We would like to thank both reviewers for their detailed and constructive reviews. In addition to the changes requested by the reviewers, explained in the responses below, in the revised paper the following additions to the manuscript have been made:

- Inclusion of 4 additional models: ACCESS-CM2, EC-Earth3, GFDL-ESM4 and NorESM2-MM
- A new version of UKESM1-0-LL has been uploaded to ESGF which supersedes the old version. This mostly affects the land use and anthropogenic experiments.
- A new version of GISS-E2-1-G r1i1p1f1 that corrects a bug in the old version is uploaded to ESGF, and results are in line with other models. See responses to comments from reviewer #2.
- Cloud adjustments in the LW using the offline method are better tuned to observed TOA fluxes and now use all available years (except for the first few years described in table 1). For some models, adjustments and IRF have changed slightly in tables 3, 4, 5, 7 and 8.
- Inclusion of the GISS-E2-1-G r1i1p3f1 model variant for the aerosol forcing experiment as a separate model. This physics setup exhibits very different behaviour to the r1i1p1f1 variant as cloud droplet number is calculated prognostically rather than specified offline. In addition, atmospheric chemistry and aerosols are prognostic. The p3 version was used to prescribe composition and aerosols in the p1 version.
- Table 1: inclusion of “residual” column, representing anthropogenic minus (WMGHG + aerosols + land-use). The largest contribution to this residual is likely to be from ozone, to the extent that the forcing components add linearly.
- ERF_reg is figure 1 is now calculated using the first 20 years of the abrupt-4xCO2 experiment. This is in line with the original definition from Gregory et al. (2004). The 150-year regression values are given in supplementary table S1, along with the other methods of calculating ERF in Figure 1 for all experiments in tables S1 to S5.
- Adjustments to the 4xCO2 experiment, not scaled down to present day concentrations, are given in table S6.

There is one last minute change following the posting of the Author Comments to the reviews. We have had confirmation from EC-Earth3 that their model setup was not specifically tuned for surface temperature. The comments below and the updated manuscript reflect this.

Anonymous Referee #1:

General Comments:

The paper presents a comprehensive analysis of the diagnosed values of effective radiative forcing (ERF) for the CMIP6 models, and breaks down the contributions of this forcing from greenhouse gases, aerosols, and land-use. The use of ERF has continued to grow and it is now at least as widely-used, if not more so, than traditional metrics of forcing such as instantaneous forcing or stratospheric-adjusted forcing. It is an important paper for benchmarking the performance of CMIP6 models, and its

findings will hopefully be used in upcoming assessment reports. That being said, there are many different findings in this paper and it would likely be more digestible if the findings were explored in more detail in separate papers. This is something of an omnibus paper. However, despite its size, the paper has important findings and only requires minor revisions before being acceptable for publication.

Thank you for your positive comments on the paper. We agree that it is a long paper and we appreciate your time and effort given to the thorough review.

It is always a philosophical point whether to split papers such as this into more than one volume. We were of the belief that just reporting the headline fixed-SST ERF results from the CMIP6 ensemble "as-is" did not constitute enough analysis to merit a paper on its own, although these figures will possibly be the most widely-used. I felt that Smith et al. (2018b) would have benefitted from the regional analysis, so I was keen to include it in this submission. It sounds like this motivated interesting questions from both reviewers, so was a useful addition, at the expense of a longer paper.

Specific Comments:

The last point made by the authors in the abstract, which appears to be supported in Figure 8, namely that they see no evidence that aerosols are contributing to the spread in ECS, is striking and bears more discussing. The range and change in ECS for CMIP6 is and should be of great concern to the large numbers of individuals involved in CMIP6 and to the scientific community as a whole, since an explanation is required. I hope that a body of literature will emerge (and quickly) to develop this explanation, and to the extent that this paper can contribute to that body, it is important that the lack of correlation between present-day aerosol forcing and ECS is promulgated. Is it fair to say, then, that the mystery of CMIP6 ECS persists or, perhaps, deepens?

Thank you for picking out this key point which we could highlight more. The last sentence of the abstract has been extended:

> Therefore, there is no evidence to suggest that the increasing spread in climate sensitivity in CMIP6 models, particularly related to high-sensitivity models, is a consequence of a stronger negative present-day aerosol forcing, and little evidence that modelling groups are systematically tuning climate sensitivity or aerosol forcing to recreate observed historical warming.

From the model description papers cited in table 1, with one exception ~~two exceptions~~ (MPI-ESM1-2~~ and EC-Earth3~~), there is either an explicit mention that historical temperatures were not used as a model performance indicator for their CMIP6 configuration, or the model paper was silent on this. Indeed it is evident from a number of models' historical temperature evolutions that they were not a target of model tuning:

- CanESM5 (high climate sensitivity, moderately low present-day aerosol forcing) shows more warming than observations over the historical period (Swart et al., 2019);
- NorESM2-LM and NorESM2-MM (lower Gregory sensitivity, stronger aerosol forcing) shows less warming than observed (Seland et al., 2020). (The NorESM2 models actually have high equilibrium sensitivity but low sensitivity as measured from a 150-year Gregory regression due to strongly increasing feedbacks over time).
- UKESM1-0-LL (high sensitivity, about average aerosol forcing) has approximately the correct level of present-day warming but is too cool in the 1960-2000 period (Sellar et al., 2019).

A footnote was added in section 5.3.2:

> MPI-ESM1-2 (Mauritsen et al., 2019) ~~and EC-Earth3 (Wyser et al., 2019) are~~ is the only documented exception~~s~~. MIROC6 (Tatebe et al., 2019) did tune the aerosol forcing to better correspond to the AR5 best estimate but explicitly did not tune for surface temperature.

While the magnitude of year-2014 aerosol ERF may not constrain climate sensitivity, potentially the evolution of aerosol forcing since 1970 may provide some constraint on transient climate response. As different models include different aerosol processes, the time history of aerosol forcing can be quite different in different models even when driven with the same emissions. 6 models so far have performed the RFMIP Tier 2 aerosol forcing transient experiment and future work will investigate this.

