphere defined for this purpose? Would Dobson Units be a better measure? Figure 5a included a latitude-height plot of ozone change due to N2O, and a similar plot would be helpful here.

Response: This was a typo, as the data are in Dobson Units. The tropospheric column was calculated by applying a monthly-mean tropospheric mask, derived from the PV+380K level, to the total ozone column. (Figure 9, page 36 of marked-up document)

565-566: I wasn't sure why the "South Pacific" was highlighted and here and when "Western Pacific" is mentioned, I presume in both cases it means tropical Pacific?

Response: This has been amended to Eastern tropical pacific. (See line 767 of the marked-up document)

588: By applying this mask, how much forcing (due to the precursors' impact on stratospheric ozone) is excluded from the analysis?

Response: We used a tropospheric O_3 radiative kernel and hence, masked off the stratosphere. However, we acknowledge that this may have the effect of underestimating the resulting tropospheric O_3 RF from methane and the other O_3 precursor gases. For example, Sovde et al. (2011a, b) estimate that of the O_3 RF from methane and the other precursors, approximately 85 % is due to changes in tropospheric O_3, with the remaining 15 % due to changes in the stratosphere.

We have now added the following statement to the discussion:

"However, in masking off the stratosphere, the UKESM1 estimate for the O_3 RF attributable to O_3 precursors is likely to be underestimated. For example, Sovde et al. (2011a; b) estimates that approximately 15 % of the O_3 response from changes in CH_4 and other O_3 precursors may be in the stratosphere, and hence not considered here. " (See lines 823-825 of the marked-up document)

594: I found this paragraph hard to follow, and in particular whether differences in the literature were due to radiation code differences (or the effect of applying a kernel) or process level differences in determining the ozone change. I ended up confused as to whether the new result was higher or lower than Stevenson et al. and for what (dominant) reason it was so.

Response: We apologise for any lack of clarity. We have now removed any discussion on the radiation code. Although it contributes to the uncertainty of the tropospheric O_3 RF in Stevenson et al. (2013), it is only one of the contributing factors. Other factors include inter-model spread in O_3 response, emissions uncertainty, definition of tropopause etc. Stevenson et al. estimate an overall uncertainty of ± 30 % on their central estimate. As a result, our estimate is within the uncertainty of the best 1850-2010 estimate from Stevenson et al. (2013) and is consistent with other estimates. The updated paragraph is as follows:

"Figure 10 shows the global distribution of the PD tropospheric O_3 RF from piClim-CH4, piClim-NOx, and piClim-VOC experiments and their sum using the kernel approach. It shows that the tropospheric O_3 RF is strongest over the northern hemisphere tropics and weakest over the southern hemisphere high latitudes. The strongest RF occurs in regions of warm surface temperatures and high albedo, coinciding with the largest tropospheric O_3 change (Shindell et al., 2013a). As was the case in Stevenson et al. (2013), the forcing is weaker over regions of high altitude (e.g., Tibetan Plateau) due to there being less O_3 column aloft to absorb in the LW. The tropospheric O_3 RF from the kernel approach is 414 mW m^{-2}. However, the increase in tropospheric O_3 at the PD (and its resulting RF) are offset by decreases due to ODSs (e.g., Søvde et al., 2011a; Søvde et al., 2011b; Shindell et al., 2013a). Applying the kernel method to the diagnosed decrease in tropospheric O_3 from the piClim-HC experiment (Sect. 4.2.3), we find a RF offset of -101 mW m^{-2}. Although larger in magnitude by nearly a factor of 2 than the estimates from Søvde et al. (2011a; 2011b) and Shindell et al. (2013a) due to the strong O_3 depletion in UKESM1 (Keeble et al., 2019; Morgenstern et al., 2020), it reduces our original estimate to 313 mW m^{-2}. This revised estimate is within the 30 % uncertainty of the Stevenson et al. (2013) best estimate, albeit lower than their central estimate by 13 %. It is also consistent with a number of other estimates: the CMIP6 historical O_3 dataset (312 mW m^{-2} from Checa-Garcia et al., 2018), a recent study in which observational isotopic data was used as a constraint on historical increases in tropospheric O_3 (330 mW m^{-2} derived from the GEOSChem model in Yeung et al., 2019), and a parametric model based on multi-model source-receptor relationships (290 ± 3 mW m^{-2} from Turnock et al., 2019). However, in masking off the stratosphere, the UKESM1 estimate for the PD O_3 RF attributable to O_3 precursors is likely to be underestimated. For example,

Søvde et al. (2011a; b) estimates that approximately 15 % of the O_3 response from changes in CH_4 and other O_3 precursors may be in the stratosphere, and hence not considered here. "

See lines 794-818 of the marked-up manuscript

607: "Section 3.2.3" – I presume this means Section 4.2.3 but, as noted above, there is essentially no detail in 4.2.3, and certainly nothing on the tropospheric ozone change in piClim-HC. This is frustrating when trying to make sense of this section.

Response: Thank you for spotting this error – now corrected. Any discussion on the tropospheric O3 change from ODSs is calculated and discussed within this section – see our response above.

616: It would be useful to have the global-mean values, perhaps in the same way that they are shown on Figure 4 in each frame. Also, some comment about how different Fig 10d and Fig 3c are (i.e. RF versus ERF) would seem useful.

Response: In updating Figure 10 due to the updated simulations following the LAI bugfix, we now include global mean values – See page 39 of marked-up manuscript. However, we do not add any discussion of the differences between Figure 10d and Figure 3c for a number of reasons:

1. *Figure 10d includes the impact of CH_4 on the tropospheric O_3 RF whereas methane is part of the GHG ERF in Figure 3*
2. *We find that the ozone precursors perturb aerosols and hence, are responsible for significant cloud adjustments, which are included in the ERF plotted in Figure 3c. These adjustments are excluded from the tropospheric O_3 SARF plotted in Figure 10d. These adjustments have a considerable impact on the ERF – see the CRE component in Table 3 (page 14 of the marked-up document).*

666: Are the methane concentrations, the change (i.e. NOx – control, as indicated in Column 1) or the absolute values for NOx. It might be better to show the change?

Response: The concentrations shown for methane were the equilibrium concentration that would have occurred had the model been emissions-driven rather than concentration-driven. However, we have now changed this table to include this as a change in methane concentration, i.e., equilibrium concentration minus prescribed concentration. See Table 5 (page 41 of the marked-up manuscript).

680: Make clear that this is the OH at 1 km.

Response: This has been done – See line 899 of the marked-up document

686: These do not appear large (e.g. compared to NOx)

Response: The word large has been removed – see line 904 of the marked-up document

695: "eastern Pacific" – does this mean the eastern tropical Pacific, as I see changes of both signs in the E Pacific?

Response; Reference to the eastern Pacific has been removed so that the sentence now reads "The impact of NOx emissions appears to dominate, given the similarity between piClim-O3 and piClim-NOx" - see lines 913-915 of the marked-up document

699: AOD is meaningless without specifying a wavelength. I'm curious why the VOC/CO run has an increase in AOD but a decrease in CDNC (e.g. over the W Pacific). Is this because one compares column integrated amounts and the other is at 1 km?

Response: The AOD was calculated for 550nm. The discussion of the second part is somewhat moot as we now believe that many of the AOD changes to not pass significance testing. We agree that column-integrated AOD is not necessarily correlated with CDNC changes at 1km, having chosen this altitude for its relevance to cloud formation processes, particularly in marine environments.

712: "negligible" – is this because of cancellation of regions of positive and negative, but the local values can be large?

Response: On a global mean basis, the aerosol IRF is less than 0.01 ± 0.01 W m^{-2}. Although regionally, the IRF varies between -0.31 and 0.25 W m^{-2}, it is only statistically significant over about 5 % of the globe.

726: Are the NOx sources just from the surface, or are aviation emissions also included? I know it is beyond the scope of this work, but it is of interest whether this net negative applies to all sources or likely just surface sources.

Response: The perturbation experiment piClim-NOx perturbs all anthropogenic emission sources of NOx – both at the surface and from aircraft. A sentence to this effect has now been added to the Section on experimental set up – see line 277 of marked-up manuscript

739: Again, a reference to an "in preparation" paper is unhelpful here.

Response: This paper has been published as a discussion paper and can be found at:

https://www.atmos-chem-phys-discuss.net/acp-2019-1205/

The corresponding citation has been updated (Lines 1731-1735 of marked-up manuscript).

750: I think this would be better labelled as Aerosol/Ozone, rather than NTCF since it clearly isn't all NTCFs (which would avoid the "excluding methane" repetition). Since both aerosol and ozone are treated separately anyway, I am not quite sure of the need for this section, especially as they, together, only form a subset of NTCFs.

Response: As indicated previously, we have chosen to retain the name "Near Term Climate Forcers" to be consistent with the experiment name from the AerChemMIP protocol. However, we note in Table 1 that piClim-NTCF is also known as piClim-AerO3 in RFMIP and we re-iterate that in Section 3 (Lines 283 of marked-up manuscript). We have also chosen to retain this section to indicate that the aerosol and ozone precursor ERFs do not add linearly.

769-771: These lines seem to contradict each other about whether it is cloud fraction or cloud optical depth, or maybe I misunderstand the point being made.

Response: Any discussion about the LWcre' component and high-level cloud fraction has been removed. The SWcre' component is correlated with the cloud fraction.

776: Units are missing for the temperature and cloud parameters. Are temperature changes really in K?

Response: The updated Figure 12 now only shows the different components of the ERF – See Page 48 of the marked-up document

810-811 "too strong" and "bias still exists" – I haven't read the Robertson paper, but could it be said what the too strong albedo response is relative to (other models? observational constraints?), and

whether this indicates that UKESM is definitely incorrect (as this text implies) rather than being a plausible outlier?

Response: Robertson (2019) evaluates the HadGEM2-ES response to land-use change against observational estimates and shows that the model has an incorrect albedo response. This bias was not addressed in the development of UKESM1 although it may have inadvertently been fixed.

Lines 1069-1072 have been changed to:

"Robertson (2019) showed that, even in the absence of snow cover, the albedo response to land-use change in HadGEM2-ES is stronger than observed and it is likely that this bias still exists in UKESM1."

Typos etc (but note that these are not exhaustive)

25: If "forcing" means "ERF" it should say this.

Response: Done (See examples on lines 187, 194, etc. of marked-up manuscript)

83-89: Long sentence 64 and elsewhere: "it's . . . doesn't . . . weren't" – it doesn't worry me much, but such contractions are not usually used in formal scientific writing

Response: These have now been removed (e.g., Lines 95, 893, etc. of revised manuscript)

216: machine's

Response: Corrected (Line 307 of marked-up manuscript)

236: These equations use the subscripts cs and clear – do they mean the same thing?

Response: Yes, they do, and we have now replaced "clear" with "cs" – See lines 332 and 341 in marked-up manuscript

552: refer to Figure 3 here?

Response: Done - See line 751 in marked-up manuscript

592: "Garcia"?

Response: Thanks for spotting this error – Now corrected (See line 798 in marked-up manuscript)

608: I think it important to also refer to the Sovde corrigendum https://doi.org/10.5194/acp-12-7725-2012.

Response: Agreed; it has now been added – See lines 104 and 823-825 in marked-up manuscript. Reference now added – See line 1721 of the marked-up manuscript

Anonymous Reviewer #3:

To the editor,

I think this work serves two main purposes, i.e., (1) documenting (and disentangling) the ERF from short-lived emissions, long-lived GHGs, land-use, and ODSs in UKESM1, and (2) indicating the relevance of ESM-interactions for estimating the anthropogenic forcing.

I think the work is solid, the subject is well explained, comparisons with earlier studies are included, and the information shared in figures and tables is appropriate.

Presenting estimates of ERF is important to understand the evolution of climate (modeled or observed). Models participating in CMIP6 were motivated to perform specific experiments in AerChemMIP and RFMIP, which would allow a better characterization of the forcings felt by ESMs, e.g., over the historical period. It is important that those ERF estimates are well-documented, as they facilitate the interpretation of the fully coupled behaviour of models (and climate). I think this study achieves presenting ERF estimates for the UKESM1 model in a nice, coherent, and attractive way. In general, the study is well written and attractive to read.

In addition, as UKESM1 contains a relatively extensive description of atmospheric chemistry, aerosols, and aerosol-cloud processes, the interactions between different species and processes are included. This makes this model an interesting tool to estimate the ERF of various short-lived emissions or longer-lived GHGs.

I think this work is valuable and of sufficient quality to be published in ACP. However, I think the manuscript could be considerably improved. I have listed several points on which I think the manuscript should be changed.

I have grouped my comments in three categories, i.e., main remarks, technical remarks, and additional detailed remarks.

The authors acknowledge Anonymous Reviewer #3 for their positive response to the manuscript and for the very detailed and comprehensive suggestions for improvements. Thank you! We aim to address them in our responses below and in the changes applied to the manuscript.

Main remarks

1. TABLE 3

Table 3 is very interesting and plays to my opinion a rather central role in this study. A lot of the values in the table are being referred to in the text. However, some parts of Table 3 are not discussed. This is the case for two large and two small groups of experiments : SO2-BC-OC-Aer, NOx-VOC-O3, LU and anthro. For these groups, only the NET ERF is discussed, but not the individual contributions. I assume this is related to the fact that for those 4 groups another split than the one in Eqs. (6)-(8) is followed, which is probably assumed to be more appropriate.

To make the study more consistent, I would at least mention that the non-discussed values exist in the Table 3, and are given for comparison/completeness/consistency. Explicitly presenting the numbers from the other approach followed for some experiments in a table would be interesting. Now these numbers are only mentioned in the text.

Response: Thank you for this constructive comment. Firstly, we would like to confirm that we have used Eqn. (8) for estimating all of the components of the anthropogenic ERFs for consistency. We state this in Section 4.1 (Line 353 of marked-up manuscript).

There are a lot of data values in Table 3 and the text draws out the most interesting and important points from these. Given the size of the paper, we do not think it would help to duplicate that information even further with additional tables or discussions. However, we do refer to Table 3 in the section on land use (Lines 1058-1061 of marked-up manuscript), the total anthropogenic ERF (Line 1128 of marked-up manuscript, and we also include a statement that the clear-sky and cloud radiative effect components can be found in Table 3 in the aerosol section (Lines 637-638 of marked-up manuscript). And it is mentioned in the context of non-linearity of the ERFs for the ozone precursors (Lines 925-936 of the marked-up document).

2. EXPLANATION OF RADIATIVE FORCING

Although radiative forcing and effectve radiative forcings are very useful concepts, it is not always so easy to explain them. I think the authors have done a nice effort in trying to explain it. However, an extra effort should be done to make it more precise, fully coherent, and more illustrative. The final link why it is a useful concept should be elaborated more. I list here below also some specific places where I think the description should be improved.

- page 2, line 61-63 : Including ... . As a result ... : I do not think that "As a result" is appropriate, as I do not think there is a causal relationship.

Response: In response to Reviewer #1, we have altered that section as follows:

"Although the error uncertainty associated with an ERF tends to be larger than that of RF (or SARF), climate sensitivity parameters (i.e. the degree of warming per unit forcing) are less dependent on the forcing agent and it is more representative of the climate response than the traditional RF (Hansen et al., 2005; Forster et al., 2016). As a result, it is now the preferred metric of choice for ranking the drivers of climate change (Boucher et al., 2013; Forster et al., 2016)."

This removes the incorrect causal link – See lines 158-162 of marked-up manuscript.

- page 2, line 63-64 : RF is a framework : this is a rather vague explanation.

Response: The reference to "framework" has been removed and replaced with "concept", i.e., that forcing is the "rapid" radiative response to a change in activity (e.g., emissions change) which occurs before the longer-term climate response occurs.

See lines 74 and 95 of the marked-up manuscript

- page 4, line 115-117 : I think that the sentence on line 115-117 illustrates that ERF can be very different from IRF, however, it does not illustrate that it is better (as suggested in the former sentence). It might be an option to reverse the order of these sentences.

Response: Now done – see lines 153-162 of marked-up manuscript.

- page 4, line 113-115 : although it is correct what is written here, I think it should be explained better.

Response: We have now included a reference to the seminal work by Hansen et al. (2005), in which they found that the efficacies of climate forcings were less dependent on the forcing agent and closer to 1 for ERFs than for SARFs - See lines 156-158 of marked-up manuscript.

3. FEEDBACKS, INTERACTIONS, NON-LINEARITY, INDIRECT EFFECTS

I think that feedbacks, interactions and indirect effects are an important part of this paper. In the abstract, it is stated that "... by quantifying ..., it enables the role of various climate-chemistry-aerosol-cloud feedbacks to be quantified." However, despite this sentence in the abstract, not all feedbacks have been quantified (it would be a very large task to quantify them all). The fact that they are active in UKESM1, and impact therefore ERF estimates, is already an important step. I think the text will benefit from more precisely describing what is quantified, and what one is not able to quantify.

Response: We agree that feedbacks, interactions and indirect effects are an important part of the paper and have a large impact on estimates of climate forcing. As the reviewer rightly points out, we

cannot quantify all of them and hence, "quantified" was perhaps a poor choice of word. We have now replaced this with "investigated".

I do not ask the authors to do additional analysis, or perform more comparisons with a model which contains less interactions (e.g., HadGEM-GC3.1). For quite some of the forcings, the authors did a very nice job in unraveling several contributions. However, it would be nice if the authors could be more careful and precise in some of the general expressions about feedbacks and interactions.

Response: In the Introduction, we state that "feedbacks are related to global mean temperature change and forcings are not (Sherwood et al., 2015)". However, in some cases, the term "feedback" was used where it was not associated with a global mean temperature change. In these cases, we have replaced "feedbacks" with "interactions", where the term "interaction" refers to a coupling (e.g., chemistry-aerosol-cloud coupling) in the ESM that results in an indirect contribution to the quantified ERF, and is considered part of the forcing. See examples on lines 20, 46, 336, 1157, 1172, and 1239 of the marked-up manuscript.`

Another example is the term "aerosol-mediated cloud feedback". This term has been replaced with "aerosol-mediated cloud adjustment" and is more consistent with the ERF definition. See, for example, lines 1038 and 1145 of the marked-up manuscript

I have in addition the following comments.

1) I think the authors should explain what according to them is the difference between a feedback and an interaction.

Response: See response above.

2) It is a pitty that the studies of Morgenstern et al. [2019, in preparation] on ozone and O'Connor et al. [2019, submitted] on methane have not been published yet.

Response: As indicated above in response to Reviewer #1, this is because these papers were part of a large cohort of papers submitted ahead of the paper submission deadline to be considered for inclusion in the 6th assessment report of the Intergovernmental Panel on Climate Change. Unfortunately, neither paper has been published yet but we chose to keep them as separate papers to avoid this manuscript becoming excessively long. However, we now include a longer paragraph on the ERF from piClim-HC (Section 4.2.3: Page 26-27 of the marked-up manuscript) and include some discussion on the comparison between the ERF estimated from piClim-CH4 here with estimates from the AerChemMIP multi-model ensemble (Thornhill et al. 2019) – See lines 580-588 of marked-up manuscript.

3) Some of the formulations, referring to feedbacks and interactions, are often rather vague (less firm than the abstract suggests). Examples of these are :

- page 15, line 354 : ... "likely" ... the effect of adjustments ... including O3 depletion and fast cloud adjustments

Response: The discussion has now been removed. The updated simulations no longer show the big effect discussed here previously. The leading effect now clearly is the direct RF due to the GHG property of N2O.

- page 15, line 359 : which "might" be due to an aerosol effect

Response: As for previous response directly above

4) the abstract stresses well the non-additivity (or non-linearity) of several ERF estimates. However, the study sometimes remains vague in reasons for it, and just uses expressions like “the non-linearity between the GHG”, or “the internal mixing of aerosols”. Other examples of not so clear usage of the concepts :

Response: As indicated in a previous response, we had to re-run all the experiments due to a bug in the seasonal cycle of the prescribed leaf area index. The non-linearity previously seen in the GHG ERFs is no longer evident and any reference to it has been removed (e.g. Lines 613-618 of marked-up manuscript).

In the case of aerosols, Section 4.3 includes a few sentences that explain why aerosol forcings do not add up linearly:

“The main reasons for this lack of linearity are: (i) Cloud droplet numbers do not increase linearly …”. (ii) The absorption of upwelling shortwave radiation…

However, it didn’t explain why internal mixing contributes to non-linearity. This additional sentence has been added as a point (iii) into that paragraph:

“Internal mixing creates an interdependency between different aerosol sources, meaning that the aerosol size distributions, optical scattering efficiency and hygroscopicity evolve differently depending on the absolute and relative abundance of different mass components (sulphate, OC, BC and sea salt mass) and differing rates of new particle production via primary emission or nucleation. “ – See lines 713-716 of marked-up manuscript

- page 2, line 36 : due to the inclusion of non-linear feedbacks and ES interactions : vague

Response: Now removed – See Lines 42-44 of marked-up manuscript

- page 2, line 38 : By including feedbacks between GHGs (isn’t it interactions?)

Response: Changed to “interactions” – See line 46 of marked-up manuscript

- page 2, line 38-39 : some of which act non-linearly

Response: Now removed – See Line 47 of marked-up manuscript

4. UNCERTAINTIES

The ERF estimates in the manuscript are accompanied by an uncertainty range in large parts of the main text and in Table 3, but not in the abstract, the conclusions, or some parts of the main text. This should be made consistent.

Response: Thank you for this comment. We have now added the model estimates for the ERF uncertainties in the Abstract (See Pages 1-2 of marked-up manuscript) and the conclusions (See Pages 54-57 of marked-up manuscript) and in other relevant sections, e.g., CO2 – See lines 440-454, N2O – See lines 473-485, ODSs – See lines 529-541, CH4 – See lines 562-578, GHG – See lines 606-610, Aerosol – See lines 649, 658-659, 670, 673, and 684. For ozone precursors – See line 750, NTCFs – See lines 986-1000, and land use – See lines 1058-1061, and total anthropogenic – See line 1126 of the marked-up manuscript.

Also, one should maybe say something about the real estimate of the uncertainty. There is something on aerosol ERF uncertainty on page 20, line 453-455, and maybe this can be expanded. In general, UKESM1 shows uncertainties of around 0.02 to 0.04 W/m2 on global 30-year mean

averages of ERF – it is the uncertainty on the ERF in the model. However, if one wants to see these numbers as best guesses for the ERF as experienced by the Earth, then the uncertainty is probably much larger. Emission uncertainty, lack of process understanding, too low spatial resolution, biases in the mean state of UKESM1 and others factors might contribute to the uncertainty. It would be interesting if the authors could also shed some light on this.

Response: Thank you for this comment. You are indeed correct that the quoted uncertainties are solely based on the signal to noise of the modelled ERF and do not represent the true uncertainty. As you say, the true uncertainty will be larger due to a number of factors, e.g., uncertainties in emissions, systematic biases and/or lack of sensitivity to emissions due to representation of chemistry, pre-industrial background aerosol state, missing and/or unresolved processes, etc..

Although a comprehensive assessment of the true uncertainty in ERF estimates is beyond the scope of the paper, it is indeed an interesting question and we acknowledge this in Section 4.1 (Lines 354-359 of marked-up manuscript).

5. VALIDATION OF THE MODEL

There is very little information on the validation of the model, except via references to Mulcahy et al. [2018, 2019] and Archibald et al. [2019]. There are a few paragraphs (e.g., page 20, line 454-457 : aerosol comparison; page 21, line 491-492 : AOD comparison), mentioning some comparisons with observations. We do not ask for a detailed comparison, but some qualitative main findings on the performance of the model might be mentioned, both related to distributions of forcing agents (aerosols, ozone, ...) as to the general behaviour of the model (mean model state).

Response: In relation to aerosols, we do indeed make use of the findings from Johnson et al. (2019) and Mulcahy et al. (2020) in the discussion of the aerosol forcing. We also now include more discussion on the ERF from ODSs (in response to a reviewer comment) and comment on how the performance of stratospheric ozone in UKESM1 may be impacting on that ERF – this is further discussion in Morgenstern et al. (2019) and Skeie et al. (2020). We also include some discussion on the impact of strong O3 depletion on the tropospheric O3 RF. And how the performance of the model may be affecting estimates of the CO2 ERF and the land use ERF is also discussed.

See Lines 549-560, 815-819, 456-472, 1068-1071 of the marked-up manuscript.

Technical remarks

1. POSITIVE NUMBERS

Some positive numbers have a "+", some not. This should be made consistent.

Response: All those numbers with a "+" have had it removed for consistency – See, for example, Table 3 (Page 14-16) and lines 362, 366, 367, 370, and 405, etc.

2. REFERENCES TO THE PHYSICAL MODEL

It might be relevant to explain or highlight the relevant differences between UKESM1 and HadGEM3-GC3.1 better (both models are very close and have both participated in CMIP6 I assume). In the text quite often comparisons with HadGEM3-GC3.1 are made, and it would illustrate the role of having more/less interactions in a coupled model. I would think that possible differences are related to fixed ozone, fixed oxidants for secondary aerosol formation, fixed methane profile, fixed CO2 profile, ...

Response: A table describing the main differences between the atmosphere-only configurations of HadGEM3-GC3.1 (called HadGEM3-GA7.1) and UKESM1 has been added to the manuscript as an appendix – See page 58 of the marked-up manuscript.

3. REFERENCE TO FORSTER ET AL. [2016] FOR STANDARD ERROR

The way to calculate the standard error is mentioned three times. It should be mentioned when the uncertainty is met for the first time (both in the text and in the tables). The standard error and its reference are now mentioned in :

- page 12, line 272 (Table 3).

Response: Now removed in multiple places and it is now only mentioned once in Section 4.1 (See line 354-355 in the *marked-up document*

- page 18, line 402 (related to CH4) : where the 0.04 W m-2 is the standard error following Forster et al. (2016) : should be mentioned once (the first time).

Response: Now removed (line 564 of marked-up document)

- page 31, line 760-761.

Response: Now removed (*line 987 of marked-up document*)

4. MULTI-ANNUAL MEAN / ANNUAL MEAN / MEAN

It is mentioned in the beginning of the manuscript that almost all values and maps will be 30-year averages. Some sentences and figure captions in the text treating those values or maps once more mention explicitly that the averages are "multi-annual", whereas other figures captions or sentences do not. Please describe it in a consistent ways. An option might be to describe it clearly at some point in the, and say that it is valid for all the remaining text. Examples of differences in the description :

- page 18, line 395 : multi-annual mean (while page 14, line 323 : global mean)

- page 24, line 562 : The multi-annual mean ...

- page 25, line 599 : on a global annual mean basis

- page 29, line 699 : Global distributions of the multi-annual distributions

- page 33, line 795 : multi-annual global mean

Response: The manuscript already states the following: "Therefore, all simulations were 45 years in length. Using the latter 30 years of the paired simulations, the ERFs from the PI-to-PD perturbations were diagnosed as the time-mean global-mean difference in the TOA net radiative fluxes. " in Section 3 (Lines 252-254 of manuscript).

We removed all reference to "multi-annual mean", "global annual mean basis", "multi-annual distributions", "multi-annual global mean" etc.., as suggested. See examples at lines 386, 417, 529, 690, 759, 763, 806, 919, 1047, 1174, 1184, and 1191 of marked-up document.

5. WRITING OF LAND USE

Both "land-use" and "land use" are used in the text.

Response: Changed all occurrences of "land-use" to "land use" for consistency (e.g. Section 4.5.2, lines 1057, 1062, 1068, and 1070, as examples, in the marked-up document).

6. CONSISTENCY

- page 10, line 261-265 : why is a difference of 0.2 W/m2 for GHG called "consistent", but for the aerosol mentioned "less then the estimate from HadGEM3-GC3.1"?

Response: This has now been altered. With the updated experiments, we now state that the aerosol forcing is consistent with HadGEM3-GC3.1 but the GHG ERF is lower i.e., 2.92 ± 0.04 cf. 3.09 W m^{-2}. See lines 361--364 of marked-up manuscript.

- page 22, line 502 : -0.12 W m-2 (-0.4 to +0.1), whereas on page 10, line 267 : 2.3 (+1.7 to +3.0) [5-95%] W m-2 : the position of Wm-2 is different.

Response: These have been made consistent with each other. See lines 368-370 of revised manuscript

- page 22, line 515-519 : only here "(i)" and "(ii)" are used. I would not use it just for this single occasion

Response: In response to your previous comment, we have now added a third point in relation to internal mixtures. As a result, we have opted to retain the numbering (i) (ii), and (iii) – See lines 705-712 of marked-up manuscript.

- page 23, line 536 : only here "&" is used. I would not use it just for this single occasion.

Response: Replaced with "and" – see line 734 of marked-up manuscript

7. NAMING of SWcs, Lwcs, SWcre and LWcre COMPONENT OF ERF

There are different abbreviations used to express the same physical quantity. One should make these coherent. An example of this is the LW clear-sky component of ERF, which appears in different ways in the text, tables, and equations :

- page 11, Table 3 : LWcs

- page 10, Eq. 8 : ERFcs

- in the text : a CS LW component

- page 16, Fig. 4 : (a) clear sky SW, (c) clear sky LW, and (f) net (SW+LW) CRE : it sounds as something is missing after SW and LW

- page 31, line 762-770 : NETcs, LWcs, SWcre

- page 32, Fig. 12, caption : SWcs, LWcs, SWcre, and LWcre

Response: In all cases, we have adopted consistent names for the different components of the ERF: SWcs', LWcs', SWcre', LWcre', NETcs', and NETcre', as given in Eqn. (8). These have been used throughout the manuscript, e.g., Lines 449 onwards for CO2, Lines 475 onwards for N2O, Lines 527-531 onwards for ODSs, Lines 564 onwards for CH4, etc. Figure captions have also been updated, e.g., for Figures 4, 5, 6, etc.

8. STATISTICAL SIGNIFICANCE OF DIFFERENCE BETWEEN IDENTICAL SIMULATIONS

Several experiments have been performed on different machines. Related to this I have the following remarks :

- Table 4 : I assume there are three piClim-SO2 experiments, so one could show the difference R1-R2 but also R1-R3.

Response: As detailed in Table 2 (Page 12 of the marked-up manuscript), there are only two piClim-SO2 experiments, realisation 1 (R1) was performed on the NIWA XC50 and realisation 2 (R2) was performed on the Met Office Cray XC40. To make this clearer, we have changed the title of the last column in Table 2 from "Realisation" to "Realisation ID" and placed "R" in front of each number to indicate more clearly that there are only 2 realisations of the piClim-SO2 experiment i.e. R1 and R2. The Realisation ID is unique for each realisation of a particular simulation and forms part of the CMIP6 metadata.

- Table 4 : Some differences are a bit large compared to their uncertainty for SO2 : NET ERF, SW CRE, NET CRE. Is there an explanation for this?

Response: The non-linearity of the equations being solved in Earth System Models makes them sensitive to the propagation of small perturbations making the models sensitive to change in HPC platform. As the model integrates (through time), these perturbations grow randomly changing the climate state of the experiment (e.g. cloud development, precipitation, temperature, etc.). The results in Table 4 show the impact in terms of ERF from the change in HPC platform and their associated standard error (as a measure of uncertainty). Even though some values in the original manuscript appeared large (as pointed out by the reviewer), the Kolmogorov-Smirnov statistical test applied showed that these differences in the time series were not statistically significant in terms of their distribution (Please refer to lines 424-432 of the marked-up manuscript). The updated results in Table 4 (Page 21 of the marked-up manuscript) from correcting the seasonal cycle of leaf area index no longer appear to be large and any difference in ERF is within the uncertainty.

- Table 4 : Why is everything 0 for the NTCF experiment?

Response: The 2 piClim-NTCF realisations were performed on similar machines, a Cray XC40 using similar compilers: Cray compiling environment 8.3.7 for R1 and 8.3.4 for R2. Because of this, the results of experiments produced on these machines were similar and no differences were found between these realisation pairs. The updated results in Table 4 (Page 21 of the marked-up manuscript) from correcting the seasonal cycle of leaf area index no longer appear as zero, and are now similar to what was found for other realisations.

9. NOx EMISSIONS AND SO2

Figure 11 : is the impact of NOx emissions on AOD not mainly in regions with SO2 emissions from volcanoes? Indonesia, west coast of South America, Etna region, existing ship-lanes in 1850 over North Atlantic? Might probably NOx have then also a large effect if it is co-emitted with anthropogenic SO2?

Response: This is an interesting observation, and it likely is the case that coupling between NOx and SO2 emissions drives stronger AOD changes. The UKESM1 model includes climatological SO2 emissions from non-explosive volcanic degassing (following Dentener et al., 2006; https://www.atmos-chem-phys.net/6/4321/2006/acp-6-4321-2006.pdf). However, while the AOD changes are largest in regions where these SO2 emissions occur, they are not truly co-located – NOx emissions occur at the surface while SO2 emissions are released at higher levels. We feel that exploring this coupling is beyond the scope of this study, and so do not specifically refer to this point in the manuscript.

10. FORMULAS

The "." or "," at the end of formulas can be closer to the actual equations (currently there is a 1.5 cm large gap). There should not always be a ":" at the end of the text just before a formula. This depends on the context. One can also have a ",", or nothing.

Response: Done – see lines 68, 146, 324, 333, 341, and 343 marked-up manuscript

11. REFERENCES TO SECTIONS

The references to specific sections are not always correct. Please check and correct them. A list of incorrect section references :

- page 8, line 192 : Sect 3.1 -> 4.1

- page 12, line 287 : Sect. 3.2 -> Sect 4.2

- page 13, line 291 : Sect. 3.3 -> Sect 4.3

- page 13, line 296 : Sect. 3.4 -> Sect4.4

- page 14 ,line 307 : Sect. 2 -> Sect 3.

- page 25, line 602 : Sect. 3.2.3 -> Sect 4.2.3

- page 29, line 705 : Sect. 3.4.2 -> Sect 4.4.2

Response: Thank you for spotting these errors. These have been corrected. See lines 255, 395, 401, 407, 424, 815, and 926 of marked-up manuscript.

12. USE OF e.g.

There should be a "," before and after "e.g.".

Response: Now corrected – See lines 72, 78, 86, 92, and 98, etc. We also corrected similar errors with "i.e." – See lines 117, 157, 211, 236, 247, etc.

13. REFERENCE TO EQUATIONS

The way one refers to equations is not homogeneous in the text (sometimes capital letters, small letters, abbreviations, or no abbreviations). Please make it coherent. A list of the differences found in the text :

- page 4, line 111 : Eqn. 1

- page 10, line 251-252 : in equation (5)

- page 15, line 348 : following the equations (6)-(8) described above

- page 19, line 425 : using Equations (6)-(8)

- page 27, line 643 : From Eqns. (10) and (11)

Response: All references to equations are now consistent – See lines 347, 353, 376, 383, 475, etc. of marked-up manuscript.

14. USE OF THE WORD "FORCINGS"

I have the impression that the word "forcing" (outside the context of radiative forcing and effective radiative forcing) is possibly not used in a completely coherent way. I have listed below a few

locations where it is used, and it seems not to always have the same meaning. I think the text should be more careful and precise in how and where it is used.

- page 1, line 16 : climate forcings

Response: Changed to "effective radiative forcings" – See line 16 of the marked-up manuscript

- page 1, line 31 : a positive forcing due to ozone

Response: Changed to "a positive radiative forcing due to increases in O3" – See line 35 of the marked-up manuscript

- page 3, line 65 : various mechanisms, both anthropogenic and natural. Here I would have used the word forcing and not "mechanism".

Response: Changed – See line 77

- page 3, line 65 : ... use of RF ... it is often used inconsistently ... : is RF meant in the first part of the sentence, because that is well-defined (page 2, line 54-56)?

Response: We agree that RF is well defined but here, we are saying that it is often calculated inconsistently, e.g., GHGs vs ozone. There are different methods for calculating RF, e.g., line-by-line radiative transfer calculations, formulae, GCM radiative transfer calculations, Gregory regression from coupled models, timeslice atmos-only experiments, different inputs (e.g., observed concentrations, modelled concentrations) etc.

- page 3, line 67-69 : This sentence mentions three things, and it is not clear whether the reader should see a causal relationship between "the inconsistent calculation of forcing between drivers", and "the large differences in forcing between CMIP5 models". For the last part of the sentence, it is not clear whether the authors mean that models are very different or that the methods of estimating forcing can be very different.

Response: There is no causality intended. We have now added that the large differences in forcing between CMIP5 models for CO2 can be due to either calculation method and/or model diversity – See line 81 of the marked-up manuscript.

- page 3, line 82 : the resulting RF : is not the same as meant by the definition on page 2 line 54-56.

*Response: Agreed, changed as follows: "*The simulations used the corresponding sea surface temperatures (SSTs) and sea ice (SI) conditions for the time periods of interest (PI and PD), therefore allowing some climate response and feedbacks at the PD, implying that the resulting estimate is not consistent with the RF definition; it does not fit into the simple forcing-feedback concept, whereby feedbacks are related to global mean temperature change and forcings are not (Sherwood et al., 2015). " *– See lines 92-96 of the marked-up manuscript*

- page 7, line 171 : in which SSTs and SI and all forcings

Response: Changed "forcings" to "boundary conditions" – See line 235

- page 33, line 782 : between the aerosol and O3 forcings

Response: Changed to "As a result, this study also attempts to estimate the effects of the non-linear interactions between chemistry and aerosols on the combined aerosol and O_3 precursor ERF" – See lines 1017-1018

- page 35, line 840-845 : “forcing” used four times. Also “anthropogenic” forcings and “natural” forcings, whereas it is not so clear what to understand here under natural forcing. In the text there is a reference “As summarized above”, but I think there was not a large focus on “natural forcings”.

Response: Changed to “As noted above, historical climate change has been driven by a wide range of anthropogenic activities that act together, alongside natural changes, to perturb the Earth’s radiation balance. The total anthropogenic ERF is, therefore, a key metric in understanding observed and modelled changes in the climate system since the PI era. These various anthropogenic drivers are not necessarily independent of each other and it is therefore worthwhile calculating the total anthropogenic ERF from a separate timeslice simulation including all PD perturbations together (piClim-anthro; Pincus et al., 2016). “ – See lines 1113 onwards of the marked-up manuscript.

15. THE USE OF ABBREVIATIONS

The definition of an abbreviation is often given several times, whereas that should be

limited to :

- defined once in the abstract,

- defined once in the conclusions,

- defined once in the figure caption, and

- defined once in the main text (but not more).

Once a definition is given in the main text, it should not be defined again.

Response: Done for Abstract – See Page 1-2 of the marked-up manuscript; Done for Conclusions – See Pages 54-56 of the marked-up manuscript; We could find no occurrence of where an abbreviation was defined twice within a figure caption – no action taken. Done for main text – See lines 64, 141, 143, 148, 149, 190, 218, 220, 236, 241, 288, 289, 298, 339, and 400, for examples.

16. PD AND PI ARE ADJECTIVES

- page 1, line 21 : at the PD, or relative to the pre-industrial (PI) : PD and PI are both adjectives, and not substantives. So they cannot be used on their own. E.g., on page

1, line 15 it is used correctly as an adjective : "... a wide range of present-day (PD) anthropogenic climate forcings ...".

Response: Any reference to PI and PD have been corrected. See Line 375, 383, 388, 418 etc. in the marked-up manuscript.

17. PD-PI DIFFERENCE

The way the ERF is described as a difference between TOA fluxes in PD and PI times, is not coherent in the text. Sometimes one finds a "PD forcing" type of expression, sometimes a "PD to PI forcing" type of expression, and sometimes neither PD nor PI are mentioned. Some examples of the varying usage are given below. It would be nice if the description could be more coherent. An option would be to write in the text that all values shown in text from that point onwards are always PD-PI differences, and then PD and PI should not be repeated every time.