The paper notes that the spread in ERF between models is narrowed relative to CMIP5. This is a most welcome finding, given the poor specification of forcings in CMIP5. I recommend that the paper indicate that the result is consistent with a high-level recommendations from the Stouffer et al, 2017 paper (doi: 10.1175/BAMS-D-15-00013.1).

Thank you for the suggestion. In the Conclusion (line 491) we add a sentence:

> This has helped to address a concern from CMIP5: that forcing was poorly characterised in CMIP5 models and inconsistently determined (Stouffer et al., 2017).

We should caveat, and we discuss later in the paragraph, that there are more models in CMIP6 that did not submit the RFMIP Tier 1 experiments in which aerosol forcing is probably stronger than the lower bound from the 17 models for which we have data:

> Although 17 models is a reasonable sample size of the CMIP6 population, more models may submit forcing results to CMIP6 that would widen this range (and indeed, we would encourage modelling groups to do so). One example is E3SM which did not perform the RFMIP aerosol forcing experiment but where it would be likely that the 1850--2014 aerosol forcing would be more negative than -1.37 W m^{-2} (fig. 25 in Golaz et al., 2019).

This may also have been true in CMIP5 in which the sstClimAerosol experiment was from a subset of models.

The narrowing of the range in aerosol forcing is particularly notable and welcome. That being said, the authors should point out in the abstract the importance of the aerosol forcing adjustments and large range in model results, especially with respect to clouds. The finding is included in the paper already and is notable in that it highlights challenges for the scientific community that studies aerosol-cloud interactions.

In combination with one of your comments further below, this sentence has been added to the abstract:

> In most cases, the largest contributors to the spread in ERF is from the instantaneous radiative forcing (IRF) and from cloud responses, particularly aerosol-cloud interactions to aerosol forcing.

Along with the previous comment, in order to more clearly illustrate comparisons with CMIP5 which were discussed in the text but not graphically compared in the first submission, a new Figure 5 has been added which detail these comparisons, clearly showing the increased $4xCO_2$ forcing in CMIP6 and slightly reduced range of aerosol forcing.

The limited importance of land use for forcing is surprising, and the spatial patterns there appears to be strong, with some overlap with aerosol forcing. Is there cancellation or reinforcement for these effects?

To attempt to explain this further, the change in aerosol optical depth in the nine models that output this diagnostic in the land-use experiment is shown in fig. S4. Aerosol-induced changes in some models are notable but not large. For example CanESM5 has the greatest increase in aerosol optical depth over the Northern Hemisphere land regions but is relatively weak in terms of land use forcing. The exception to this is for the models that include ice nucleation effects from biogenic aerosol which is coupled to the land surface scheme. This is clearly the case for NorESM2-LM, where the cloud adjustment dominates the forcing and results in a positive land-use change ERF, and is already documented in fig. S3. All of the other 13 models have a negative land use forcing, which are in line with an observationally-constrained estimate from CMIP5 models (Lejeune et al., 2020).

As you assert, regionally land-use change is important. In regions experiencing lots of deforestation (North America, Western Eurasia and South America) albedo is increased causing a negative ERF (fig. 12). In these regions land-use forcing determines the multi-model mean ERF, however, while models agree on the negative land-use ERF, there is no model consensus on the net forcing (fig. 13). Compared to land-use change, aerosols cause a slightly weaker negative forcing in the deforested regions, which in the northern hemisphere is caused by an increase in SW reflectance (fig. 8). The aerosol effect will reduce downward SW at the surface and so reduce the effect of surface albedo changes caused by land-use change. The cause of the negative aerosol ERF in the South American

deforested region is more complex and here there is no obvious non-linearity in the combined land-use and aerosol ERF.

It should also be noted that land-use change does not only impact climate via radiative forcing and that the temperature impacts of other mechanisms are often larger and of opposite sign (e.g. Bright et al., 2017).

The final point made by the authors in the conclusion, which is that there is a need to constrain cloud responses to forcing since they contribute to the largest uncertainty in forcing, is well-taken but disturbing. Clouds appear to be not just a problem for feedbacks, as is widely accepted by the community and has motivated a sustained focus on constraining cloud feedbacks, but they are a problem for forcing as well. This point should also be in the abstract and discussed in the abstract.

From tables 3, 4, 5, 7 and 8 it is apparent the largest contributions to the spread in ERF are from IRF and cloud adjustments. In addition to improving cloud processes, there is still some way to go in radiative transfer modelling. Work is in progress under RFMIP to do this, and we show in Figure 5 for 4xCO2 that the inter-model spread is much reduced compared to CMIP5 which may be indicative of an improvement in model radiative transfer.

In conclusion:

> The instantaneous radiative forcing and cloud adjustments are generally the largest sources of inter-model spread in the forcing component in climate models. Since IRF is not directly calculated in this study, some of this spread may be from residuals in the kernel decomposition and the true spread in IRF may be smaller than reported here. One strand of RFMIP will include benchmarking of GCM radiative transfer against line-by-line codes. Radiative transfer is a well-grounded theoretical problem where the diversity in line-by-line codes is small (Pincus et al., 2015), so this component of inter-model diversity has a measurable yardstick for improvement.

However, the recommendation of the authors is vague and it is highly unclear to me how on how cloud responses can be constrained. Through process studies? Developing observational constraints? There are strikingly strong spatial patterns of ERF. Can some type of fingerprinting be used? The authors should indicate in the paper whether or not there even is a path forward for actually constraining these cloud responses or if the community needs to develop one before even being able to go down it to actually develop those constraints.

We agree that the last sentence of the conclusion was vague in the first submission. We have added some suggestions at the end of the conclusion (following on from the passage above) and a final summary sentence that links back to the introduction.