- page 3, line 71 : between the pre-industrial (PI) and the present day (PD)

- page 3, line 76 : at the PD relative to the PI

- page 5, line 130 : the PI to PD effective radiative forcings (ERFs)

- page 5, line 132 and 133 : PD forcings will be quantified relative to the PI [in addition is "the" strange before PI which is just an adjective]

- page 5, line 139 : PD anthropogenic forcings relative to PI

- page 6, line 157 : To calculate the pre-industrial to present-day effective radiative forcings ... due to a PI-to-PD perturbation

- page 10, line 260 : the global mean PD ERF [here without any reference to PI]

- page 12, line 271 : PD effective radiative forcings (ERFs) of climate relative to PI

- page 12, Fig. 2 : y axis PI-to-PD ERF

- page 12, line 280-281 : PD ERFs from changes in ... since PI

- page 12, line 281 : PD ERF from GHG [but no reference to PI]

- page 13, line 303-304 : effective radiative forcing (ERF) relative to the pre-industrial

- page 14, line 319 : cloud radiative effect (CRE) at the present day [no reference to PI]

- page 15, line 338 : this forcing is small for the PD [no reference to PI]

- page 15, line 347 : due to changes in N2O from PI to PD

- page 16, line 370 : for the present relative to the pre-industrial

- page 18, line 400 : resulting in a PD ERF [no reference to PI]

- page 19, line 423 : of the PD methane ERF relative to PI

- page 19, line 430 : This PD GHG ERF

- page 20, line 462-463 : of PI to PD aerosol ERF

- page 21, line 472 (Fig. 7) : Aerosol ERF at TOA ... [no PD or PI mentioned]

- page 22, line 501 : The PD OC ERF relative to PI

- page 23, line 551 : The global mean ERF from ... [no PD or PI mentioned]

- page 25, line 591 : the PI-to-PD change in tropospheric O3

- page 27, line 651-652 : the tropospheric O3 RF between PI and PD

- page 28, line 679 : and the PD aerosol ERF

- page 30, line 723 : ERF from PI-to-PD changes

- page 31, line 744 : the PD ERF from piClim-O3

- page 33, line 795 : of changes (PD-PI) in

Response: Thank you for pointing this out. In Section 4, we have the following statement: "Therefore, all simulations were 45 years in length. Using the latter 30 years of the paired simulations, the ERFs were diagnosed as the time-mean global-mean PD-PI difference in the TOA net radiative fluxes.

Details on how the ERF was further decomposed can be found in Sect. 4.1." – See lines 252-254 of marked-up manuscript.

*We also write "*The aim of the current study is to quantify PD (Year 2014) ERFs from anthropogenic drivers of climate change with an atmosphere-only configuration of an ESM. Using the experimental protocol recommended for the Radiative Forcing Model Intercomparison Project (RFMIP; Pincus et al., 2016), PD ERFs will be quantified relative to the PI period from PD-PI changes in emissions, concentrations, and/or land use due to anthropogenic activities." – See lines 184-188 of the marked-up manuscript.

In response to the lack of consistency, we have therefore removed most references to PI and PD when referring to the quantified ERFs in the subsequent sections – See lines 25, 40, 89, 194, 361, etc. of the marked-up manuscript for examples.

18. TIMESLICE EXPERIMENTS

Maybe it would be nice to better define what a timeslice experiment is. I would possibly describe it as an experiment with fixed boundary conditions (possibly having seasonal cycles), which one runs for several years to reduce the noise-to-signal ratio (the noise is caused by inter-annual variability). The longer one runs, the better estimate for the mean one can obtain. The locations were it is used are :

- page 6, line 158 : maybe explain timeslice here (forcings are kept constant)

- page 7, line 171 : time slice

- page 15, line 604 : timeslices : in ACCMIP they were different from here ...

Response: In Section 3, we introduce the term "timeslice" and we have added further details as suggested. For example, we now state that "SSTs, SI and all other boundary conditions were fixed at year-1850 levels" for the pre-industrial control simulation – See lines 234 onwards of marked-up manuscript. We also note that "Running for 30 years when the model has reached steady state reduces the uncertainty associated with meteorological variability (e.g., Shindell et al., 2013a) and improves the estimate of the ERF. " – See lines 254 onwards of marked-up manuscript.

19. FIGURE 4

Fig. 4, panel (a) : Why is there only some spatial variability north of 30N?

Response: This figure has now been updated following the leaf area index bugfix and the new figure no longer shows this feature.

20. CONSISTENCY OF THE FIGURES

The figures and their captions should be more coherent throughout the manuscript. Below I list some places where improvements should be made.

- page 16, Fig. 4 : at the top of the atmosphere (TOA) : this is however not mentioned in Fig. 3.

- page 17, Fig. 5 : why mentioning "annual mean" or "multi-annual mean", whereas it is not mentioned in Fig. 4.

- page 19, Fig. 6 : it would be nice to have the global mean values given in the figure heading.

- page 19, Fig. 6 : mentions "according to Ghan (2013)". Why not in Fig. 4?

- units in figures : Fig 7 uses "(W/m2)", whereas other figures use "/ W m-2".

- page 24, Fig. 9 : of the “multi-annual” distributions, whereas other figures do not mention “mult-annual”.

- page 26, Fig. 10 : it would be nice to add the global mean RF.

- page 32, Fig. 12 : it would be nice to add the global mean values

Response: Figures and Figure captions have been made more consistent, as indicated above to a previous comment. Figures also include global mean values, where appropriate (e.g., Figure 3, Figure 4, Figure 5, Figure 6, Figure 10, Figure 12). We have also removed all references to “multi-annual annual mean”, “multi-annual mean”, etc., in response to a previous comment. We have also made the units in Figure 7 more consistent (i.e., W m^{-2}) with the other figures.

Additional detailed remarks

ABSTRACT :

- page 1, line 23 and 28 : "larger than the sum of the individual GHG ERFs" and "less than the sum of the individual speciated aerosol ERFs" : I don’t know whether these aspects should be mentioned in the abstract.

Response: Now removed – See line 26

- page 1, line 15 : "In this paper" : I would not mention "In this paper" in the abstract.

Response: Removed – See line 15

- page 1, line 18-19 : by quantifying ..., it enables ... : this sentence seems not completely coherent. There is also twice "quantify" in the same sentence (in addition there was already “quantify” on line 15).

Response: Replaced with “investigated” – See line 20

- page 1, line 21-22 : I would put the numbers at the end of the sentence, and "carbon dioxide, nitrous oxide, ..." at the beginning of the sentence. Now, one first reads numbers, but one does not know what their meaning is.

Response: Good idea – Done! Line 22

- page 1, line 19 : by this sentence, one suggests that climate feedbacks can be quantified by fixed-SST simulations. However, some of them are strongly suppressed in fixed-SST simulations.

Response: We change “feedbacks” to “interactions” and say that they can be investigated.

- page 1, line 25-26 : is the "BC absorption" not part of the "instantaneous forcing from aerosol-radiation interactions"?

Response: Yes, it is. In Section 4.4, we state that “the IRF is rather small and negative (-0.15 ± 0.01 W m^{-2}) (Fig. 7a) due to scattering by sulphate and OC that is partially offset by absorption from BC.” For the abstract, we modify the sentence to make it clearer: A relatively strong negative forcing from aerosol-cloud interactions and a small negative instantaneous forcing from aerosol-radiation interactions from sulphate and organic carbon are partially offset by a substantial forcing from black carbon absorption.”

- page 1, line 27 : mean -> "imply" or "cause".

Response; Done – See line 31

INTRODUCTION

- page 2, line 42-45 : necessary ... detailed ... all aspects : this is probably exaggerated. I suggest to formulate it differently.

Response: Done – See line 52

- page 2, line 42 : attribute it and its impacts : it (refers to climate change) and its impacts. It is not clear whether the authors mean something different with "climate change" and "impacts".

Response: Modified as suggested – See line 51

- page 2, line 44 : climate response and its impacts : (same comment as above).

Response: Modified as suggested – See line 53

- page 2, line 44 : a key mechanism : I would not call CMIP6 a "mechanism". Also "key" is possibly exaggerated - other initiatives (if CMIP would not have existed) might have also had good outcome.

Response: Agreed, modified accordingly. – See line 54

- page 2, line 45-46 : which designs and distributes data : "designs data" sounds strange. Possibly one could say that "experiments are designed".

Response: Done – See line 55

- page 2, line 47 : "these important climate science questions" : it is not clear which questions one refers to.

Response: Removed – See line 57

- page 2, line 50 : quantifying changes to the Earth's radiation budget, often termed radiative forcing : maybe radiative forcing needs a bit more explanation.

Response: Added an explanatory sentence. – See lines 60-61

- page 2, line 64 : It's been -> It has been.

Response: Corrected – See line 76

- page 3, line 65 : the strength of the various mechanisms : of various mechanisms

Response: Corrected – See line 77

- page 3, line 65 : mechanisms, both anthropogenic and natural. I do not think that "mechanism" is the most appropriate word to be used here, especially in the context of "anthropogenic" and "natural".

Response: Changed to "forcings" – See line 77

- page 3, line 70-71 : is typically based on .. and using -> uses.

Response: Corrected – See line 83

- page 3, line 72 : I suggest to put "e.g." before "based on Myhre et al. (1998) and Ramaswamy et al. (2001)", as there might exist other expressions.

Response: Done – See response above

- page 3, line 75 : Skeie et al. [2011] : I think that study is not so much about observational-based estimates of forcing. Is this paper very relevant in the discussion of forcing strength of GHGs?

Response: Removed – See line 86

- page 3, line 81 : including -> therefore allowing.

Response: Done – See line 94

- page 3, line 82 : meaning -> implying.

Response: Done – See line 94

- page 3, line 82 : doesn't -> does not.

Response: Done – See line 95

- page 3, line 84-85 : of a robust ... constraint -> of robust ... constraints.

Response: Done – See line 97

- page 3, line 84 and 89 : twice "additional uncertainties".

Response: Removed duplicate – see line 102

- page 3, line 85-86 : across multi-model ensemble : is that really what the authors want to stress? Might "across models" be sufficient?

Response: Changed as suggested – See line 99

- page 3, line 87 : chemistry models : does one mean CTMs or CCMs?

Response: Changed to "chemistry-climate models" – See line 100

- page 3, line 91-93 : three times the word "uncertainty" in one sentence. Maybe it can be reduced to two.

Response: Done! See line 105

- page 3, line 91-93 : it looks like aerosols get only very limited text attributed.

Response: A section on aerosol forcing has now been added to the Introduction – See lines 108 onwards of marked-up manuscript

- page 4, line 96 : is "schematic" the correct wording?

Response: Changed to "Vertical profiles of temperature" – See line 135

- page 4, line 101 : Although -> Because/As.

Response: Done! See line 140

- page 4, line 102 : maybe "also" can be skipped. I don't know if it really reflects well the meaning of the sentence.

Response: Done – See line 141

- page 4, line 109 : andAi -> and Ai (blanco needed).

Response: Done – See line 148

- page 4, line 110 : or over the land : vague.

Response: Changed to "land surface" and added an extra example of a land surface adjustment, i.e., land surface temperature – See line 151

- page 4, line 112-113 : but global mean surface temperatures or global ocean conditions remain unchanged. However, in reality with fixed-SST simulations, land surface temperature (and thus global mean surface temperatures) can still change a bit.

Response: Correct – have now removed any reference to global mean surface temperatures – See line 152

- page 4, line 113 : error -> uncertainty (I assume the authors mean "uncertainty").

Response: Changed – See line 159

- page 5, line 129 : I would skip "including" because the list mentioned seems rather complete.

Response: Removed – See line 180

- page 5, line 130 : forcings from anthropogenic drivers -> forcing from anthropogenic drivers.

Response: Done! See line 184

- page 5, line 131 : with a fully coupled : the "with" gives the impression that one uses the model here in its "fully-coupled" configuration. But here it is not used in its fully-coupled configuration.

Response: Corrected – changed to "an atmosphere-only configuration of an Earth System Model" – See line 185

SECTION 2

- page 5, line 142 : "is" the atmospheric and land components -> consists of.

Response: Done – See line 197

- page 5, line 150 : "is determined by prescribed oxidant fields" : maybe describe differently, as oxidants are not the only determining factor.

Response: Now changed – See line 205

SECTION 3

- page 7, table 1 : piClim-VOC : add that also CO is perturbed.

Response: Done – See Table 1 (page 8-9)

- page 7, line 166-167 : twice ERF : I think mentioning "fixed-SST" is enough to describe the experiments. Obtaining the ERF is the result of such an experiment.

Response: Caption now changed – See line 227

- page 7, line 171 : Effectively, this involves ... : it is a bit a strange way to mention that also a reference simulation is needed.

Response: I would respectfully argue that it is important to explain the experiment setup of piClim-control, as this underpins all of the other perturbation experiments.

- page 7, line 171 : SSTs and SI and all forcings : one should not have two "and"s in a row.

Response: Corrected – See line 234

- page 7, line 171 : the abbreviations SST and SI have not been defined.

Response: They were defined at the very beginning of Section 3 – See line 93

- page 7, line 173-176 : One uses two different expressions, i.e., "monthly time-varying climatologies derived from 30 years of output" and "30-year monthly mean climatologies", to describe the same thing.

Response: Dropped the "monthly time-varying climatologies" – See line 235

- page 7, line 177-184 : is it not in agreement with RFMIP and AerChemMIP?

Response: Interactive vegetation was specified as part of the RFMIP protocol. However, I could find no discussion about vegetation in the AerChemMIP protocol.

- page 8, line 194 : emissions and/or GHG concentrations : I think this can just be "and".

Response: Now removed

- page 8, line 199 : fixed SST ERF experiments : I think fixed-SST is enough to describe the experiments.

Response: Now removed – line 258

- page 8, line 200 : ammonium nitrate : but other forms of nitrate are probably also not present (e.g., nitrate on dust and seasalt).

Response: Replaced with "nitrate aerosol" – See line 291

- page 8, line 208 : fixed SST timeslice ERF experiments : I think fixed-SST is enough to characterize the experiments.

Response: Replaced with "fSST" which has been previously defined – See line 298

- page 8, line 210 : This makes ... to changes in platform that cannot guarantee bitreproducible results : this can probably be expressed more precisely.

Response: Now changed – See lines 300 onwards

- page 8, line 211-213 : was scientifically consistent with each other : one should be more clear in what is meant by "scientifically consistent".

Response: Now edited – See line 3030

- page 8, line 220 : Further to this -> in addition to this.

Response: Done – See line 311

SECTION 4

- page 9, line 234 : Cloud-Radiative Effect (CRE): in the rest of the text, no capital letters are used when defining an abbreviation.

Respone: Done – See line 329

- page 10, line 240 : either ... and/or : I do not think that it is common to combine "either" with "and/or".

Response: Changed sentence to: "However, many of the experiments in this study either directly perturb aerosol emissions or indirectly alter aerosol concentrations via chemical and dynamical feedbacks." – See line 355

- page 10, line 241 : what is meant by "dynamical feedbacks" : changing meteorology which changes lifetime and thus burden of aerosols?

Response: Yes

- page 10, line 251-252 : and any non-aerosol changes in CS flux : is, e.g., the deposition on snow of BC included in this term?

Response: In UKESM1, deposition of BC acts as a sink for BC and it does not influence the radiation budget once deposited. The non-aerosol changes in CS flux is referring to absorption by gas-phase constituents, e.g., ozone. We have now included this example.

- page 10, line 252 : "The effective radiative forcing (ERF), clear-sky CS), and cloud radiative (CRE) contributions" : as I assume that "contributions" also relates to "clear sky", I would write : "The effective radiative forcing (ERF), and its clear-sky CS) and cloud radiative (CRE) contributions".

Response: Corrected as suggested – See line 352

- page 10, line 257-258 : following equations (6) to (8) : maybe only (8)? (As that is the final split which is presented in Table 3).

Response: Agreed. Now changed to just say Eqn. (8). – See line 353

- page 10, line 262 : HadGEM3 GC3.1 -> HadGEM3-GC3.1

Response: This had been previously removed due to the updated results from the simulations with the LAI bugfix included. See line 368

- page 12, line 271 : effective radiative forcings (ERFs) of climate : I do not think "of climate" is needed here.

Response: Removed, as suggested. See line 375

- page 12, line 273 : use Realisations 2. -> use realisation 2.

Response: Corrected – See line 377

- page 12, Fig. 2, caption : "diagnosed from paired fixed SST timeslice simulations with an atmosphere-only configuration of UKESM1". It is not clear why the fact that paired simulations are needed to estimate ERFs is mentioned here. It is, e.g., not mentioned in Table 3 (although also paired simulations are the bases for the results of Table 3).

Response: Now removed from Figure 2 caption – See line 382.

- page 12, line 283-284 : ERF from the piClim-HC perturbation experiment ... positive forcing from the other LLGHGs : in the first part of the sentence one talks about the ERF from experiments, and in the second part about the ERF of physical things (in this case LLGHGs). The sentence should be improved.

Response: One of those sentences has already been removed and the other was made more consistent. See line 391 onwards

- page 12, line 287 : The aerosol forcing is ... due to their ... : "aerosol" is singular, but "their" refers to something plural. So it sounds a bit strange.

Response: Replaced with "the shorter aerosol lifetime" – See line 397

- page 13, line 294 : "weakly positive in comparison with other forcings" : this sounds a bit strange.

Response: Removed "in comparison with other forcings" – See line 406

- page 13, line 298 : their combined : sounds strange. I would suggest "the combined".

Response: Done – See line 410

- page 13, Fig. 3 : maybe add the global mean values in the figure headings.

Response: Done in response to a previous comment

- page 14, line 308-309 : Despite ..., ... produce slightly different results. Isn't it what one should expect? As it is expected, I would not use "despite".

Response: Changed sentence as follows: "Statistical methods ensure that the model is not scientifically different on the different HPC platforms, but such duplicate experiments still produce slightly different results." – See line 425

- page 14, line 324 : 1.83 Wm-2 (but 1.82 Wm-2 in Table 3).

This value has changed slightly due to the LAI bugfix but it is now consistent between this section (Line 441) and Table 3 (Page 14-15).

- page 14, line 325-328 : It is informative to stress the absolute difference in CO2 ppm. However, maybe one could add that the relative change in CO2 concentration is more relevant for the forcing.

Response: Added that CO2 has a logarithmic dependency on concentration – See line 444

- page 15, line 337 : low-clouds -> low clouds.

Response: Corrected – line 458

- page 15, line 338 : piClim-CO2phys : the "2" should not be an index.

Response: Corrected – See line 459

- page 15, line 340-341 : low-level cloud -> low-level clouds.

Response: Corrected – See line 463

- page 15, line 337-341 : The first sentence gives the impression that the balance comes from two terms. However, the second sentence adds that the balance (or closure) comes from other terms.

Response: These have now been combined into one sentence. See line 449 onwards

- page 15, line 342 : found a much larger effect : this looks like a dramatic message, but it is just because the forcing is stronger (it is a 4xCO2 experiment). Therefore the reader is a bit in doubt whether he captures what the authors want to say.

Response: Correct, the physiological forcing simply scales in the same way as the total CO2 ERF. All reference to it being stronger has been removed – See line 460.

- page 15, line 354 : weren't -> were not.

Response: Done – See line 483

- page 16, line 366 : correlated to -> correlated with.

Response: This section has been removed now, because similar responses were not found following the LAI bugfix – See line 496

- page 16, line 378 : SAM : this abbreviation has not been defined.

Response: No longer referring to the SAM – See lines 509 onwards

- page 17, line 380-381 : "and a reduction of associated ... " : this sentence is slightly confusing, as through the reduction in high clouds, the outgoing LW radiation can be stronger again.

Response: This section has been removed, as indicated above – See lines 509 onwards.

- page 17, line 386 : near surface wind : maybe one can specify the altitude. Is it at 10 m?

Response: This figure has now been replaced with one from the piClim-HC experiment

- page 18, line 408 : are of the order -> are in the order.

Response: Corrected – See line 571

- page 18, line 417 : The major driver ... is greenhouse gases (GHGs) -> are [although I am not sure].

Response: Replaced with "major drivers ….. are" – See line 590

- page 18, line 417 : which is offset by aerosol : This is a slightly unlogical construction: I would think that forcings can be offset, but not GHGs.

Response: Now corrected this to specify that it is the forcing from GHGs that is offset – See line 590

- page 18, line 418 : key metrics -> key values.

Response: Changed – See line 591

- page 19, line 424 : "in e)" -> I would advance that slightly.

Response: Updated – See line 508 onwards

- page 19, line 433-437 : is the value 2.82 Wm-2 representing the 1850-2011 estimate? (does it already include the correction for going from 1750 to 1850?)

Response: It does. The sentence has been modified to make this clearer: "This latter estimate of 2.82 W m^{-2} has been adjusted …." – See line 164

- page 19, line 436 : e.g. CH4 : I assume this is also valid for CO2. Why not mentioning CO2?

Response: Added – See line 616

- page 19, line 440 : However, there is a discrepancy in ERF of 0.35 W m-2 ... which cannot ... -> However, the discrepancy of 0.35 Wm-2 ... cannot ...

Response: This discrepancy no longer exists, following the re-runs of the simulations with the LAI bugfix. See line 619 onwards

- page 20, line 443 : is this non-linearity similar to (or larger/smaller than) the one which one sees in the RF formulas of AR3 for N2O and CH4?

Response: No longer applicable, as indicated above.

- page 20, line 450 : The rapid adjustments (RA) ... includes -> The rapid adjustments (RAs) ... include.

Response: Corrected – See line 633

- page 20, line 453-454 : are there no other sources of uncertainty : the lifetime of aerosols? Their vertical profile?

Response: Now include other sources of uncertainty – See line 640

- page 20, line 454-455 : twice "sources" in this sentence.

Response: Altered - See line 644

- page 21, line480 : AEROCOM II -> AEROCOM Phase II.

Response: Now added – See line 672

- page 21, line 481 : The BC ERF was +0.32 W m-2 -> The BC ERF is +0.32 W/m2.

Response: Corrected – See line 672

- page 21, line 481 : and small negative offset -> and a small negative offset.

Response: Corrected – See line 673

- page 21, line 484 : in upper-level cloud -> in upper-level clouds.

Response: Changed – See line 676

- page 21, Fig 7b : should the orange bar represent -0.14 Wm-2? It looks larger.

Response: We checked and they are consistent

- page 23, line 525-526 : Are there two experiments with prescribed CDNC : piClimcontrol-fixedCDNC and piClim-aer-fixedCDNC? "By comparison with the main piClimaer" : shouldn't be added "and piClim-control"?

Response: These additional simulations were paired experiments. To add clarity, we altered the text as follows: "To further understand which processes contribute most to the aerosol ERF, a series of additional control and perturbation experiments were conducted …"

and

"By comparison with the main piClim-aer/piClim-control experiment pair, this indicated a …."

See lines 718-719, 723, and 730 of the marked-up manuscript

- page 23, line 532-533 : To complete the breakdown, ... : I assume this was on the main piClim-control and piClim-aer simulations, and not on the ones with fixed CDNC. This is maybe not so clear from the text.

Response: Text changed to: "To complete the breakdown, the method in Ghan (2013) was applied to the main piClim-aer/piClim-control experiment pair to derive …" – See line 730

- page 24, line 566 : "over the South Pacific" : looking at the figure, it is not so clear that the South Pacific stands out more than other regions.

Response: *Response: This has been amended to Eastern tropical pacific. (See line 767 of the marked-up document)*

- page 24, line 573 : is increased -> are increased.

Response: Corrected (Line 775 of marked-up document)

- page 25, line 596 : tropics -> subtropics.

Response: Corrected (Line 803 of marked-up document)

- page 25, line 602-603 : due to experimental setup -> due to the experimental setup.

Response: Part of that section has been removed to improve clarity (Lines 807-812 of marked-up document)

- page 25, line 609 : 15 % -> 15% (there is a blanco space between "15" and "%" in the text).

Response: This is intentional and follows the recommendations of the Bureau international des poids et mesures. From the SI Brochure, §5.3.7 states "When it is used, a space separates the number and the symbol %." We use this consistently throughout the manuscript.

- page 26, Eq. 9, and line 630 : CH should not be written in italic.

Response: Corrected although not showing as a tracked change – see Eqn. (9) on page 40 and see line 843 of the marked-up document

- page 26, line 630 : where ... is ... the concentrations -> concentration.

Response: Corrected - see line 841 of the marked-up document

- page 26, line 631 : there is apparently no blanco space after "piClim-control,".

Response: Corrected although not showing as a tracked change – see line 843 of the marked-up document

- page 28, line 669-670 : I suggest to write "hydroxyl radical" and "nitrate radical".

Response: Added - see line 887 of the marked-up document

- page 28, line 674: it doesn't -> it does not.

Response: Corrected - see line 893 of the marked-up document

- page 28, line 680-681 : maybe add at which altitude. It is mentioned in the caption of the figure, but it is maybe informative to mention it also in the text.

Response: Added - see line 898 of the marked-up document

- page 29, line 699 : Global distributions of the multi-annual distributions : twice "distributions".

Response: Corrected - see line 919 of the marked-up document

- page 29, line 707 : is changes to OH -> are changes to OH.

Response: Corrected - see line 929 of the marked-up document

- page 30, line 710-714 : how is the ari from NOx calculated? How in general are the ari/aci from NOx and VOC calculated?

Response: We use the Ghan (2013) approach to calculate the aerosol instantaneous radiative effect (IRE) in a simulation by taking the difference in radiative fluxes between two radiation calls, i.e., one call including aerosols (F) and the other without aerosols (Fclean), i.e., F - Fclean. The aerosol IRF is then calculated from the difference in the aerosol IRE between 2 simulations, where one simulation is piClim-control and the other is, for example, piClim-NOx, i.e., Δ(F - Fclean). We also use this approach to estimate the forcing due to changes in the cloud radiative effect (CRE) by calculating the CRE in each simulation using "clean" radiation calls and the difference between all-sky and clear-sky fluxes, i.e., Fclean – Fcs,clean. The cloud forcing is then calculated as the difference in the CRE between 2 simulations, i.e., Δ(Fclean – Fcs,clean). This breakdown is detailed in Eqns. (6)-(8) on pages 12-13 of the marked-up document.

- page 30, line 731 : as was the case with NOx : is meant here that the same mechanisms are active related to OH? As the change of OH is however opposite, the forcing is also opposite.

Response: Yes, the reviewer is correct. It is the same mechanism but the effect is in the opposite sense for COC/CO as for NOx. This has been re-written as follows:

*"*Via the same mechanisms as was the case with NOx*, …." - see line 954 of the marked-up document*

- page 30, line 735 : and CO2 -> and CO2 response.

Response: Corrected - see line 959 of the marked-up document

- page 30, line 735-739 : maybe one can mention explicitly that part of this message was already given earlier (on page 30, line 718-720).

Response: It wasn't clear that we could do this without affecting the flow of that paragraph so, we have opted not to change the text.

- page 31, line 753-754 : Is this true : are these the two main reasons (the fact that there are interactions, and the fact that there is non-linearity)?

Response: This has been changed as follows: "This is due to the uncertainty in the individual forcings (e.g., Bellouin et al., 2019) but the interaction between individual forcings, as well as the non-linear response of climate feedbacks due to aerosol-cloud interactions (Feichter et al., 2004; Deng et al., 2016; Collins et al., 2017; Shim et al., 2019*) may play a role." - see line 978 onwards of the marked-up document*

- page 31, line 764 : closely correlated : the correlation seems not that high (-0.44). In addition, from Figs. 12a and 12b it is not so easy to see that there is an anti-correlation. So I would not write "closely" correlated.

Response: Removed "closely" - see line 992 of the marked-up document

- page 31, line 767 : "may be" : can this not be said with more certainty? It seems like a sound explanation.

Response: Changed, as suggested - see line 997 of the marked-up document

- page 31, line 770 : with good correlation with -> correlating well with.

Response: This has been removed due to an error with the high-level cloud fraction diagnostic from the model simulations - see line 1000 onwards of the marked-up document

- page 31, line 771-772 : cloud -> clouds (twice).

Response: This whole sentence has now been removed, as indicated above - see line 998 onwards of the marked-up document

- page 32, line 778 : ", (h)" -> ", and (h)".

Response: The number of panels in Figure 12 has now changed but the caption does follow this recommendation, albeit no longer for panel (h) - see page 48 of the marked-up document

- page 33, line 782 : This study also attempts ... : it seems to be a bit a sudden introduction. Maybe one can first introduce the topic in general, and then say that it is also a focus of this study.

Response: This has now been addressed by including more of an introduction and referring to what has been done in other sections, e.g., GHGs, aerosols, and O3 precursor gases. See line 1018 onwards of the marked-up document

- page 33, line 782 : interaction between the aerosol and O3 forcings : is it really an interaction between the forcings which causes this?

Response: This has been corrected to say that the non-linear interactions between chemistry and aerosol may lead to non-linearities in the forcings - See lines 1021 onwards of the marked-up document

- page 33, line 784 : particularly in the net CS components : it is a bit strange to focus on the the net CS component, as the difference in SE CRE is even bigger.

Response: Agreed, this has been changed and the non-linearity in the SW CRE explicitly mentioned. See line 1031 onwards of the marked-up document

- page 33, line 784 : Firstly we calculate the aerosol IRFs. How are the IRFs calculated?

Response: We now refer to Eqn. (7) - See line 1025 of the marked-up document

- page 33, line 792 (not shown) and Fig. 13: it might be interesting to add a figure panel with the differences in O3 profiles.

Response: Given there is little impact on the LWcs' component of the ERF, we do not think that the differences in O3 are contributing to the non-linearity. As a result, we do not include a plot of differences in O3 profiles.

- page 34, line 805-807 : when reading this sentence, is seems that more attention is given to the +0.07 Wm-2 effect (i.e., the change from -0.39 to -0.32 Wm-2), than on the value of -0.32 Wm-2 itself.

Response: The apparent strong land use ERF in UKESM1 did indeed warrant more attention. However, this somewhat changed as a result of the new simulations following the discovery of the bug in the LAI seasonal cycle. Nevertheless, we now include some discussion of the different components to the newly quantified ERF of -0.17 ± 0.04 W m^{-2}. See lines 1057 onwards

- page 34, line 817 : and its cloud-free, aerosol-free, component -> and its cloud-free and aerosol-free component.

Response: Done - See line 1084 of the marked-up document

- page 35, line 822 : The seasonality of ... cause -> causes.

Response: Done - See line 1092 of the marked-up document

- page 35, line 830 : hist -> historical (the official name of this CMIP6 experiment).

Response: Done - See line 1102 of the marked-up document

- page 35, line 840 : As summarized above ... : maybe this is not a very good introduction - I don't think that "summarized" is the best way to refer to earlier text.

Response: Changed - See line 1113 of the marked-up document

- page 35, line 848-850 : What is the motivation for this sentence? Why is it mentioned that ozone is prescribed, but not, e.g., that oxidants like OH, NO3, and H2O2 are sometimes prescribed in models?

Response: Oxidants are now also mentioned - See line 1124 of the marked-up document

- page 36, line 857 : AF19 : this abbreviation is used only three times - I would think that it does not make so much sense to define it.

Response: Abbreviation now removed – See lines 1132 onwards of the marked-up document

- page 36, line 866 : what is GC3.1? Should probably be HadGEM-GC3.1.

Response: Now corrected – See line 1142 of the marked-up document

- page 36, line 866 : (well within the uncertainty range) : maybe one can add which uncertainty range is meant.

Response: Added – See line 1143 of the marked-up document

SECTION 5

- page 37, line 887 : paper -> study.

Response: Changed – See line 1166 of the marked-up document

- page 37, line 900 : may result : cannot it be expressed more firmly?

Response: No longer applicable with the updated simulations – See line 1182

- page 37, line 900 : coming from cloud top -> coming from cloud tops.

Response: No longer applicable with the updated simulations

- page 38, line 942 : reproduces -> reproduces well.

Response: Added - See line 1123 of the marked-up document

- page 38, line 949 : we consider -> we suggest.

Response: Changed - See line 1242 of the marked-up document

Assessment of pre-industrial to present-day anthropogenic climate forcing in UKESM1

Fiona M. O'Connor[1], N. Luke Abraham[2,3], Mohit Dalvi[1], Gerd Folberth[1], Paul Griffiths[2,3], Catherine Hardacre[1], Ben T. Johnson[1], Ron Kahana[1], James Keeble[2,3], Byeonghyeon Kim[4], Olaf Morgenstern[5], Jane P. Mulcahy[1], Mark ~~G.~~ Richardson[6], Eddy Robertson[1], Jeongbyn Seo[4], Sungbo Shim[4], Joao C. Teixeira[1], Steven Turnock[1], Jonny Williams[5], Andy Wiltshire[1], and Guang Zeng[5]

[1]Met Office, Exeter, United Kingdom
[2]National Centre for Atmospheric Science, University of Cambridge, United Kingdom
[3]Department of Chemistry, University of Cambridge, United Kingdom
[4]National Institute of Meteorological Sciences, Seogwipo-si, Jeju-do, Korea
[5]National Institute, for Water and Atmospheric Research, Wellington, New Zealand
[6]Centre for Environmental Modelling And Computation, University of Leeds, United Kingdom

Correspondence to: Fiona M. O'Connor (fiona.oconnor@metoffice.gov.uk)

Abstract. Quantifying forcings from anthropogenic perturbations to the Earth System (ES) is important for understanding changes in climate since the pre-industrial (PI) period. ~~In this paper~~ Here, we quantify and analyse a wide range of present-day (PD) anthropogenic ~~climate~~ effective radiative forcings (ERFs) with the UK's Earth System Model (ESM), UKESM1, following the protocols defined by the Radiative Forcing Model Intercomparison Project (RFMIP) and the Aerosol and Chemistry Model Intercomparison Project (AerChemMIP). In particular, by quantifying ~~effective radiative forcings (~~ERFs~~)~~ that include rapid adjustments (RAs) within a full ESM, it enables the role of various climate-chemistry-aerosol-cloud interactions to be investigated.~~feedbacks to be quantified~~.

Global mean ERFs for the PD (Year 2014) relative to the PI (Year 1850) period for carbon dioxide (CO_2), nitrous oxide (N_2O), ozone-depleting substances (ODSs), and methane (CH_4) are ~~1.83, 0.13, 0.33, and 0.93~~ 1.89 ± 0.04, 0.25 ± 0.04, -0.18 ± 0.04, and 0.97 ± 0.04 W m^{-2} ~~at the PD (Year 2014) relative to the pre-industrial (PI; Year 1850) for carbon dioxide, nitrous oxide, ozone-depleting substances, and methane~~, respectively. The ~~PD~~ total greenhouse gas (GHG) ERF is ~~2.89~~ 2.92 ± 0.04 W m^{-2}~~, larger than the sum of the individual GHG ERFs~~.

UKESM1 has an aerosol ~~forcing~~ ERF of ~~-1.13~~ -1.09 ± 0.04 W m^{-2}. A relatively strong negative forcing from aerosol-cloud interactions (aci) and a small negative instantaneous forcing from aerosol-radiation interactions (ari) from sulphate and organic carbon (OC) are partially offset by a substantial forcing from black carbon (BC) absorption. Internal mixing and chemical interactions ~~mean~~ imply that neither the forcing from ~~aerosol-radiation interactions~~ ari nor ~~aerosol-cloud interactions~~ aci are linear, making the ~~total~~ aerosol ERF less than the sum of the individual speciated aerosol ERFs.

~~Tropospheric ozone~~ Ozone (O_3) precursor~~s~~ gases consisting of volatile organic compounds (VOCs), carbon monoxide (CO), and nitrogen oxides (NOx) but excluding CH_4, ~~in addition to~~ exert~~ing~~ a positive radiative forcing due to increases in ~~ozone~~O_3. However, they also~~,~~ lead to oxidant changes, which in turn cause an indirect aerosol ERF. The net effect is that the ERF from PD-PI changes in NOx emissions is negligible at 0.03 ± 0.04 W m^{-2}~~, altering the sign of the net ERF from nitrogen oxide emissions~~, while the ERF from changes in VOC and CO emissions is 0.33 ± 0.04 W m^{-2}. Together, aerosol and ~~tropospheric ozone~~ O_3 precursors (called near-term climate forcers (NTCFs) in the context of AerChemMIP) exert an ~~global mean~~ ERF of ~~-1.12~~ -1.03 ± 0.04 W m^{-2}, mainly due to changes in the cloud radiative effect (CRE). There is also a negative ~~PD~~ ERF from land use change (~~-0.32~~ -0.17 ± 0.04 W m^{-2}). ~~It~~ When adjusted from Year 1850 to 1700, it is outside the range of previous estimates, and is most likely due to too strong an albedo response. In combination, the net anthropogenic ERF ~~is potentially biased low (1.61 W m^{-2}) relative to~~ (1.76 ± 0.04 W m^{-2}) is consistent with other estimates~~, due to the inclusion of non-linear feedbacks and ES interactions~~.

By including ~~feedbacks~~ interactions between ~~greenhouse gases~~GHGs, stratospheric and tropospheric ~~ozone~~O_3, aerosols, and clouds, ~~some of which act non-linearly,~~ this work demonstrates the importance of ES interactions when quantifying ERFs~~climate forcing~~. It also suggests that ~~rapid adjustments~~RAs need to include chemical as well as physical adjustments to fully account for complex ES interactions.

1 Introduction

In order to have a quantitative understanding of past and future climate change, and attribute ~~it~~ climate change and its impacts to different anthropogenic and natural drivers, it is ~~necessary~~ important to have a ~~detailed~~ process-based understanding of ~~all~~ critical aspects of the pathway from anthropogenic (or natural) activity through to climate response and its impacts. A recent international effort for understanding climate change ~~key mechanism for addressing this overarching objective~~ is the 6th Coupled Model Intercomparison Project (CMIP6; Eyring et al., 2016), which designs experiments and distributes data from multi-model simulations. These simulations, with state-of-the-art climate models or Earth System Models (ESMs), are aimed at addressing ~~these important~~ climate science questions directly or via dedicated CMIP6-endorsed model intercomparison projects (MIPs) such as the Aerosol and Chemistry Model Intercomparison Project (AerChemMIP; Collins et al., 2017). ~~The first~~ An important part of this cause-effect chain from activity to climate response, mediated through the atmosphere and the land surface, is quantifying changes to the Earth's radiation budget, often termed radiative forcing (RF). RF is a direct measure of the response in the Earth's radiation budget by changes in anthropogenic (or natural) activities.

Successive assessment reports of the Intergovernmental Panel on Climate Change (IPCC) have used the concept of ~~radiative forcing (~~RF~~)~~ as a metric to quantify the effects of different anthropogenic and natural drivers on the Earth's radiation balance.

For this purpose, ~~radiative forcing (~~RF~~)~~, or more precisely, the stratospherically-adjusted ~~radiative forcing~~RF (SARF) is defined at the tropopause (Myhre et al., 2013a) as:

$$RF = IRF + A_{strattemp}\text{———}, \quad (1)$$

where IRF is the instantaneous radiative forcing and $A_{strattemp}$ is the additional change in the net downward radiative fluxes at the tropopause solely due to stratospheric temperature adjustment (Hansen et al., 1997), while holding all other variables fixed. Including the stratospheric temperature adjustment can significantly affect the magnitude of a forcing (e.g., Smith et al., 2018) and even change the sign of the forcing in the case of ozone (O_3) depletion (Shine et al., 1995). As a result, RF rather than IRF is a better predictor of the drivers of global mean temperature response. RF is a concept~~part of a framework~~, based around energy budget analyses, that has split perturbations to the Earth's radiative balance~~forcing~~ from climate response (Boucher et al., 2013; Myhre et al., 2013a; Sherwood et al., 2015). ~~It's~~ It has been used extensively to evaluate and compare the strength of ~~the~~ various forcings~~mechanisms~~, both anthropogenic and natural, affecting the Earth's radiation balance and hence, their contribution to climate change (e.g., Hansen et al., 1997; Shine and Forster, 1999). However, despite the extensive use of RF as a metric for climate change, it is often calculated inconsistently between the different drivers of climate change (e.g., Myhre et al., 2013a). ~~and p~~Participating models in the 5th Coupled Model Intercomparison Project (CMIP5; Taylor et al., 2012) can show large differences in CO_2 forcing (Andrews et al., 2012a; Forster et al., 2013) due to model diversity and/or calculation method. For example, the RF attributed to long-lived greenhouse gases (LLGHGs) is typically based on changes in observed concentrations between the pre-industrial (PI) and the present day (PD) periods and ~~using~~ uses line-by-line radiative transfer calculations and/or simple, yet justified, expressions for RF based on, e.g., Myhre et al. (1998) and Ramaswamy et al. (2001). These expressions have been updated for some LLGHGs recently (Etminan et al., 2016). However, the observed concentrations themselves may be subject to biogeochemical feedbacks (e.g., Arneth et al., 2010; O'Connor et al., 2010) ~~and observational-based estimates of forcing are also relatively uncertain (Skeie et al., 2011)~~.

In contrast to the quantification of the LLGHG RF, the RF from PD-PI changes in tropospheric ~~ozone (~~O_3~~)~~ ~~changes at the PD relative to the PI~~ (Stevenson et al., 2013) is based solely on models. It has been calculated using the ensemble of models (Young et al., 2013) participating in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP; Lamarque et al., 2013) providing input to offline radiative transfer models (e.g., Edwards and Slingo, 1996). The simulations used the corresponding sea surface temperatures (SSTs) and sea ice (SI) conditions for the time periods of interest (PI and PD), ~~including~~ therefore allowing some climate response and feedbacks at the PD, ~~meaning~~ implying that the resulting ~~RF~~ estimate is not consistent with the RF definition; ~~doesn't~~ it does not fit into the simple forcing-feedback ~~framework~~concept, whereby feedbacks are related to global mean temperature change and forcings are not (Sherwood et al., 2015). There are also additional uncertainties associated with the estimate of tropospheric O_3 RF due to the lack of ~~a~~ robust and reliable observational constraints for PI O_3 concentrations (e.g., Stevenson et al., 2013), the diversity in modelled PD tropospheric O_3 burden across

~~multi-model ensembles~~ models (e.g., Young et al., 2013; Young et al., 2018), uncertainties in historical emissions of ~~tropospheric~~ O_3 precursors, and the apparent inability of current state-of-the-art chemistry-climate models to replicate near-recent observed trends in tropospheric O_3 (Parrish et al., 2014; Young et al., 2018) although recently, isotopic measurements seem to corroborate the modelled trends (Yeung et al., 2019). ~~Additional~~ Other uncertainties in Stevenson et al. (2013) arise from neglecting the change in O_3 in the lower stratosphere attributable to changes in ~~tropospheric~~ O_3 precursors and the contribution from stratospheric O_3 depletion on the modelled changes in tropospheric O_3 (e.g., Søvde et al., 2011; 2012). Despite these ~~uncertainties~~difficulties in estimating tropospheric O_3 RF, the even larger uncertainty in aerosol forcing (Myhre et al., 2013a; Bellouin et al., 2020~~2019~~) accounts for the majority of the uncertainty in the total anthropogenic forcing.

Aerosol forcing involves a wide range of physical processes. These include (i) direct changes to the radiation budget through scattering and absorption of both shortwave (SW) and longwave (LW) radiation (e.g., Haywood and Boucher, 2000), (ii) indirect impacts on the radiation budget by changing the microphysical properties of clouds (Twomey et al., 1977), and (iii) changes in the distribution of cloud cover or condensate that follow on from perturbations in cloud microphysics (Albrecht 1989) or radiative heating by aerosols (Hansen et al., 1997). Direct aerosol RF can be calculated using offline radiative transfer models in a similar manner to greenhouse gas (GHG) and O_3 forcing, whereas assessing impacts of aerosols on clouds requires simulations in atmospheric models. The 5th assessment report (AR5) of the IPCC recommended the effective radiative forcing (ERF) framework as a suitable metric for assessing the overall aerosol forcing as it enables the more complex cloud impacts to be evaluated as part of the climate's rapid adjustments (RAs) (Myhre et al., 2013a). To simplify terminology, AR5 also made a clear distinction between components of the forcing driven by aerosol-radiation interactions (ari; i.e., the direct or IRF) and aerosol-cloud interactions (aci) (that include all indirect or semi-direct cloud-related forcings). Despite wide-ranging and on-going research, the role of aerosols remains the leading source of uncertainty in estimates of climate forcing, due to the difficulty in constraining the sensitivity of clouds to changing microphysical processes (Bellouin et al., 2020).

In the case of land use, RF estimates have been made using single general circulation model (GCM) simulations with a double-call to the radiation scheme (e.g., Betts et al., 2007) or by comparing paired simulations that include RAs (e.g., Andrews et al., 2016). However, the choice of RF calculation is not the major source of differences in RF estimates. Similar to O_3, uncertainty in PI land cover is a major source of uncertainty in land use RF (e.g., de Noblet-Ducoudré et al., 2012). Historically, deforestation has been the dominant type of land use change, and this causes a positive RF due to increased carbon dioxide (CO_2) emissions and a negative RF due to increased surface albedo. Here, we include the effects of land use CO_2 emissions in the CO_2 ERF estimates and the land use ERF is due to biophysical changes, predominately albedo. Deforestation has a much larger effect on albedo in snowy regions and model biases in snow cover also contribute to uncertainty in land use RF (Pitman et al., 2011). Land use RF estimates also vary due to different time periods being considered (Myhre et al., 2013a), because, unlike many other forcing agents, there was substantial land use change before the industrial revolution.

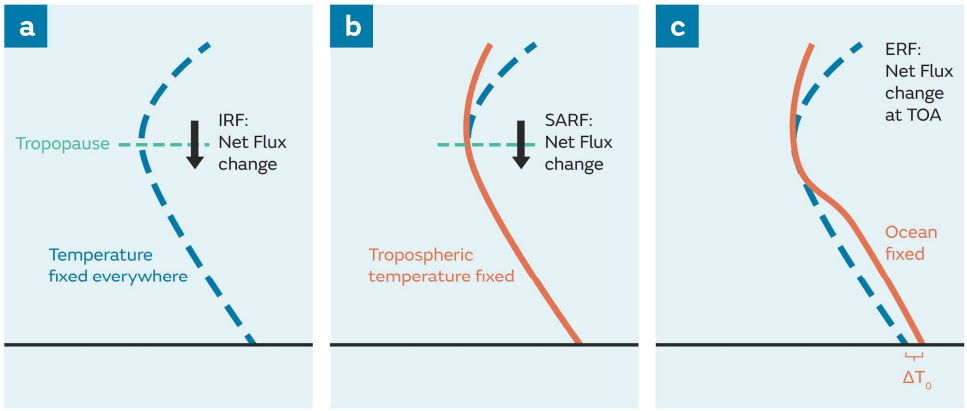


Figure 1: ~~Schematic~~ Vertical profiles of temperature, showing the different definitions of radiative forcing: a) instantaneous radiative forcing (IRF), b) stratospherically-adjusted radiative forcing (SARF), and c) effective radiative forcing (ERF). IRF and SARF are ~~typically~~ defined at the tropopause whereas ERF is defined at the top of atmosphere (TOA). Adapted from Fig. 8.1 of the IPCC AR5~~5th assessment report~~ (Myhre et al., 2013a), which in turn was updated from Hansen et al. (2005).

~~Although~~ As~~RF has been calculated inconsistently between the different drivers of climate change to date~~ indicated previously, the IPCC AR5 ~~fifth assessment report (AR5) of the IPCC also adopted~~recommended an extension to the definition of ~~radiative forcing~~RF to include ~~fast feedbacks or rapid adjustments~~RAs other than the stratospheric temperature adjustment. This updated definition, the ~~effective radiative forcing (~~ERF~~), is now the preferred metric of choice for ranking the drivers of climate change. ERF~~ is defined at the top of the atmosphere (TOA), following Chung and Soden (2015), as:

$$\mathrm{ERF} = \mathrm{IRF} + \sum_i A_i \quad \text{——}, \tag{2}$$

where IRF is ~~the instantaneous radiative forcing~~ now defined at the TOA ~~following a perturbation~~ and A_i is a ~~rapid adjustment~~ RA in the atmosphere or over the land surface that alters the net downward radiative flux at the TOA either positively or negatively. These ~~adjustments~~RAs include changes in stratospheric temperatures (as included in the definition of RF or SARF above; Eqn. (1)) as well as adjustments such as tropospheric temperatures, water vapour, clouds, land surface temperature, and land surface albedo, as examples, but ~~global mean surface temperatures or~~ global ocean conditions remain unchanged. When comparing the ERF and IRF for black carbon (BC), as an example, RAs lead to the ERF being half that of its IRF (Stjern et al., 2017; Smith et al., 2018). On the other hand, Xia et al. (2016) found in their model that cloud and sea ice adjustments driven by stratospheric O_3 recovery included in the ERF definition lead to the ERF being different in both sign and magnitude from the SARF. And when comparing different forcing metrics, Hansen et al. (2005) found that the efficacy of a climate forcing (i.e., the change in global mean temperature relative to that from CO_2 for an equivalent forcing at the TOA or the tropopause) is less sensitive to the forcing agent and closer to 1 for ERFs than for IRFs or SARFs. Although the ~~error~~

uncertainty associated with an ERF tends to be larger than that of RF ~~(or SARF)~~, ~~climate sensitivity parameters (i.e. the degree of warming per unit forcing) are less dependent on the forcing agent and~~ it is more representative of the climate response than the traditional RF (Hansen et al., 2005; Forster et al., 2016). As a result, it is now the preferred metric of choice for ranking the drivers of climate change (Boucher et al., 2013; Forster et al., 2016). ~~For example, rapid adjustments lead to the ERF for black carbon (BC) being half that of its IRF (Stjern et al., 2017; Smith et al., 2018). On the other hand, Xia et al. (2016) found that cloud and sea ice feedbacks driven by stratospheric O_3 recovery included in the ERF definition lead to a significant adjustment that changes the sign and magnitude of the stratospheric O_3 forcing from the SARF. Although the definitions of~~

The SARF and ERF differ in where the change in radiative fluxes is diagnosed. The SARF, diagnosed at the tropopause, requires the tropopause to be defined but the ERF has the advantage of being diagnosed at the TOA, with no need for a tropopause. While the SARF at the tropopause and TOA are, by definition, identical, this is not the case for the ERF; climate forcings can lead to adjustments in stratospheric circulation and therefore changes to dynamical heating above the tropopause. Changes in dynamical heating are thus balanced by radiative divergence across the stratosphere, explaining the change in net fluxes between the tropopause and TOA.~~, ERFs at the tropopause are nearly identical to those at the TOA for tropospheric forcing agents; this is also the case for SARFs. An overview of the historical evolution of the RF concept, its quantification for different forcing agents, and applications of RF can be found in Ramaswamy et al. (2019).~~

The use of ERF as a metric also offers the advantage that it can be readily calculated using a pair of parallel simulations with standard model TOA radiative flux diagnostics (Forster et al., 2016) albeit with a requirement to run for relatively long periods (~30 years) to reduce the uncertainty associated with meteorological variability (e.g., Shindell et al., 2013a). ~~Despite this requirement, it has now become the metric of choice over RF (Boucher et al., 2013) to rank the drivers of climate change.~~ Figure 1 illustrates the various definitions of ~~radiative forcing~~RF: ~~, including those of~~ IRF, SARF ~~(or RF)~~, and ERF. An overview of the historical evolution of the RF concept, its quantification for different forcing agents, and applications of RF can be found in Ramaswamy et al. (2019).

The aim of the current study is to quantify ~~the PI (Year 1850) to~~ PD (Year 2014) ~~effective radiative forcings (~~ERF~~s)~~ from anthropogenic drivers of climate change with ~~a fully coupled~~ an atmosphere-only configuration of an ~~Earth System Model (~~ESM~~)~~. Using the experimental protocol recommended for the Radiative Forcing Model Intercomparison Project (RFMIP; Pincus et al., 2016), PD ~~forcings~~ ERFs will be quantified relative to the PI period from PD-PI changes in emissions, concentrations, and/or land use due to anthropogenic activities. This approach offers a consistent methodology for diagnosing the ERFs of all forcing agents or Earth System (ES) perturbations and, if applied consistently across all models as part of ~~the 6th Coupled Model Intercomparison Project (~~CMIP6~~;~~ (Eyring et al., 2016), should help to address some of the deficiencies and uncertainties associated with previous estimates of forcing (e.g., Myhre et al., 2013a) and improve our understanding of how the ES responds to forcing (Pincus et al., 2016). The paper is organised as follows. Section 2 provides a brief description

of the UK's ~~Earth System Model~~ESM, UKESM1, used in this study. Section 3 outlines the experimental setup and the simulations carried out. ~~PD~~ Aanthropogenic ~~forcings~~ ERFs ~~relative to PI~~ are presented in Sect. 4, while conclusions can be found in Sect. 5.

2 Model Description

The model used in this study ~~is~~ consists of the atmospheric and land components of the UK's ~~Earth System Model~~ESM, UKESM1 (Sellar et al., 2019). UKESM1 is based on the Global Atmosphere 7.1/Global Land 7.0 (GA7.1/GL7.0; Walters et al., 2019) configuration of the Hadley Centre Global Environment Model version 3 (HadGEM3; Hewitt et al., 2011), herein referred to as HadGEM3-GA7.1, to which terrestrial carbon/nitrogen cycles (Sellar et al., 2019) and interactive stratosphere-troposphere chemistry (Archibald et al., ~~2019~~2020) from the UK Chemistry and Aerosol (UKCA; Morgenstern et al., 2009; O'Connor et al., 2014) model have been coupled. The model resolution is N96L85; this is equivalent to a horizontal resolution of roughly 135 km and the 85 terrain-following model levels cover an altitude range up to 85 km above sea level. The physical atmosphere model, HadGEM3-GA7.1, already includes the UKCA prognostic aerosol scheme called GLOMAP-mode (Mann et al., 2010; Mulcahy et al., 2018), in which secondary aerosol formation ~~is determined by~~ makes use of prescribed oxidant fields. Here, the UKCA chemistry and aerosol schemes are coupled, with oxidants from the stratosphere-troposphere chemistry scheme (Archibald et al., ~~2019~~2020) ~~determining~~ influencing secondary aerosol formation rates. A full description and evaluation of the UKCA chemistry and aerosol schemes in UKESM1 can be found in Archibald et al. (~~2019~~2020) and Mulcahy et al. (~~2020~~2019), respectively. Mulcahy et al. (2018) also implemented a number of aerosol process improvements in HadGEM3-GA7.1 which helped to reduce the aerosol ERF at the present day (Year 2000) from -2.75 ± 0.06 W m^{-2} in ~~address the strong negative PD aerosol forcings from~~ its predecessor model (i.e., HadGEM3-GA7.0; Walters et al., 2019) to -1.45 ± 0.04 W m^{-2}; the aerosol ERF in HadGEM3-GA7.0 previously led to an unrealistic negative total anthropogenic ERF of -0.60 W m^{-2}.

Differences between HadGEM3-GA7.1 and the atmosphere-only configuration of UKESM1 can be found in Table A1 in the Appendix.

3 Model Setup and Experiments

To calculate the ~~pre-industrial (PI; Year 1850) to present-day (~~PD~~;~~ (Year 2014) ~~effective radiative forcings (~~ERFs~~)~~ relative to the PI (Year 1850) period due to a ~~PI-to-PD~~ PD-PI perturbation (e.g., change in emissions), timeslice experiments with fixed ~~sea surface temperatures (~~SSTs~~)~~ and ~~sea ice (~~SI~~)~~ were carried out, following the protocol defined by ~~the Radiative Forcing Model Intercomparison Project (~~RFMIP~~;~~ (Pincus et al., 2016). The experimental setup was also consistent with

recommendations from Forster et al. (2016) and the protocol for timeslice ERF experiments defined in ~~the Aerosol and Chemistry Model Intercomparison Project (~~AerChemMIP~~;~~ (Collins et al., 2017).

225

Experiment ID	**MIP**	**CO_2 concn.**	**N_2O concn.**	**ODS ~~HCs~~ concns.**	**Other HC concns.**	**CH_4 concn.**	**O_3 Precursor (VOC, CO, NOx) emissions**	**Aerosol and aerosol precursor (BC, OC, SO_2) emissions**	**Land use**
piClim-control	AerChemMIP ; RFMIP	1850	1850	1850	1850	1850	1850	1850	1850
piClim-CO2	N/A	2014	1850	1850	1850	1850	1850	1850	1850
piClim-CO2phys[a]	N/A	2014 (land surface) 1850 (radiation)	1850	1850	1850	1850	1850	1850	1850
piClim-4xCO2	RFMIP	4x1850	1850	1850	1850	1850	1850	1850	1850
piClim-4xCO2phys[a]	N/A	4x1850 (land surface) 1850 (radiation)	1850	1850	1850	1850	1850	1850	1850
piClim-N2O	AerChemMIP	1850	2014	1850	1850	1850	1850	1850	1850
piClim-HC	AerChemMIP	1850	1850	2014	1850	1850	1850	1850	1850
piClim-CH4	AerChemMIP	1850	1850	1850	1850	2014	1850	1850	1850
piClim-GHG	RFMIP	2014	2014	2014	1850	2014	1850	1850	1850
piClim-SO2	AerChemMIP	1850	1850	1850	1850	1850	1850	2014 (SO2); 1850 (~~non-SO2~~BC, OC)	1850
piClim-BC	AerChemMIP	1850	1850	1850	1850	1850	1850	2014 (BC); 1850 (~~non-BC~~OC, SO_2)	1850
piClim-OC	AerChemMIP	1850	1850	1850	1850	1850	1850	2014 (OC);	1850

								1850 (~~non-OC~~BC, SO_2)	
piClim-aer	AerChemMIP	1850	1850	1850	1850	1850	1850	2014	1850
piClim-NOx	AerChemMIP	1850	1850	1850	1850	1850	2014 (NOx); 1850 (VOC, CO)	1850	1850
piClim-VOC	AerChemMIP	1850	1850	1850	1850	1850	2014 (VOC, CO); 1850 (NOx)	1850	1850
piClim-O3	AerChemMIP	1850	1850	1850	1850	1850	2014 (VOC, CO, NOx)	1850	1850
piClim-NTCF[b]	AerChemMIP	1850	1850	1850	1850	1850	2014	2014	1850
piClim-LU	RFMIP	1850	1850	1850	1850	1850	1850	1850	2014
piClim-Anthro	RFMIP	2014	2014	2014	2014	2014	2014	2014	2014

Table 1: List of all the ~~fixed sea surface temperature (fSST) effective radiative forcing (ERF)~~ atmosphere-only experiments carried out with the UK's Earth System Model, UKESM1, to quantify ~~diagnose the pre-industrial to present-day~~ PD ERFs from PD-PI changes in emissions, concentrations and/or land use. Each simulation was 45 years in length, with analysis based on the last 30 years. [a]In the *piClim-CO2phys* and *piClim-4xCO2phys* simulations, only the land surface sees the ~~PI-to-PD~~ perturbation in CO_2, the radiation scheme still sees the PI concentration. [b]The AerChemMIP experiment *piClim-NTCF* is also known as *piClim-aerO3* in RFMIP.

Effectively, this involves running a PI timeslice experiment, called *piClim-control* here, in which SSTs, ~~and~~ SI and all ~~forcings~~ other boundary conditions were fixed at year-1850 levels. The SSTs and SI used in *piClim-control* were monthly mean ~~time-varying~~ climatologies derived from 30 years (i.e., Years 2156-2185 inclusive) of output from the UKESM1 ~~pre-industrial~~PI coupled control experiment (*piControl*) characterised in Sellar et al. (2019) and one of an underpinning set of coupled experiments for CMIP6 (Eyring et al., 2016). It also used 30-year monthly mean climatologies for the vegetation distribution, canopy height, leaf area index, surface seawater dimethyl sulphide (DMS) and chlorophyll concentrations derived from the same period of *piControl*. Fixing the vegetation distribution was not ~~included in~~ part of the RFMIP protocol and any potential vegetation ~~rapid adjustments~~RAs will be somewhat constrained. This is due to the simulations being based on the configuration of UKESM1 used for the Atmosphere Model Intercomparison Project (AMIP) simulation (which prescribed vegetation characteristics). Although the AerChemMIP protocol (Collins et al., 2017) requested use of the maximum capability possible,

interactive vegetation was not a model requirement. The extra RFMIP experiments carried out here were only done as a late addition and the same experimental setup was kept for internal consistency.

The model was initialised using output from the start of the 30-year period used to produce the PI climatologies (i.e., January 2156 of *piControl*). All the other experiments are perturbation experiments, parallel to *piClim-control*, in which selected emissions, concentrations and/or land use were changed from year-1850 to year-2014 values. Although AerChemMIP and RFMIP recommend 30 years for fixed SST (fSST) timeslice ERF experiments, the perturbations to the LLGHGs took up to 15 years to propagate fully into the stratosphere due to the turnover timescale associated with the Brewer-Dobson circulation (e.g., Butchart, 2014). Therefore, all simulations were 45 years in length. Using the latter 30 years of the paired simulations, the ERFs ~~from the PI-to-PD perturbations~~ were diagnosed as the time-mean global-mean PD-PI difference in the TOA net radiative fluxes. Running for 30 years when the model has reached steady state reduces the uncertainty associated with meteorological variability (e.g., Shindell et al., 2013a) and improves the estimate of the ERF. Details on how the ERF was further decomposed can be found in Sect. ~~3.1~~4.1.

In all cases, the ~~emissions and/or~~ GHG concentrations for 1850 and/or 2014 were taken from ~~Hoesly et al. (2018), van Marle et al. (2017), and~~ Meinhausen et al. (2017). However, the recommended concentrations for the different GHGs in UKESM1 are implemented differently. In the case of CO_2, the prescribed concentration is uniform in mass mixing ratio throughout the model domain. For methane (CH_4) and nitrous oxide (N_2O), the recommended concentrations are treated as lower boundary conditions (LBCs); their 3D distributions are modelled interactively by the UKCA chemistry scheme (Archibald et al., 2020) and coupled to radiation. For O_3 depleting substances (ODSs) in *piClim-HC*, their concentrations are prescribed separately (and consistently) in UKCA and in the radiation scheme. For the UKCA chemistry scheme, LBCs are prescribed for trichlorofluoromethane (CFC11), dichlorodifluoromethane (CFC12), and methyl bromide (CH_3Br), all of which include contributions from other chlorine- and bromine-containing source gases not explicitly treated in UKCA. This approach ensures that the correct stratospheric chlorine and bromine loadings are used for the PD period. Further details on the species included in the CFC11, CFC12, and CH_3Br LBCs can be found in Archibald et al. (2020). For the radiation scheme, the radiative effects of ODSs are handled by prescribing the mass mixing ratio of a lumped species (CFC12-eq) uniformly throughout the atmosphere, consistent with the UKCA LBCs, Finally, the *piClim-GHG* experiment collectively perturbs all the GHGs from PI to PD levels, including non O_3-depleting halocarbons (HCs). For these latter GHGs, a uniform mass mixing ratio of a lumped species (HFC134a-eq), provided by Meinhausen et al. (2017), is prescribed in the radiation scheme.

In all the experiments in Table 1, the emissions of primary aerosol and aerosol precursors (BC, Organic carbon (OC), and sulphur dioxide (SO_2)) and O_3 precursors (volatile organic compounds (VOCs), carbon monoxide (CO), and nitrogen oxides (NOx)) excluding CH_4 for 1850 and/or 2014 were taken from Hoesly et al. (2018) and van Marle et al. (2017). In the case of the NOx emissions perturbation experiment (*piClim-NOx*), both aircraft and surface anthropogenic emissions were changed to

PD levels. In *piClim-VOC*, both VOC *and* CO anthropogenic emissions were changed to PD levels while the experiment *piClim-O3* perturbs emissions of VOC, CO, and NOx *only*, with the CH_4 concentration remaining at PI levels. Finally, although near-term climate forcers (NTCFs) include CH_4 and short-lived halocarbons (e.g., Myhre et al., 2013a), in the context of the AerChemMIP protocol (Collins et al., 2017), the experiment *piClim-NTCF* does not perturb concentrations of CH_4 or other short-lived GHGs. It *only* changes anthropogenic emissions of aerosol and aerosol precursors (BC, OC, SO_2) and O_3 precursors (VOC, CO, NOx) to PD levels; it is also referred to as *piClim-aerO3* in the RFMIP protocol (Table 1).

For prescribing the anthropogenic land use change at 2014, the difference in vegetation between 1850 and 2014 was taken from a UKESM1 coupled historical simulation, in which the only transient forcing was anthropogenic land use change. Natural volcanic and solar forcings were fixed in all simulations at 1850 levels (Arfeuille et al., 2014; Thomason et al., 2018; Matthes et al., 2017) using those specified for CMIP6 (Eyring et al., 2016). Table 1 gives a full list of the ~~fixed SST~~ fSST ERF experiments carried out with UKESM1 for this study. The only experiment omitted from the ~~fixed SST~~ fSST ERF experiments specified in the RFMIP and AerChemMIP protocols is *piClim-NH3*; this is because UKESM1's aerosol scheme, GLOMAP-mode (Mann et al., 2010; Mulcahy et al., 2018; Mulcahy et al., 2020~~2019~~), does not include any treatment for ~~ammonium~~ nitrate aerosol.

Through a partnership between the Met Office Hadley Centre (MOHC; https://www.metoffice.gov.uk/climate-guide/science/science-behind-climate-change/hadley), the UK's National Centre for Atmospheric Science (NCAS; https://www.ncas.ac.uk/en/), New Zealand's National Institute for Water and Atmospheric Research (NIWA; https://www.niwa.co.nz/), and National Institute of Meteorological Science/Korean Meteorological Administration (NIMS-KMA; http://nims.go.kr/MA/main.jsp), the ~~fixed SST~~ fSST ~~ERF~~ experiments carried out with UKESM1 were spread across multiple high performance computing (HPC) platforms. Due to the non-linearity of the equations being solved, ESMs are sensitive to the propagation of small perturbations, resulting in a lack of bit reproducibility. ~~This makes the simulations sensitive to changes in platform that cannot guarantee bit reproducible results. Considering that the RFMIP/AerChemMIP experiments were spread across 3 different HPC platforms~~As a result, we aimed to verify that the differences in model output ~~from the atmosphere-only configuration of UKESM1 was scientifically consistent with~~ were not statistically significant from each other.

To test this, we created an ensemble of short runs on each machine by perturbing selected variables in their initial conditions, using a perturbation with a numerical value comparable to the ~~machines~~ machine's precision. The spread of results (at each point in time and space) on each platform was then used to determine whether they could each have been sampled from the same ensemble of results generated on either machine. A permutation method was used to ensure statistical independence between neighbouring points according to the work described by Wilks (1997). A paper, describing this protocol in more detail, is in preparation (Teixeira et al., 2019~~2020~~). ~~Further~~ In addition to this, a number of perturbation experiments were

carried out in duplicate to test the sensitivity of the ~~global mean~~ ERFs to differences in HPC platform. These duplicate experiments are listed in Table 2 and will be available through the Earth System Grid Federation (ESGF; https://esgf.llnl.gov/) archive as different realisations of the same experiment.

HPC	Compiler	Experiment ID	Realisation ID
Met Office CrayXC40	Cray compiling environment 8.3.4	*piClim-control*	R1
		piClim-SO2	R2
		piClim-OC	R2
		piClim-NTCF	R2
NIWA XC50	Intel Compilers 17.0.4 20170411	*piClim-SO2*	R1
		piClim-OC	R1
KMA Cray XC40	Cray compiling environment 8.3.7	*piClim-NTCF*	R1

Table 2: List of atmosphere-only PI control (*piClim-control*) and duplicate ~~fixed sea surface temperature (fSST)~~ perturbation experiments (*piClim-X*) carried out with ~~the UK's Earth System Model,~~ UKESM1, on different ~~high performance computing (~~HPC~~)~~ platforms.

4 ~~Present Day~~ Anthropogenic Effective Radiative Forcings (ERFs)

The ERF has been calculated from the difference (Δ) in the net TOA radiative flux (F) between the perturbed simulation (e.g., *piClim-CH4;* Table 1) and the *piClim-control* simulation as follows:

$$ERF = \Delta F \quad (3)$$

where ΔF is in response to whole-atmosphere PD-PI changes in composition and/or other RAs; no masking of the response is applied. For example, the ERF quantified from *piClim-HC* minus *piClim-control* includes the direct radiative effect from the increase in ODS concentrations, the indirect radiative effect of the whole-atmosphere O_3 response as well as other changes to the TOA radiative fluxes due to whole-atmosphere and land surface RAs. ~~This~~ The ERF can be decomposed into the clear-sky

(CS) ERF ($ERFcs$) and the change in the ~~Cloud Radiative Effect~~ cloud radiative effect (ΔCRE) using the diagnosed CS radiative flux ($Fcs\sout{lear}$):

$$ERF = \Delta Fcs\sout{lear} + \Delta(F - Fcs\sout{lear}), \quad (4)$$

$$= ERFcs + \Delta CRE. \quad (5)$$

However, many of the experiments in this study either directly perturb aerosol emissions and/or alter aerosol concentrations via chemical and dynamical ~~feedbacks~~interactions. Changes in aerosol can bias the diagnosed CRE as aerosol scattering and absorption typically reduce the contrast in ~~shortwave (~~SW~~)~~ reflection between cloudy and ~~clear-sky~~ CS scenes; a process termed "cloud masking" (e.g., Zelinka et al., 2014). In consideration of this, we have calculated the change in the CRE from "clean" radiation calls that exclude ~~aerosol-radiation interactions (~~ari~~)~~, as recommended in Ghan (2013):

$$ERF = \Delta(F - Fclean) + \Delta Fcs\sout{lear}, clean + \Delta(Fclean - Fcs\sout{lear}, clean), \quad (6)$$

$$= Aerosol\ IRF + ERFcs, clean + \Delta CRE', \quad (7)$$

$$= ERFcs' + \Delta CRE'. \quad (8)$$

The ERF is thus separated into a component due to cloud property changes ($\Delta CRE'$) and the non-cloud forcing (*ERFcs'*). Here, *ERFcs'* is the sum of the aerosol IRF and any non-aerosol changes in CS fluxes and differs slightly from *ERFcs* in ~~equation~~Eqn. (5), in that it can include the impact of aerosol scattering and absorption in the clear-air above or below clouds. One acknowledged limitation is that variations in gaseous absorption and emission between clear and cloudy scenes also lead to cloud masking effects (e.g., Soden et al., 2008). Although Ghan's method removes the very prominent influence of aerosols, cloud-masking from ~~ozone (~~O_3~~)~~ and GHGs may still affect the separation of ERF into the CS and CRE components.

4.1 Overview of ~~Present-Day Effective Radiative Forcings (~~ERFs~~)~~

The ~~effective radiative forcing (~~ERF~~)~~, and its ~~clear-sky~~CS (*ERFcs'* ~~CS~~), and ~~cloud radiative effect ($\Delta CRE'$ CRE)~~ contributions, following ~~equations (6) to~~Eqn. (8) are listed in Table 3 for all perturbation experiments relative to *piClim-control*, and are further decomposed into the SW (solar), LW (terrestrial) and net (SW + LW) components. Table 3 also includes estimates of the modelled-derived uncertainty in the different components by calculating the standard error based on Forster et al. (2016). Although the errors quoted are small (i.e., less than 0.04 W m^{-2}), they do not represent the true uncertainty in the quantified ERFs; uncertainties due to emissions (e.g., Hoesly et al., 2018), model biases (e.g., Archibald et al., 2020), incorrect sensitivity to changing emissions (e.g., Archibald et al., 2010; Wild et al., 2020), radiative transfer schemes (e.g., Pincus et al., 2020), and/or missing or unresolved processes are not considered.

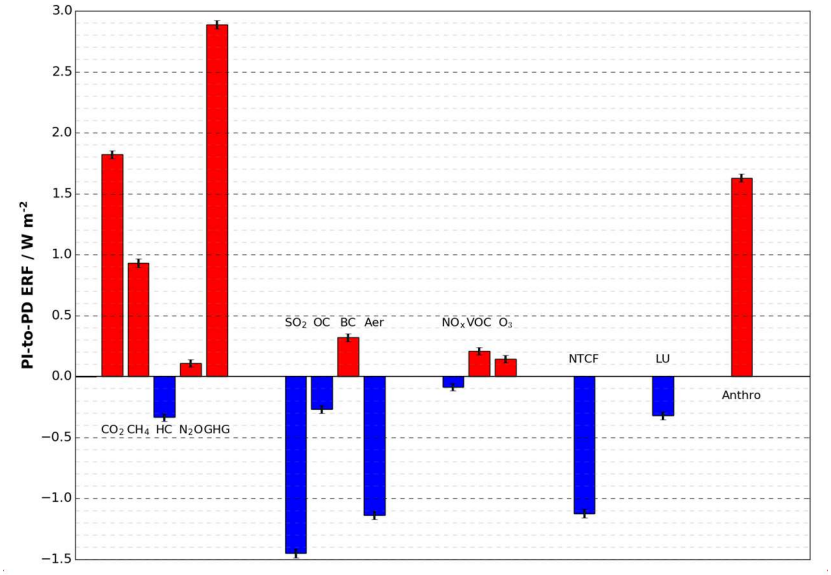
The ~~global mean~~ ERFs are also plotted in Fig. 2. Together, they show that the ~~global mean PD~~ ERF from~~by greenhouse gases~~ ~~(~~GHGs~~)~~ is ~~+2.89 ± 0.04~~ 2.92 ± 0.04 W m^{-2}, which is offset by an aerosol ERF of ~~-1.13 ± 0.04~~ -1.09 ± 0.04 W m^{-2}. The~~se~~ GHG ERF estimate~~s~~ is~~are~~ lower and the aerosol ERF is consistent with ~~ERFs~~estimates of ~~+~~3.09 and -1.10 W m^{-2}, respectively (Andrews et al., 2019), from the HadGEM3 GC3.1 (Williams et al., 2017) physical model (herein referred to as HadGEM3-GC3.1) upon which UKESM1 is based ~~and~~ but are consistent with the range of previous estimates (e.g., Myhre et al., 2013a). The net anthropogenic ERF is ~~+1.61~~ 1.76 ± 0.04 W m^{-2}, again consistent with the range of estimates from AR5 (Myhre et al., 2013a) ~~but less than~~ and the estimate from HadGEM3-GC3.1 (~~+~~1.81 W m^{-2}; Andrews et al., 2019)~~ by nearly 0.2 W m^{-2}; this is largely due to the land use ERF in UKESM1 being more strongly negative than in HadGEM3-GC3.1 by 0.21 W m^{-2}~~. The net anthropogenic ERF quantified here is also narrowly ~~outside~~ inside the range of net anthropogenic ~~total historical radiative forcing~~ERF of 2.3 W m^{-2} (~~+~~1.7 to ~~+~~3.0 W m^{-2}; 5 – 95 % confidence interval) ~~[5-95%] W m^{-2} (Andrews and Forster, 2019)~~ derived from a top-down energy budget constraint~~s~~ based on measurements of historical global mean temperature change and the Earth's heat uptake, and model estimates of the Earth's radiative response (Andrews and Forster, 2020).

~~PI-to-PD~~PD-PI Perturbation	Present day (PD; Year 2014) e~~E~~ffective radiative forcings (ERFs) ~~from 1850 to 2014~~ relative to the pre-industrial (PI; Year 1850) period (W m^{-2})						
	NET ERF	**LW*cs'***	**SW*cs'***	**LW ΔCRE'**	**SW ΔCRE'**	**NET*cs'***	**NET ΔCRE'**
CO_2 concn.	~~+1.82 ±0.04~~ 1.89 ±0.04	~~+1.56 ±0.02~~ 1.61 ±0.02	~~+0.12 ±0.02~~ 0.09 ±0.02	~~-0.35 ±0.02~~ -0.31 ±0.02	~~0.50 ±0.03~~ 0.50 ±0.02	~~+1.68 ±0.03~~ 1.70 ±0.02	~~+0.14 ±0.03~~ 0.19 ±0.02
CO_2 phys concn.	~~+0.008 ±0.04~~ 0.03 ±0.04	~~-0.065 ±0.03~~ -0.07 ±0.03	~~-0.019 ±0.02~~ -0.03 ±0.02	~~-0.021 ±0.02~~ -0.02 ±0.02	~~+0.112 ±0.02~~ 0.16 ±0.02	~~-0.083 ±0.03~~ -0.11 ±0.03	~~+0.09 ±0.02~~ 0.14 ±0.02
4xCO_2 concn.	~~+7.88 ±0.04~~ 7.97 ±0.04	~~+6.80 ±0.03~~ 6.83 ±0.03	~~+0.49 ±0.02~~ 0.46 ±0.02	~~-1.55 ±0.02~~ -1.51 ±0.02	~~+2.14 ±0.02~~ 2.18 ±0.02	~~+7.29 ±0.02~~ 7.29 ±0.03	~~+0.60 ±0.02~~ 0.68 ±0.02
4xCO_2 phys concn.	0.13 ±0.03	-0.30 ±0.02	-0.10 ±0.02	0.00 ±0.01	0.53 ±0.02	-0.40 ±0.02	0.53 ±0.02
N_2O concn.	~~+0.13 ±0.03~~	~~+0.25 ±0.03~~	~~-0.03 ±0.02~~	~~-0.13 ±0.02~~	~~+0.03 ±0.03~~	~~+0.22 ±0.03~~	~~-0.10 ±0.03~~

	0.25 ±0.04	0.28 ±0.03	-0.04 ±0.02	-0.08 ±0.01	0.09 ±0.03	0.25 ±0.03	0.01 ±0.02
ODS ~~HC~~ concns.	~~-0.33 ±0.04~~ -0.18 ±0.04	~~+0.37 ±0.02~~ 0.45 ±0.02	~~-0.43 ±0.02~~ -0.45 ±0.02	~~+0.23 ±0.02~~ 0.23 ±0.02	~~-0.50 ±0.03~~ -0.40 ±0.03	~~-0.06 ±0.02~~ 0.00 ±0.02	~~-0.27 ±0.02~~ -0.18 ±0.02
CH_4 concn.	~~+0.93 ±0.04~~ 0.97 ±0.04	~~+0.73 ±0.02~~ 0.74 ±0.02	~~+0.12 ±0.01~~ 0.11 ±0.02	~~-0.39 ±0.02~~ -0.39 ±0.02	~~+0.48 ±0.03~~ 0.50 ±0.02	~~+0.85 ±0.03~~ 0.85 ±0.03	~~+0.08 ±0.03~~ 0.12 ±0.02
GHG (CO2, N2O, ODSs, Other HCs, CH4) concns.	~~+2.89 ±0.04~~ 2.92 ±0.04	~~+3.08 ±0.02~~ 3.08 ±0.02	~~-0.16 ±0.02~~ -0.18 ±0.02	~~-0.65 ±0.02~~ -0.63 ±0.02	~~+0.62 ±0.03~~ 0.65 ±0.03	~~+2.92 ±0.02~~ 2.90 ±0.02	~~-0.01 ±0.03~~ 0.02 ±0.03
SO_2 emissions	~~-1.45 ±0.03~~ -1.37 ±0.03	~~+0.14 ±0.03~~ 0.15 ±0.03	~~-0.60 ±0.02~~ -0.61 ±0.02	~~0.18 ±0.01~~ 0.17 ±0.02	~~-1.17 ±0.03~~ -1.08 ±0.03	~~-0.46 ±0.03~~ -0.46 ±0.03	~~-0.99 ±0.02~~ -0.91 ±0.02
BC emissions	~~+0.32 ±0.04~~ 0.37 ±0.03	~~-0.00 ±0.02~~ -0.02 ±0.02	~~+0.38 ±0.02~~ 0.36 ±0.02	~~-0.18 ±0.01~~ -0.16 ±0.01	~~+0.12 ±0.03~~ 0.15 ±0.02	~~+0.38 ±0.03~~ 0.38 ±0.02	~~-0.06 ±0.02~~ -0.01 ±0.02
OC emissions	~~-0.27 ±0.03~~ -0.22 ±0.04	~~-0.01 ±0.02~~ 0.03 ±0.02	~~-0.17 ±0.02~~ -0.18 ±0.02	~~+0.01 ±0.02~~ -0.02 ±0.02	~~-0.09 ±0.03~~ -0.05 ±0.03	~~-0.18 ±0.02~~ -0.14 ±0.03	~~-0.09 ±0.02~~ -0.07 ±0.02
Aerosol and aerosol precursor (BC, OC, SO_2) emissions	~~-1.13 ±0.04~~ -1.09 ±0.04	~~+0.17 ±0.02~~ 0.16 ±0.03	~~-0.28 ±0.02~~ -0.26 ±0.02	~~-0.02 ±0.01~~ 0.01 ±0.02	~~-1.00 ±0.03~~ -1.00 ±0.03	~~-0.11 ±0.02~~ -0.10 ±0.02	~~-1.02 ±0.03~~ -1.00 ±0.02
NOx emissions	~~-0.08 ±0.04~~ 0.03 ±0.04	~~+0.02 ±0.03~~ 0.05 ±0.03	~~+0.02 ±0.02~~ 0.03 ±0.02	~~-0.02 ±0.01~~ -0.03 ±0.01	~~-0.11 ±0.02~~ -0.02 ±0.02	~~+0.04 ±0.03~~ 0.08 ±0.03	~~-0.13 ±0.02~~ -0.05 ±0.02

VOC and CO emissions	+0.21 ±0.04 0.33 ±0.04	+0.08 ±0.03 0.10 ±0.03	+0.01 ±0.02 0.03 ±0.02	-0.08 ±0.02 -0.09 ±0.01	+0.20 ±0.03 0.28 ±0.02	+0.09 ±0.03 0.13 ±0.03	+0.12 ±0.03 0.20 ±0.02
O3 precursor (VOC, CO, NOx) emissions	+0.15 ±0.04 0.21 ±0.04	+0.11 ±0.02 0.07 ±0.03	+0.05 ±0.02 0.06 ±0.02	-0.10 ±0.02 -0.07 ±0.02	+0.10 ±0.03 0.15 ±0.03	+0.16 ±0.03 0.13 ±0.02	-0.01 ±0.03 0.08 ±0.02
NTCF (BC, OC, SO_2, VOC, CO, NOx) emissions	-1.12 ±0.03 -1.03 ±0.04	+0.24 ±0.02 0.23 ±0.03	-0.29 ±0.02 -0.26 ±0.02	-0.06 ±0.02 -0.08 ±0.02	-1.02 ±0.03 -0.92 ±0.03	-0.04 ±0.03 -0.03 ±0.03	-1.08 ±0.02 -1.00 ±0.02
Land Use	-0.32 ±0.04 -0.17 ±0.04	+0.05 ±0.03 0.02 ±0.03	-0.46 ±0.02 -0.30 ±0.02	+0.01 ±0.01 0.03 ±0.01	+0.07 ±0.02 0.09 ±0.03	-0.40 ±0.03 -0.28 ±0.03	+0.09 ±0.02 0.11 ±0.03
Anthro (GHG concns., Aer. and aerosol precursor (BC, OC, SO_2) ems., O_3 precursor (VOC, CO, NOx) ems., and land use	+1.61 ±0.03 1.76 ±0.04	+3.38 ±0.02 3.34 ±0.02	-0.87 ±0.02 -0.72 ±0.02	-0.66 ±0.02 -0.64 ±0.01	-0.24 ±0.02 -0.22 ±0.03	+2.50 ±0.03 2.63 ±0.03	-0.89 ±0.02 -0.86 ±0.03

Table 3: Present-day (PD; (Year 2014) effective radiative forcings (ERFs) of climate relative to the pre-industrial (PI; (Year 1850) period in W m^{-2} derived from Equations (6) to Eqn. (8), and including an estimate of the standard error (Forster et al., 2016). Where duplicate experiments exist (e.g., *piClim-SO2*), the quoted ERFs use rRealisations 2. Units in W m^{-2}.