> Cloud responses are more difficult to constrain and exhibit a wide range of behaviour to both greenhouse gas and aerosol forcing. However, progress is beginning to be made. For greenhouse gas forcing, techniques from the climate feedback literature

> that have observational parallels, such as analysing cloud-controlling factors (Klein et al., 2017), can be applied to adjustments. Use of the ISCCP simulator diagnostics with the ISCCP cloud kernel, another method conceptualised by climate feedback investigations (Zelinka et al., 2012), allows cloud adjustments to be calculated directly facilitating better inter-model comparison. For aerosol forcing, observational methods exist to determine RFari and RFaci using satellite and reanalysis data (Bellouin et al., 2013; Bellouin et al., 2020a). Ultimately, reducing uncertainty in effective radiative forcing will reduce uncertainty in climate projections due to the central role of forcing in driving the Earth's global mean temperature response.

Minor points:

The x-axes on Figure 4 need fixing.

Rotated x-axis labels to make clearer.

Figure 5 has lots of information but is confusing in there is concurrence between models in the spatial patterns of ERF but there appears to be little concurrence in some of the spatial patterns of adjustments and cloud contributions, even though when summed up, they are significant across models. This is even more the case for Figures 7 and 11, and some explanation of how this is achieved is needed for readers.

You make a good point here. Forcing adjustments are in many cases robust in sign in the global mean change but less so spatially between models. This highlights the point that forcing and adjustments are best considered globally averaged quantities. At the end of section 5 (line 481) we have added the following:

> For all forcings, but particularly for land-use, aerosol and total anthropogenic, many of the forcing and adjustment terms do not show robust signals regionally. This indicates that adjustments are best considered as global-mean quantities that affect the globally-resolved forcing-feedback framework (Eq (1)).

Line 386: Should be "equivalent" not "equalivent"

Typo corrected - thank you.

References used in response to reviewer #1 not in manuscript:

Bright, R. M., Davin, E., O'Halloran, T., Pongratz, J., Zhao, K. and A. Cescatti, 2017: Local temperature response to land cover and management change driven by non-radiative processes. Nat. Climate Change, 7, 296-302, https://doi.org/10.1038/nclimate3250.

Anonymous Referee #2:

This paper investigates the effective radiative forcing (ERF) from 13 CMIP6 models, and contributes to the RFMIP project. It presents contributions of particular climate forcers to anthropogenic forcing, including greenhouse gases, aerosols, and land-use. Results show a smaller anthropogenic ERF compared to AR5, and it is contributed by a stronger aerosol ERF. Additionally, the range of aerosol ERF from CMIP6 is narrower than CMIP5. This work introduces a range of methods to calculate ERF and adjustments as well. It is certainly a very comprehensive work and would make a valuable contribution to IPCC next assessment report. However, I feel there still could be some more interpretations of the work presented here. For these reasons, I am recommending this paper to be accepted for publications with minor revisions.

Thank you for your thorough review of this paper and positive comments. We try to address the further interpretations that you mention in the comments below.

General comments:

1. It is interesting to see the range of aerosol ERF is narrower in CMIP6. However, it could be better if the authors can demonstrate which part (e.g., ERFari or ERFaci) contributes to the improvement most, and why is it.

Table 6 shows the ERFari and ERFaci components, which can be compared with Zelinka et al. (2014) for CMIP5 models. The standard deviation of model estimates has reduced for ERFari, ERFaci and ERFari+aci (0.19 W m^{-2}, 0.30 W m^{-2} and 0.20 W m^{-2} respectively compared to 0.22 W m^{-2}, 0.34 W m^{-2} and 0.30 W m^{-2} for CMIP5 in Zelinka et al. (2014)). The reduced standard deviation is a consequence of a slightly smaller spread in each of the components with more models. It is not clear that one component of the aerosol forcing has shown a greater reduction in spread than the other between CMIP5 and and CMIP6. However, in aid of easier comparison of the total aerosol forcing (and CO_2), we have included a new figure 5 that compares CMIP5 and CMIP6.

We have been careful to point out that the models submitting results to Tier 1 of RFMIP are only a subset of all CMIP6 models, and inclusion of more models could extend the range of aerosol ERF. In the conclusion we have added:

> Although 17 models is a reasonable sample size of the CMIP6 population, more models may submit forcing results to CMIP6 that would widen this range (and indeed, we would encourage modelling groups to do so). One example is E3SM which did not perform the RFMIP aerosol forcing experiment but where it would be likely that the 1850--2014 aerosol forcing would be more negative than -1.37 W m^{-2} (fig. 25 in Golaz et al., 2019).

This may also have been true in CMIP5 in which the sstClimAerosol experiment was from a subset of 10 models.

2. The Introduction part could add some a short paragraph to talk about the contribution of aerosols, GHGs, and land-use to anthropogenic ERF, in terms of sign and magnitude. For example, how aerosols ERF counteracts a large part of the warming effect from GHGs meanwhile has the largest uncertainty.

We agree that this would give the study some additional context. At the end of the first paragraph we have included the following:

> Since the start of the Industrial Era until the present-day, anthropogenic forcing has typically been increasing, and has been the dominant component of the total forcing on the Earth system except for brief periods following large volcanic eruptions (Myhre et al., 2013). The main constituents of anthropogenic ERF are a positive forcing from greenhouse gases and a partially offsetting negative forcing from aerosols. While greenhouse gas forcing is reasonably well-known, aerosol forcing is more uncertain due to the spatial variation of aerosols, their short atmospheric lifetime, and their complex interactions with clouds (Boucher et al., 2013; Bellouin et al., 2020b).

3. Fig 8: Not fully understand why do the correlation between aerosol ERF and ECS/TCR. According to the definition, ECS and TCR are directly related to CO2, so it won't be a surprise to me that the correlation is bad. Can you give more explanations here?

One goal of researchers analysing CMIP6 models is to try and understand the drivers of increased climate sensitivity compared to CMIP5, and whether the very high sensitivity models are realistic and under what circumstances. If realistic, this means that high sensitivity cannot be ruled out, which is an important result for policy communication. In order to reproduce historically-observed warming a high ECS and high TCR requires a strongly negative aerosol forcing. I wasn't the first to do this but this is shown in an energy balance framework with a large probabilistic ensemble in Figure 7 of Smith et al. (2018a). In coupled models, Kiehl (2007) showed that ECS and aerosol forcing were negatively correlated in CMIP3, as well as in CMIP5 for models that included an aerosol-cloud interaction (Chylek et al., 2016) but not for the model population as a whole (Forster et al., 2013). The lack of correlation in CMIP6 suggests that climate modelling groups are not using historical warming observations as a model tuning constraint, an assertion that is on the whole verified from a review of available CMIP6 model description papers listed in table 1 except for the ~~two~~ counterexample~~s~~ specifically noted (footnote on page 22). It does mean we cannot constrain climate sensitivity using model-diagnosed aerosol forcing in the present day.

You touch on a relevant point here on the contribution of CO_2 forcing to climate sensitivity. As Zelinka et al. (2020) showed, the increase in (150-year Gregory) CO_2 forcing is a small but substantial contributor to the increase in climate sensitivity. We also show this to be the case in the fixed-sea surface temperature definition of ERF and this is highlighted in Figure 5.

4. This paper provides a number of methods to examine ERF from different climate forcers by using several climate models. It is certainly a very comprehensive work. However, it is easy to get lost when I am trying to understand the results. It would be of interest if some further work can be done to help the audience to better understand the results (not necessarily in this paper). For example, the adjustment from clouds contributes to most of the uncertainties. Are these uncertainties caused by different methods or different models? If it is caused by model variability, then what are the essential parameterization of clouds been used in these models? The geographical patterns shown in Fig 5, 6, 7, 11, and 12 are interesting, and it would be nice if the authors can explore more on them.

Addressing the point of ERF methods first, there is some discussion as to how to best define effective radiative forcing with the goal that it should be a convertible currency for measuring long-term surface temperature changes from different forcers. This was motivated originally by Hansen et al. (2005) and discussed recently in Tang et al. (2019) and Richardson et al. (2019), where correcting for the land surface warming improves the ERF to temperature relationship slightly, so we provide the alternatives in Figure 1. Although it extends the discussion slightly, it puts these alternative ERF definitions in the open. We focus the rest of the paper on the fixed SST results, one reason being the separation into adjustments is difficult or ill-defined for some of the other methods.

Addressing the point on cloud uncertainties, this is a very good suggestion for future work. One question that was raised when I presented this at the virtual EGU conference was whether there was any relationship between the complexity of cloud parameterisation and forcing, and it is one that we do not know the answer to. For climate sensitivity results, a more realistic cloud water phase parameterisation has led to an increase in sensitivity in CESM2 (Gettleman et al., 2019), and I believe also UKESM1-0-LL (compared to HadGEM2-ES). It would be interesting to determine whether this also applies to adjustments. At the request of the other reviewer, we have added some more explanation near the end of the discussion around how cloud changes could possibly be constrained:

> Cloud responses are more difficult to constrain and exhibit a wide range of behaviour to both greenhouse gas and aerosol forcing. However, progress is beginning to be made. For greenhouse gas forcing, techniques from the climate feedback literature that have observational parallels, such as analysing cloud-controlling factors (Klein et al., 2017), can be applied to adjustments. Use of the ISCCP simulator diagnostics with the ISCCP cloud kernel, another method conceptualised by climate feedback investigations (Zelinka et al., 2012), allows cloud adjustments to be calculated directly facilitating better inter-model comparison. For aerosol forcing, observational methods exist to determine RFari and RFaci using satellite and reanalysis data (Bellouin et al., 2013; Bellouin et al., 2020a).

Specific comments:

1. Figure 2: maybe put this figure in the supplementary file? It is an interesting figure in terms of methodology, but not very necessarily related to the story and may distract readers.

We can see arguments for and against moving this figure to the Supplement and agree with the reviewer that it is probably not central to the discussion of forcing. However, we believe that it is an important point to highlight to others. We have already in fact obtained one citation for this paper by pointing out that the forcing behaviour of CNRM-ESM2-1 is different to other models (Williams, Ceppi & Katavouta, accepted in Environmental Research Letters), so believe it is a useful reference for the community.

2. Figure 3: GISS-E2-1-G is acting very differently to other models, especially on adjustments from aerosols. Why is that?

An error was discovered in the first version on ESGF of the GISS-E2-1-G 4xCO_2 simulation. This has now been corrected on ESGF and the ERF and adjustments have changed for this model and in line with other 4xCO_2 experiments. Several parts of the text discussing the old treatment have been modified or deleted.

3. Fig 4, x axis: It is hard to read the rightmost labels as they are overlaid.

Rotated x-axis labels to make distinction clearer.

4. Line 276 and the following paragraph: "This effect is traced to a slight cooling in the mid-troposphere in this model whereas other models show a distinct warming." It is interesting, but why GISS-E2-1-G shows a cooling which is apparently different from other models.

As referenced above, an error in the GISS model was corrected and this tropospheric cooling no longer exists. This paragraph has been deleted.

5. Line 331: "Atmospheric adjustments are small in magnitude in the aerosol forcing experiment, but large enough such that there is a noticeable difference between ERF and RF." I assume this conclusion is derived from table 5?

This follows from table 5 (as the sum of IRF and ta_st) and figure 1. The table does not list RF separately. This has now been done for all experiments, and added to the Supplement as tables S1 to S5. A reference to figure 1 and the supplementary table S3 has been inserted.

6. Line 338 and the following paragraph: I agree with the explanations. However, it is possible that absorbing aerosols play a minor role compared to non-absorbing aerosols just due to the smaller BC emissions than sulphate?

Good point, also shown by the magnitude of the individual forcings in AerChemMIP (Thornhill et al., 2020). The text has been updated as follows:

> This also implies that absorbing aerosols play only a minor role in most models, as BC induces strong adjustments that cause a general increase in cloud height in PDRMIP models from an increasing tropospheric stability (Smith et al. (2018b); Stjern et al. (2017); fig. S2). There is no evidence of this in the RFMIP aerosol forcing experiment, although some models do also include aerosol-cloud interactions from BC, and the effect may be due to the BC forcing being a smaller fraction of the total aerosol forcing than sulfate (Thornhill et al., 2020).