380

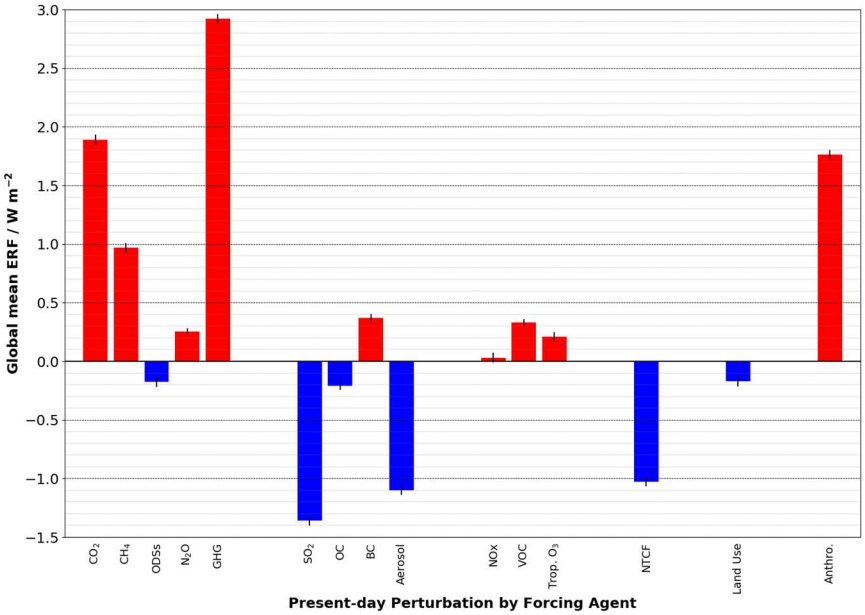


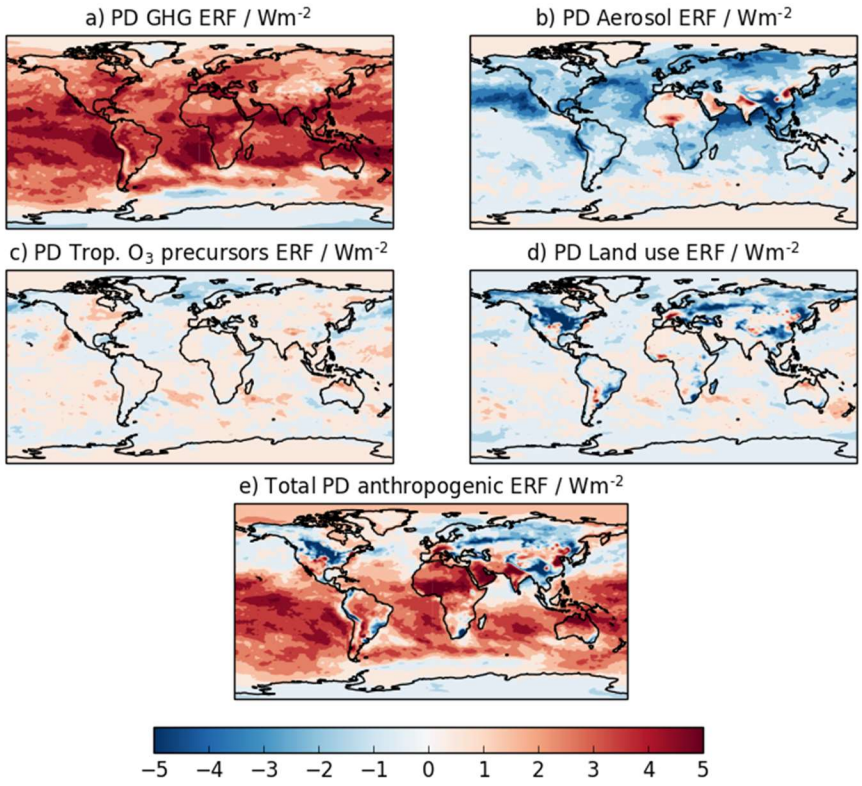
Figure 2: ~~**Bar chart showing global mean**~~ Present day (Year 2014) **anthropogenic** ~~**effective radiative forcings (**~~**ERFs**~~**)**~~ ~~**at the present day (Year 2014)**~~ **relative to the pre-industrial (Year 1850)** period derived from Eqn. (8), and including an estimate of the standard error.~~**, diagnosed from paired fixed SST timeslice simulations with an atmosphere-only configuration of UKESM1.**~~ Units in W m^{-2}.

385

Figure 3 shows global and zonal mean distributions of the ~~multi-annual mean PD~~ ERFs from PD-PI changes in GHG concentrations, aerosol and aerosol precursor (BC, OC. SO_2) emissions, ~~tropospheric ozone (~~O_3~~)~~ precursor (VOC, CO, NOx)

emissions, land use change, and total anthropogenic sources ~~since PI~~. It shows that the ~~PD~~ ERF from GHGs (Fig. 3a) has a robust signal over 93 % of the globe; it is ~~strongly~~ positive everywhere except for the southern high latitudes (Fig. 3b). ~~although there is some evidence of negative forcing (up to –2 W m^{-2}) in the high latitudes, particularly in the southern hemisphere (SH).~~ This negative ERF is due to ~~a~~the negative ERF from the *piClim-HC* perturbation experiment (Table 3) which offsets the positive ~~forcings~~ ERFs from the other LLGHG experiments. ~~Indeed, Morgenstern et al. (2019) found that the negative ERF from ozone depletion and cloud feedbacks (including aerosol-mediated cloud feedbacks) offsets the direct positive ERF from the ozone-depleting substances (ODSs).~~ The breakdown of the GHG ERF~~forcing~~ into its individual speciated contributions will be discussed further in Sect. 4.2~~3~~.2 and results from *piClim-HC* will be discussed in Sect. 4.2.3. The aerosol ERF~~forcing~~ (Figure 3~~b~~c and d) is robust over a smaller area (52 %) of the globe and is more spatially heterogeneous than the GHG ~~forcing~~ ERF due to the~~ir~~ shorter aerosol lifetime. However, the distribution of the aerosol ~~forcing~~ERF is ~~still largely~~ mostly negative ~~everywhere~~, except for over bright surfaces and regions where the positive forcing from black carbon (BC) emissions outweighs the negative forcing from scattering aerosols such as sulphate and ~~organic carbon (~~OC~~)~~. A breakdown of the aerosol ~~forcing~~ ERF between constituents and between ~~aerosol-radiation interactions (~~ari~~)~~ and ~~aerosol-cloud interactions (~~aci~~)~~ will be presented and discussed in Sect. ~~3.3~~4.3.

Figures ~~3c~~ 3e and 3f shows the global and zonal mean distributions of the ~~PD~~ ERF from emissions of ~~tropospheric~~ O_3 precursors (VOC, CO, and NOx), excluding CH_4. It shows that the ~~forcing~~ERF from changes in VOC, CO, and NOx emissions alone is ~~spatially heterogeneous, has regions of both weak positive and weak negative contributions, but on a global mean basis is weakly~~ positive (~~+0.15~~ 0.21 ± 0.04 W m^{-2}; Table 3), with only 10 % of the globe showing a robust signal. ~~in comparison with other forcings (e.g. GHGs).~~ Further analysis of the ERF from ~~tropospheric~~ O_3 precursor emissions and their contribution, along with ~~methane (~~CH_4~~)~~, to ~~forcing by~~ tropospheric O_3 RF, can be found in Sect. ~~3.4~~4.4. The distribution of ~~the~~a robust land use ERF (Fig. ~~3d~~3g; Fig. 3h) is also limited in spatial extent (~12 % of the globe); ~~spatially heterogeneous although~~ much of the negative ~~forcing~~ERF is concentrated over the northern hemisphere (NH) continental regions (e.g., North America, South East Asia). Together with aerosols, the~~ir~~ combined ERF outweighs the positive ERF from GHGs, leading to a negative total anthropogenic ERF over ~~much~~parts of the NH continents (Fig. ~~3e~~3i). As was the case with the GHG ~~forcing~~ERF, the negative ERF over the SH high latitudes is still evident in the total anthropogenic ~~forcing~~ERF (Fig. 3i; Fig. 3j).


a) PD GHG ERF / Wm-2
b) PD Aerosol ERF / Wm-2
c) PD Trop. O3 precursors ERF / Wm-2
d) PD Land use ERF / Wm-2
e) Total PD anthropogenic ERF / Wm-2
−5 −4 −3 −2 −1 0 1 2 3 4 5

415

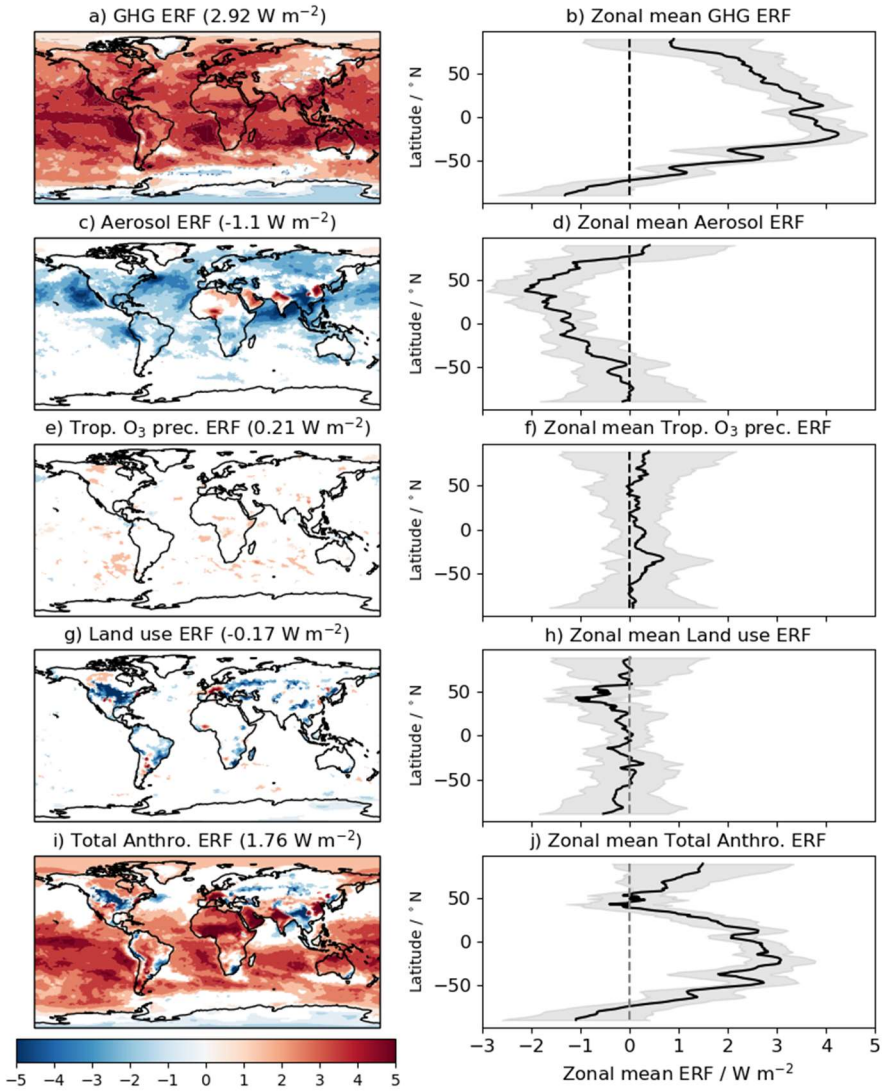


Figure 3: Geographical and zonal mean distributions of the ~~multi-annual mean~~ present-day (PD; Year 2014) ~~effective radiative forcing (~~ERFs~~)~~ relative to the pre-industrial (PI; Year 1850) period for ~~(a)~~ GHGs in a) and b), ~~(b)~~ aerosols in c) and d), (c) ~~tropospheric ozone (~~O_3~~)~~ precursors in e) and f), ~~(d)~~ land use change in g) and h), and ~~(e)~~ total anthropogenic in i) and j), respectively. In the global distribution plots, global mean ERFs are included and areas are masked out in white, where the ERF is not statistically significant at the 95 % confidence interval. In the zonal mean plots, the grey shading shows the ± 1 standard deviation in the zonal mean ERF. Units in W m^{-2}.

As discussed in Sect. ~~2~~3, the UKESM1 ~~fixed SST~~ fSST experiments were run on multiple ~~High Performance Computing (~~HPC~~)~~ platforms. ~~Despite s~~Statistical methods ensur~~cing~~ that the model is not scientifically different on the different HPC platforms, but such duplicate experiments still produce slightly different results. This raises the question of the impact of such differences on the quantification of ~~PD~~ ERFs with UKESM1. To address this question, the experiments described in Table 2

were used to compare the difference in TOA radiative fluxes of equivalent realisation experiment pairs. The two-sample Kolmogorov–Smirnov test between the monthly TOA radiative fluxes from the realisation pairs (null hypothesis that the samples are drawn from the same distribution) shows that the TOA radiative fluxes between the two realisation pairs are statistically identical at a confidence level (α) of 5 % (Table 4). Despite the fact that one cannot conclude that the distributions are identical, for each pair of experiments there is no evidence suggesting that the two distributions are different.

Pair	Differences in ERF and its components ~~the TOA radiative fluxes~~ (W m^{-2})						
	NET ERF	**LW*cs'***	**SW*cs'***	**LW ΔCRE'**	**SW ΔCRE'**	**NET*cs'***	**NET ΔCRE'**
piClim-SO2 R1 vs R2	~~+0.06 ±0.03~~ 0.01 ±0.03	~~+0.02 0.02~~ -0.01 0.02	~~+0.02 ±0.02~~ 0.01 ±0.02	~~0.01 ±0.02~~ 0.01 ±0.02	~~0.04 ±0.02~~ -0.01 ±0.02	~~-0.008 ±0.02~~ 0.01 ±0.02	~~+0.05 ±0.02~~ 0.01 ±0.02
piClim-OC R1 vs R2	~~-0.03 ±0.03~~ 0.01 ±0.03	~~-0.03 ±0.02~~ 0.01 ±0.02	~~+0.007 ±0.01~~ -0.01 ±0.01	~~-0.004 ±0.01~~ 0.03 ±0.01	~~-0.004 ±0.02~~ -0.02 ±0.02	~~-0.03 ±0.02~~ -0.01 ±0.02	~~-0.008 ±0.02~~ 0.01 ±0.02
piClim-NTCF R1 vs R2	~~0.00 ±0.00~~ 0.01 ±0.03	~~0.00 ±0.00~~ -0.01 ±0.02	~~0.00 ±0.00~~ 0.01 ±0.02	~~0.00 ±0.00~~ 0.01 ±0.02	~~0.00 ±0.00~~ -0.01 ±0.03	~~0.00 ±0.00~~ 0.01 ±0.02	~~0.00 ±0.00~~ 0.01 ±0.02

Table 4: Differences in the ~~TOA radiative fluxes (± standard error) in the~~ present-day ERF and its components ~~clear-sky (CS) shortwave (SW), CS longwave (LW), and the SW and LW cloud radiative effect (CRE) at the present day (PD; Year 2014)~~ between two realisations (R1 and R2) of the same perturbation experiment relative to *piClim-control*., Units in W m^{-2}.

4.2 Long-Lived Greenhouse Gases (LLGHGs)

4.2.1 Carbon Dioxide (CO_2)

Atmospheric ~~carbon dioxide (~~CO_2~~)~~ concentrations have risen from 284.3 to 397.5 ppm between 1850 and 2014 (Meinshausen et al., 2017) resulting in an ~~global mean~~ ERF of ~~+1.83 ± 0.04~~ 1.89 ± 0.04 W m^{-2} (Table 3) and is the largest individual contribution to the total historical forcing. Myhre et al. (2013a) report a ~~stratospherically adjusted radiative forcing (~~SARF~~) or RF~~ of ~~+~~1.82 ± 0.19 W m^{-2} for the period from 1750 to 2011 which, coincidentally, has a near identical rise in CO_2 of 113 ppm to that assessed here, although CO_2 forcing has a logarithmic dependency on concentration. An updated ~~radiative forcing~~ assessment based on line-by-line calculations (Etminan et al., 2016) increased SARF for 2015 relative to 1750 to 1.95 W m^{-2}

but for a larger CO_2 rise of 121 ppm. Applying Etminan et al., (2016) to our case reveals a SARF of 1.80 W m^{-2}. Our ERF estimate for 2014 is, therefore, ~~consistent with~~ larger than the SARF by 0.09 W m^{-2}. ~~from line-by-line calculations and confirms that for CO_2, the SARF is a useful metric for evaluating the radiative effects of increasing CO_2 (Smith et al., 2018).~~ As expected for a GHG, the ERF is dominated by the LWcs' component ~~absorption in the clear-sky (CS) longwave (LW)~~ (~~+1.56 ± 0.02~~ 1.61 ± 0.02 W m^{-2}) with an additional contribution from the SWcs' component ~~CS shortwave (SW) forcing~~ (~~+0.122 ± 0.02~~ 0.09 ± 0.02 W m^{-2}) coming from direct effects and a ~~rapid adjustment~~RA in snow- cover across the northern latitudes. ~~and~~ There is also a ~~net cloud radiative effect~~NET △CRE' contribution (~~0.144 ± 0.03~~ 0.19 ± 0.02 W m^{-2}) from a reduction in cloud cover. The SW direct effect occurs in the near-infrared and is small, equivalent to around 2 % of the LW forcing (Pincus et al., 2020). On that assumption, the SWcs' component is dominated by the rapid albedo adjustment over the direct effect.

Rising atmospheric CO_2 also exerts an indirect forcing through rapid changes in plant stomatal conductance (Doutriaux-Boucher et al., 2009; Richardson et al., 2018) enhancing plant water use efficiency and reducing evapotranspiration leading to an increase in sensible heating at the surface and corresponding drying of the boundary layer and reduction in low clouds~~low-clouds~~. This mechanism is known as a physiological forcing. In UKESM1 (*piClim-$CO_2$2phys*; Table 3), this forcing is small ~~for~~ at ~~the~~ PD and 4xCO_2 levels (~~0.008 ± 0.04~~ 0.03 ± 0.04 cf. 0.13 ± 0.03 W m^{-2}) ~~but~~ and scales in line with the total CO_2 ERF. It results from a balance between a negative ~~CS LW~~ LWcs' component (~~-0.065 ± 0.03~~ -0.07 ± 0.03; -0.30 ± 0.02 W m^{-2}) associated with surface warming and a positive SW ~~cloud radiative effect (CRE)~~ △CRE' component (~~0.112 ± 0.02~~ 0.14 ± 0.02; 0.53 ± 0.02 W m^{-2}) associated with a reduction in low-level clouds. ~~The balance comes~~ and from smaller terms including a negative ~~CS SW contribution~~ SWcs' component from reduced water vapour. Previous Hadley Centre models at 4xCO_2 found a ~~much~~ similar but larger effect: 1.1 W m^{-2} in HadCM3LC (Doutriaux-Boucher et al., 2009) and 0.25 W m^{-2} in HadGEM2-ES (Andrews et al., 2012b). ~~Assuming the physiological effect scales linearly with atmospheric CO_2, we~~ We find that UKESM1 has a ~~stronger~~ similar ~~CS LW~~ LWcs' component compared to HadGEM2-ES implying that UKESM1 has a ~~more pronounced surface warming adjustment associated with the physiological effect~~ similar surface warming adjustment associated with the physiological effect but offset by a weaker ~~SW CRE~~ SW △CRE' component. The inclusion of the physiological effect and rapid albedo adjustment in the CO_2 ERF acts to increase the forcing slightly. These additional adjustments likely account for our slightly higher ERF relative to the SARF from line-by-line estimates (Etminan et al., 2016).

4.2.2 Nitrous Oxide (N_2O)

The ~~global mean~~ ERF due to changes in N_2O concentration from the PI period (~~Year1850 value of~~ 273 ppbv in 1850) to ~~PD~~the present day (~~Year 2014 value of~~ 327 ppbv in 2014) is calculated as ~~+0.13±0.03~~ 0.25 ± 0.03 W m^{-2} (Table 3), following ~~the equations (6)~~ Eqn. (8) described above. The predominant contribution to the N_2O ERF is the ~~CS LW forcing~~LWcs' component (~~+0.25 ± 0.03~~ 0.28 ± 0.03 W m^{-2}), with a small ~~and non-significant cancellation~~ offset by the ~~CS SW forcing~~SWcs' component (~~-0.03 ± 0.02~~ -0.04 ± 0.02 Wm^{-2}); the ~~net CS forcing~~ NETcs' component sums up to ~~+0.22 ± 0.03~~ 0.25 ± 0.03 W m^{-2}. The ~~net cloud radiative effect (CRE) forcing~~NET △CRE' component is insignificant (~~-0.10 ± 0.03~~ 0.01 ± 0.03 W m^{-2}), with ~~the SW~~

and LW SW △CRE' and LW △CRE' contributions of +0.03 ± 0.030.09 ± 0.03 W m^{-2} and -0.13 ± 0.02 -0.08 ± 0.02 W m^{-2}, respectively. In comparison, the net ERF calculated here (+0.13 ± 0.03 0.25 ± 0.03 W m^{-2}) is slightly lower higher than the SARF values of +0.17 ± 0.03 W m^{-2} from AR5 for 2011 (Myhre et al., 2013a) and of +0.18 W m^{-2} for 2014 based on the updated expression from Etminan et al. (2016). This is likely to be due to the effect of adjustments associated with changing N_2O that weren'twere not considered as part of the SARF in AR5 (Myhre et al., 2013a) or Etminan et al. (2016), including O_3 depletion and fast cloud adjustments. The UKESM1 estimate agrees well with the AerChemMIP multi-model mean of 0.23 ± 0.05 W m^{-2} (Thornhill et al., 2020). Previously, Hansen et al. (2005) calculated an IRF and a SARF of 0.15 W m^{-2} due to the change in N_2O from 278 to 316 ppbv.

The global distribution of the individual components contributing to the N_2O ERF (Fig. 4) shows that the SW forcing SWcs' component is negligible under CS conditions and is largely made up of noise in cloudy conditions which might be due to an aerosol effect. CS LW forcingThe LWcs' component is predominantly positive, due mainly to N_2O direct forcing. The SW △CRE' and LW △CRE' components of the ERF are largely noise (at 95 % confidence level; Fig 4), with a net contribution to the ERF of close to zero. There are relatively large regional changes in the LW due to clouds leading to a significant net negative LW CRE of -0.13 W m^{-2}. The positive LW CRE over the southern high latitudes appears to be correlated to the change in total column ozone (TCO) and the shift of the surface wind shown in Fig. 5. Zonal mean ozone (O_3) and temperature changes (Fig. 5) show the warming of the upper troposphere and lower stratosphere due to increased O_3 in that region and the cooling of the stratosphere due to O_3 depletion associated with the N_2O increase. As a result, the atmospheric circulation has been modified, reflected in a shift of the upward velocity (w component of the wind) (Fig. 5d), which is correlated to the water content of the clouds. Such changes in clouds drive the LW forcing, which shows a pattern of significant regional changes with cancelling effects.

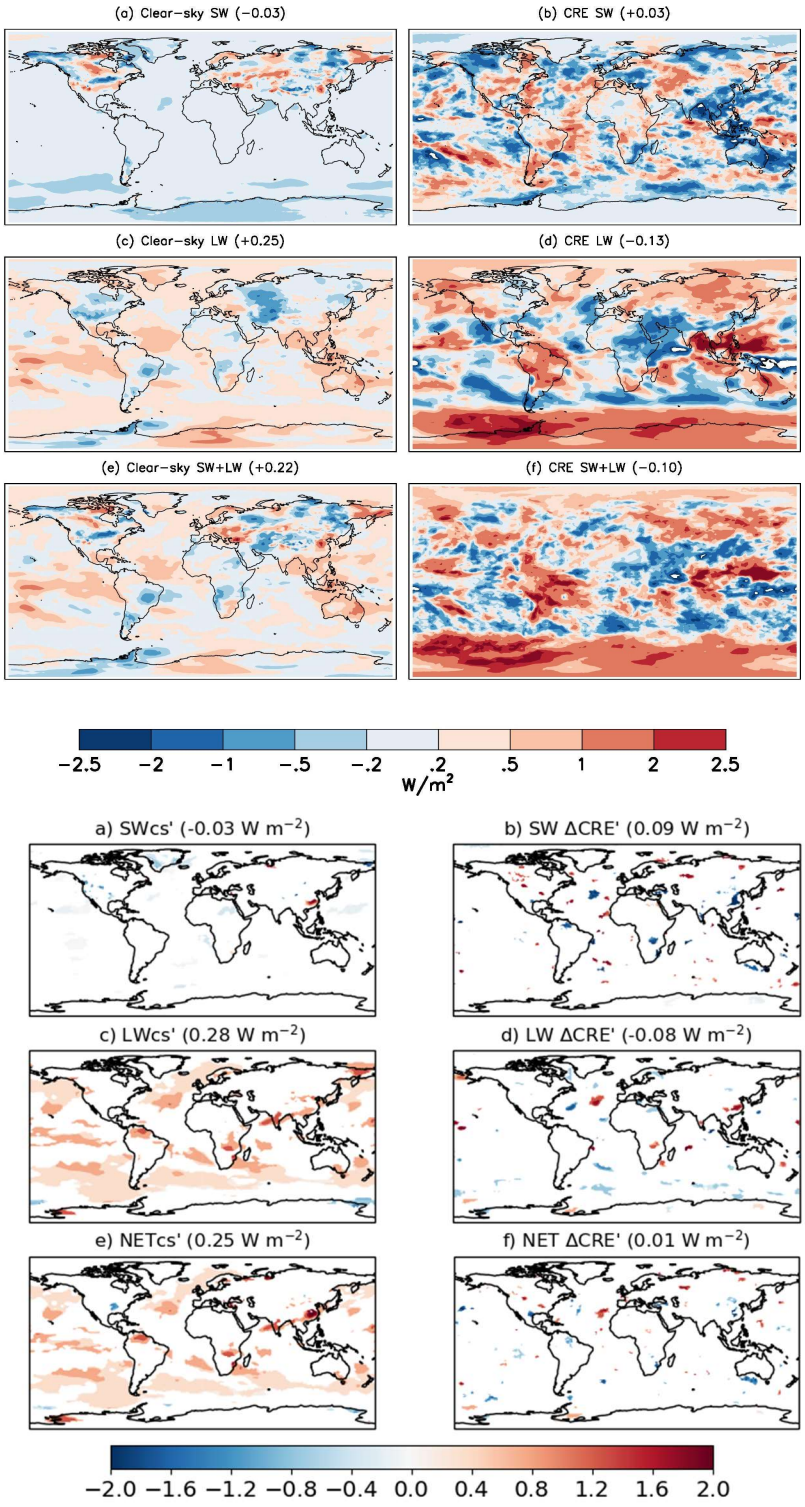
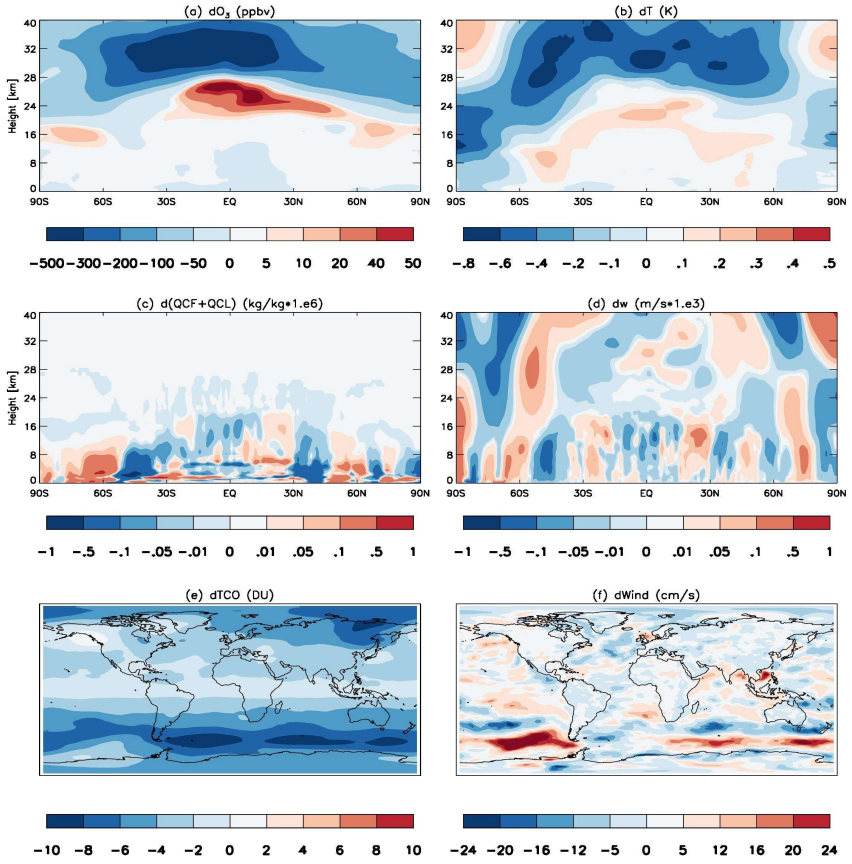

(a) Clear−sky SW (−0.03)
(b) CRE SW (+0.03)
(c) Clear−sky LW (+0.25)
(d) CRE LW (−0.13)
(e) Clear−sky SW+LW (+0.22)
(f) CRE SW+LW (−0.10)
−2.5 −2 −1 −.5 −.2 .2 .5 1 2 2.5
W/m²
a) SWcs' (-0.03 W m⁻²)
b) SW ΔCRE' (0.09 W m⁻²)
c) LWcs' (0.28 W m⁻²)
d) LW ΔCRE' (-0.08 W m⁻²)
e) NETcs' (0.25 W m⁻²)
f) NET ΔCRE' (0.01 W m⁻²)
−2.0 −1.6 −1.2 −0.8 −0.4 0.0 0.4 0.8 1.2 1.6 2.0

Figure 4: Global distributions of the PD N_2O ERF components ~~at the top of atmosphere (TOA) for the present day (PD; Year 2014)~~ relative to the ~~pre-industrial (PI; Year 1850)~~ period, i.e., *piClim-N2O* minus *piClim-Control*; (a) ~~clear sky SW~~ SWcs', (b) ~~cloud-radiative effect (CRE)~~ SW △CRE', (c) ~~clear sky LW~~LWcs', (d) ~~CRE~~ LW △CRE', (e) ~~net (SW+LW) for clear sky~~NETcs', and (f) ~~net (SW+LW) CRE~~NET △CRE', based on Eqn. (8). Global mean values are shown in brackets. Regions where the ERF components are outside the 95 % confidence level are masked out in white. Units in W m^{-2}.

~~The change in TCO is largely negative due to a reduction in stratospheric O_3 (Fig. 5a). That would lead to a negative indirect forcing by O_3; indeed, the ERF calculated here is 0.05 W m^{-2} smaller than the direct N_2O forcing in AR5 (Myhre et al., 2013a). N_2O-induced O_3 depletion causes a strengthening of the Southern Annular Mode (evidenced in a poleward shift of the westerlies associated with the SAM). With that shift comes a redistribution of clouds with cloud structures also moving poleward. Tropical effects include a warming of the upper troposphere/lower stratosphere (UTLS) region associated with O_3 increases there, resulting in a regional suppression of convection and a reduction of associated LW radiation coming from cloud tops.~~



~~**Figure 5: Global distribution of changes in zonal and annual mean a) O_3 (ppbv), b) temperature (K), c) cloud liquid water (qcl) and cloud frozen water (qcf), and d), vertical wind velocity, and global distributions of the multi-annual mean e) total column ozone (TCO) and f) near surface wind.**~~

~~Prather and Hsu (2010) also noted that coupling of N_2O and methane (CH_4), via stratospheric O_3 depletion and photolysis, can reduce the global warming potential (GWP) of N_2O by 5 %, based on PD concentrations. Here, the response of CH_4 is constrained due to being concentration-driven. However, the whole-atmosphere CH_4 lifetime changes from 8.1 years in *piClim-control* to 8.0 years in *piClim-N2O*, which would lead to a small offset to the N_2O ERF of 0.01 W m^{-2} due to the direct radiative effect of CH_4 if the model was emissions-driven; indirect CH_4 forcings (e.g. O_3) could enhance this offset even further (O'Connor et al., 2019).~~

4.2.3 Ozone (O_3) Depleting Substances (ODSs) ~~Halocarbons~~

The ~~global multi-annual mean~~ ERF from the ~~PI to PD~~ PD-PI change in ~~ozone depleting substances (~~ODSs~~)~~ is quantified, using *piClim-HC* relative to *piClim-control*, as ~~-0.33 ± 0.04~~ -0.18 ± 0.04 W m^{-2}, which is dominated by the NET △CRE' component ~~the net ERF under cloudy conditions~~ (-0.18 ± 0.02 W m^{-2}; Table 3). Figure 5 shows global distributions of the SWcs', SW △CRE', LWcs', LW △CRE', NETcs', and NET △CRE' components, respectively. For CS conditions, the SW component (SWcs') is characterised by negative values over the southern high latitudes (to a lesser extent in the northern high latitudes) which is linked to pronounced Antarctic O_3 depletion and some decreases in Arctic O_3 (not shown). The LWcs' component is predominantly positive, reflecting the direct effect of ODSs acting as GHGs in the *piClim-HC* simulation. The positive LWcs' component at high latitudes contains an offset caused by O_3 depletion. Overall, the global mean NETcs' component is negligible (0.00 ± 0.02 W m^{-2}; Table 3). Under cloudy conditions, the negative SW △CRE' component and the positive LW △CRE' component of the ERFs are anti-correlated, especially over the Southern Ocean, summing up to a net contribution to the ERF of -0.18 ± 0.02 W m^{-2}. ~~A quantitative and process-based understanding of why the ERF is negative, despite our understanding to date of the relative roles of ODSs and O_3 depletion in climate forcing, is presented in Morgenstern et al. (2019).~~

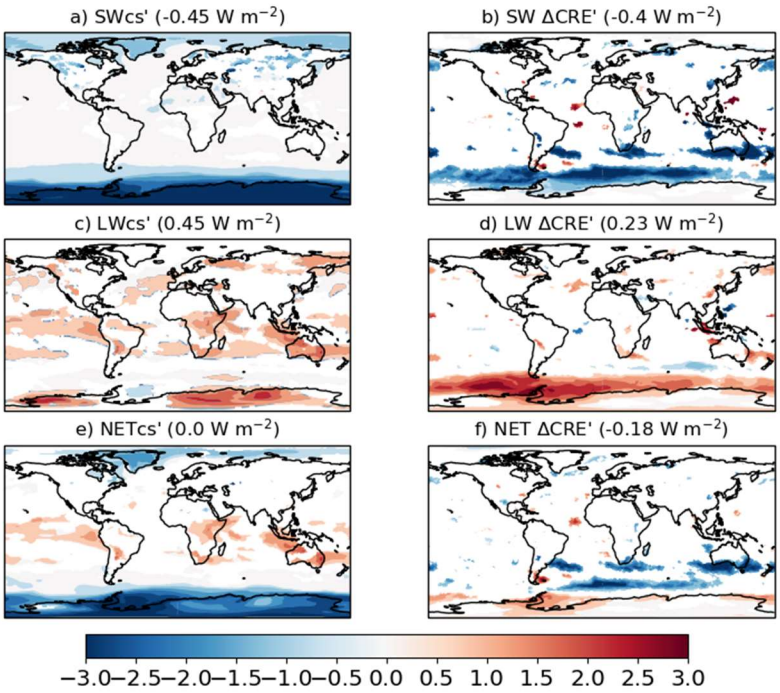


Figure 5: Global distributions of the PD ODS ERF components relative to the PI period, i.e., *piClim-HC* minus *piClim-Control*; (a) SWcs', (b) SW ΔCRE', (c) LWcs', (d) LW ΔCRE', (e) NETcs', and (f) NET ΔCRE', based on Eqn. (8). Global mean values are shown in brackets. Regions where the ERF components are outside the 95 % confidence level are masked out in white. Units in W m^{-2}.

As mentioned previously, the ERF includes the direct effect of the ODSs acting as GHGs, the indirect effect of the O_3 depletion that they cause, and any resulting TOA changes due to other RAs. Quantifying the historical evolution of the total O_3 RF *alone* from the CMIP6 models, Skeie et al. (2020) found that UKESM1 was the only model with both tropospheric and stratospheric chemistry that had a negative total O_3 RF at the present day. This is likely due to UKESM1 having a stronger global O_3 decline during the period of increasing ODSs than other models (Keeble et al., 2020). Likewise, in the AerChemMIP multi-model ODS ERF assessment, Morgenstern et al. (2020) used observed O_3 trends as a constraint on the modelled ODS ERF; they found that the ERF is likely to be between -0.05 and 0.13 W m^{-2}, with the UKESM1 estimate of -0.18 ± 0.04 W m^{-2} outside of this range. Although Keeble et al. (2020), Skeie et al. (2020) and Morgenstern et al. (2020) suggest that the negative contribution from O_3 depletion to the ODS ERF is too strong in UKESM1, there is an additional negative offset to the direct radiative effect of the ODSs through the NET ΔCRE' component (Figure 5). Furthermore, the ODSs will have an impact on tropospheric O_3; their contribution to the tropospheric O_3 RF will be quantified in Section 4.4.2.