7. Line 410: Why LW ERFari+aci from the double call method doesn't always equal the total ERF? According to equation 8 and 9, it should be closed. And. Additionally, is this only for LW or both SW and LW?

This section has been simplified a little. Going back to the original Ghan (2013) reference for the aerosol forcing double call, the aerosol ERF can be broken down into direct radiative forcing (RFari in AR5 terminology), cloud radiative forcing (ERFaci plus the semi-direct effect AKA the adjustment part of ERFari) and surface albedo adjustment. In this case, the breakdown given in the paper has slightly incorrect terminology (see equations below) but it is the same bias as for APRP as discussed in Zelinka et al. (2014) which is why these methods are compared in Figure 11. Because the semi-direct is small due to the small influence of black carbon in these results, any adjustments to RFari that are being counted as part of ERFaci will not affect the results too much. Spelling out these equations for the double call with reference to Ghan (2013) and focusing on the SW:

Direct radiative forcing (RFari) = -Δrsut - (-Δrsutaf)
Cloud radiative forcing (ERFaci + semi-direct effect) = -Δrsutaf - (-Δrsutcsaf)
Surface albedo forcing = -Δrsutcsaf

Summing up the three components, everything except -Δrsut cancels out, which is identically the SW ERF. The same breakdown is done for the LW. So you are correct that the full breakdown should be exact.

By reporting ERFari+aci, we neglect the surface albedo component, which could usefully be compared to the surface albedo adjustment calculated by the kernel method (although we do not do so). The "surface albedo" forcing is not zero in the LW decomposition using the double call either, which may be compared to the surface temperature or other LW adjustments.

As an additional confidence that our results are correctly calculated, they agree with results obtained for CNRM-ESM2-1 and CNRM-CM6-1 independently by Séférian et al. (2019) and Michou et al. (2020).

8. Fig 11: I am a bit confused about land-use ERF results here. I can understand that it is small on global averages. However, I am surprised that it is still insignificant in

some regions (e.g., North America, China), even though the regional ERF there is large (~ -6 W m-2) (Fig 11). How's the significance been calculated?

Originally, significance in these plots was defined as the multi-model mean being different from zero at the one standard-deviation level in each grid cell. In the case of land use, the standard deviation can be quite large in some of the North American grid cells (and larger than the absolute mean) as some models show strong negative ERF and some show weak negative ERF, but the ERF is still robustly regative. This was, probably incorrectly, leading to a lack of "significance" under this definition. For this reason, and to be consistent with the ISCCP simulator plot in fig. 4, the definition is now that 75% or more of models should agree on sign. In addition, the hatching scheme on these figures has been improved so it should be clearer which regions are shaded out.

In the caption to fig. 7, added

> Hatched regions are where less than 75% of models agree on the sign of the change.

Where spatial plots first introduced in the text in line 300, added

> Hatched areas are defined where less than 75% of models agree on the sign of the change.

References used in response to reviewer #2 not in manuscript:

Michou, M., Nabat, P., Saint-Martin, D., Bock, J., Decharme, B., Mallet, M., et al. (2020). Present-day and historical aerosol and ozone characteristics in CNRM CMIP6 simulations. Journal of Advances in Modeling Earth Systems, 12, e2019MS001816. https://doi.org/10.1029/2019MS001816

Gettelman, A., Hannay, C., Bacmeister, J. T., Neale, R. B., Pendergrass, A. G., Danabasoglu, G., et al. (2019). High climate sensitivity in the Community Earth System Model Version 2 (CESM2). Geophysical Research Letters, 46, 8329– 8337. https://doi.org/10.1029/2019GL083978

Effective radiative forcing and adjustments in CMIP6 models

Christopher J. Smith[1,2], Ryan J. Kramer[3,4], Gunnar Myhre[5], Kari Alterskjær[5], William Collins[6], Adriana Sima[7], Olivier Boucher[8], Jean-Louis Dufresne[7], Pierre Nabat[9], Martine Michou[9], Seiji Yukimoto[10], Jason Cole[11], David Paynter[12], Hideo Shiogama[13,14], Fiona M. O'Connor[15], Eddy Robertson[15], Andy Wiltshire[15], Timothy Andrews[15], Cécile Hannay[16], Ron Miller[17], Larissa Nazarenko[17], Alf Kirkevåg[18], Dirk Olivié[18], Stephanie Fiedler[19], Anna Lewinschal[20], Chloe Mackallah[21], Martin Dix[21], Robert Pincus[22,23], and Piers M. Forster[1]

[1]School of Earth & Environment, University of Leeds, LS2 9JT, UK
[2]International Institute for Applied Systems Analysis (IIASA), Laxenburg A-2361, Austria
[3]Climate and Radiation Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[4]Universities Space Research Association, 7178 Columbia Gateway Drive, Columbia, MD 21046, USA
[5]CICERO, Oslo, Norway
[6]Department of Meteorology, University of Reading, UK
[7]LMD/IPSL, Sorbonne Université, ENS, PSL Université, École polytechnique, Institut Polytechnique de Paris, CNRS, Paris France
[8]Institut Pierre-Simon Laplace, Sorbonne Université / CNRS, Paris, France
[9]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
[10]Meteorological Research Institute, Tsukuba, Japan
[11]Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, British Columbia, Canada
[12]Geophysical Fluid Dynamics Laboratory, Princeton University Forrestal Campus, 201 Forrestal Road, Princeton, NJ 08540-6649
[13]Center for Global Environmental Research, National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan
[14]Atmosphere and Ocean Research Institute, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8564, Japan
[15]Met Office Hadley Centre, FitzRoy Road, Exeter, EX1 3PB, UK
[16]NCAR/UCAR, Boulder, Colorado, USA
[17]NASA Goddard Institute for Space Studies, New York, NY 10025 USA
[18]Norwegian Meteorological Institute, Oslo, Norway
[19]University of Cologne, Institute of Geophysics and Meteorology, Cologne, Germany
[20]Department of Meteorology, Stockholm University, Sweden
[21]CSIRO Oceans and Atmosphere, Aspendale, Australia
[22]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, CO, USA
[23]NOAA/ESRL Physical Sciences Division, Boulder, CO, USA

Correspondence: C.J. Smith (c.j.smith1@leeds.ac.uk)

Abstract. The effective radiative forcing, which includes the instantaneous forcing plus adjustments from the atmosphere and surface, has emerged as the key metric of evaluating human and natural influence on the climate. We evaluate effective radiative forcing and adjustments in ~~13~~ 17 contemporary climate models that are participating in CMIP6 and have contributed to the Radiative Forcing Model Intercomparison Project (RFMIP). Present-day (2014) global mean anthropogenic forcing relative to pre-industrial (1850) from climate models stands at ~~1.97~~ 2.00 (± ~~0.26~~0.23) W m^{-2}, comprised of ~~1.80~~ 1.81 (± ~~0.11~~0.09) W m^{-2} from CO_2, ~~1.07~~ 1.08 (± 0.21) W m^{-2} from other well-mixed greenhouse gases, ~~−1.04~~ −1.01 (± 0.23) W m^{-2} from

aerosols and ~~−0.08~~ −0.09 (± ~~0.14~~0.13) W m^{-2} from land use change. Quoted uncertainties are one standard deviation across model best estimates, and 90% confidence in the reported forcings, due to internal variability, is typically within 0.1 W m^{-2}. The majority of the remaining ~~0.17~~ 0.21 W m^{-2} is likely to be from ozone. In most cases, the largest contributors to the spread in ERF is from the instantaneous radiative forcing (IRF) and from cloud responses, particularly aerosol-cloud interactions to aerosol forcing. As determined in previous studies, cancellation of tropospheric and surface adjustments means that the ~~"traditional"~~ stratospherically adjusted radiative forcing is approximately equal to ERF for greenhouse gas forcing, but not for aerosols, and consequentially, not for the anthropogenic total. The spread of aerosol forcing ranges from -0.63 to -1.37 W m^{-2}, exhibiting a less negative mean and narrower range compared to 10 CMIP5 models. The spread in $4\times CO_2$ forcing has also narrowed in CMIP6 compared to 13 CMIP5 models. Aerosol forcing is uncorrelated with ~~equilibrium~~ climate sensitivity. Therefore, there is no evidence to suggest that the increasing spread in climate sensitivity in CMIP6 models, particularly related to high-sensitivity models, is a consequence of a stronger negative present-day aerosol forcing, and little evidence that modelling groups are systematically tuning climate sensitivity or aerosol forcing to recreate observed historical warming.

1 Introduction

The effective radiative forcing (ERF) has gained acceptance as the most ~~appropriate~~ useful measure of defining the impact on the Earth's energy imbalance to a radiative perturbation (Myhre et al., 2013; Boucher et al., 2013; Forster et al., 2016). These perturbations can be anthropogenic or natural in origin, and include changes in greenhouse gas concentrations, aerosol burdens, land use characteristics, solar activity, and volcanic eruptions. ~~A desirable aspect of ERF is that long-term equilibrium temperatures correspond better to ERF than to the long-used stratospherically-adjusted radiative forcing (RF) in the simple~~ Since the start of the Industrial Era until the present-day, anthropogenic forcing has typically been increasing, and has been the dominant component of the total forcing on the Earth system except for brief periods following large volcanic eruptions (Myhre et al., 2013) . The main constituents of anthropogenic ERF are a positive forcing from greenhouse gases and a partially offsetting negative forcing from aerosols. While greenhouse gas forcing is reasonably well-known, aerosol forcing is more uncertain due to the spatial variation of aerosols, their short atmospheric lifetime, and their complex interactions with clouds (Boucher et al., 2013; Bellouin et al., 2020b) .

ERF is useful because equilibrium temperatures are more closely related to surface warming in the forcing-feedback relationship of the Earth's atmosphere:

$$\Delta N = F - \lambda \Delta T \tag{1}$$

where ΔN, F, λ and ΔT are the top-of-atmosphere (TOA) energy imbalance, (effective) radiative forcing, climate feedback parameter, and change in global-mean surface air temperature respectively. Richardson et al. (2019) showed that using ERF

rather than RF reduces the need for forcing-specific efficacy values (the temperature response per unit forcing), first introduced by Hansen et al. (2005) as an observation that different values of λ better predicted ΔT for different forcing agents under RF. Conversely, evaluating ERF is less straightforward than RF, requiring climate model integrations, and numerous different methods of calculating ERF exist with their own benefits and drawbacks (Shine et al., 2003; Gregory et al., 2004; Hansen et al., 2005; Forster et al., 2016; Tang et al., 2019; Richardson et al., 2019).

The difference between ERF and RF is that ERF includes all tropospheric and land-surface adjustments whereas RF only includes the adjustment due to stratospheric temperature change (Sherwood et al., 2015; Myhre et al., 2013). Adjustments are often termed "rapid" (Myhre et al., 2013; Smith et al., 2018b), however, there is no formal separation of adjustments and climate feedbacks based on timescale alone (Sherwood et al., 2015). It is conceptually more appropriate to divide adjustments as those changes in state that occur purely as a result of the action of a forcing agent from slow feedbacks that occur as a result of a change in global mean surface temperature. The instantaneous radiative forcing (IRF) is the initial perturbation to the Earth's radiation budget and unlike the RF and ERF does not include adjustments. By analysing atmosphere-only climate simulations using fixed climatological sea-surface temperatures (SSTs) and sea ice distributions, surface temperature driven feedbacks are largely suppressed except for a small contribution from land surface warming or cooling (Vial et al., 2013; Tang et al., 2019), allowing for adjustments to be diagnosed from atmospheric state changes (Forster et al., 2016; Smith et al., 2018b). This provides insight into the mechanisms contributing to the effective radiative forcing. For example, the ERF of black carbon is half of the impact estimated from its IRF as a consequence of its strong atmospheric absorption and adjustments arising from how it perturbs tropospheric heating rates, affecting the distribution of tropospheric temperatures, water vapour and clouds (Stjern et al., 2017; Smith et al., 2018b; Johnson et al., 2019; Allen et al., 2019).