4.2.4 Methane (CH_4)

The global mean ~~methane (~~CH_4~~)~~ concentration changed from 808.3 ppbv in the PI (Year 1850) period to 1831.5 ppbv in the PD (Year 2014) period, resulting in an ~~PD~~ ERF of ~~+0.93±0.04~~ 0.97 ± 0.04 W m^{-2} (Table 3; Fig. 6)~~, where the 0.04 W m^{-2} is the standard error following Forster et al. (2016).~~ Most of the ~~forcing~~ERF (~~+0.73±0.02~~ 0.74 ± 0.03 W m^{-2}~~)~~; Table 3) is due to the LWcs' component~~occurs in the CS LW~~, with an additional positive contribution (~~+0.12±0.01~~ 0.11 ± 0.02 W m^{-2}) from the ~~CS SW~~SWcs' component, consistent with the growing recognition of the importance of the SW absorption bands in CH_4 forcing (Collins et al., 2006; Li et al., 2010; Etminan et al., 2016). There are additional SW and LW ~~contributions from the cloud radiative effect (CRE)~~ △CRE' components but these ~~largely~~partly cancel out, leading to a small ~~net positive forcing~~ NET △CRE' (~~+0.08 ± 0.03~~ 0.12 ± 0.03 W m^{-2}) in addition to the ~~net CS forcing~~NETcs' component (~~+0.85±0.03~~ 0.85 ± 0.03 W m^{-2}). Estimates of the direct CH_4 ERF ~~at the PD~~ from HadGEM2 model simulations (Andrews, 2014) and the updated RF expression for CH_4 based on line-by-line calculations (Etminan et al., 2016) are ~~of~~in the order of 0.50-0.56 W m^{-2}. But the ~~forcing~~ERF calculated here is higher by more than 0.4 W m^{-2}. However, it is consistent with other studies (e.g.~~,~~ Hansen et al., 2005; Shindell et al., 2009; Myhre et al., 2013a), who concluded that the total climate forcing by CH_4 is almost double that of the direct forcing and are due to indirect effects. The UKESM1 estimate is also larger than the ~~±~~0.69 W m^{-2} radiative impact of an increase in CH_4 concentration of 1800 ppbv above ~~present-day~~ PD levels quantified by Winterstein et al. (2019) with the ECHAM/MESSy Atmospheric Chemistry (EMAC) coupled model. Although the Winterstein et al. (2019) estimate included indirect forcings from O_3 and stratospheric water vapour (SWV), their direct CH_4 forcing in the LW is low relative to other models (Lohmann et al., 2010).

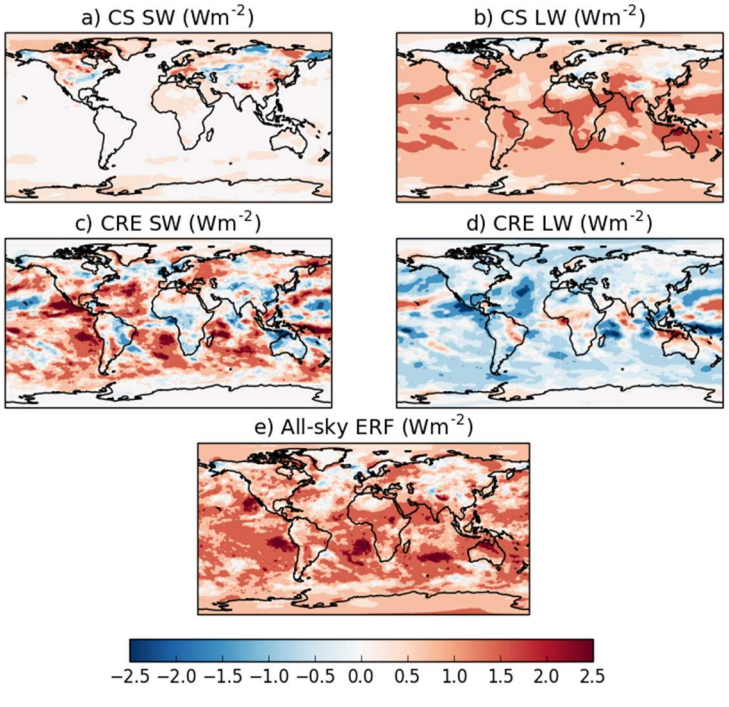
The UKESM1 ERF quantified here is at the upper end of estimates from the recent study of AerChemMIP multi-model ERFs by Thornhill et al. (2020). They found that the multi-model mean ERF was 0.70 W m^{-2}, with a standard deviation of 0.22 W m^{-2}. They attributed part of the inter-model spread to different complexities in the representation of interactive chemistry in the respective models, i.e., some models only captured the direct radiative effect of CH_4 (e.g., NorESM2) while others (e.g., UKESM1) also included indirect contributions from CH_4-driven changes in O_3 and SWV. However, the contribution to the ERF from tropospheric adjustments differed in both magnitude and sign between the models, with UKESM1 being the only model with a positive contribution to the ERF from tropospheric RAs. The relative contributions of the direct and indirect ~~forcings~~ contributions to the total CH_4 ERF quantified here and the mechanism behind the positive tropospheric RA ~~an emissions-based perspective of the CH_4 forcing~~ can be found in O'Connor et al. (2019).

4.2.5 Total Greenhouse Gases (GHGs)

The major driver~~s~~ of anthropogenic climate change ~~is~~are ~~greenhouse gases (~~GHGs~~)~~, ~~which~~whose forcing is offset by aerosols (Myhre et al., 2013a). Therefore, the total ~~greenhouse~~GHG ERF and the ~~total~~ aerosol ERF are key ~~metrics~~values in understanding observed and modelled changes in the climate system since the PI period. As a result, a separate timeslice

simulation with all GHG concentrations (*piClim-GHG*; Table 1) at PD levels was conducted following the RFMIP protocol (Pincus et al., 2016).

595



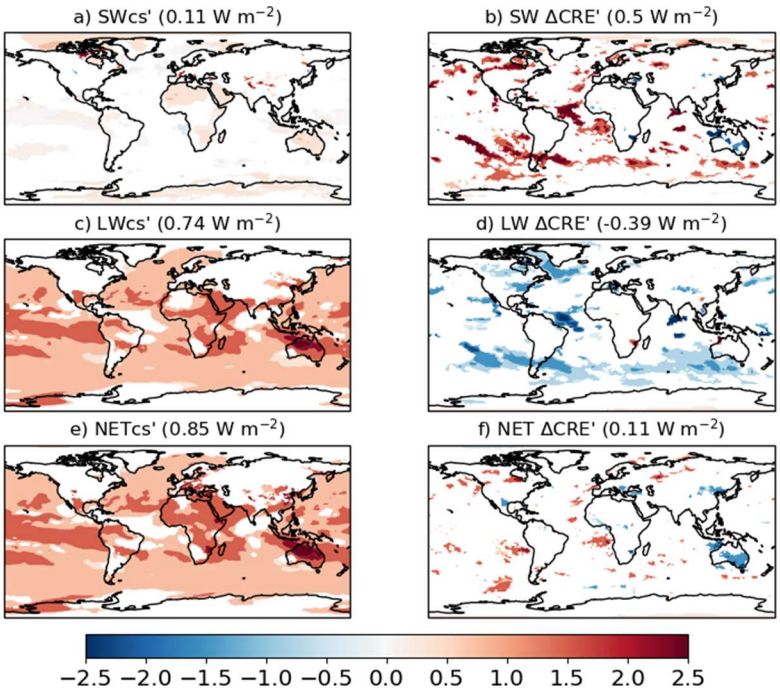


Figure 6: Global distributions of the PD CH_4 ERF components relative to the PI period, i.e., *piClim-CH4* minus *piClim-Control*; (a) SWcs', (b) SW △CRE', (c) LWcs', (d) LW △CRE', (e) NETcs', and (f) NET △CRE', based on Eqn. (8). Global mean values are shown in brackets. Regions where the ERF components are outside the 95 % confidence level are masked out in white. Units in W m^{-2}.

~~**Figure 6: Global distribution of the components of the PD (Year 2014) methane ERF relative to PI (Year 1850): a) Clear-sky (CS) shortwave (SW), b) CS longwave (LW), and the cloud radiative effect (CRE) in the c) SW and d) LW, and the all-sky ERF in e) calculated using Equations (6)-(8) according to Ghan (2013). Forcings in W m^{-2}.**~~

The UKESM1 *piClim-GHG* experiment leads to an ERF of ~~2.89 ± 0.04~~ 2.92 ± 0.04 W m^{-2} (Table 3), which is dominated by a positive ~~CS LW contribution~~LWcs' component (~~3.08~~ 3.08 ± 0.02 W m^{-2}) that is partially offset by a negative ~~CS SW~~ SWcs' component (~~-0.16~~ -0.18 ± 0.02 W m^{-2}). There are significant positive and negative contributions from ~~the CRE in the SW and LW~~ SW △CRE' (-0.63 ± 0.02 W m^{-2}) and LW △CRE' (0.65 ± 0.03 W m^{-2}) components but these largely cancel out and contribute little (~~-0.01~~ 0.02 ± 0.03 W m^{-2}) to the total ERF (Table 3). This ~~PD~~ GHG ERF is lower than the ERF of 3.09 W m^{-2} estimated from the physical model HadGEM3-GC3.1 (Andrews et al., 2019). Part of this discrepancy may be due to the inclusion in UKESM1 of indirect forcings by O_3 and/or aerosols from O_3-depleting substances (Morgenstern et al., ~~2019~~2020) and CH_4 (O'Connor et al., ~~2019~~2020), for example. The UKESM1 ~~1850-2014~~ estimate, however, is consistent with the AR5 year-2011 SARF estimate of 2.82 W m^{-2}. The latter estimate of 2.82 W m^{-2} has been adjusted to an 1850 baseline from 1750, and taking stratospheric O_3 depletion, CH_4-driven SWV, and half of the tropospheric O_3 forcing (based on the attribution by

Stevenson et al., 2013) into account, although some GHG concentrations (e.g., CH_4, CO_2) have increased between 2011 and 2014 (Nisbet et al., 2016) while others (e.g., ODSs) have declined (Engel et al., 2018).

~~Although, in principle, one could use the combination of GHG simulations to test for linearity, this isn't feasible here; *piClim-GHG* perturbs non O_3-depleting halocarbons whereas *piClim-HC* does not. However, there is a discrepancy in ERF of 0.35 W m^{-2} between *piClim-GHG* and the sum of the individual GHG simulations which cannot be fully accounted for by the small positive RF from non O_3-depleting halocarbons (HFCs, PFCs, SF6) of 0.02 W m^{-2} at 2011 (Myhre et al., 2013a). This suggests that there is a potential non-linearity in the GHG forcing, perhaps involving O_3 chemistry and aerosol-mediated cloud feedbacks, which may warrant further investigation.~~

The simulation *piClim-GHG* perturbs non O_3-depleting halocarbons but *piClim-HC* does not. Nevertheless, the small positive RF from non O_3-depleting halocarbons of 0.02 W m^{-2} in 2011 (Myhre et al., 2013a) allows one to use the combination of GHG simulations to test for linearity. The sum of the ERFs from *piClim-CO2*, *piClim-CH4*, *piClim-N2O,* and *piClim-HC* is 2.93 ± 0.08 W m^{-2}, indicating that the ERFs add linearly and agree with the ERF from *piClim-GHG* (2.92 ± 0.04 W m^{-2}).

4.3 Aerosols and Aerosol Precursors

Figure 7a summarizes the results from the anthropogenic aerosol experiments, including a breakdown of the ERF into ~~instantaneous radiative forcing (~~IRF~~)~~ and ~~rapid adjustments (~~RAs~~)~~ for each of the anthropogenic aerosol experiments (*piClim-SO2*, *piClim-OC*, *piClim-BC*, and *piClim-aer*). The IRF was calculated using the double-call system where ~~aerosol radiation interactions (~~ari~~)~~ (i.e., scattering and absorption) are withdrawn from the second call to the radiation scheme, as in Ghan et al. (2012). The ~~rapid adjustments (~~RAs~~)~~ have then been derived as the residual between the ERF and IRF and include~~s~~ all aerosol-induced changes in cloud radiative effects. For completeness, Table 3 also summarises the contributions to the ERF from the SWcs', SW △CRE', LWcs', LW △CRE', NETcs', and NET △CRE' components quantified using Eqn. (8).

Uncertainties in aerosol ERF are driven partly by uncertainties in aerosol lifetime, spatial and temporal distributions, the historical change in aerosol loading, and the cloud response to aerosols (Bellouin et al., 2020). Other contributing factors include ~~are likely to be driven by~~ uncertainties in the PI aerosol state (Carslaw et al., 2013) as well as the oxidising capacity of the atmosphere (Karset et al., 2018). The additional interactive sources of natural aerosol in UKESM1 ~~for~~from marine ~~dimethyl sulphide (~~DMS~~)~~, terrestrial and marine biogenic ~~sources~~emissions, along with the inclusion of a fully interactive chemistry scheme (Archibald et al., ~~2019~~2020) are generally found to improve the evaluation of PD aerosol in UKESM1 (Mulcahy et al., 2020~~2019~~). This provides some confidence in the underlying physical processes driving the PI aerosol state in this model.

The ~~total~~ anthropogenic aerosol ERF evaluated from the "all" (*piClim-aer*) experiment is ~~-1.13~~ -1.09 ± 0.04 W m^{-2}, which is ~~almost~~ identical to the ERF of -1.10 W m^{-2} from the physical model HadGEM3-GA7.1 (Andrews et al., 2019) and ~~slightly~~ lower in magnitude than the ERF of -1.45 W m^{-2} derived from HadGEM3-GA7.1 with CMIP5 emissions (Mulcahy et al., 2018). The estimate fits well within the likely range of -1.60 to -0.65 W m^{-2} (16-84 % confidence level) provided by a recent major assessment of ~~PI to PD~~ aerosol ERF (Bellouin et al., 2020~~2019~~) and the -1.5 to -0.4 W m^{-2} likely range previously assessed by AR5 (Myhre et al., 2013a). The ~~annual mean~~ global distribution of aerosol ERF, IRF and aci are shown in Fig. 8. The ~~total~~ aerosol ERF (Fig. 8a) is negative over most regions that have a robust signal, and is strongest over the cloudy ocean regions of the Northern Hemisphere where the forcing is dominated by the RA term (Fig. 7a) driven by ~~aerosol-cloud interactions (~~aci~~)~~ (Fig. 8c). Some areas of Asia and North Africa have a positive ~~total~~ aerosol ERF due to BC-rich aerosol loadings that give locally positive aerosol IRF (Fig. 8b). However, globally, the IRF is rather small and negative (~~-0.16~~ -0.15 ± 0.01 W m^{-2}) (Fig. 7a) due to scattering by sulphate and OC that is partially offset by absorption from BC.

Figure 7: Aerosol ERF ~~at TOA for the period 1850 - 2014~~, broken down by species and process. (a) Results are from *piClim-BC*, *piClim-OC*, *piClim-SO2*, from summing those three experiments (Sum), and from *piClim-aer* (All aerosol). (b) ~~Total a~~Aerosol ERF decomposed into various contributing processes from the *piClim-aer* and accompanying sensitivity tests: IRF = instantaneous radiative forcing, RA = rapid adjustments, aci~~ACI~~ = aerosol-cloud interactions, RA_cs = clear-sky rapid adjustments (e.g., surface albedo, atmospheric temperature and water vapour). Units in W m^{-2}.

The SO_2 emissions are the largest individual contributor to the aerosol ERF and have a strong RA (aci) component leading to an SO_2 (equivalent to sulphate) ERF of ~~-1.45~~ -1.37 ± 0.03 W m^{-2} (Table 3, Fig. 7a). The IRF component of the sulphate ERF is ~~-0.48~~ -0.46 ± 0.01 W m^{-2}, which agrees well with the best estimate from AR5 (-0.40 ± 0.2 W m^{-2}) and from AEROCOM Phase II (-0.58 to -0.11 W m^{-2}) (Myhre et al., 2013b). The BC ERF ~~was~~is ~~+0.32~~ 0.37 ± 0.03 W m^{-2}, coming mostly from the IRF (~~0.38~~ 0.38 ± 0.01 W m^{-2}) and a small negative offset of ~~-0.06~~ -0.01 ± 0.02 W m^{-2} from the RA term~~rapid adjustments~~. As noted in Johnson et al. (2019), BC absorption leads to strong cloud adjustments but the SW and LW components of these almost cancel in HadGEM3-GA7.1 and UKESM1 (see Table 3). This contrasts with many other models where the combination of low cloud enhancements and reductions in upper-level clouds typically result in more substantial negative adjustments, making the BC ERF on average about half the magnitude of the IRF (Stjern et al., 2017). The BC ERF given by UKESM1 is however well within the range assessed by AR5 (0.05 -to 0.8 W m^{-2}), which took into consideration the possibility that BC emissions and/or absorption efficiency were underestimated in CMIP5 models (Bond et al., 2013). The anthropogenic emissions of BC were specified as 5 Tg yr^{-1} in CMIP5 (Year 2000 as PD) but have increased to 8 Tg yr^{-1} in CMIP6 (Year 2014 as PD). With CMIP5 emissions, the BC ERF from HadGEM3-GA7.1 was found to be 0.17 W m^{-2} (Johnson et al., 2019) and is comparable to direct BC forcing from other CMIP5 model estimates (Myhre et al., 2013a). It is worth noting that the aerosol absorption was in fairly good agreement with AERONET observations in HadGEM3-GA7.1 simulations that used the CMIP5 emission set (Mulcahy et al., 2018). The slightly higher CMIP6-based estimate of ~~0.32~~ 0.37 ± 0.03 W m^{-2} provided in the present study could therefore be an overestimate, although this is difficult to judge given the uncertainties in comparing models

with absorption measurements. ~~It is also worth noting that BC forcing can have proportionately larger impacts on surface radiation, precipitation and regional circulation due to the localised absorption of solar radiation in the atmosphere.~~

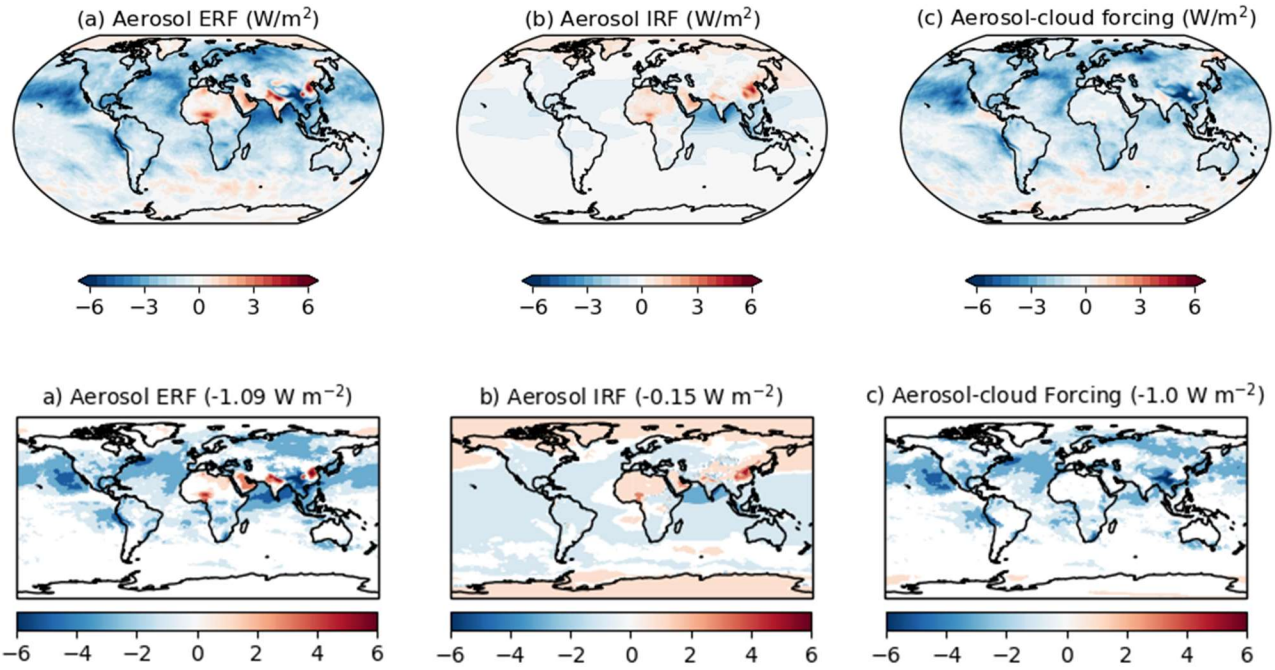


Figure 8: ~~Multi-annual mean changes~~Changes in TOA net radiation from piClim-aer (all anthropogenic aerosols), including (a) ~~T~~the ERF, (b) the IRF, (c) aerosol-cloud forcing due to aci~~erosol-cloud interactions~~ and the semi-direct aerosol effect. Global mean values are shown in brackets. Regions where the ERF, IRF, and the aerosol-cloud forcing are outside the 95 % confidence level are masked out in white. Units in W m^{-2}.

The ~~PD~~ OC ERF ~~relative to PI~~ is ~~-0.27~~ -0.22 ± 0.04 W m^{-2}, with ~~-0.16~~ -0.14 ± 0.01 W m^{-2} from the IRF and ~~-0.11~~ -0.07 ± 0.02 Wm^{-2} from RAs (Fig. 7a, Table 3). The OC IRF estimate agrees fairly well with AR5 that assessed the RF of primary and secondary OC to be -0.12 W m^{-2} (-0.4 to ~~+~~0.1 W m^{-2}). Note that OC is non-absorbing in UKESM1 and, as such, neglects the role of brown carbon. This potentially misses a small positive contribution to the aerosol forcing (e.g., Feng et al., 2013), although biomass burning emissions are the dominant global source of brown carbon and these do not change significantly from 1850 to 2014 (Saleh et al., 2014). The contribution of brown carbon to aerosol ERF is, therefore, likely to be small compared to the overall uncertainty in modelling aerosol absorption by BC (Bond et al., 2013). Nitrate aerosols are also not represented in HadGEM3-GA7.1 (Mulcahy et al., 2018; Walters et al., 2019) or UKESM1 (Sellar et al., 2019; Mulcahy et al., 2020~~2019~~) so the ERF associated with ammonia (NH_3) emissions could not be evaluated here. The ~~total~~ aerosol ERF in UKESM1 would presumably be more strongly negative if the role of nitrate aerosol was included (e.g., Bellouin et al., 2011) and this should be borne in mind when making comparisons with other models or estimates based on observational constraints.

The "all" aerosol forcing experiment combines the increases in BC, OC and SO_2 emissions together in one simulation, and interestingly, the ERF is ~~0.27~~ 0.13 W m^{-2} weaker (less negative) than the sum of the ERFs from the experiments that perturb those emissions separately. The main reasons for this lack of linearity are: (i) Cloud droplet numbers do not increase linearly with aerosol loading (e.g., Jones et al., 1994) and begin to saturate in the "all" experiment (*piClim-aer*), which means OC and BC emissions no longer contribute significantly to aci once co-emitted with year-2014 levels of SO_2. (ii) The absorption of upwelling ~~shortwave~~ SW radiation by BC is enhanced by increases in aerosol scattering and cloud brightness due to aci, making the IRF less negative in the "all" experiment. And (iii) Internal mixing creates an interdependency between different aerosol sources, meaning that the aerosol size distributions, optical scattering efficiency and hygroscopicity evolve differently depending on the absolute and relative abundance of different mass components (sulphate, OC, BC and sea salt mass) and differing rates of new particle production via primary emission or nucleation.

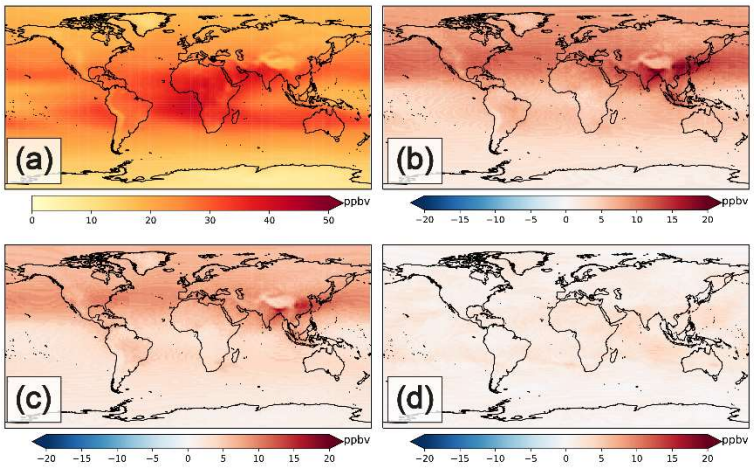
To further understand which processes contribute most to ~~total~~the aerosol ERF, a series of additional ~~sensitivity~~ control and perturbation experiments were conducted with aci processes selectively disabled. In these tests, the cloud droplet number concentrations (CDNCs) used for the calculation of cloud droplet effective radius (Reff) and/or autoconversion were prescribed via a 3D monthly-mean PI climatology constructed from the final 30 years of the *piClim-control* simulation. The resulting ~~global mean~~ ERFs are summarized in Fig. 7b. In one pair of simulations, CDNCs were prescribed for the Reff calculation to disable the so-called Twomey effect (Twomey, 1977). By comparison with the main *piClim-aer*/*piClim-control* experiment pair, this indicated a Twomey effect (ACI_Twomey) of ~~-0.73~~ -0.70 W m^{-2}. A similar pair with CDNCs prescribed only for the autoconversion process led to an estimate of ~~-0.34~~ -0.31 W m^{-2} for the cloud lifetime effect (ACI_lifetime) (Albrecht, 1989). By prescribing CDNCs for both Reff and autoconversion both microphysical aci processes are disabled and only ~~aerosol-radiation interactions (~~ari~~)~~ are included. A pair of simulations with this setup provided an estimate for ERF_{ari} of ~~-0.14~~ -0.15 W m^{-2}. The change in CRE in the ari-only experiment was only ~~+0.04~~ 0.02 W m^{-2}, which is not statistically significant at the 95 % confidence level and indicates that the semi-direct aerosol effect is small or approximately neutral in this model. To complete the breakdown, the method in Ghan (2013) was ~~further used~~ applied to the main *piClim-aer*/*piClim-control* experiment pair to derive the contribution from changes in "clean" (aerosol-free) CS radiation (RA_cs). This term was found to be ~~+0.05~~ 0.05 W m^{-2}, and arises due to changes in surface albedo and atmospheric temperature and humidity. This overall breakdown suggests an aerosol-cloud forcing of ~~-1.03~~ -0.99 W m^{-2}, with a roughly 70/30 split between the Twomey effect and aerosol effects on cloud cover and water content (lifetime ~~&~~and semi-direct effects).

The estimated aerosol-cloud forcing sits well within the 90 % likelihood ranges assessed by AR5 (-1.2 to 0.0 W m^{-2}) and Bellouin et al. (2020~~2019~~) (-2.0 to -0.35 W m^{-2}), although these broad ranges reflect the large uncertainties involved in constraining global estimates with observations (e.g., Ghan et al., 2016). At present there are no definitive constraints on the proportionate contributions that the Twomey and other aerosol-cloud effects make towards the ~~total~~ aerosol-cloud forcing, but recent observational evidence (Malavelle et al., 2017; Toll et al., 2017) supports the supposition that the Twomey effect is the

dominant process and that some models overestimate cloud lifetime effects. Toll et al. (2017) indicated that HadGEM3 can indeed overestimate the cloud lifetime effect in marine stratocumulus, whereas Mallavelle et al. (2017) found HadGEM3 to correctly simulate only weak cloud lifetime effects for mixed cloud regimes over the North Atlantic. The UKESM1 estimate of ERF_{ari} of ~~-0.14~~ -0.15 W m^{-2} is within the 5-95 % confidence range (-0.45 ~~+/-~~ ± 0.5 W m^{-2}) from AR5 but slightly weaker than the range of values estimated by Bellouin et al. (2020~~2019~~) (-0.60 to -0.25 W m^{-2}). Possible reasons include the lack of nitrate, the relatively strong BC forcing compared to CMIP5 models, and a slight underestimation of aerosol optical depth (AOD) in PD simulations relative to some satellite products (Mulcahy et al., 2020~~2019~~).

4.4 ~~Tropospheric ozone~~Ozone (O_3) precursor (VOC, CO, and NOx) gases

The ~~global mean~~ ERF from emissions of ~~tropospheric ozone (~~O_3~~)~~ precursors (VOC, CO, and NOx) excluding CH_4 is weakly positive (~~+0.15~~ 0.21 ± 0.04 W m^{-2}; Table 3; Figure 3), although spatially heterogeneous, with sparse regions of the globe showing a statistically significant ERF (Figure 3)~~regions of positive and negative forcing~~. ~~Tropospheric~~ O_3 precursor emissions affect the ERF both through changes in ~~to tropospheric~~ O_3, a GHG, and by changing tropospheric oxidants such as the hydroxyl (OH) radical, which in turn affect aerosols (Karset et al., 2018) and CH_4. Here, we explore the composition and forcings resulting from O_3 precursor emissions changes ~~between the PI and PD~~ by comparing the *piClim-O3* simulation with *piClim-control* (Table 1) and the separate effects of ~~nitrogen oxide (~~NOx~~)~~ and ~~volatile organic compound (~~VOC~~)~~/~~carbon monoxide (~~CO~~)~~ emissions changes using the *piClim-NOx* and *piClim-VOC* simulations, respectively.



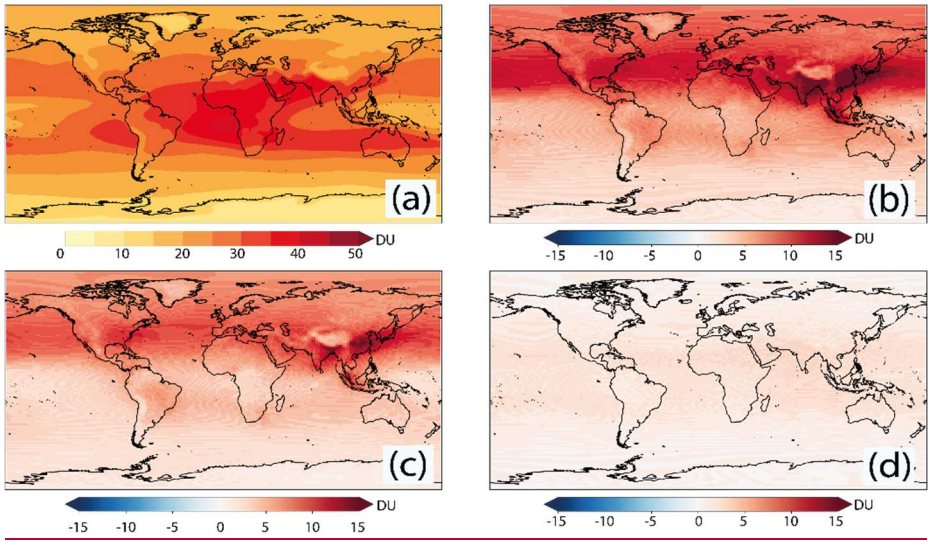


Figure 9: Global distributions ~~of the multi-annual distributions~~ of tropospheric O_3 column (a) in *piClim-control*. Differences with respect to *piClim-control* from *piClim-O3*, *piClim-NOx* and *piClim-VOC* can be seen in panels b), c) and d), respectively. Units in Dobson Units (DU).

4.4.1 Tropospheric Ozone (O_3) Changes

The ~~multi-annual mean~~ tropospheric O_3 column for the *piClim-control* simulation, and tropospheric O_3 column differences between *piClim-control* and the *piClim-O3*, *piClim-NOx* and *piClim-VOC* simulations are shown in Fig. 9. O_3 levels in the troposphere are the result of competing production and loss processes. Production occurs in the presence of NOx and VOC/CO, with most regions of the troposphere being NOx-limited. PI tropospheric column O_3 values (Fig. 9a) show maxima over central Africa and the ~~South~~ Eastern tropical Pacific, the result of relatively large emissions of NOx from soil, biomass burning and lightning in this region, and minima over regions remote from NOx sources, such as the Western equatorial Pacific (where O_3 loss is efficient)~~,~~ ~~(~~ consistent with Young et al.~~,~~ ~~(~~2018). Tropospheric column O_3 ~~columns are~~is also low over regions of high surface elevation where the atmospheric column is shallower.

Tropospheric O_3 column values increase when the model is perturbed with the larger PD O_3 precursor emissions in *piClim-O3*, with the largest increases in the northern hemisphere, particularly in southern and eastern Asia. The tropospheric O_3 burden increases from ~~280.1~~ 280.9 Tg in *piClim-control* to ~~356.1~~355.5 Tg in *piClim-O3*. As can be seen in Fig. 9, the dominant driver of these changes ~~is~~are increased NOx emissions, and there is a similar pattern of changes in the *piClim-O3* and *piClim-NOx* simulations, although the change in O_3 is smaller in the latter; the burden in *piClim-NOx* is ~~335.7~~337.5 Tg. Small increases in O_3 are ~~observed~~modelled in *piClim-VOC*, with some hotspots in regions such as South East Asia and the O_3 burden increases to ~~297.2~~296.8 Tg.

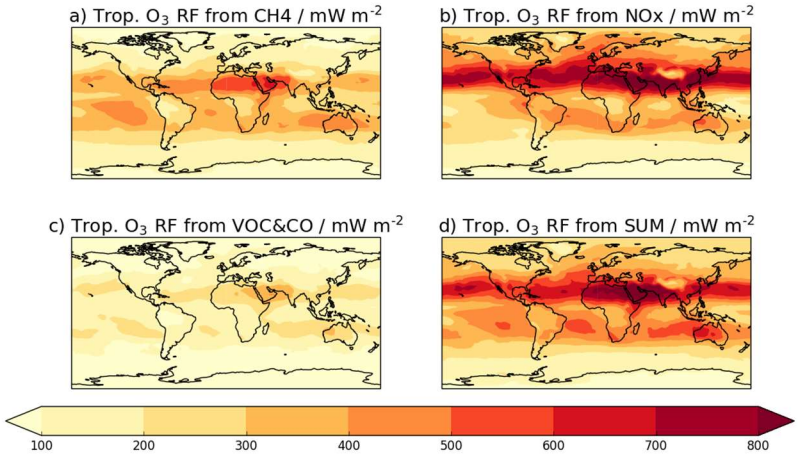
The tropospheric O_3 ~~column~~ burden difference between *piClim-O3* and *piClim-control* of ~~76.0~~74.6 Tg is slightly larger than the sum of the individual O_3 burden changes in *piClim-NOx* (~~55.6~~56.6 Tg) and *piClim-VOC* (~~17.1~~15.9 Tg) which total ~~72.7~~72.5 Tg. However, these differences due to the non-linear nature of tropospheric chemistry are small, on the order of < ~~10~~5 %. Similar behaviour is seen in the patterns of the tropospheric O_3 column difference between *piClim-control* and other experiments.

4.4.2 Tropospheric ~~Ozone (~~O_3~~)~~ Stratospherically-Adjusted Radiative Forcing (SARF)

As seen from earlier sections, it can be difficult to compare ERFs estimated from these experiments against estimates of SARF (e.g., Stevenson et al., 2013; Myhre et al., 2013a), due to the inclusion of indirect forcings and/or ~~rapid adjustments~~RAs other than the stratospheric temperature adjustment although Shindell et al. (2013b) and Skeie et al. (2020) noted that for O_3, SARF and ERF estimates are comparable. Nevertheless, in order to compare against Stevenson et al. (2013), we estimate the tropospheric O_3 SARF ~~for~~from the *piClim-CH4*, *piClim-VOC*, and the *piClim-NOx* ~~tropospheric O_3 precursor~~ experiments by adopting a radiative kernel approach (e.g., Soden et al., 2008). This involves applying the tropospheric O_3 radiative kernel from Rap et al. (2015) to the diagnosed change in tropospheric O_3 (using the 150 ppbv O_3 isoline in *piClim-control* as a tropospheric mask, as used in Young et al., 2013, Stevenson et al., 2013, and Rap et al., 2015) to calculate a ~~stratospherically-adjusted radiative forcing (~~SARF~~) or RF~~. While the ERF captures changes in radiative fluxes at the TOA due to whole-atmosphere responses, here, we mask off the stratosphere and focus *solely* on the tropospheric O_3 response. In this way, we can directly compare against the best estimate of the 1850-2010 tropospheric O_3 SARF of 364 mW m^{-2} by Stevenson et al. (2013) and quantify the contribution of different ~~tropospheric~~ O_3 precursors (including ~~methane~~CH_4) to the ~~PI-to-PD~~ change in tropospheric O_3 and its SARF. We can also compare with more recent estimates from ~~Checa-Garcier~~ Checa-Garcia et al. (2018) and Yeung et al. (2019).

Figure 10 shows the global distribution of the ~~PD~~ tropospheric O_3 SARF from *piClim-CH4*, *piClim-NOx*, and *piClim-VOC* experiments and their sum using the kernel approach. It shows that the tropospheric O_3 SARF is strongest over the northern hemisphere sub-tropics and weakest over the southern hemisphere high latitudes. The strongest SARF occurs in regions of warm surface temperatures and high albedo, coinciding with the largest tropospheric O_3 change (Shindell et al., 2013a). As was the case in Stevenson et al. (2013), the ~~forcing~~ SARF is weaker over regions of high altitude (e.g., Tibetan Plateau) due to there being less O_3 column aloft to absorb in the ~~longwave (~~LW~~)~~. ~~On a global annual mean basis, the~~The tropospheric O_3 SARF from the kernel approach is ~~407~~ 414 mW m^{-2}~~, higher than the best estimate of Stevenson et al. (2013) by approximately 11 %. However, the Edwards and Slingo (1996) radiation scheme, which was used to generate the kernel, can lead to forcing estimates that are higher than other comparable radiation schemes (Stevenson et al., 2013). Other potential differences are due to experimental setup. The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP; Lamarque et al., 2013) simulations were timeslices and included historical climate change although Stevenson et al. (2013) found that to have a minimal impact on the RF~~. ~~More importantly~~However, the increase in tropospheric O_3 ~~at the PD~~ (and its resulting SARF)

~~were~~are offset by decreases due to ~~the downward transport of O_3-depleted stratospheric air~~ ODSs (e.g., Søvde et al., 2011; Søvde et al., 2012; Shindell et al., 2013a). Applying the kernel method to the diagnosed decrease in tropospheric O_3 from the *piClim-HC* experiment (Sect. ~~3.2.3~~4.2.3), we find a SARF offset of ~~-101~~ -101 mW m^{-2}~~.~~. Although larger in magnitude by nearly a factor of 2 than the estimates from Søvde et al. (2011; 2012) and Shindell et al. (2013a) due to the strong O_3 depletion in UKESM1 (Keeble et al., 2019; Skeie et al., 2020; Morgenstern et al., 2020), ~~thus reducing~~ it reduces our original estimate to ~~306~~ 313 mW m^{-2}. This revised estimate is within the 30 % uncertainty of ~~15 % lower than~~ the Stevenson et al. (2013) ~~best~~ estimate, albeit lower than their central estimate by 13 %. ~~However, it~~It is also consistent with a number of other estimates: the CMIP6 historical O_3 dataset (312 mW m^{-2} from Checa-Garcia et al., 2018), a recent study in which observational isotopic data was used as a constraint on historical increases in tropospheric O_3 (330 mW m^{-2} derived from the GEOSChem model in Yeung et al., 2019), and a parametric model based on multi-model source-receptor relationships (290 ± 3 mW m^{-2} from Turnock et al., 2019). However, in masking off the stratosphere, the UKESM1 estimate for the O_3 SARF attributable to O_3 precursors is likely to be underestimated. For example, Søvde et al. (2011; 2012) estimate that approximately 15 % of the O_3 response from changes in CH_4 and other O_3 precursors may be in the stratosphere, and hence not considered here.