The experimental protocol for determining (effective) radiative forcing in models has been extended since Phase 5 of the Coupled Model Intercomparison Project (CMIP5). CMIP5 included experiments for present-day (year 2000) all-aerosol and sulfate-only forcing (Zelinka et al., 2014, CMIP5 experiment labels sstClimAerosol and sstClimSulfate), and $4\times CO_2$ forcing (sstClim4xCO2; Andrews et al., 2012; Kamae and Watanabe, 2012) with respect to a pre-industrial baseline with climatological SSTs and sea ice distributions (sstClim). A handful of IRF outputs from quadrupled CO_2 experiments (Chung and Soden, 2015) were also obtained. For CMIP6, the Radiative Forcing Model Intercomparison Project (RFMIP; Pincus et al., 2016) provides a number of present-day time-slice and historical-to-future transient experiments designed to evaluate the ERF in climate models for different forcing agents, providing insight into why climate models respond the way they do to particular forcings. This is important when diagnosing climate feedbacks (Forster et al., 2013), given the role of forcing in the Earth's energy budget (eq. (1)), and knowledge of forcing is required for attribution of historical temperature change (Haustein et al., 2017), evaluating non-CO_2 contributions to remaining carbon budgets (Tokarska et al., 2018) , and in future scenario projections (Gidden et al., 2019). Effective radiative forcings derived from models can be used to validate assumptions derived from other lines of evidence, particularly for aerosol forcing, as is done by the Intergovernmental Panel on Climate Change (IPCC) in their periodic Assessment Reports.

2 Models and experimental protocol

We use results from ~~13~~ 17 state-of-the-art atmospheric general circulation models (GCMs) and Earth system models (ESMs) contributing to Tier 1 of RFMIP (table 1) as part of CMIP6 (Eyring et al., 2016). In addition, GISS-E2-1-G provided two physics variants, r1i1p1f1 and r1i1p3f1, with aerosol treatments that are different enough to justify treating the variants as separate models, bringing the total to 18. Models with diagnostics available on the Earth System Grid Foundation (ESGF) up until ~~19 December 2019~~ 13 May 2020 have been analysed, with the addition of one models (ACCESS-CM2) where experiments have been performed but not yet published on ESGF. Each model is run in atmosphere-only mode using pre-industrial climatologies of sea-surface temperatures (SSTs) and sea-ice distributions from at least 30 years of the same model's corresponding coupled pre-industrial control run (piControl, Eyring et al. (2016)). RFMIP's Tier 1 calls for 30-year timeslice experiments forced with $4\times$ pre-industrial CO_2 concentrations (RFMIP name piClim-4xCO2), all present-day anthropogenic forcers (piClim-anthro), present-day well-mixed greenhouse gases (piClim-ghg), present-day aerosols (piClim-aer) and present-day land use (piClim-lu) in this fixed-SST configuration. All forcing components that are not perturbed in a particular experiment remain at pre-industrial (year 1850) values, and "present-day" is defined as year 2014 conditions. A 30-year experiment with pre-industrial conditions, piClim-control, is also performed as a reference case, and all results presented in this paper are with reference to piClim-control, accounting for the ~~fact~~ possibility that models may ~~not be in precise radiative equilibrium~~have a non-zero pre-industrial TOA flux imbalance. Results from the $4\times CO_2$ experiment are also rescaled to the ratio of 2014 to 1850 CO_2 concentrations of approximately $1.4\times$ pre-industrial by a factor of 0.2266, being the ratio of RF from $1.4\times CO_2$ to $4\times CO_2$ from the Etminan et al. (2016) formula. This is performed to isolate an estimate of the CO_2-only contribution to the present-day forcing, and is based on year-1850 and year-2014 CO_2 concentrations of 284.32 and 397.55 ppm respectively (Meinshausen et al., 2017) along with the 1850 concentrations of 808.25 ppb for CH_4 and 273.02 ppb for N_2O. Except where explicitly stated, we present results from this experiment as $1.4\times CO_2$.

The experiments and results presented in this study follow on from the assessment of ERF and adjustments in 11 models contributing to the Precipitation Driver and Response Model Intercomparison Project (PDRMIP, see Myhre et al., 2017) in Smith et al. (2018b). In Smith et al. (2018b) idealised experiments of $2\times CO_2$ concentrations, $3\times CH_4$ concentrations, $10\times$ black carbon (BC) emissions or burdens, $5\times SO_4$ emissions or burdens and a 2% solar constant increase were analysed from CMIP5-era and interim models. Only the $4\times CO_2$ experiment has a similar experiment for comparison in Smith et al. (2018b), whereas the RFMIP protocol focuses more on combinations of anthropogenic forcers. In addition, extended model diagnostics allow us to determine cloud responses and aerosol forcing in more detail in this study.

3 Effective radiative forcing

Using climatological SSTs allows for ERF to be diagnosed as the difference of top-of-atmosphere net radiative flux between a given forcing experiment and a pre-industrial control simulation (Hansen et al., 2005). Using 30 year timeslices generally results in standard absolute errors of less than 0.1 W m^{-2} (Forster et al., 2016). Although inter-annual variability affects the diagnosed ERF using this climatological SST method, the standard error in the estimates obtained is much smaller than using

Table 1. Contributing climate models to RFMIP-ERF Tier 1. The adjustment time is based on approximately how long stratospheric temperatures take to equilibriate in the $4\times CO_2$ experiment (fig. 2). ISCCP simulator diagnostics are indicated where existent.