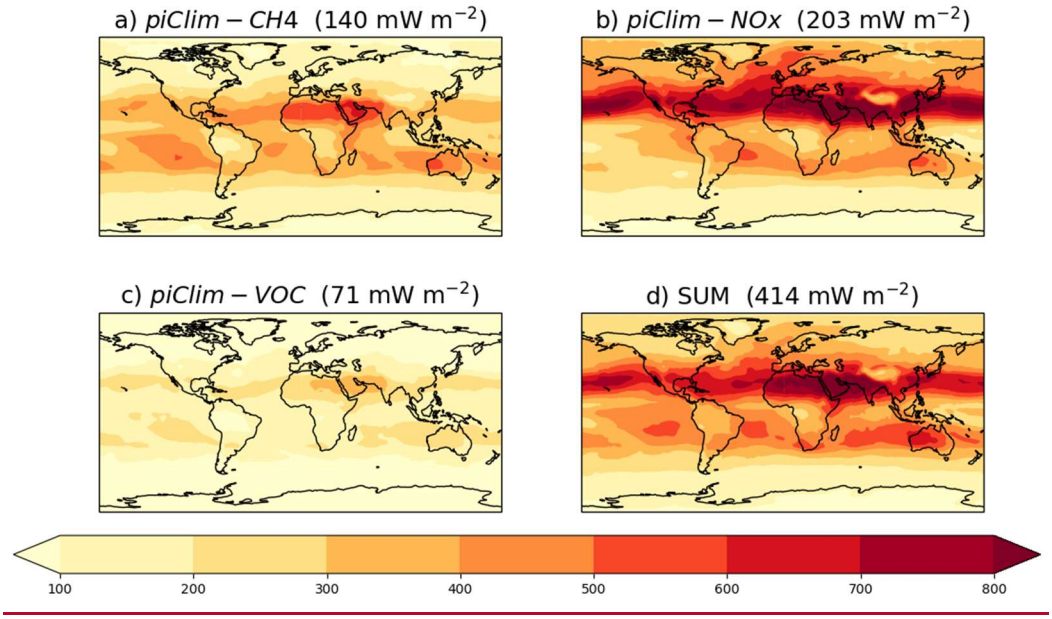


Figure 10: Global distribution of ~~PD~~ tropospheric O_3 SARF diagnosed from a) *piclim-CH4*, b) *piClim-NOX*, c) *piClim-VOC*, and d) their sum relative to *piClim-control*, based on the diagnosed change in tropospheric O_3 and the tropospheric O_3 radiative kernel of Rap et al. (2015). Global mean values are included. Units in mW m^{-2}.

In considering the attribution of the tropospheric O_3 SARF to its precursors, initial estimates (Fig. 10; Table 5) indicate that it is predominantly NOx driven (~~48~~ 49 %), followed by CH_4 (~~34~~ 34 %), with the smallest contribution from VOCs and CO (~~18~~ 17 %). This is qualitatively consistent with Stevenson et al. (2013). However, as outlined in detail in that paper, these estimates do not take account of potential CH_4 (and O_3) changes that would occur were these experiments driven by CH_4 emissions rather than concentrations (e.g., Shindell et al., 2005). Taking the same approach as Stevenson et al. (2013), we calculate an equilibrium CH_4 concentration for the perturbation experiments, using the total CH_4 lifetime relative to *piClim-control* based on the following (Fiore et al., 2009):

$$[\mathrm{CH_4}]_{piClim-X} = [\mathrm{CH_4}]_{piClim-control} * \left(\frac{\tau_{piClim-X}}{\tau_{piClim-co}}\right)^{f}, \qquad (9)$$

where $[\mathrm{CH_4}]_{piClim-X}$ is the global mean equilibrium CH_4 concentration~~s~~ in the *piClim-X* experiments, $[\mathrm{CH_4}]_{piClim-control}$ is the prescribed global mean CH_4 concentration in *piClim-control*, $\tau_{piClim-contr}$ and $\tau_{piClim-X}$ are the whole-atmosphere CH_4 lifetimes in the *piClim-control* and *piClim-X* perturbation experiments, respectively. The CH_4-OH feedback factor (Prather, 1996) is denoted by *f* and is defined as:

$$f = \frac{1}{1-s}, \quad (10)$$

where s, called the sensitivity coefficient, is calculated from the following:

$$s = \frac{\delta ln\, \tau}{\delta\, ln\, [CH_4]}. \quad (11)$$

Using the whole-atmosphere CH_4 burden and its removal by OH, the whole-atmosphere CH_4 lifetime is ~~8.1~~ 8.1 and ~~9.8~~ 9.8 years in *piClim-control* and *piClim-CH4,* respectively, when adjusted for stratospheric removal (~~160~~120 yr lifetime) and soil uptake (~~120~~160 yr lifetime). From Eqns. (10) and (11), the UKESM1 feedback factor f is ~~1.28~~ 1.28, consistent with the range of other estimates (Prather et al., 2001; Shindell et al., 2005; Fiore et al., 2009; Stevenson et al., 2013; Voulgarakis et al., 2013; Turnock et al., 2018) and within 5 % of the observationally constrained best estimate of 1.34 (Holmes et al., 2013). Then, using the equilibrium minus prescribed difference ~~between the global mean prescribed and equilibrium~~ in surface CH_4 concentrations (Table 5), we calculate the additional tropospheric O_3 ~~forcing~~SARF in the *piClim-CH4*, *piClim-NOx* and *piClim-VOC* experiments by applying the tropospheric O_3 radiative kernel (Rap et al., 2015) to the O_3 response derived from scaling the (*piClim-CH4* minus *piClim-control*) O_3 response based on the relationship in Turnock et al. (2018); this additional contribution to the tropospheric O_3 SARF is shown in Table 5.

Changing from a concentration-based perspective to an emissions-based view increases the tropospheric O_3 SARF ~~between PI and PD~~ from ~~407~~ 414 to ~~461~~ 469 mW m^{-2} (an increase of ~~13~~ 13 %), which agrees better with the central estimate from Stevenson et al. (2013) once the offset by ~~O_3 depletion~~ODSs is accounted for (~~-101~~ -101 mW m^{-2}). It also changes the relative contributions of the different ~~tropospheric~~ O_3 precursors. The contribution of CH_4 now dominates (~~45~~ 45 %), with NOx playing a smaller role (~~36~~ 37 %), while the contribution from VOC/CO emission increases is relatively unchanged (~~19~~ 18 %). These emissions-based contributions are well within the spread of estimates from Stevenson et al. (2013) who quantified contributions from 6 of the ACCMIP models: CH_4 (44 ± 12 %), NOx (31 ± 9 %), and VOC/CO (25 ± 3 %) although there are some differences with Shindell et al. (2005) and Shindell et al. (2009). In both Shindell et al. studies, the contribution from NOx is lower at 15 ± 8 and 11 %. The difference may be due to the strong sensitivity of the CH_4 lifetime to NOx in the GISS model compared with other models (Wild et al., ~~2019~~2020) and/or could be due to differences in VOC chemistry; Archibald et al. (2010) showed that the response of OH to increasing NOx strongly depends on the treatment of VOC chemistry. Nevertheless, this approach demonstrates the importance of an emissions-based view of climate forcing and is more directly relevant to policy makers than a concentration-based view (Shindell et al., 2005; Shindell et al., 2009).

Pair	Trop. O_3 SARF / mW m^{-2}	~~Equilibrium~~ ΔCH_4 concentration / ppbv	Additional Trop. O_3 SARF from ΔCH_4 / mW m^{-2}	Total Trop. O_3 SARF / mW m^{-2}
piClim-NOx minus *piClim-control*	~~196~~ 203	~~563.4~~ -246.2	~~31~~ -31	~~165~~ 172
piClim-VOC minus *piClim-control*	~~73~~ 71	~~927.8~~ 123.2	~~15~~ 16	~~88~~ 87
piClim-CH4 minus *piClim-control*	~~138~~ 140	~~2364.9~~ 533.0	~~70~~ 70	~~208~~ 210
SUM minus *piClim-Control*	~~407~~ 414	N/A	N/A	~~461~~ 469

~~**Table 5: Equilibrium CH_4 concentrations in the different perturbation experiments (*piClim-NOx*, *piClim-VOC*, and *piClim-CH4*) relative to the pre-industrial control (*piClim-control*), and the resulting additional tropospheric O_3 radiative forcing.**~~

Table 5: Contribution to the tropospheric O_3 SARF from the different perturbation experiments (*piClim-NOx*, *piClim-VOC*, and *piClim-CH4*) relative to the pre-industrial control (*piClim-control*). Also shown is the absolute difference between the equilibrium and prescribed CH_4 concentrations in the different experiments and the resulting additional contribution to the total tropospheric O_3 SARF from the CH_4-driven response in O_3.

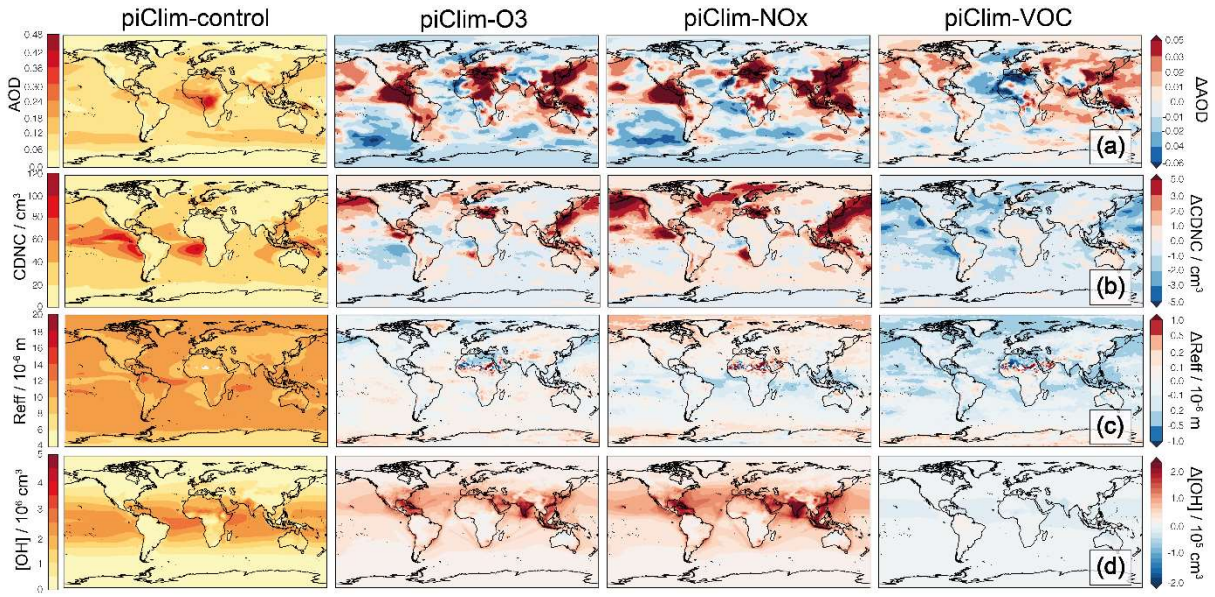
4.4.3 ERF: Role of Other Oxidants and Aerosols

In addition to O_3 being a GHG, it is also important for secondary aerosol formation along with other oxidants: the ~~hydroxyl (~~OH~~)~~ radical, hydrogen peroxide (H_2O_2) and the nitrate (NO_3) radical. The OH radical in UKESM1 (Sellar et al., 2019) is involved in aerosol nucleation of gas-phase sulphuric acid (H_2SO_4) via the reaction with sulphur dioxide (SO_2), leading to new particle formation (Mulcahy et al., 20202~~019~~). O_3 and H_2O_2 are important for SO_2 oxidation in cloud and aerosol droplets, creating sulphate aerosol mass but not number. Likewise, oxidation of monoterpenes by O_3, OH, and NO_3 determines the rate of formation of secondary organic aerosol (SOA) in UKESM1 (Kelly et al., 2018; Mulcahy et al., 20202~~019~~) although it ~~doesn't~~ does not lead to new particle formation. Thus, sulphate aerosol alone, through changes in O_3, OH, and H_2O_2, has potentially important impacts on cloud and aerosol radiative properties. Indeed, Karset et al. (2018) found that ~~PD to PI~~ oxidant changes in the CAM5.3-Oslo model alter the relative importance of different chemical reactions, leading to changes in aerosol size distribution, cloud condensation nuclei (CCN), and the ~~PD~~ aerosol ERF.

Figure 11 (bottom row) shows a global distribution of OH at 1 km altitude in *piClim-control* and changes in the perturbation experiments relative to *piClim-control*. The *piClim-control* experiment shows an OH maximum in the equatorial humid regions, where photolytic production of OH from excited oxygen atoms (O^1D) and water vapour (H_2O) is at a maximum. In *piClim-O3*, OH increases throughout the northern hemisphere due to increases in O_3, which is the precursor of O^1D. These

increases in OH are driven largely by increases in NOx. However, the *piClim-VOC* experiment shows the opposite behaviour - while VOC and ~~carbon monoxide (~~CO~~)~~ emissions increases serve to increase O_3, they also remove OH via direct reaction with OH. This latter effect outweighs the small increase in O_3 in *piClim-VOC* and there are ~~large~~ decreases in OH throughout the troposphere.

When OH is lower, we anticipate a decrease in the number of CCN, a decrease in ~~cloud droplet number concentration (~~CDNC~~)~~, leading to larger cloud droplets (Twomey, 1977), and an increase in ~~the effective radius of cloud droplets (~~Reff~~)~~. The middle rows of Fig. 11 show these effects at work. In *piClim-control*, the distribution of CDNC shows large values in equatorial regions, in regions of continental outflow and regions of deep convection. Large increases in CDNC are seen in *piClim-NOx*, with large decreases in *piClim-VOC*. These results reflect the changes in OH in these experiments - increases in OH lead to increases in CDNC, and vice versa, but it should be noted that the effect of OH on CDNC is seen over a larger region downwind, particularly in East Asia and over the North Atlantic. The impact of NOx emissions appears to dominate, given the similarity between *piClim-O3* and *piClim-NOx*~~, but the eastern Pacific is a region where all experiments show a decrease in CDNC~~.



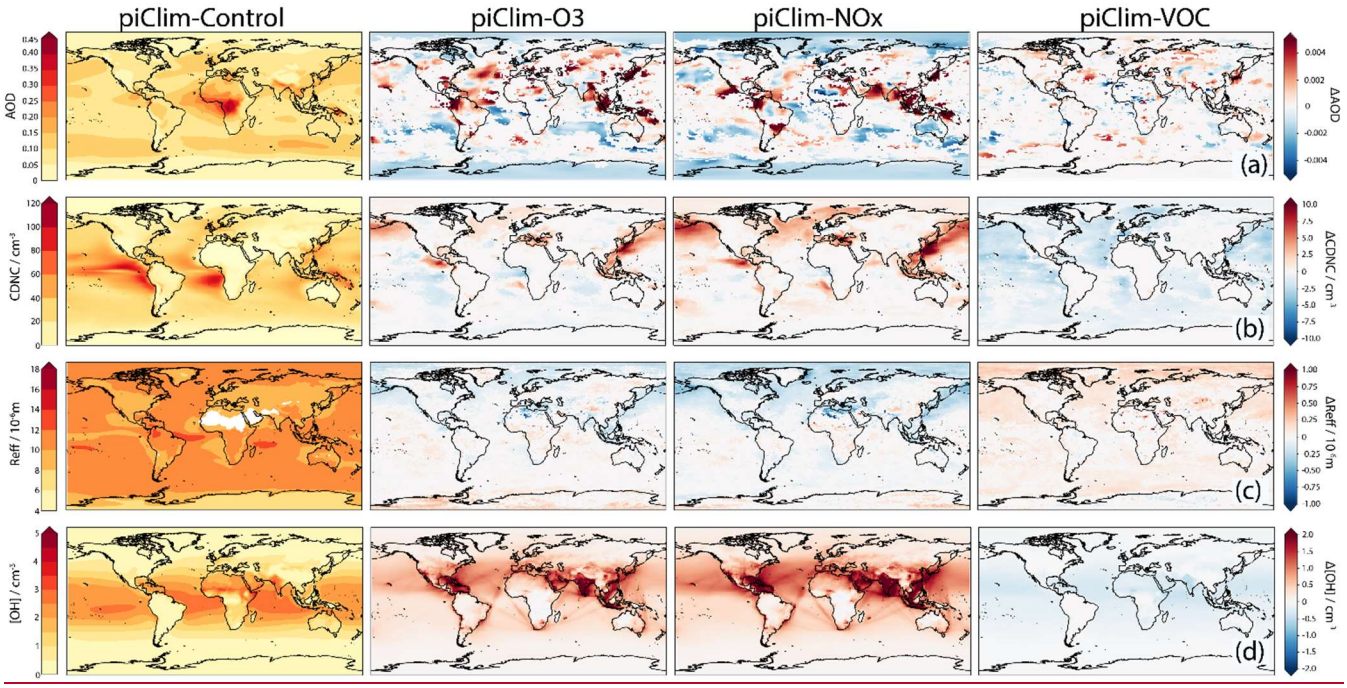


Figure 11: Global distributions of ~~the multi-annual distributions of~~ AOD at 550 nm (row (a); left column), cloud droplet number concentration (CDNC; cm^{-3}) at 1 km altitude (row (b); left column), cloud droplet effective radius (Reff; µm) at 1 km altitude (row (c); left column) and OH number concentration (cm^{-3}) at 1 km altitude (row (d); left column) in *piClim-control*. Differences with respect to *piClim-control* from *piClim-O3*, *piClim-NOx* and *piClim-VOC* can be seen in the second, third and right hand columns, respectively.

While tropospheric column O_3 increases in both *piClim-NOx* and *piClim-VOC*, leading to positive contributions to the ~~PD~~ tropospheric O_3 SARF (Sect. ~~3.4.2~~4.4.2), the ~~global mean PD~~ ERFs are~~is~~ very different between the two simulations (~~+0.21~~ 0.33 ± 0.04 W m^{-2} for *piClim-VOC* and ~~-0.08~~ 0.03 ± 0.04 W m^{-2} for *piClim-NOx* relative to *piClim-control*). Further, the spatial pattern of the ERF does not match those regions of largest O_3 changes. Instead, the dominant driver of the ERF differences ~~is~~are changes to OH and the subsequent impacts on aerosol particle formation and clouds. In particular, the positive tropospheric O_3 ~~forcing~~SARF in *piClim-NOx* (~~+0.2~~ 0.2 W m^{-2}) appears to be ~~outweighed~~ nearly completely offset by a negative aerosol forcing, largely from ~~aerosol-cloud interactions (~~aci~~)~~ driven by changes in oxidants and aerosol nucleation; the contribution to the global mean ERF from ~~aerosol-radiation interactions (~~ari~~)~~ is only ~~-0.02~~ -0.03 ± 0.01 W m^{-2}, despite the strong regional changes in ~~aerosol optical depth (~~AOD at 550 nm; (Fig. 11; top row). Similarly, the ~~forcing~~ aerosol IRF in *piClim-VOC* ~~due to ari~~ is negligible (~~0.00~~ less than 0.01 ± 0.01 W m^{-2}). However, the forcing in *piClim-VOC* due to aci, particularly from the SW ΔCRE' component (Table 3), enhances the positive tropospheric O_3 ~~forcing~~SARF (~~+0.07~~ 0.07 W m^{-2}), leading to an ~~global mean~~ ERF of ~~+0.21 ± 0.04~~ 0.33 ± 0.03 W m^{-2}.

A negative ~~PD~~ SARF~~forcing~~ attributable to NOx emissions has been found in other studies (Shindell et al., 2009; Collins et al., 2010) from a balance between the direct O_3 response (positive SARF), the NOx-driven CH_4 response (negative SARF), and the subsequent O_3 response to CH_4 changes (negative SARF). Inclusion of ari and aci from sulphate and nitrate aerosol further increases the magnitude of the net negative forcing or cooling (Shindell et al., 2009; Collins et al., 2010). However, a study by Fry et al. (2012) found that the chemistry response is sensitive to the location of the emissions, with so large an uncertainty that it is difficult to determine whether NOx emissions cause a warming or cooling. Indeed, other indirect effects such as NOx deposition to the terrestrial biosphere leading to fertilisation (Collins et al., 2010) and/or NOx-driven O_3 damage (Sitch et al., 2007; Collins et al., 2010) increase the uncertainty further through changes in ~~carbon dioxide (~~CO_2~~)~~. Thornhill et al. (2020~~2019~~) also shows that the ERF from ~~PI-to-PD~~ changes in NOx emissions among the AerChemMIP models differ in both sign and magnitude. Here, the longer time-scale CH_4 response to NOx emissions (Collins et al., 2010) is constrained, nitrate aerosol is neglected (Sellar et al., 2019; Mulcahy et al., 2020~~2019~~), and CO_2 is concentration-driven (Sellar et al., 2019). Nevertheless, in UKESM1, the negative forcing due to aci from sulphate aerosol ~~outweighs~~offsets the positive NOx-driven O_3 SARF, leading to a ~~net negative~~ negligible ERF overall.

Previous work has also found a positive ~~PD~~ SARF ~~or RF~~ from VOC and CO emissions, due to the combined indirect forcings by O_3 and CH_4 (Shindell et al., 2005; Forster et al., 2007); they estimate a global mean SARF of ~~+~~0.21 ± 0.10 W m^{-2} at the PD (Year 1998) relative to the PI (Year 1750) period. ~~As~~ Via the same mechanisms as was the case with NOx, the magnitude of this SARF increases (to ~~+~~0.25 ± 0.04 W m^{-2} for the year 2000) when additional indirect forcings from sulphate, nitrate, and CO_2 are included (Shindell et al., 2009). More recently, Stevenson et al. (2013) found the SARF from VOC/CO emissions to be marginally higher, at ~~+~~0.29 W m^{-2}, excluding aerosols, with contributions of 0.09, 0.08, and 0.12 W m^{-2} from O_3, CH_4, and CO_2, respectively. The SARF contribution from O_3 alone quantified here (~~88~~ 87 mW m^{-2}; Table 5) is consistent with the Stevenson et al. (2013) estimate. Despite excluding the longer-term CH_4 and CO_2 responses ~~and CO_2~~, the ~~total global mean~~ ERF from VOC/CO emissions of ~~+0.21 ± 0.04~~ 0.33 ± 0.03 W m^{-2} is ~~consistent with~~ higher than previous estimates of SARF due to the additional positive contribution from aci driven by OH changes. However, Fry et al. (2014) found that the SARF from VOCs is sensitive to the location of emissions and could influence the strength of the contribution from aerosols. Interestingly, other AerChemMIP models show a negative ERF from VOC and CO emissions (Thornhill et al., 2020~~2019~~); these differences in sign of the ERF warrant further investigation.

~~Despite~~ As a result of the very different chemical response between NOx and VOC/CO emissions, both in terms of the magnitude of the O_3 changes and the different impacts on OH, aerosols and clouds, a comparison of the ~~global~~ ERFs ~~changes~~ (Table 3) indicates that the ~~PD~~ ERF from *piClim-O3* is not a linear combination of that from *piClim-NOx* and *piClim-VOC* for the NET, ~~clear-sky (~~CS~~)~~ and the ~~cloud radiative effect (~~CRE~~)~~ components. In particular, there are differences in the LWcs' and SW ΔCRE' components (Table 3). ~~Nevertheless, the~~These results clearly suggest that Earth System (ES) interactions, particularly chemistry-aerosol coupling, can strongly affect estimates of climate forcing. Here, these interactions alter the ~~PD~~

ERF from ~~tropospheric~~ O_3 precursor emissions, while other studies (e.g., Shindell et al., 2009; Karset et al., 2018) show that they also affect estimates of ~~PD~~ anthropogenic aerosol forcing.

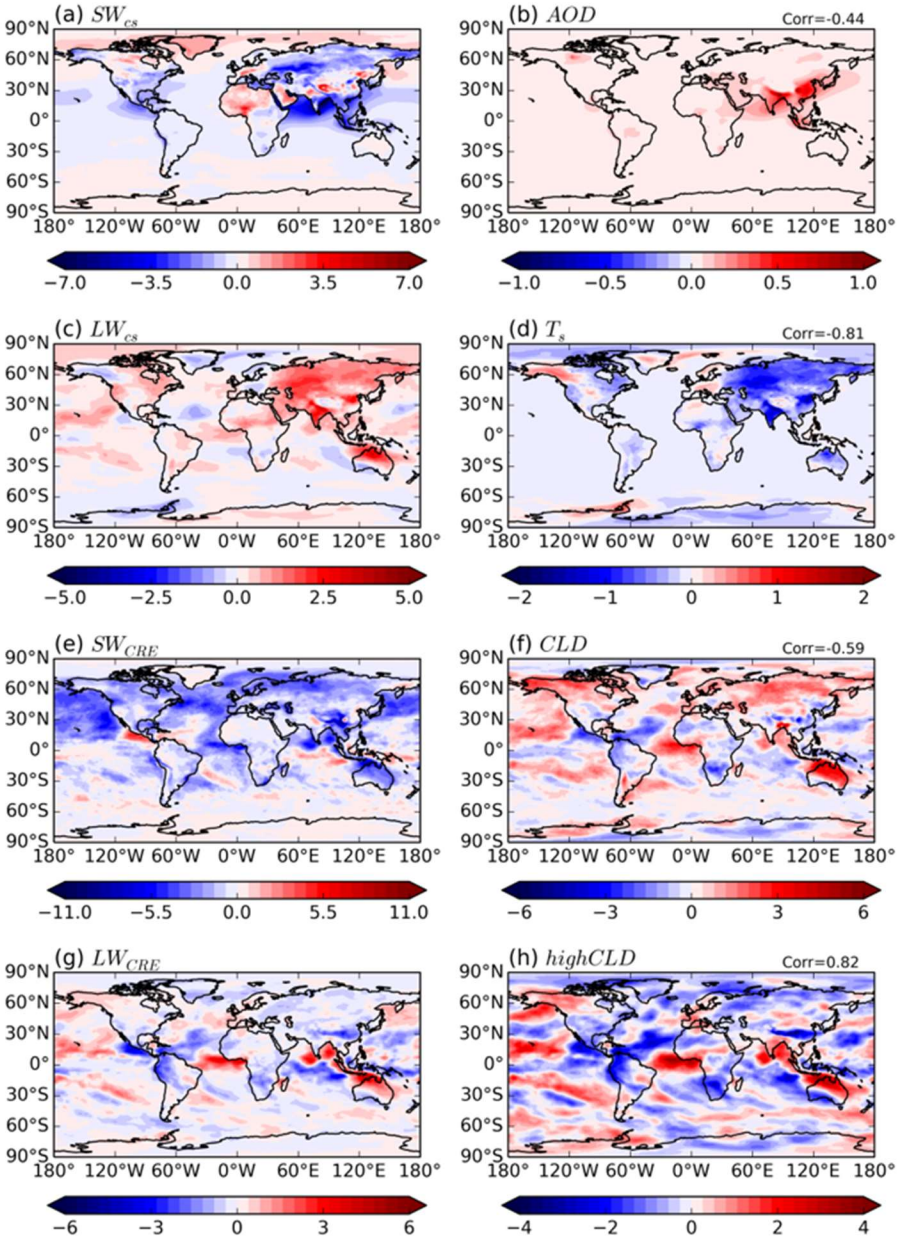
4.5 Other Forcings

4.5.1 Non-Methane Near-Term Climate Forcers (NTCFs)

The ~~PD~~ anthropogenic ERF due to CH_4, aerosols and O_3 abundances ~~(also known as near-term climate forcers, NTCFs)~~ was identified as the main source of uncertainty in the total anthropogenic ERF since ~~pre-industrial~~ PI times (Myhre et al., 2013a). This is due to the uncertainty in the individual forcings (e.g., Bellouin et al., 2020) but the interaction between individual forcings, as well as the non-linear response of climate feedbacks due to ~~aerosol-cloud interactions~~aci (Feichter et al., 2004; Deng et al., 2016; Collins et al., 2017; Shim et al., 2019) may play a role. In this section, three experiments related to non-CH_4 NTCFs are discussed: the combined simulation (*piClim-NTCF*) is identical to *piClim-control* except that ~~the~~ aerosol and ~~tropospheric~~ O_3 precursor emissions ~~(excluding CH_4)~~ are set to PD (Year 2014) levels. The single-perturbation~~forcing~~ runs change aerosol and aerosol precursor emissions only (*piClim-aer*), and ~~tropospheric~~ O_3 precursor (VOC, CO, NOx) emissions only (*piClim-O3*) from PI to PD levels. More details are described in Table 1.

The ~~global mean~~ ERF of non-CH_4 NTCFs ~~(excluding CH_4) at the PD relative to the PI~~ is ~~-1.12 ± 0.03~~ -1.03 ± 0.04 W m^{-2} (Table 3; Figure 12)~~, where the 0.03 W m^{-2} is the standard error (Forster et al., 2016)~~. The negative ERF results from the combination of a weak negative NETcs' component ~~forcing in CS conditions~~ (~~-0.04 ± 0.03~~ -0.03 ± 0.03 W m^{-2}) and a strong ~~forcing~~ contribution due to the NET ΔCRE' component ~~CRE~~ (~~-1.08 ± 0.02~~ -1.00 ± 0.02 W m^{-2}). The weak negative ~~forcing~~contribution in the CS (NETcs') is due to the negative ~~CS SW forcing~~SWcs' component (~~-0.29 ± 0.02~~ -0.26 ± 0.02 W m^{-2}) being largely offset by the ~~CS LW forcing~~ positive LWcs' component (~~+0.24 ± 0.02~~ 0.23 ± 0.03 W m^{-2}). The negative ~~CS SW forcing~~SWcs' component is ~~closely~~ correlated with changes in AOD at 550 nm (spatial correlation coefficient of -0.44) and is predominantly due to the aerosol IRF (-0.29 ± 0.01 W m^{-2}). ~~(Figs. 12a, 12b).~~ The ~~CS LW component of the ERF (~~LWcs' component~~)~~ is positive (~~+0.24 ± 0.02~~ 0.23 ± 0.03 W m^{-2}) due to both aerosols and O_3, but with the aerosol IRF only contributing 0.04 ± 0.01 W m^{-2}. The spatial variations in the global distribution of the LWcs' component are closely related to land surface temperature (Ts) changes ~~(Figs. 12c, 12d)~~; the correlation coefficient is ~~-0.81~~ -0.78 with a statistical significance well over 99 %. Considering this good correlation, the increased LWcs' component ~~may be~~ is in response to the Ts change due to NTCFs. The negative ~~SWcre~~SW ΔCRE' component (~~-1.02 ± 0.03~~ -0.92 ± 0.03 W m^{-2}), however, dominates the ERF and is largely correlated with changes in cloudiness, with the spatial pattern correlation of -0.59 between the ~~SWcre~~SW ΔCRE' component and ~~total~~ cloud fraction ~~(Figs. 12e, 12f). The LW CRE (-0.06 ± 0.02 -0.08 ± 0.02 W m^{-2}) is the smallest in magnitude, with good correlation with the changes in high level cloud fraction (Figs. 12g, 12h). This is because the impact of cloud in the SW primarily depends on the cloud optical thickness, but the impact in the LW depends on the optically thin cloud for which both the optical depth and the cloud top temperature matter.~~

1005


(a) SW_{cs}
(b) AOD
Corr=-0.44
(c) LW_{cs}
(d) T_s
Corr=-0.81
(e) SW_{CRE}
(f) CLD
Corr=-0.59
(g) LW_{CRE}
(h) $highCLD$
Corr=0.82
90°N 60°N 30°N 0° 30°S 60°S 90°S
180° 120°W 60°W 0°W 60°E 120°E 180°
−7.0 −3.5 0.0 3.5 7.0
−1.0 −0.5 0.0 0.5 1.0
−5.0 −2.5 0.0 2.5 5.0
−2 −1 0 1 2
−11.0 −5.5 0.0 5.5 11.0
−6 −3 0 3 6
−6 −3 0 3 6
−4 −2 0 2 4

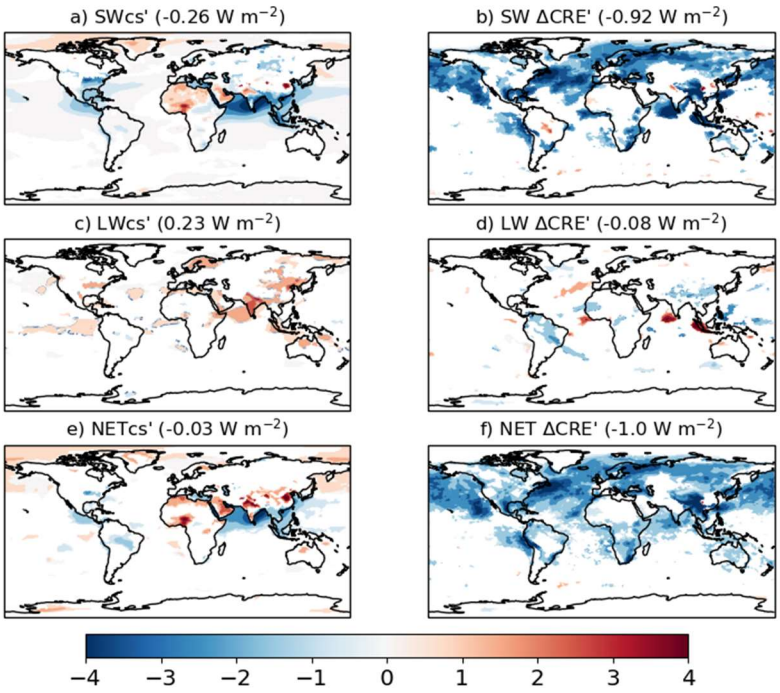


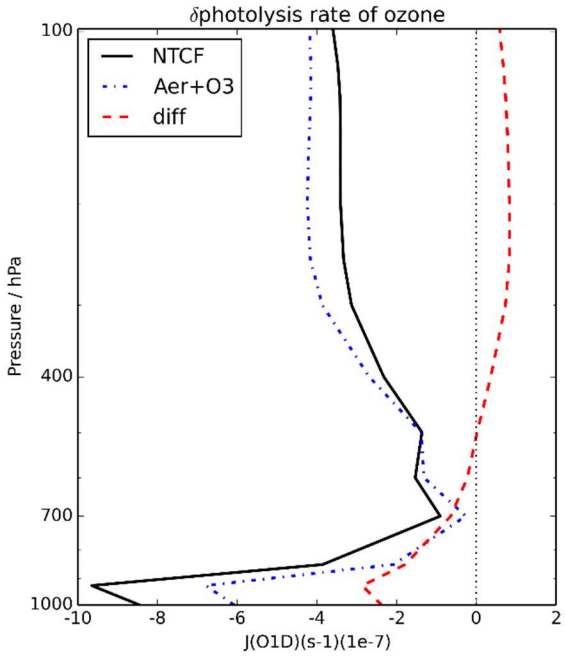
Figure 12: Global distributions of the PD NTCF ERF components at the top of atmosphere (TOA) relative to the PI period, i.e., *piClim-NTCF* minus *piClim-Control*; (a) SWcs', (b) SW △CRE', (c) LWcs', (d) LW △CRE', (e) NETcs', and (f) NET △CRE', based on Eqn. (8). Global mean values are shown in brackets. Regions where the ERF components are outside the 95 % confidence level are masked out in white. Units in W m^{-2}.

~~**Figure 12: Spatial distribution of PD ERF relative to PI in W m^{-2} derived from calculating the mean difference in the TOA radiative fluxes, (a) SWcs, (c) LWcs, (e) SWcre, (g) LWcre, and the changes (PD-PI) in (b) AOD at 550 nm, (d) surface temperature, (f) total cloud fraction, (h) high-level cloud fraction. The spatial pattern correlation between the left and right panel is shown in the right corner.**~~

As was the case for other ERFs (e.g., GHGs, aerosols, and O_3 precursor gases), the range of perturbation simulations carried out with UKESM1 enables the role of non-linear interactions to be investigated. ~~This~~The total GHG ERF was found to be equal to the sum of the individual GHG ERFs (Section 4.2.5) but non-linearities were evident for the aerosol (Section 4.3) and the O_3 precursor (Section 4.4) ERFs. As a result, this study also attempts to estimate the effects of the non-linear interactions between chemistry and aerosols on the combined aerosol and O_3 precursor ERF~~forcings~~. When combined (aerosol and aerosol and O_3 precursor emissions), their interaction may induce an effect that differs from the sum of the individual single ~~forcings~~ERFs. The ERFs do not add linearly, particularly in the ~~net CS~~ SWcs' and NETcre' components (Table 3). Firstly, we calculate the aerosol IRFs using Eqn. (7). In *piClim-NTCF,* ~~it~~the net (SW + LW) aerosol IRF is ~~-0.25~~ -0.24 ± 0.01 W m^{-2}, which is more negative than the sum of the aerosol IRFs in *piClim-aer* (~~-0.16~~ -0.15 ± 0.01 W m^{-2}) and *piClim-O3* (~~-0.02~~ -0.02 ± 0.01 W m^{-2}). This is due to the sulphate aerosol loading being higher in *piClim-NTCF* relative to *piClim-aer* by up to 3.4

(27) % globally (regionally), and driven by changes in oxidants due to the PD levels of O_3 precursors. Secondly, the NET △CRE' component contributes to the non-linearity in the ERFs. The NET △CRE' component is more negative in *piClim-NTCF* (-1.0 ± 0.02 W m^{-2}) than in the sum of *piClim-aer* (-1.0 ± 0.02 W m^{-2}) and *piClim-O3* (0.08 ± 0.02 W m^{-2}). This is primarily the result of differences in the SW ΔCRE' component, again driven by the higher sulphate loading in *piClim-NTCF* relative to *piClim-aer*. ~~Secondly, the non-linearity may also be due to O_3. Photochemistry in the atmosphere is a well-known source of tropospheric O_3 and is determined by ambient level of tropospheric O_3 precursors (i.e. NOx and VOC/CO), and photolysis rates, which are largely influenced by meteorological factors such as solar irradiance and temperature (Xing et al., 2017).~~

~~Despite~~ Although there is no direct coupling between aerosols and photolysis in UKESM1 (Archibald et al., ~~2019~~2020), aerosol-mediated cloud ~~feedback~~adjustments result in SW reduction and surface cooling ~~(Figs. 12a, 12d, 12e).~~, ~~which~~ These changes impact thermal and photochemical reactions leading to reduced photolysis rates ($O_3 + h\nu \rightarrow O(^1D) + O_2$) in the lower troposphere, while enhancing the photolysis rate of O_3 in the upper troposphere (black and blue dotted lines in Figure 13)~~.~~ leading to a reduction in surface O_3 in *piClim-NTCF* relative to the sum. However, despite differences in O_3 (not shown), the LWcs' components of the ERFs appear to add linearly. ~~This is consistent with the vertical distribution of O_3 (not shown) and changes in surface O_3 (Fig. 14).~~



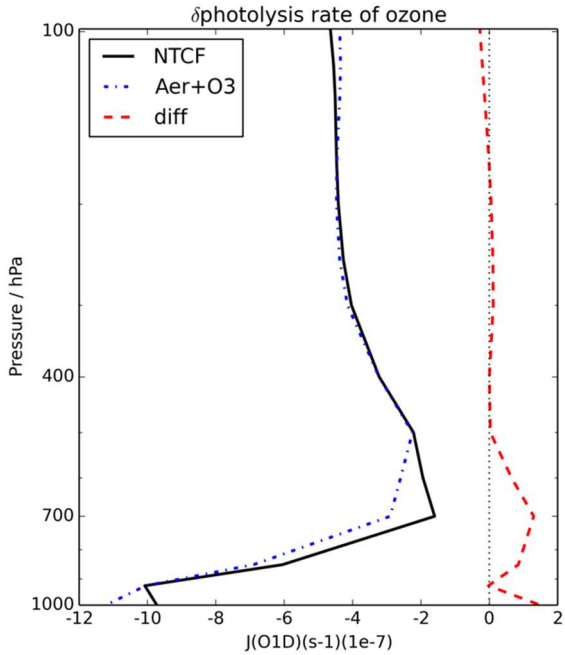
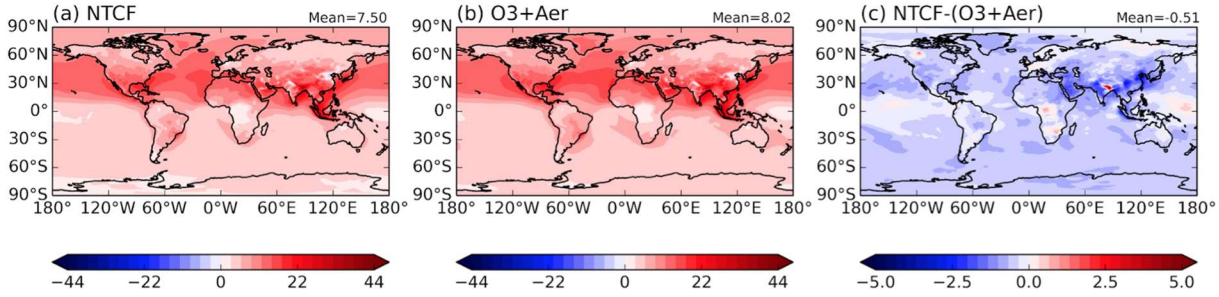


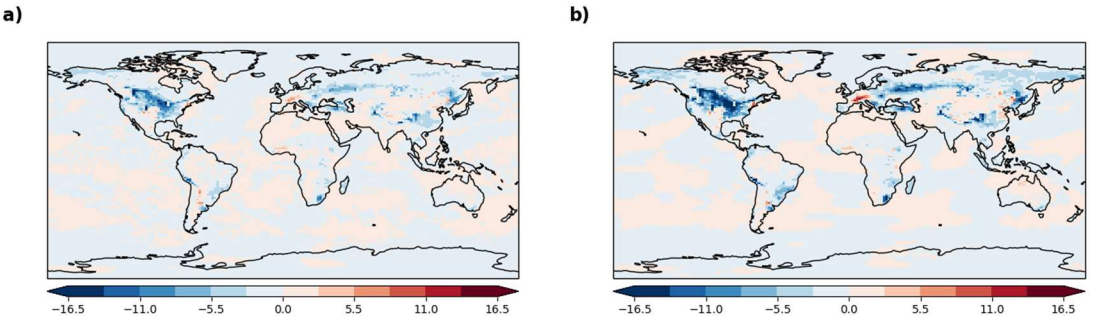
Figure 13: ~~Multi-annual global mean vertical~~Vertical **distribution of changes** ~~(PD-PI)~~ **in j(O^1D) photolysis rates (10^{-7} s^{-1})** in *piClim-NTCF* (black solid line) and the sum (blue dotted line) of *piClim-aer* and *piClim-O3* relative to *piClim-control* and the difference between them (red dashed line). ~~The black line indicates the effect of NTCFs in *piClim-NTCF* and the blue line indicates the effect of the sum of (*piclim-aer* + *piclim-O3*) relative to *piClim-control*. The red line indicates the difference between them.~~

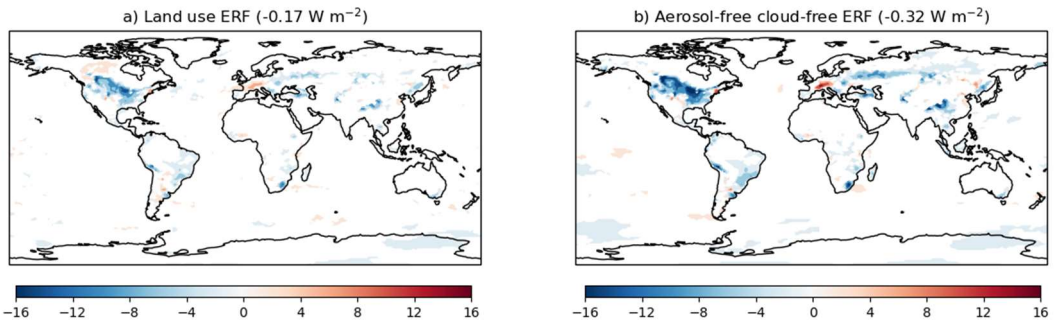


~~**Figure 14: Spatial distribution of the PI-to-PD change in surface O3 volume mixing ratio (ppbv]) in (a) *piClim-NTCF* (b) the sum (*piClim-Aer* + *piClim-O3*) relative to *piClim-control*, (c) nonlinear interaction effects [*piClim-NTCF* − (*piClim-Aer* + *piClim-O3*)] between the aerosol and ozone forcing. The global average is shown in the right corner.**~~

4.5.2 Land Use

~~Land-use~~ Land use change causes radiative forcing primarily by changing surface albedo; croplands and pastures have higher albedos than forests and are less able to mask the high albedo of snow cover. Increased albedo leads to a negative SWcs' component (-0.30 ± 0.02 W m^{-2}; Table 3) that is damped by the SW △CRE' (0.09 ± 0.03 W m^{-2}; Table 3) and the inclusion of small positive LWcs' (0.02 ± 0.03 W m^{-2}; Table 3) and LW △CRE' (0.03 ± 0.01 W m^{-2}; Table 3) terms leads to a net ERF of -0.17 ± 0.04 W m^{-2}. ~~In UKESM1, historical land-use change causes surface cooling and this reduces outgoing LW radiation which acts to damp the negative SW ERF, reducing the magnitude of the ERF from -0.39 to -0.32 Wm^{-2}.~~ The ~~land-use~~ land use ERF of UKESM1 ~~is more negative than other estimates and outside~~ falls within the 'very likely' range (~~=0.25 to -0.05~~ -0.15 to -0.10 W m^{-2}) of AR5 (~~Andrews et al., 2017;~~ Myhre et al., 2013a). However, the AR5 estimate is valid for land use change since 1700, not 1850 as in our estimate. Andrews et al. (2017) used HadGEM2-ES, UKESM1's predecessor, to calculate a 1700-1860 land use ERF of -0.1 W m^{-2}. If we apply this adjustment to our estimate, we find an ERF of -0.27 W m^{-2}, which is outside of the AR5 range. ~~However, its magnitude~~ The UKESM1 land use ERF (-0.17 ± 0.04 W m^{-2}) is reduced relative to ~~UKESM1's predecessor,~~ HadGEM2-ES ~~(Collins et al., 2011),~~ which produced ~~a~~ an 1860-2005 ~~land-use~~ land use ERF of -0.40 W m^{-2} (Andrews et al., 2017). Robertson (2019) showed that, even in the absence of snow cover, the albedo response to land use change in HadGEM2-ES is stronger than observed ~~had too strong an albedo response to land-use change~~ and it is likely that this bias still exists in UKESM1.





~~**Figure 15: Global distribution of a) the total land-use ERF and b) the cloud-free and aerosol-free land-use ERF at the present-day (Year 2014) relative to the pre-industrial (Year 1850) in W m^{-2}.**~~

Figure 14: Global distribution of the present-day (PD; Year 2014) ERF relative to the pre-industrial (PI; Year 1850) period of a) land use ERF and b) the contribution from the cloud-free and aerosol-free component. Global mean values are included and areas are masked out in white, where the ERF and/or its component is not statistically significant from zero at the 95 % confidence interval. Units in W m^{-2}.

While the global mean ~~land-use~~ land use ERF is small, regionally it can be the dominant source of the anthropogenic ERF~~radiative forcing~~. Figure 14~~15~~ shows the ~~time-mean~~ distribution of the ~~land-use~~ land use ERF and its cloud-free and~~,~~ aerosol-free, component which emphasizes the surface flux contribution. The ERF is mostly ~~constrained~~ confined to regions of ~~land-use~~ land use change, with deforestation in North America and western Eurasia causing a negative ERF and increased tree cover in central Europe causing a positive ERF. Alongside the distribution of ~~land-use~~ land use change itself, the magnitude of the ERF is larger in the mid-latitudes than the tropics, because the masking of snow cover by trees greatly increases the albedo response to changes in tree cover. The large regions of negative ERF in the northern mid-latitudes are consistent with previous model studies (e.g., de Noblet-Ducoudre et al., 2012); they are caused by agricultural expansion and the forcing is expected to have gradually increased from about 1850-1970 in North America and from 1900-1980 in Eurasia. The seasonality of leafiness, snow cover and insolation causes the ~~land-use~~ land use ERF to be largest in northern hemisphere spring~~summer~~ (~~June to August~~March to May mean of ~~-0.53~~ -0.28 W m^{-2}) and smallest in autumn (September to November mean of ~~-0.16~~ -0.08 W m^{-2}).

The ~~land-use~~ land use ERF is calculated by modifying three land-surface fields: land-cover, leaf area index (LAI) and canopy height. The values of the modified fields are taken from a coupled UKESM1 simulation that is ~~forced~~ only ~~by~~subject to historical ~~land-use~~ land use change. In the coupled configuration of UKESM1 (Sellar et al., 2019), the three land surface fields are prognostic fields calculated using a dynamic global vegetation model (DGVM), while in the atmosphere-only ERF configuration they are prescribed fields. This choice was made because the land-surface fields can be very slow to respond to changes in ~~land-use~~ land use change~~forcing~~ and ~~radiative forcing~~RF. In the coupled simulation of historical climate change

including all forcings (*hist*orical), the land surface fields respond to changes in CO_2, climate and ~~land-use~~land use, but if we use the land surface fields from this simulation, we find no substantial change in the ~~land-use~~ land use ERF (~~-0.35~~ -0.20 W m^{-2}).

In addition to the ~~radiative forcing~~ ERF caused by albedo changes, ~~land-use~~ land use change can alter climate via a number of other mechanisms. ~~Land-use~~ Land use change alters the land carbon sink. In particular, deforestation emits CO_2 to the atmosphere, and so some fraction of the CO_2 ERF is attributable to land-use change. ~~Land-use~~ Land use change also affects surface climate via changes in roughness length and transpiration. These non-radiative mechanisms usually drive larger temperature changes than the albedo response and both the non-radiative mechanisms and the CO_2 forcings tend to oppose the albedo response.

4.5.3 Total Anthropogenic ERF

As ~~summarized~~noted above, historical climate change has been driven by a wide range of anthropogenic ~~forcings~~activities that act together, alongside natural ~~forcings~~changes, to perturb the Earth's radiation balance. The total anthropogenic ERF is, therefore, a key metric in understanding observed and modelled changes in the climate system since the PI era. These various anthropogenic ~~forcing mechanisms~~drivers are not necessarily independent of each other and it is therefore worthwhile calculating the total anthropogenic ERF from a separate timeslice simulation including all ~~forcings~~perturbations together (*piClim-anthro*; Pincus et al., 2016). We completed a dedicated timeslice simulation with all ~~anthropogenic forcings~~concentrations, emissions, and land use set to 2014 levels, including all GHG concentrations (CO_2, N_2O, CH_4, HCs – both ODSs and non-ODSs), ~~tropospheric~~ O_3 precursors (VOC, /CO, ~~and~~ NOx), ~~land-use~~ land use changes, and anthropogenic aerosol or aerosol precursor emissions (SO_2, OC, BC). This experiment was not proposed in AerChemMIP (Collins et al., 2017) but is included here as part of RFMIP (Pincus et al., 2016). The main difference is that atmospheric chemistry in UKESM1 is fully interactive whereas other models participating in RFMIP (e.g., Andrews et al., 2019) use the CMIP6 ~~ozone~~O_3 dataset to represent changes in tropospheric and stratospheric O_3 and prescribe oxidants for secondary aerosol formation.

The UKESM1 *piClim-anthro* experiment leads to an ERF of ~~1.61~~ 1.76 ± 0.04 W m^{-2} (Table 3), which is dominated by a positive LWcs' component ~~CS LW forcing~~ due to GHGs and partially offset by a negative SWcs' component due to aerosols and negative contributions from the SW ΔCRE' and LW ΔCRE' components~~adjustments to LW and SW cloud-radiative effects~~ (Table 3). This ERF is a little lower than the equivalent estimate from HadGEM3-GA7.1, which was 1.81 W m^{-2} (Andrews et al., 2019). The UKESM1 estimate is also ~~somewhat~~ lower than the median estimate from CMIP5 models assessed in AR5, which equates to approximately 1.9 W m^{-2} after adjustment to the reference period of 1861-1880 to 2010-11 (Andrews and Forster, 2020~~2019; hereafter AF19~~). AR5 also provided an overall central estimate of 2.2 W m^{-2} and a 5-95 % confidence range of 1.0 ~~-~~to 3.2 W m^{-2} (after adjustment to the same reference period as above in ~~AF19~~Andrews and Forster, 2020) taking into account multiple streams of evidence. ~~AF19~~Andrews and Forster (2020) re-evaluated this as 2.3 W m^{-2} with a narrower

range of 1.7 - 3.0 W m^{-2} [5 - 95 % confidence] using a combination of atmospheric model outputs and observational constraints. The lower bound of this range was reduced to 1.5 W m^{-2} if larger uncertainties were assumed for the climate feedback parameter or for the global-mean surface temperatures anomalies used to constrain the forcing. The UKESM1 estimate of ~~1.61~~ 1.76 ± 0.04 W m^{-2} is therefore within the original uncertainty range given by AR5 and just within ~~but on the edge of~~ the range proposed in ~~AF19~~Andrews and Forster (2020).

There are several factors that contribute to the relatively low estimate of anthropogenic ERF in UKESM1. Firstly, the anthropogenic aerosol ~~forcing~~ERF in UKESM1 (and HadGEM3-GC3.1) is ~~-1.13~~ -1.09 ± 0.04 W m^{-2} (well within the uncertainty range of Bellouin et al., 2020) and offsets a major portion of the positive GHG ~~forcing~~ERF (~~2.89~~ 2.92 ± 0.04 W m^{-2}). Secondly, the ERF from *piClim-HC* is negative (~~-0.33~~ -0.18 ± 0.04 W m^{-2}) due to a strong O_3 response and connected aerosol-mediated cloud ~~feedbacks~~adjustments (Morgenstern et al., 2020~~2019~~). Thirdly, adjustments in vegetation lead to an appreciable negative ERF from ~~land-use~~ land use changes (~~-0.32~~ -0.17 ± 0.04 W m^{-2}). The stronger negative ~~forcing~~ERF from *piClim-HC* balanced by a stronger positive ERF from *piClim-CH4* (0.97 ± 0.04 W m^{-2}) due to indirect effects and the positive tropospheric RA term (Thornhill et al., 2020) ~~and *piClim-LU*~~ largely explain why the UKESM1 estimate is only ~~~0.2~~ ~0.05 W m^{-2} lower than HadGEM3-GA7.1. As shown in Fig. 3ie, the negative contributions more than offset the positive GHG ERF~~forcing~~ in certain regions. For instance, the anthropogenic ERF is negative over large parts of North America and Asia (from a combination of land use change and aerosol ~~forcing~~ERFs; see Figs. 3cd and 3g~~b~~~~), over the North Pacific (due to aerosol-cloud forcing)~~ and at southern high latitudes (from O_3 depletion due to ~~HC~~ODSs, see Fig. 3a; Section 4.2.3). In contrast, the anthropogenic ~~forcing~~ERF is strongly positive over the tropics and southern hemisphere sub-tropics (Fig. 3e) where the direct radiative effect of GHGs dominates.

The couplings between chemistry, aerosol and land surface processes included in UKESM1 increase the possibilities for non-linear ~~feedbacks and~~ interactions among the various anthropogenic forcing agents. However, these apparently have little net overall effect on the total forcing (~~1.61~~ 1.76 ± 0.04 W m^{-2}), which is ~~almost identical to~~within the uncertainty of the sum of the forcings (~~1.59~~ 1.87 ± 0.08 W m^{-2}) from the four separate groups that it includes (GHGs, aerosol, ~~tropospheric~~ O_3 precursors, and land use). The two estimates are not statistically different given the standard error on these is around 0.03 - 0.04 W m^{-2} (Table 3). This does not imply that the forcings act independently as it is possible that competing non-linear interactions cancel in this particular case. As shown in Sect. 4.5.1, the aerosol and ~~tropospheric~~ O_3 precursor emission ERFs~~forcings~~ did not add linearly~~, and neither did the sum of individual GHGs forcings (Sect. 4.2.4)~~.

5 Conclusions

Quantifying effective radiative forcings (ERFs) from anthropogenic perturbations to the Earth System (ES) is important for understanding changes in climate since the pre-industrial (PI) period. In this ~~paper~~study, we have quantified and analysed a

wide range of present-day (PD) anthropogenic ~~forcings~~ERFs with the UK's Earth System Model (ESM), UKESM1 (Sellar et al., 2019)~~, using the concept of effective radiative forcing (ERF)~~. ERFs have been shown to be a more useful metric for evaluating and comparing the relative roles of diverse forcing agents due to the relationship to global-mean temperature and other impacts that scale with it. In particular, by quantifying ERFs within a full ESM, this study addresses gaps in previous assessments in which rapid adjustments (RAs) were neglected and enables the role of indirect contributions to ERF estimates~~forcings~~ and various climate-chemistry-aerosol-cloud ~~feedbacks~~interactions to be ~~quantified~~investigated.

We find that the change in carbon dioxide (CO_2) concentration since the PI period exerts an ~~global mean~~ ERF of ~~1.82~~ 1.89 ± 0.04 W m^{-2} ~~at the present-day (PD)~~, consistent with previous estimates, making it the single largest contributor to the total anthropogenic ~~PD~~ ERF~~climate forcing~~. However, UKESM1 appears to have a more pronounced surface warming adjustment associated with the physiological forcing by CO_2 than its successor, HadGEM2-ES. The nitrous oxide (N_2O) ~~PD~~ ERF quantified here (~~0.13~~ 0.25 ± 0.04 W m^{-2}) is ~~lower~~higher than previous estimates (e.g., Hansen et al., 2005; Myhre et al., 2013a) but is consistent with the more recent Aerosol and Chemistry Model Intercomparison Project (AerChemMIP) multi-model ERF assessment (Thornhill et al., 2020). ~~due to indirect effects from ozone (O_3) depletion, and fast cloud adjustments; a shift of circulation results in a redistribution of clouds with clouds structure moving poleward. A warming associated with an O_3 increase in the tropical upper troposphere/lower stratosphere (UTLS) region may result in a regional suppression of convection and a reduction of associated LW radiation coming from cloud top.~~

The ~~PI to PD~~ PD-PI change in methane (CH_4) concentration leads to an ~~global mean~~ ERF of ~~+0.93~~ 0.97 ± 0.04 W m^{-2}, with the majority of the ~~forcing~~ERF due to ~~in~~ the clear-sky longwave (LWcs') component. Given the inclusion of interactive chemistry in UKESM1, the ERF is larger than other estimates of direct CH_4 forcing as a result of indirect effects. It is also at the high end of the range of estimates from AerChemMIP (Thornhill et al., 2020), partly due to these indirect effects but partly due to the additional positive contribution from the tropospheric RA term. O'Connor et al. (2019) apportion the CH_4 ERF between direct and indirect contributions as well as considering an emission-based perspective.

The ~~global annual mean~~ ERF from the ~~PI to PD~~ change in ozone (O_3) depleting substances (ODSs) is ~~-0.33~~ -0.18 ± 0.04 W m^{-2}. ~~A quantitative and process-based understanding of what's driving the negative ERF, despite our understanding to date of the relative roles of ODSs and O_3 depletion in climate forcing, is presented in Morgenstern et al. (2019).~~ Using a range of AerChemMIP models and observed O_3 trends as a constraint, Morgenstern et al. (2020) estimate that the UKESM1 ERF is too strongly negative; this is the result of a high O_3 bias in the PI period and a strong response to increasing ODSs in UKESM1 relative to other models (Keeble et al., 2020; Morgenstern et al., 2020). Considering all greenhouse gases (GHGs) together, we quantify an ~~global mean~~ ERF of ~~2.89~~ 2.92 ± 0.04 W m^{-2}, less than the 3.09 W m^{-2} estimate from the physical model HadGEM3-GC31, due to indirect effects. There is also ~~some~~ no evidence of non-linearity between the combined GHG ERF and the sum of the individual GHG ERFs~~, potentially due to interactive O_3 and aerosol-mediated cloud feedbacks~~adjustments~~, which may warrant further investigation~~.

The new GLOMAP-mode aerosol scheme in UKESM1 (Mulcahy et al., 20202019) leads to an intermediate sized negative ERF (-1.13 -1.09 ± 0.04 W m^{-2}) due to strong aerosol-cloud interactions (aci), despite strong absorption by black carbon (BC) and relatively weak negative aerosol instantaneous radiative forcing (IRF) from aerosol-radiation interactions (ari). Internal mixing and chemical interactions included in the new aerosol scheme mean that neither aerosol IRF nor aci are linear (sulphate, organic carbon (OC), BC interact with one another) making the total aerosol ERF less than the sum of the individual speciated aerosol ERFs.

Examining tropospheric O_3 radiative forcingstratospherically-adjusted radiative forcing (SARF) alone, results from UKESM1 suggest that the contribution from CH_4 dominates (45 45 %), with nitrogen oxides (NOx) and volatile organic compound (VOC)/carbon monoxide (CO) contributing 3637 and 1918 %, respectively. These emissions-based contributions are well within the spread of estimates from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP; (Stevenson et al., 2013): CH_4 (44 ± 12 %), NOx (31 ± 9 %), and VOC/CO (25 ± 3 %) although there is disagreement with other studies. Changes in oxidants, driven by the PI-to-PD changes in tropospheric O_3 precursor emissions, lead to an indirect aerosol ERF from aci, which either supplements or offsets the positive ERFSARF from tropospheric O_3, leading to global mean ERFs of 0.210.33 ± 0.04 and -0.080.03 ± 0.04 W m^{-2} for VOC/CO and NOx emissions changes, respectively. However, there appears to be disagreement across the AerChemMIP models on the sign and/or magnitude of the tropospheric O_3 precursor ERFs and further analysis to understand what's driving these differences is required.

The aerosol and tropospheric O_3 precursors (called near-term climate forcers, (NTCFs) in the context of AerChemMIP) together exert an global mean ERF of -1.12 -1.03 ± 0.04 W m^{-2}, which is mainly due to changes in the cloud radiative effect (CRE). There is also evidence of non-linearity in the global mean ERF between the combined *piClim-NTCF* experiment and the sum of the individual *piClim-aer* and *piCelim-O3* experiments; this is mainly evident in the shortwave clear-sky (SWcs') and shortwave cloud radiative effect (SW ΔCRE') componentsCS and is driven by changes in aerosol optical depth (AOD) and O_3. Land use (LU) change since the PI period has also exerted a negative forcingERF at the PD, estimated to be -0.32 -0.17 ± 0.04 W m^{-2} on a global mean basis. However, this estimate is outside the likely range from previous estimates, and is most likely due to too strong an albedo response.

Historical climate change has been driven by a wide range of anthropogenic forcingsactivities that act together, alongside natural forcingschanges, to perturb the Earth's radiation balance. As a result, the total anthropogenic ERF is a key metric in understanding observed and modelled changes in the climate system since the PI era. The estimate of the total anthropogenic ERF from UKESM1 is 1.61 1.76 ± 0.04 W m^{-2}, which is relatively low compared to previous assessments; this is mainly due to an intermediate negative aerosol ERF, a modest negative LUland use forcingERF and strong stratospheric O_3 depletion.

Although it may be biased low, that combined with high climate sensitivity (Andrews et al., 2019) means that UKESM1 reproduces well the historical global mean warming over the 1850 – 2014 period (Sellar et al., 2019).

In addition to quantifying anthropogenic ~~PD~~ ERFs with a fully coupled ESM, this study and other studies (e.g., Morgenstern et al., ~~2019~~2020; O'Connor et al., 2019) show the importance of indirect contributions to the ERFs~~forcings~~ and climate-aerosol-chemical ~~feedbacks~~interactions~~, and quantified their role within the ERF framework~~. There are substantial ~~feedbacks~~interactions between GHGs, stratospheric and tropospheric O_3, and aerosols, some of which act non-linearly. These effects demonstrate the importance of including ~~Earth System (~~ES~~)~~ interactions when quantifying ~~PD climate forcing~~ERFs. In particular, we ~~consider~~suggest that ~~rapid adjustments~~RAs included in the definition of ERF should include chemical as well as physical adjustments, consistent with Ramaswamy et al. (2019). They concluded in their recent assessment that although the radiative forcing concept is simple, it needs to increasingly account for the complex relevant processes in the Earth System.

Appendix A

Table A1 shows the main differences between the atmosphere-only configurations of HadGEM3-GC3.1 (called HadGEM3-GA7.1) and UKESM1 used to calculate present day effective radiative forcings (ERFs) in Andrews et al. (2019) and in this study, respectively. The implementation of the CMIP6 inputs, as applied to both models, is described in detail in Sellar et al. (2020).

Model Feature	HadGEM3-GC3.1 (Kuhlbrodt et al., 2018; Williams et al., 2018)	UKESM1 (Sellar et al., 2019)
Atmosphere Resolution	N96L85[a] or N216L85	N96L85
Vegetation, land use, and dust	Prescribed using 9 surface types, including 5 plant functional types (PFTs)	Prescribed from UKESM1 climatology using 17 surface types, including 9 natural PFTs and 4 crop/pasture PFTs; Dust tuning
Biogenic volatile organic compound (VOC) emissions	Prescribed	Monoterpene and isoprene emissions interactive; Other biogenic VOCs prescribed
Primary Marine Organic Aerosol (PMOA) emissions	Not included	Interactive, using prescribed surface water chlorophyll climatology from UKESM1
Dimethyl Sulphide (DMS) surface water concentration	Prescribed from Lana et al. (2011)	Prescribed from UKESM1 climatology
DMS emissions	Interactive, with scaling of 1.7	Interactive, with no scaling
Sulphur dioxide (SO_2) anthropogenic emissions	Split between surface and "high level" (0.5 km) dependent on sector	Added at surface only
Oxidants for secondary aerosol formation	Prescribed OH, HO_2, H_2O2, NO_3, and O_3 fields	Interactive
Long-lived greenhouse gases (LLGHGs) in radiation scheme	Uniform mass mixing ratio prescribed	Differs depending on the LLGHG – see Section 3
O_3 in radiation scheme	Prescribed, with vertical re-distribution scheme	Interactive

Table A1: Differences between the atmosphere components of the UK's models for CMIP6: HadGEM3-GC3.1 (called HadGEM3-GA7.1) and UKESM1. [a]Atmosphere resolution of HadGEM3-GC3.1 used in the Tier 1 RFMIP simulations (Andrews et al., 2019).

Acknowledgements

The development of the UK's Earth System Model, UKESM1, was funded by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra (GA01101) and by the National Environmental Research Council (NERC) national capability grant for the UK Earth System Modelling project, grant number NE/N017951/1. MD, GAF, CH, BJ, JPM, FMO'C, ER, and AW were funded by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra (GA01101). GF, FMO'C and JCT also acknowledge the EU Horizon 2020 Research Programme CRESCENDO project, grant agreement number 641816. OM and GZ were supported by the NZ Government's Strategic Science Investment Fund (SSIF) through the NIWA programme CACV. JW acknowledges support by the Deep South National Science Challenge (DSNSC), funded by the New Zealand Ministry for Business, Innovation and Employment (MBIE). BK, JS and SS were supported by the Korea Meteorological Administration Research and Development Program "Development and Assessment of IPCC AR6 Climate Change Scenario", grant agreement number 1365003000.

The authors acknowledge the contribution of NeSI high-performance computing facilities to the results of this research. New Zealand's national facilities are provided by the New Zealand eScience Infrastructure (NeSI) and funded jointly by NeSI's collaborator institutions and through MBIE's Research Infrastructure programme. This work used Monsoon2, a collaborative High Performance Computing facility funded by the Met Office and the Natural Environment Research Council. This work also used the NEXCS High Performance Computing facility funded by the Natural Environment Research Council and delivered by the Met Office. This work used JASMIN, the UK collaborative data analysis facility; the authors are grateful for the provision of these facilities.

The authors thank T. Andrews for constructive comments on the manuscript. Last but not least, the authors wish to acknowledge the huge effort from the UKESM1 core group in building and evaluating UKESM1 and making it available for use in RFMIP and AerChemMIP.

Data Availability

~~Data from all the simulations which underpin this paper are in the process of being post-processed in readiness for uploading to the Earth System Grid Federation (ESGF).~~ All of the data from the RFMIP and AerChemMIP simulations analysed in this study have been published on the Earth System Grid Federation and the model Source ID is UKESM1-0-LL. Below is a table listing all of the corresponding dataset citations. In some cases, data from 2 realisations of the same experiment are available. ~~The authors endeavour to use the precise data citation for each piece of data once all of it is citable.~~

Experiment ID	MIP	Data Citation(s)
piClim-control	AerChemMIP/RFMIP	doi:10.22033/ESGF/CMIP6.6276
piClim-4xCO2	RFMIP	doi:10.22033/ESGF/CMIP6.11061
piClim-N2O	AerChemMIP	doi:10.22033/ESGF/CMIP6.9434
piClim-HC	AerChemMIP	doi:10.22033/ESGF/CMIP6.9433
piClim-CH4	AerChemMIP	doi:10.22033/ESGF/CMIP6.6229
piClim-GHG	RFMIP	doi:10.22033/ESGF/CMIP6.11094
piClim-SO2	AerChemMIP	Realisation 1: doi:10.22033/ESGF/CMIP6.9440 Realisation 2: doi:10.22033/ESGF/CMIP6.6261
piClim-BC	AerChemMIP	doi:10.22033/ESGF/CMIP6.6225
piClim-OC	AerChemMIP	Realisation 1: doi:10.22033/ESGF/CMIP6.9439 Realisation 2: doi:10.22033/ESGF/CMIP6.6257
piClim-aer	AerChemMIP	doi:10.22033/ESGF/CMIP6.6270
piClim-NOx	AerChemMIP	doi:10.22033/ESGF/CMIP6.6246
piClim-VOC	AerChemMIP	doi:10.22033/ESGF/CMIP6.6266
piClim-O3	AerChemMIP	doi:10.22033/ESGF/CMIP6.6254
piClim-NTCF/piClim-aerO3	AerChemMIP/RFMIP	Realisation 1: doi:10.22033/ESGF/CMIP6.8418 Realisation 2: doi:10.22033/ESGF/CMIP6.6249
piClim-LU	RFMIP	doi:10.22033/ESGF/CMIP6.11104
piClim-Anthro	RFMIP	doi:10.22033/ESGF/CMIP6.11090

Author Contributions

The model simulations were set up, reviewed, and/or ran by FMO'C, NLA, MD, GF, PG, CH, BTJ, BK, JK, JPM, ER, SS, ST, AW, JW, and GZ. Data processing and upload to the Earth System Grid Federation (ESGF) was carried out by MD, BK,

RK, MR, SS, JCT, and JW. Analysis was carried out by FMO'C, PG, BTJ, JK, JPM, ER, JS, SS, JCT, AW, and GZ. The manuscript was prepared by FMO'C, PG, BTJ, JK, OM, ER, JS, SS, JCT, AW, and GZ, with additional contributions from all co-authors.

Completing Interests

The authors declare that they have no conflict of interest.

References

Albrecht, B. A.: Aerosols, Cloud Microphysics, and Fractional Cloudiness, *Science*, 245, 1227–1230, DOI: 10.1126/science.245.4923.1227, 1989.

Andrews, T., J. M. Gregory, M. J. Webb, and K. E. Taylor: Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models, *Geophys. Res. Lett.*, 39, L09712, doi:10.1029/ 2012GL051607, 2012a.

Andrews, T., M. A. Ringer, M. Doutriaux-Boucher, M. J. Webb, and W. J. Collins: Sensitivity of an Earth system climate model to idealized radiative forcing, *Geophys. Res. Lett.*, *39*(10), 2012b.

Andrews, T.: Using an AGCM to Diagnose Historical Effective Radiative Forcing and Mechanisms of Recent Decadal Climate Change, *J. Climate*, DOI: 10.1175/JCLI-D-13-00336.1, 2014.

Andrews, T., R. A. Betts, B. B. B. Booth, C. D. Jones, and G. S. Jones: Effective radiative forcing from historical land use change. *Climate Dyn.*, 48, 3489, doi.org/10.1007/s00382-016-3280-7, 2017.

Andrews, T., M. B. Andrews, A. Bodas-Salcedo, G. S. Jones, T. Kulhbrodt, J. Manners, M. B. Menary, J. Ridley, M. A. Ringer, A. A. Sellar, C. A. Senior, and Y. Tang: Forcings, feedbacks and climate sensitivity in HadGEM3-GC3.1 and UKESM1, *J. Adv. Modeling Earth Sys*., 11, 4377-4394, https://doi.org/ 10.1029/2019MS001866~~Submitted~~, 2019.

Andrews, T., and P. M. Forster: Energy budget constraints on historical radiative forcing, *Nature Climate Change*, 10, 313-316, doi.org/10.1038/s41558-020-0696-1, 2020~~Submitted, 2019~~.

Archibald, A. T., M. C. Cooke, S. R. Utembe, D. E. Shallcross, R. G. Derwent, and M. E. Jenkin: Impacts of mechanistic changes on HOx formation and recycling in the oxidation of isoprene, *Atmos. Chem. Phys.*, 10, 8097–8118, doi:10.5194/acp-10-8097-2010, 2010.

Archibald, A. T., O'Connor, F. M., N. L. Abraham, S. Archer-Nicholls, M. P. Chipperfield, M. Dalvi, G. A. Folberth, F. Dennison, S. S. Dhomse, P. T. Griffiths, C. Hardacre, A. J. Hewitt, R. Hill, C. E. Johnson, J. Keeble, M. O. Köhler, O. Morgenstern, J. P. Mulcahy, C. Ordóñez, R. J. Pope, S. Rumbold, M. R. Russo, N. Savage, A. Sellar, M. Stringer, S. Turnock, O. Wild, and G. Zeng: Description and evaluation of the UKCA stratosphere-troposphere chemistry scheme (StratTrop vn 1.0) implemented in UKESM1, *Geosci. Model Dev.* ~~*Disc.*~~, doi.org/10.5194/gmd-2019-246, 2020~~2019~~.

Arfeuille, F., D. Weisenstein, H. Mack, E. Rozanov, T. Peter, and S. Brönnimann: Volcanic forcing for climate modeling: a new microphysics-based data set covering years 1600–present, *Clim. Past*, 10, 359–375, doi:10.5194/cp-10-359-2014, 2014.

Arneth, A., S. P. Harrison, S. Zaehle, K. Tsigaridis, S. Menon, P. J. Bartlein, J. Feichter, A. Korhola, M. Kulmala, D. O'Donnell, G. Schurgers, S. Sorvari, and T. Vesala: Terrestrial biogeochemical feedbacks in the climate system, *Nature Geoscience*, 3, 525-532, 2010.

Bellouin, N., J. Rae, A. Jones, C. Johnson, J. Haywood, and O. Boucher: Aerosol forcing in the Climate Model Intercomparison Project (CMIP5) simulations by HadGEM2-ES and the role of ammonium nitrate, *J. Geophys. Res.*, 116, D20206, doi:10.1029/2011JD016074, 2011.

Bellouin, N., J. Quaas, E. Gryspeerdt, S. Kinne, P. Stier, D. Watson-Parris, O.Boucher, K. S. Carslaw, M. Christensen, A.-L. Daniau, J.-L. Dufresne, G. Feingold, S. Fiedler, P. Forster, A. Gettelman, J. M. Haywood, F. Malavelle, U. Lohmann, T. Mauritsen, D. T. McCoy, G. Myhre, J. Mülmenstädt, D. Neubauer, A. Possner, M. Rugenstein, Y. Sato, M. Schulz, S. E. Schwartz, O. Sourdeval, T. Storelvmo, V. Toll, D. Winker, and B. Stevens: Bounding aerosol radiative forcing of climate change, *Rev. Geophys.*, 58, e2019RG000660. doi.org/10.1029/2019RG000660, 2020~~In press, 2019~~.

Bond, T. C., S. J. Doherty, D. W. Fahey, P. M. Forster, T. Berntsen, B. J. DeAngelo, M. G. Flanner, S. Ghan, B. Kärcher, D. Koch, S. Kinne, Y. Kondo, P. K. Quinn, M. C. Sarofim, M. G. Schultz, M. Schulz, C. Venkataraman, H. Zhang, S. Zhang, N. Bellouin, S. K. Guttikunda, P. K. Hopke, M. Z. Jacobson, J. W. Kaiser, Z. Klimont, U. Lohmann, J. P. Schwarz, D. Shindell, T. Storelvmo, S. G. Warren, and C. S. Zender: Bounding the Role of Black Carbon in the Climate System: A Scientific Assessment, *J. Geophys. Res.-Atmos*., 118, 5380–5552, doi.org/10.1002/jgrd.50171, 2013.

Boucher, O., D. Randall, P. Artaxo, C. Bretherton, G. Feingold, P. Forster, V.-M. Kerminen, Y. Kondo, H. Liao, U. Lohmann, P. Rasch, S.K. Satheesh, S. Sherwood, B. Stevens and X.Y. Zhang: Clouds and Aerosols. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Butchart, N.: The Brewer-Dobson circulation, *Rev. Geophys.*, 52, 157–184, doi:10.1002/2013RG000448, 2014.

Carslaw, K. S., L. A. Lee, C. L. Reddington, K. J. Pringle, A. Rap, P. M. Forster, G. W. Mann, D. V. Spracklen, M. T. Woodhouse, L. A. Regayre, and J. R. Pierce: Large contribution of natural aerosols to uncertainty in indirect forcing, *Nature*, 503, 67-71, DOI:10.1038/nature12674, 2013.

Checa-Garcia, R., Hegglin, M. I., Kinnison, D., Plummer, D. A., and K. P. Shine: Historical tropospheric and stratospheric ozone radiative forcing using the CMIP6 database, *Geophys. Res. Lett*., 45, 3264–3273, doi.org/10.1002/2017GL076770, 2018.

Chung, E.-S., and B. J. Soden: An assessment of methods for computing radiative forcing in climate models, *Environ. Res. Lett.*, 10, 074004, 2015.

Collins, W. D., V. Ramaswamy, M. D. Schwarzkopf, Y. Sun, R. W. Portmann, Q. Fu, S. E. B. Casanova, J.-L. Dufresne, D. W. Fillmore, P. M. D. Forster, V. Y. Galin, L. K. Gohar, W. J. Ingram, D. P. Kratz, M.-P. Lefebvre, J. Li, P. Marquet, V. Oinas, Y. Tsushima, T. Uchiyama, and W. Y. Zhong: Radiative forcing by well-mixed greenhouse gases: Estimates from climate models in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4), *J. Geophys. Res*., 111, D14317, doi:10.1029/2005JD006713, 2006.

Collins, W. J., S. Sitch, and O. Boucher: How vegetation impacts affect climate metrics for ozone precursors, *J. Geophys. Res*., 115, D23308, doi:10.1029/2010JD014187, 2010.

Collins, W. J., N. Bellouin, M. Doutriaux-Boucher, N. Gedney, P. Halloran, T. Hinton, J. Hughes, C. D. Jones, M. Joshi, S. Liddicoat, G. Martin, F. O'Connor, J. Rae, C. Senior, S. Sitch, I. Totterdell, A. Wiltshire, and S. Woodward: Development and evaluation of an Earth-System model – HadGEM2, *Geosci. Model Dev.*, 4, 1051–1075, doi:10.5194/gmd-4-1051-2011, 2011.

Collins, W. J., J.-F. Lamarque, M. Schulz, O. Boucher, V. Eyring, M. I. Hegglin, A. Maycock, G. Myhre, M. Prather, D. Shindell, and S. J. Smith: AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6, *Geosci. Model Dev.*, 10, 585-607, doi.org/10.5194/gmd-10-585-2017, 2017.

Deng, J., H. Xu, and L. Zhang: Nonlinear effects of anthropogenic aerosol and urban land surface forcing on spring climate in eastern China, *J. Geophys. Res*., 121, 4581-4599, doi:10.1002/2015JD024377, 2016.

Doutriaux-Boucher, M., M. J. Webb, J. M. Gregory, and O. Boucher: Carbon dioxide induced stomatal closure increases radiative forcing via a rapid reduction in low cloud, *Geophys. Res. Lett.*, *36*(2), 2009.

Edwards, J. M. and A. Slingo: Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model, *Q. J. R. Meteorol. Soc.*, 122, pp. 689-719, 1996.

Engel, A. and M. Rigby (Lead Authors), J.B. Burkholder, R.P. Fernandez, L. Froidevaux, B.D. Hall, R. Hossaini, T. Saito, M.K. Vollmer, and B. Yao: Update on Ozone-Depleting Substances (ODSs) and Other Gases of Interest to the Montreal Protocol, Chapter 1 in Scientific Assessment of Ozone Depletion: 2018, Global Ozone Research and Monitoring Project–Report No. 58, World Meteorological Organization, Geneva, Switzerland, 2018.

Etminan, M., G. Myhre, E. J. Highwood, and K. P. Shine: Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing, *Geophys. Res. Lett*., 43, 12614–12623, doi:10.1002/2016GL071930, 2016.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev*., 9, 1937-1958, doi.org/10.5194/gmd-9-1937-2016, 2016.

Feichter, J., E. Roeckner, U. Lohmann, and B. Liepert: Nonlinear aspects of the climate response to greenhouse gas and aerosol forcing, *J. Climate*, 17, 2384-2398, doi:10.1175/1520-0442(2004)017, 2004.

Feng, Y., V. Ramanathan, and V. R. Kotamarthi: Brown carbon: a significant atmospheric absorber of solar radiation?, *Atmos. Chem. Phys.*, 13, 8607–8621, doi.org/10.5194/acp-13-8607-2013, 2013.

Fiore, A. M., F. J. Dentener, O. Wild, C. Cuvelier, M. G. Schultz, P. Hess, C. Textor, M. Schulz, R. M. Doherty, L. W. Horowitz, I. A. MacKenzie, M. G. Sanderson, D. T. Shindell, D. S. Stevenson, S. Szopa, R. Van Dingenen, G. Zeng, C. Atherton, D. Bergmann, I. Bey, G. Carmichael, 1W. J. Collins, B. N. Duncan, G. Faluvegi, G. Folberth, M. Gauss, S. Gong, D. Hauglustaine, T. Holloway, I. S. A. Isaksen, D. J. Jacob, J. E. Jonson, J. W. Kaminski, T. J. Keating, A. Lupu, E. Marmer, V. Montanaro, R. J. Park, G. Pitari, K. J. Pringle, J. A. Pyle, S. Schroeder, M. G. Vivanco, P. Wind, G. Wojcik, S. Wu, and A. Zuber: Multimodel estimates of intercontinental source-receptor relationships for ozone pollution, *J. Geophys. Res.*, 114, D04301, doi:10.1029/2008JD010816, 2009.

Forster, P., V. Ramaswamy, P. Artaxo, T. Berntsen, R. Betts, D.W. Fahey, J. Haywood, J. Lean, D.C. Lowe, G. Myhre, J. Nganga, R. Prinn, G. Raga, M. Schulz and R. Van Dorland: Changes in Atmospheric Constituents and in Radiative Forcing. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Forster, P. M., T. Andrews, P. Good, J. M. Gregory, L. S. Jackson, and M. Zelinka: Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models, *J. Geophys. Res. Atmos*., 118, 1139–1150, doi:10.1002/jgrd.50174, 2013.

Forster, P. M., T. Richardson, A. C. Maycock, C. J. Smith, B. H. Samset, G. Myhre, T. Andrews, R. Pincus, and M. Schulz: Recommendations for diagnosing effective radiative forcing from climate models for CMIP6, *J. Geophys. Res. Atmos*.,121,12460–12475, doi:10.1002/2016JD025320, 2016.

Fry, M. M., V. Naik, J. J. West, M. D. Schwarzkopf, A. M. Fiore, W. J. Collins, F. J. Dentener, D. T. Shindell, C. Atherton, D. Bergmann, B. N. Duncan, P. Hess, I. A. MacKenzie, E. Marmer, M. G. Schultz, S. Szopa, O. Wild, and G. Zeng: The influence of ozone precursor emissions from four world regions on tropospheric composition and radiative climate forcing, *J. Geophys. Res*., 117, D07306, doi:10.1029/2011JD017134, 2012.

Fry, M. M., M. D. Schwarzkopf, Z. Adelman, and J. J. West: Air quality and radiative forcing impacts of anthropogenic volatile organic compound emissions from ten world regions, *Atmos. Chem. Phys*., 14, 523–535, doi:10.5194/acp-14-523-2014, 2014.

Ghan, S.J., X. Liu, R.C. Easter, R. Zaveri, P.J. Rasch, J. Yoon, and B. Eaton: Toward a Minimal Representation of Aerosols in Climate Models: Comparative Decomposition of Aerosol Direct, Semidirect, and Indirect Radiative Forcing. *J. Climate,* 25, 6461–6476, https://doi.org/10.1175/JCLI-D-11-00650.1, 2012.

Ghan, S. J.: Technical Note: Estimating aerosol effects on cloud radiative forcing, *Atmos. Chem. Phys.*, 13, 9971-9974, https://doi.org/10.5194/acp-13-9971-2013, 2013.

Ghan, S., M. Wang, S. Zhang, S. Ferrachat, A. Gettelman, J. Griesfeller, Z. Kipling, U. Lohmann, H. Morrison, D. Neubauer, D. G. Partridge, P. Stier, T. Takemura, H. Wang, K. Zhang: Constraining aerosol effects on cloud forcing, P. Natl. Acad. Sci. USA, 113 (21) 5804-5811; DOI: 10.1073/pnas.1514036113, 2016.

Hansen, J., Sato, M., and R. Ruedy: Radiative Forcing and Climate Response, *J. Geophys. Res*., 102, D6, 6831-6864, 1997.

Hansen, J., M. Sato, R. Ruedy, L. Nazarenko, A. Lacis, G. A. Schmidt, G. Russell, I. Aleinov, M. Bauer, S. Bauer, N. Bell, B. Cairns, V. Canuto, M. Chandler, Y. Cheng, A. Del Genio, G. Faluvegi, E. Fleming, A. Friend, T. Hall, C. Jackman, M. Kelley, N. Kiang, D. Koch, J. Lean, J. Lerner, K. Lo, S. Menon, R. Miller, P. Minnis, T. Novakov, V. Oinas, Ja. Perlwitz, Ju. Perlwitz, D. Rind, A. Romanou, D. Shindell, P. Stone, S. Sun, N. Tausnev, D. Thresher, B. Wielicki, T. Wong, M. Yao, and S. Zhang: Efficacy of climate forcings, J. Geophys. Res., 110, D18104, doi:10.1029/2005JD005776, 2005.

Hewitt, H. T., D. Copsey, I. D. Culverwell, C. M. Harris, R. S. R. Hill, A. B. Keen, A. J. McLaren, and E. C. Hunke: Design and implementation of the infrastructure of HadGEM3: the next-generation Met Office climate modelling system, *Geosci. Model Dev.*, 4, 223–253, doi: 10.5194/gmd-4-223-2011, 2011.

Hoesly, R. M., S. J. Smith, L. Feng, Z. Klimont, G. Janssens-Maenhout, T. Pitkanen, J. J. Seibert, L. Vu, R. J. Andres, R. M. Bolt, T. C. Bond, L. Dawidowski, N. Kholod, J. Kurokawa, M. Li, L. Liu, Z. Lu, M. C. P. Moura, P. R. O'Rourke, and Q. Zhang: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), *Geosci. Model Dev*., 11, 369-408, doi.org/10.5194/gmd-11-369-2018, 2018.

Holmes, C. D., M. J. Prather, O. A. Søvde, and G. Myhre: Future methane, hydroxyl, and their uncertainties: key climate and emission parameters for future predictions, *Atmos. Chem. Phys.*, 13, 285–302, doi.org/10.5194/acp-13-285-2013, 2013.

Johnson, B. T., J. M. Haywood, and M. K. Hawcroft: Are changes in atmospheric circulation important for black carbon aerosol impacts on clouds, precipitation, and radiation? *J. Geophys. Res.: Atmospheres*, 124, 7930– 7950, doi.org/10.1029/2019JD030568, 2019.

Jones, A., Roberts, D. L., and A. Slingo: A climate model study of indirect radiative forcing by anthropogenic sulphate aerosols, *Nature*, 370, 450–453, https://www.nature.com/articles/370450a0, 1994.

Karset, I. H. H., T. K. Berntsen, T. Storelvmo, K. Alterskjæ , A. Grini , D. Olivié , A. Kirkevåg , Ø. Seland, T. Iversen, and M. Schulz: Strong impacts on aerosol indirect effects from historical oxidant changes, *Atmos. Chem. Phys*., 18, 7669–7690, doi.org/10.5194/acp-18-7669-2018, 2018.

Keeble, J., Hassler, B., Banerjee, A., Checa-Garcia, R., Chiodo, G., Davis, S., Eyring, V., Griffiths, P. T., Morgenstern, O., Nowack, P., Zeng, G., Zhang, J., Bodeker, G., Cugnet, D., Danabasoglu, G., Deushi, M., Horowitz, L. W., Li, L., Michou, M., Mills, M. J., Nabat, P., Park, S., and Wu, T.: Evaluating stratospheric ozone and water vapor changes in CMIP6 models from 1850–2100, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-1202, in review, 2020.

Kelly, J., R. M. Doherty, F. M. O'Connor, and G.W. Mann: The impact of biogenic, anthropogenic, and biomass burning emissions on regional and seasonal variations in secondary organic aerosol concentrations, *Atmos. Chem. Phys.*, doi.org/10.5194/acp-18-7393-2018, 2018.

Kuhlbrodt, T., Jones, C. G., Sellar, A., Storkey, D., Blockley, E., Stringer, M., et al. (2018). The low-resolution version of HadGEM3-GC3.1:Development and evaluation for global climate. Journal of Advances in Modeling Earth Systems, 10, 2865–2888. https://doi.org/10.1029/2018MS001370.

Lamarque, J.-F., D. T. Shindell, B. Josse, P. J. Young, I. Cionni, V. Eyring, D. Bergmann, P. Cameron-Smith, W. J. Collins, R. Doherty, S. Dalsoren, G. Faluvegi, G. Folberth, S. J. Ghan, L. W. Horowitz, Y. H. Lee, I. A. MacKenzie, T. Nagashima, V. Naik, D. Plummer, M. Righi, S. T. Rumbold, M. Schulz, R. B. Skeie, D. S. Stevenson, S. Strode, K. Sudo, S. Szopa, A. Voulgarakis, and G. Zeng: Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics, *Geosci. Model Dev*., 6, 179–206, doi:10.5194/gmd-6-179-2013, 2013.

Lana, A., Bell, T. G., Simó, R., Vallina, S. M., Ballabrera-Poy, J., Kettle, A. J., et al. (2011). An updated climatology of surface dimethlysulfideconcentrations and emission fluxes in the global ocean. Global Biogeochemical Cycles, 25, GB1004. https://doi.org/10.1029/2010GB003850.

Li, J., C. L. Curry, Z. Sun, and F. Zhang: Overlap of solar and infrared spectra and the shortwave radiative effect of methane, *J. Atmos. Sci*., 67(7), 2372–2389, doi:10.1175/2010jas3282.1, 2010.

Lohmann, U., L. Rotstayn, T. Storelvmo, A. Jones, S. Menon, J. Quaas, A. M. L. Ekman, D. Koch, and R. Ruedy: Total aerosol effect: radiative forcing or radiative flux perturbation?, Atmos. Chem. Phys., 10, 3235–3246, www.atmos-chem-phys.net/10/3235/2010/, 2010.

Malavelle, F., J. Haywood, A. Jones, A. Gettelman, L. Clarisse, S. Bauduin, R. P. Allan, I. H. H. Karset, J. Egill Kristjánsson, L. Oreopoulos, N. Cho, D. Lee, N. Bellouin, O. Boucher, D. P. Grosvenor, K. S. Carslaw, S. Dhomse, G. W. Mann, A. Schmidt, H. Coe, M. E. Hartley, M. Dalvi, A. A. Hill, B. T. Johnson, C. E. Johnson, J. R. Knight, F. M. O'Connor, D. G. Partridge, P. Stier, G. Myhre, S. Platnick, G. L. Stephens, H. Takahashi, and T. Thordarson: Strong constraints on aerosol–cloud interactions from volcanic eruptions, *Nature,* 546, 485–491, doi:10.1038/nature22974, 2017.

Mann, G. W., K. S. Carslaw, D. V. Spracklen, D. A. Ridley, P. T. Manktelow, M. P. Chipperfield, S. J. Pickering, and C. E. Johnson: Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model, *Geosci. Model Dev*., 3, 519–551, doi:10.5194/gmd-3-519-2010, 2010.

Matthes, K., B. Funke, M. E. Andersson, L. Barnard, J. Beer, P. Charbonneau, M. A. Clilverd, T. Dudok de Wit, M. Haberreiter, A. Hendry, C. H. Jackman, M. Kretzschmar, T. Kruschke, M. Kunze, U. Langematz, D. R. Marsh, A. C. Maycock, S. Misios, C. J. Rodger, A. A. Scaife, A. Seppälä, M. Shangguan, M. Sinnhuber, K. Tourpali, I. Usoskin, M. van de Kamp, P. T. Verronen, and S. Versick: Solar forcing for CMIP6 (v3.2), *Geosci. Model Dev*., 10, 2247-2302, doi.org/10.5194/gmd-10-2247-2017, 2017.

van Marle, M. J. E., S. Kloster, B. I. Magi, J. R. Marlon, A.-L. Daniau, R. D. Field, A. Arneth, M. Forrest, S. Hantson, N. M. Kehrwald, W. Knorr, G. Lasslop, F. Li, S. Mangeon, C. Yue, J. W. Kaiser, and G. R. van der Werf: Historic global biomass burning emissions for CMIP6 (BB4CMIP) based on merging satellite observations with proxies and fire models (1750–2015), *Geosci. Model Dev*., 10, 3329-3357, doi.org/10.5194/gmd-10-3329-2017, 2017.

Meinshausen, M., E. Vogel, A. Nauels, K. Lorbacher, N. Meinshausen, D. M. Etheridge, P. J. Fraser, S. A. Montzka, P. J. Rayner, C. M. Trudinger, P. B. Krummel, U. Beyerle, J. G. Canadell, J. S. Daniel, I. G. Enting, R. M. Law, C. R. Lunder, S. O'Doherty, R. G. Prinn, S. Reimann, M. Rubino, G. J. M. Velders, M. K. Vollmer, R. H. J. Wang, and R. Weiss: Historical greenhouse gas concentrations for climate modelling (CMIP6), *Geosci. Model Dev*., 10, 2057-2116, doi.org/10.5194/gmd-10-2057-2017, 2017.

Morgenstern, O., P. Braesicke, F. M. O'Connor, A. C. Bushell, C. E. Johnson, S. M. Osprey, and J. A. Pyle: Evaluation of the new UKCA climate-composition model – Part 1: The stratosphere, *Geosci. Model Dev*., 2, 43-57, doi.org/10.5194/gmd-2-43-2009, 2009.

Morgenstern, O., F. M. O'Connor, B. Johnson, G. Zeng, J. P. Mulcahy, J. Williams, and J. Teixeira, M. Michou, P. Nabat, L. Horowitz, V. Naik, and L. Sentman: Reappraisal of the climate impacts of ozone-depleting substances, In preparation, 2019.

Mulcahy, J. P., Jones, C., Sellar, A., Johnson, B., Boutle, I. A., Jones, A., T. Andrews, S. T. Rumbold, J. Mollard, N. Bellouin, C. E. Johnson, K. D. Williams, D. P. Grosvenor, and D. T. McCoy: Improved aerosol processes and effective radiative forcing in HadGEM3 and UKESM1, *J. Adv. Modeling Earth Sys.*, 10, 2786–2805, doi.org/10.1029/2018MS001464, 2018.

Mulcahy, J. P., Johnson C., Jones C., Povey A., Sellar A., Scott C. E., Turnock S. T., Woodhouse M. T., Abraham L. N., Andrews M., Bellouin N., Browse J., Carslaw K. S., Dalvi M., Folberth G., Grosvenor D., Hardacre C., Johnson B., Jones A., Kipling Z., Mann G., Mollard J., Schutgens N., O'Connor F., Palmieri J., Reddington C., Richardson M., Stier P., Woodward S., and Yool A.: Description and evaluation of aerosol in UKESM1 and HadGEM3-GC3.1 CMIP6 historical simulations, *Geosci. Model. Dev. Disc., https://doi.org/10.5194/gmd-2019-357, in review, 2020,* ~~Submitted, 2019~~.

Myhre, G., E. Highwood, K. Shine, and F. Stordal: New estimates of radiative forcing due to well mixed greenhouse gases, *Geophys. Res. Lett.*, 25(14), 2715–2718, doi:10.1029/98GL01908, 1998.

Myhre, G., D. Shindell, F.-M. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura and H. Zhang (2013a): Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Myhre, G., Samset, B. H., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K., Bian, H., Bellouin, N., Chin, M., Diehl, T., Easter, R. C., Feichter, J., Ghan, S. J., Hauglustaine, D., Iversen, T., Kinne, S., Kirkevåg, A., Lamarque, J.-F., Lin, G., Liu, X., Lund, M. T., Luo, G., Ma, X., van Noije, T., Penner, J. E., Rasch, P. J., Ruiz, A., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., Wang, P., Wang, Z., Xu, L., Yu, H., Yu, F., Yoon, J.-H., Zhang, K., Zhang, H., and Zhou, C. (2013b): Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations, *Atmos. Chem. Phys.*, 13, 1853-1877, doi.org/10.5194/acp-13-1853-2013.

Nisbet, E. G., E. J. Dlugokencky, M. R. Manning, D. Lowry, R. E. Fisher, J. L. France, S. E. Michel, J. B. Miller, J. W. C. White, B. Vaughn, P. Bousquet, J. A. Pyle, N. J. Warwick, M. Cain, R. Brownlow, G. Zazzeri, M. Lanoisellé, A. C. Manning, E. Gloor, D. E. J. Worthy, E.-G. Brunke, C. Labuschagne, E. W. Wolff, and A. L. Ganesan (2016): Rising atmospheric methane: 2007–2014 growth and isotopic shift, *Global Biogeochem. Cycles*, 30, 1356–1370, doi:10.1002/2016GB005406.

O'Connor, F. M., O. Boucher, N. Gedney, C.D. Jones, G.A. Folberth, R. Coppell, P. Friedlingstein, W.J. Collins, J. Chappellaz, J. Ridley, and C.E. Johnson (2010): The possible role of wetlands, permafrost and methane hydrates in the future methane cycle: A review, *Rev. Geophys.*, 48, RG4005, doi:10.1029/2010RG000326.

O'Connor, F. M., C. E. Johnson, O. Morgenstern, N. L. Abraham, P. Braesicke, M. Dalvi, G. A. Folberth, M. G. Sanderson, P. J. Telford, A. Voulgarakis, P. J. Young, G. Zeng, W. J. Collins, and J. A. Pyle (2014): Evaluation of the new UKCA climate-composition model. Part II. The troposphere, *Geosci. Model Dev*., 7, 41-91, doi.org/10.5194/gmd-7-41-2014.

O'Connor, F. M., O. Jamil, T. Andrews, B. T. Johnson, J. P. Mulcahy, and J. Manners (2019): Apportionment of the Pre-Industrial to Present-Day Climate Forcing by Methane using UKESM1, J. Adv. Modeling Earth Sys., Submitted.

Parrish, D. D., J.-F. Lamarque, V. Naik, L. Horowitz, D. T. Shindell, J. Staehelin, R. Derwent, O. R. Cooper, H. Tanimoto, A. Volz-Thomas, S. Gilge, H.-E. Scheel, M. Steinbacher, and M. Fröhlich (2014): Long-term changes in lower tropospheric baseline ozone concentrations: Comparing chemistry-climate models and observations at northern midlatitudes, *J. Geophys. Res. Atmos.*, 119, 5719–5736, doi:10.1002/ 2013JD021435.

Pincus, R., P. M. Forster, and B. Stevens (2016): The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6, *Geosci. Model Dev.*, 9, 3447–3460, www.geosci-model-dev.net/9/3447/2016/.

Pincus, R., Buehler, S. A., Brath, M., Jamil, O., Evans, K. F., Manners, J., Menzel, R. L., Mlawer, E. J., Paynter, D., Pernak, R. L., Benchmark calculations of radiative forcing by greenhouse gases, Geophys. Res. Lett., Submitted, https://doi.org/10.1002/essoar.10501550.1, 2020.

Prather, M. J. (1996): Time scales in atmospheric chemistry: Theory, GWPs for CH4 and CO, and runaway growth, *Geophys. Res. Lett*., 23, 2597–2600, doi:10.1029/96GL02371.

Prather, M. J., Ehhalt, D., Dentener, F., Derwent, R., Dlugokencky, E., Holland, E., Isaksen, I., Katima, J., Kirchoff, V., Matson, P., Midgley, P. and Wang, M.: Atmospheric chemistry and greenhouse gases, in: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K., and Johnson, C. A., Cambridge University Press, Cambridge, UK, 329–287, 2001.

Prather, M. J., and J. Hsu (2010): Coupling of Nitrous Oxide and Methane by Global Atmospheric Chemistry *Science,* Vol. 330, 6006, 952-954, DOI: 10.1126/science.1196285.

Ramaswamy, V., O. Boucher, J. Haigh, D. Hauglustaine, J. Haywood, G. Myhre, T. Nakajima, G. Y. Shi, and S. Solomon (2001), Radiative forcing of climate change, in Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, edited by Y. Ding et al., Cambridge Univ. Press, Cambridge and New York.

Ramaswamy, V., W. Collins, J. Haywood, J. Lean, N. Mahowald, G. Myhre, V. Naik, K.P. Shine, B. Soden, G. Stenchikov, and T. Storelvmo (2019): Radiative Forcing of Climate: The Historical Evolution of the Radiative Forcing Concept, the Forcing Agents and their Quantification, and Applications, Accepted, American Meteorol. Soc. Centenary Monograph, DOI 10.1175/AMSMONOGRAPHS-D-19-0001.1 .

Rap, A., N. A. D. Richards, P. M. Forster, S. A. Monks, S. R. Arnold, and M. P. Chipperfield (2015): Satellite constraint on the tropospheric ozone radiative effect, *Geophys. Res. Lett.*, 42, 5074–5081, doi:10.1002/2015GL064037.

Richardson, T.B., Forster, P.M., Andrews, T., Boucher, O., Faluvegi, G., Fläschner, D., Kasoar, M., Kirkevåg, A., Lamarque, J.F., Myhre, G. and Olivié, D. (2018): Carbon dioxide physiological forcing dominates projected eastern Amazonian drying. *Geophys. Res. Lett.*, *45*(6), pp.2815-2825.

Robertson, E. (2019): The Local Biophysical Response to Land-Use Change in HadGEM2-ES, *J. Climate,* 32, 7611–7627, doi.org/10.1175/JCLI-D-18-0738.1.

Saleh, R., et al. (2014), Brownness of organics in aerosols from biomass burning linked to their black carbon content, *Nat. Geosci.*, doi:10.1038/NGEO2220.

Sellar, A. A., C. G. Jones,, J. Mulcahy, Y. Tang, A. Yool, A. Wiltshire, F. M. O'Connor, M. Stringer, R. Hill, J. Palmieri, S. Woodward, L. de Mora, T. Kuhlbrodt, S. Rumbold, D. I. Kelley, R. Ellis, C. E. Johnson, J. Walton, N. L. Abraham, M. B. Andrews, T. Andrews, A. T. Archibald, S. Berthou, E. Burke, E. Blockley, K. Carslaw, M. Dalvi, J. Edwards, G. A. Folberth, N. Gedney, P. T. Griffiths, A. B. Harper, M. A. Hendry, A. J. Hewitt, B. Johnson, A. Jones, C. D. Jones, J. Keeble, S. Liddicoat, O. Morgenstern, R. J. Parker, V. Predoi, E. Robertson, A. Siahaan, R. S. Smith, R. Swaminathan, M. Woodhouse, G. Zeng, and M. Zerroukat (2019): UKESM1: Description and evaluation of the UK Earth System Model, *J. Adv. Modeling Earth Sys.*, 11, 4513-4558, https://doi.org/10.1029/2019MS001739, 2019.~~doi.org/10.1029/2019MS001739, In press.~~

Sellar, A. A., Walton, J., Jones, C. G., Wood, R., Abraham, N.L., Andrejczuk, M., Andrews, M. B., Andrews, T., Archibald, A. T., de Mora, L., Dyson, H., Elkington, M., Ellis, R., Florek, P., Good, P., Gohar, L., Haddad, S., Hardiman, S. C., Hogan, E., Iwi, A., Jones, C. D., Johnson, B., Kelley, D. I., Kettleborough, J., Knight, J. R., Köhler, M. O., Kuhlbrodt, T., Liddicoat, S., Linova-Pavlova, I., Mizielinski, M. S., Morgenstern, O., Mulcahy, J., Erica Neininger, E., O'Connor, F. M., Petrie, R., Ridley, J., Rioual, J.-C., Roberts, M., Robertson, E., Rumbold, S., Seddon, J., Shepherd, H., Shim, S., Stephens, A., Teixiera, J. C., Tang, Y., Williams, J., Wiltshire, A., and Griffiths, P. T. (2020). Implementation of U.K. Earth system models for CMIP6. J. Adv. Modeling Earth Sys., 12, e2019MS001946. https://doi.org/10.1029/2019MS001946.

Sherwood, S., P. M. Forster, J. Gregory, S. Bony, B. Stevens, and C. Bretherton (2015): Adjustments in the forcing feedback framework for understanding climate change, *Bull. Am. Meteorol. Soc.*, 96, 217–228, doi:10.1175/BAMS-D-13-00167.1.

Shim, S., J. Kim, S.S. Yum, H. Lee, K.-O. Boo, and Y.-H. Byun (2019), Effects of anthropogenic and natural forcings on the summer temperature variations in East Asia during the 20th century, Atmosphere, 10(1), 690, doi:10.3390/atmos10110690.

Shindell, D. T., G. Faluvegi, N. Bell, and G. A. Schmidt (2005): An emissions-based view of climate forcing by methane and tropospheric ozone, *Geophys. Res. Lett*., 32, L04803, doi:10.1029/2004GL021900.

Shindell, D. T., Faluvegi, G., Koch, D. M., Schmidt, G. A., Unger, N., and Bauer, S. E. (2009): Improved attribution of climate forcing to emissions, *Science*, 326, 716–718, doi:10.1126/science.1174760.

Shindell, D., G. Faluvegi, L. Nazarenko, K. Bowman, J.-F. Lamarque, A. Voulgarakis, G. A. Schmidt, O. Pechony and R. Ruedy (2013a): Attribution of historical ozone forcing to anthropogenic emissions, *Nature Climate Change*, 3, pp. 567–570.

Shindell, D. T., Pechony, O., Voulgarakis, A.., Faluvegi, G., Milly, G., Kovari, B., Ruedy, R., and G. A. Schmidt (2013b): Interactive ozone and methane chemistry in GISS-E2 historical and future climate simulations, *Atmos. Chem. Phys*., 13, 2653–2689, doi:10.5194/acp-13-2653-2013.

Shine K. P., B. P. Briegleb, A. S. Grossman, D. Hauglustaine, Huiting Mao, V. Ramaswamy, M. D. Schwarzkopf, R. Van Dorland, W.-C. Wang (1995): Radiative forcing due to changes in ozone: a comparison of different codes. In: Wang W.-C., Isaksen I.S.A. (eds) Atmospheric Ozone as a Climate Gas. NATO ASI Series (Series I: Global Environmental Change), vol 32. Springer, Berlin, Heidelberg.

Shine, K. P., and P. M. de F. Forster (1999): The effect of human activity on radiative forcing of climate change: a review of recent developments, *Global Planet. Change*, 20, 1999, 205–225.

Sitch, S., Cox, P. M., Collins, W. J., and Huntingford, C. (2007): Indirect radiative forcing of climate change through ozone effects on the land-carbon sink, *Nature*, 448, 791–794, doi:10.1038/nature06059.

Skeie, R. B., T. K. Berntsen, G. Myhre, K. Tanaka, M. M. Kvalevåg, and C. R. Hoyle (2011): Anthropogenic radiative forcing time series from pre-industrial times until 2010, *Atmos. Chem. Phys.*, 11, 11827–11857, doi:10.5194/acp-11-11827-2011.

Skeie, R. B., Myhre, G., Hodnebrog, Ø., Cameron-Smith, P. J., Deushi, M., Hegglin, M. I., Horowitz, L. W., Kramer, R. J., Michou, M., Mills, M. J., Olivié, D. J. L., O'Connor, F. M., Paynter, D., Samset, B. H., Sellar, A., Shindell, D., Takemura, T., Tilmes, S., and Wu, T., Historical total ozone radiative forcing derived from CMIP6 simulations, npj Climate Atmos. Sci., Accepted, 2020.

Smith, C. J., R. J. Kramer, G. Myhre, G., P. M. Forster, B. J. Soden, T. Andrews, O. Boucher, G. Faluvegi, D. Fläschner, Ø. Hodnebrog, M. Kasoar, V. Kharin, A. Kirkevåg, J.-F. Lamarque, J. Mülmenstädt, D. Olivié, T. Richardson, B. H. Samset, D. Shindell, P. Stier, T. Takemura, A. Voulgarakis, and D. Watson-Parris (2018): Understanding rapid adjustments to diverse forcing agents, *Geophys. Res. Lett*., 45, 12023–12031, doi.org/10.1029/2018GL079826.

Soden, B.J., I.M. Held, R. Colman, K.M. Shell, J.T. Kiehl, and C.A. Shields (2008): Quantifying Climate Feedbacks Using Radiative Kernels. *J. Climate,* 21, 3504–3520, doi.org/10.1175/2007JCLI2110.1.

Stevenson, D. S., P. J. Young, V. Naik, J.-F. Lamarque, D. T. Shindell, A. Voulgarakis, R. B. Skeie, S. B. Dalsoren, G. Myhre, T. K. Berntsen, G. A. Folberth, S. T. Rumbold, W. J. Collins, I. A. MacKenzie, R. M. Doherty, G. Zeng, T. P. C. van Noije, A. Strunk, D. Bergmann, P. Cameron-Smith, D. A. Plummer 13 , S. A. Strode, L. Horowitz , Y. H. Lee , S. Szopa , K. Sudo , T. Nagashima , B. Josse, I. Cionni, M. Righi, V. Eyring , A. Conley , K. W. Bowman , O. Wild, and A. Archibald (2013): Tropospheric ozone changes, radiative forcing and attribution to emissions in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), *Atmos. Chem. Phys*., 13, 3063–3085, doi:10.5194/acp-13-3063-2013.

Stjern, C. W., Samset, B. H., Myhre, G., Forster, P. M., Hodnebrog, Ø. Andrews, T., O. Boucher, G. Faluvegi, T. Iversen, M. Kasoar, V. Kharin, A. Kirkevåg, J.-F. Lamarque, D. Olivié, T. Richardson, D. Shawki, D. Shindell, C. J. Smith, T. Takemura, and A. Voulgarakis (2017): Rapid adjustments cause weak surface temperature response to increased black carbon concentrations, *J. Geophys. Res.: Atmos.*, 122, 11462– 11481, doi.org/10.1002/2017JD027326.

Søvde, O., C. Hoyle, G. Myhre, and I. Isaksen (2011): The HNO3 forming branch of the HO2 + NO reaction: Pre-industrial-to-present trends in atmospheric species and radiative forcings, *Atmos. Chem. Phys.*, 11, 8929–8943.

Søvde, O., C. Hoyle, G. Myhre, and I. Isaksen (2012): Corrigendum to "The HNO3 forming branch of the HO2 + NO reaction: pre-industrial-to-present trends in atmospheric species and radiative forcings" published in Atmos. Chem. Phys., 11, 8929–8943, 2011, Atmos. Chem. Phys., 12, 7725–7725, https://doi.org/10.5194/acp-12-7725-2012, 2012.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012): An Overview of CMIP5 and the Experiment Design, *B. Am. Meteorol. Soc*., 93, 485–498.

Teixeira, J. C., et al. (2019): Resampling Hypothesis for Testing Reproducibility of Earth System models, In preparation.

Thomason, L. W., N. Ernest, L. Millán, L. Rieger, A. Bourassa, J.-P. Vernier, G. Manney, B. Luo, F. Arfeuille, and T. Peter (2018): A global space-based stratospheric aerosol climatology: 1979–2016, *Earth Syst. Sci. Data*, 10, 469–492, doi.org/10.5194/essd-10-469-2018.

~~Thornhill, G., et al. (2019): Effective Radiative forcing from emissions of reactive gases and aerosols, In preparation.~~
Thornhill, G. D., Collins, W. J., Kramer, R. J., Olivié, D., O'Connor, F., Abraham, N. L., Bauer, S. E., Deushi, M., Emmons, L., Forster, P., Horowitz, L., Johnson, B., Keeble, J., Lamarque, J.-F., Michou, M., Mills, M., Mulcahy, J., Myhre, G., Nabat, P., Naik, V., Oshima, N., Schulz, M., Smith, C., Takemura, T., Tilmes, S., Wu, T., Zeng, G., and Zhang, J.: Effective Radiative forcing from emissions of reactive gases and aerosols – a multimodel comparison, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-1205, in review, 2020.

Toll, V., Christensen, M., Gassó, S., and Bellouin, N. (2017): Volcano and ship tracks indicate excessive aerosol-induced cloud water increases in a climate model, *Geophys. Res. Lett.*, 44, 12492–12500, doi.org/10.1002/2017GL075280.

Turnock, S., O. Wild, F. Dentener, Y. Davila, L. Emmons, J. Flemming, G. Folberth, D. Henze, J. Jonson, T. Keating, S. Kengo, M. Lin, M. Lund, S. Tilmes, and F. M. O'Connor (2018):The Impact of Future Emission Policies on Tropospheric Ozone using a Parameterised Approach, *Atmos. Chem. Phys*., 18, 8953–8978, doi.org/10.5194/acp-18-8953-2018.

Turnock, S. T., Wild, O., Sellar, A., and O'Connor, F. M. (2019), 300 years of tropospheric ozone changes using CMIP6 scenarios with a parameterised approach, *Atmos. Environ.,* 213, 686-698, doi.org/10.1016/j.atmosenv.2019.07.001.

Twomey, S. (1977): The Influence of Pollution on the Shortwave Albedo of Clouds. *J. Atmos. Sci.,* 34, 1149–1152, https://doi.org/10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2.

Voulgarakis, A., Naik, V., Lamarque, J.-F., Shindell, D. T., Young, P. J., Prather, M. J., Wild, O., Field, R. D., Bergmann, D., Cameron-Smith, P., Cionni, I., Collins, W. J., Dalsøren, S. B., Doherty, R. M., Eyring, V., Faluvegi, G., Folberth, G. A., Horowitz, L. W., Josse, B., MacKenzie, I. A., Nagashima, T., Plummer, D. A., Righi, M., Rumbold, S. T., Stevenson, D. S., Strode, S. A., Sudo, K., Szopa, S., and Zeng, G. (2013): Analysis of present day and future OH and methane lifetime in the ACCMIP simulations, *Atmos. Chem. Phys.*, 13, 2563–2587, doi.org/10.5194/acp-13-2563-2013.

Walters, D., A. J. Baran, I. Boutle, M. Brooks, P. Earnshaw, J. Edwards, K. Furtado, P. Hill, A. Lock, J. Manners, C. Morcrette, J. Mulcahy, C. Sanchez, C. Smith, R. Stratton, W. Tennant, L. Tomassini, K. Van Weverberg, S. Vosper, M. Willett, J. Browse, A. Bushell, K. Carslaw, M. Dalvi, R. Essery, N. Gedney, S. Hardiman, B. Johnson, C. Johnson, A. Jones, C. Jones, G. Mann, S. Milton, H. Rumbold, A. Sellar, M. Ujiie, M. Whitall, K. Williams, and M. Zerroukat (2019), The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, *Geosci. Model Dev*., 12, 1909–1963, doi.org/10.5194/gmd-12-1909-2019.

Wild, O., A. Voulgarakis, F. O'Connor, J.-F. Lamarque, E. M. Ryan, and L. Lee (2019): Global sensitivity analysis of chemistry-climate model budgets of tropospheric ozone and OH: Exploring model diversity, *Atmos. Chem. Phys. Disc.*, doi.org/10.5194/acp-2019-774.

Wilks, D.S. (1997): Resampling hypothesis tests for autocorrelated fields, *J. Climate*, 10(1), 65-82.

Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., P. Davis, T. Graham, H. T. Hewitt, R. Hill, P. Hyder, S. Ineson, T. C. Johns, A. B. Keen, R. W. Lee, A. Megann, S. F. Milton, J. G. L. Rae, M. J. Roberts, A. A. Scaife, R. Schiemann, D. Storkey, L. Thorpe, I. G. Watterson, D. N. Walters, A. West, R. A. Wood, T. Woollings, and Xavier, P. K. (2017): The Met Office Global Coupled model 3.0 and 3.1 (GC3.0 and GC3.1) configurations, *J. Adv. Modeling Earth Systems*, 10, 357–380, doi.org/10.1002/2017MS001115.

Winterstein, F., F. Tanalski, P. Jöckel, M. Dameris, and M. Ponater (2019): Implication of strongly increased atmospheric methane concentrations for chemistry–climate connections, *Atmos. Chem. Phys.*, 19, 7151–7163, doi.org/10.5194/acp-19-7151-2019.

Xia, Y., Hu, Y., and Huang, Y. (2016): Strong modification of stratospheric ozone forcing by cloud and sea-ice adjustments, *Atmos. Chem. Phys.,* 16, 7559–7567.

Xing, J., J. Wang, R. Mathur, S. Wang, G. Sarwar, J. Pleim, C. Hogrefe, Y. Zhang, J. Jiang, D. Wong, and J. Hao (2017): Impacts of aerosol direct effects on tropospheric ozone through changes in atmospheric dynamics and photolysis rates, *Atmos. Chem. Phys*., 17, 9869-9883, doi:10.5194/acp-17-9869-2017.

Yeung, L.Y., Murray, L.T., P. Martinerie, E. Witrant, H. Hu, A. Banerjee, A. Orsi, and J. Chappellaz (2019): Isotopic constraint on the twentieth-century increase in tropospheric ozone. *Nature,* 570, 224–227, doi:10.1038/s41586-019-1277-1.

Young, P. J., A. T. Archibald, K. W. Bowman, J.-F. Lamarque, V. Naik, D. S. Stevenson, S. Tilmes, A. Voulgarakis, O. Wild, D. Bergmann, P. Cameron-Smith, I. Cionni, W. J. Collins, S. B. Dalsøren, R. M. Doherty, V. Eyring, G. Faluvegi, L. W. Horowitz, B. Josse, Y. H. Lee, I. A. MacKenzie8 , T. Nagashima, D. A. Plummer, M. Righi, S. T. Rumbold, R. B. Skeie, D. T. Shindell, S. A. Strode, K. Sudo, S. Szopa, and G. Zeng (2013): Pre-industrial to end 21st century projections of tropospheric ozone from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), *Atmos. Chem. Phys.*, 13, 2063-2090, doi:10.5194/acp-13-2063-2013.

Young, P. J., Naik, V., Fiore, A. M., Gaudel, A., Guo, J., Lin, M. Y., Neu, J. L., Parrish, D. D., Rieder, H. E., Schnell, J. L., Tilmes, S., Wild, O., Zhang, L., Ziemke, J. R., Brandt, J., Delcloo, A., Doherty, R. M., Geels, C., Hegglin, M. I., Hu, L., Im, U., Kumar, R., Luhar, A., Murray, L., Plummer, D., Rodriguez, J., Saiz-Lopez, A., Schultz, M. G., Woodhouse, M. T. and Zeng, G. (2018): Tropospheric Ozone Assessment Report: Assessment of global-scale model performance for global and regional ozone distributions, variability, and trends, *Elem. Sci. Anth.*, 6(1), p.10, doi.org/10.1525/elementa.265.

Zelinka, M. D., Andrews, T., Forster, P. M., and Taylor, K. E. (2014): Quantifying components of aerosol-cloud-radiation interactions in climate models, *J. Geophys. Res. Atmos.*, 119, 7599– 7615, doi:10.1002/2014JD021710.