Model	Atmospheric resolution (lon × lat)	Adjustment timescale (yr)	Model years	ISCCP simulator	Reference
ACCESS-CM2	1.875° × 1.25°, 85 levels to 85 km	1	30		Bi et al. (submitted)
CanESM5	2.81° × 2.81°, 49 levels to 1 hPa	1	50	all	Swart et al. (2019)
CESM2	1.25° × 0.9°, 32 levels to 2.25 hPa	1	30	~~ghg, aer, lu, anthro~~ all	~~?~~ Danabasoglu et al. (202
CNRM-CM6-1	1.4° × 1.4°, 91 levels to 0.01 hPa	5	30	CO_2, ghg, aer, ~~lu~~ anthro	Voldoire et al. (2019)
CNRM-ESM2-1	1.4° × 1.4°, 91 levels to 0.01 hPa	15	30	all	Séférian et al. (2019)
EC-Earth3	0.7° × 0.7°, 91 levels to 0.01 hPa	1	30		Wyser et al. (2019)
GFDL-CM4	1.25° × 1°, 33 levels to 1 hPa	1	30	all	Held et al. (2019)
GFDL-ESM4	1.25° × 1°, 49 levels to 1 hPa	1	30		Dunne et al. (in prep.)
GISS-E2-1-G[1]	2.5° × 2°, 40 levels to 0.1 hPa	5	31/41[2]		~~Schmidt et al. (2014)~~ Kel
HadGEM3-GC31-LL	1.875° × 1.25°, 85 levels to 85 km	1	30	all	Williams et al. (2018)
IPSL-CM6A-LR	2.5° × 1.27°, 79 levels to 80 km	10	30	all	Boucher et al. (submitted
MIROC6	1.4° × 1.4°, 81 levels up to 0.004 hPa	1	30	aer	Tatebe et al. (2019)
MPI-ESM1-2-LR	1.875° × 1.875°, 47 levels up to 0.01 hPa	1	31		Mauritsen et al. (2019)
MRI-ESM2-0	1.125° × 1.125°, 80 levels to 0.01 hPa	1	30	all	Yukimoto et al. (2019)
NorESM2-LM	2.5° × 1.875°, 32 levels to 3 hPa	1	30		Seland et al. (2020) Kirkevåg et al. (2018)
NorESM2-MM	1.25° × 0.9375°, 32 levels to 3 hPa	1	30		Seland et al. (2020)
UKESM1-0-LL	1.875° × 1.25°, 85 levels to 85 km	3	45 [2]	~~CO_2, aer, lu, anthro~~ all	Sellar et al. (2019)

1. GISS-E2-1-G produced two physics variants for piClim-control and piClim-aer; physics_version=1 (p1) includes aerosol and ozone specified by pre-computed transient fields and physics_versio
aerosol-cloud interactions. Both physics versions are analysed in this paper and treated as separate models.
3. 41 years for r1i1p3f1

a fully-coupled ocean-atmosphere model with a Gregory regression (Gregory et al., 2004), and as such fewer model years are needed to diagnose ERF. Two advantages of this is that it reduces the computational burden for modelling centres, and can also be used to diagnose forcings of the order of 0.1 W m^{-2} (Forster et al., 2016). For this reason, the climatological-SST method is implemented to derive forcing in RFMIP, and ERF in this paper (without qualifier) is taken to mean this.

The climatological-SST method of deriving ERF includes the TOA flux changes resulting from land-surface warming or cooling as part of the ERF. Conceptually, any land-surface temperature change as a response to forcing should be excluded in the same way that SST changes are (Shine et al., 2003; Hansen et al., 2005; Vial et al., 2013), but ~~in general,~~ prescribing land surface temperatures is difficult in GCMs and this has not been performed in RFMIP. In essence, the goal is to completely isolate the forcing from any surface temperature change (ΔT) or feedbacks (λ) in eq. (1). We test several methods to correct for adjustments to attempt to isolate forcing at $\Delta T = 0$ (also performed in Richardson et al. (2019); Tang et al. (2019)):

– *Effective radiative forcing (ERF)* is reserved to mean the TOA flux difference between a perturbed and control simulation, with climatological SSTs and sea ice distributions and no correction for land surface temperature change, as in Hansen et al. (2005); Myhre et al. (2013); Forster et al. (2016); Smith et al. (2018b).

– *Effective radiative forcing using a Gregory regression (ERF_reg)* is calculated from each model's CMIP abrupt4xCO2 experiment by regressing the annual temperature anomaly compared to the same model's pre-industrial control (piControl) against the annual TOA energy imbalance anomaly ΔN (eq. (1)) and finding the intercept at $\Delta T = 0$, as in Gregory et al. (2004). This is done for the first 20 years of model output to avoid the changing value of λ over time present in many models (Armour, 2017) ; using the full 150 years tends to underestimate the forcing (denoted ERF_reg150; table S1). It is only possible to determine ERF_reg for $4\times CO_2$ as coupled abrupt forcing experiments are not performed for other forcing agents as part of CMIP6.

– *Stratospherically adjusted radiative forcing (RF):* All tropospheric and surface adjustments, calculated using radiative kernels (section 4) are subtracted from the ERF, leaving just the stratospheric temperature adjustment to the IRF. The RF is included for historical comparison, although it is usually calculated using an offline method such as fixed dynamical heating (Forster and Shine, 1997). It should be noted that the stratospheric adjustment is included in all definitions of ERF.

– *Land-surface corrected effective radiative forcing (ERF_ts):* Land surface temperature change adjustment is subtracted from the climatological-SST ERF using the surface temperature radiative kernel.

– *Tropospherically corrected effective radiative forcing (ERF_trop)*: In addition to land-surface warming a proportion of tropospheric temperature and water vapour change is subtracted from the ERF using radiative kernels, by assuming a fixed lapse rate in the troposphere based on the land surface temperature change. The remaining tropospheric temperature change when the constant lapse rate is subtracted is treated as the tropospheric temperature adjustment. The water vapour correction from the land surface warming is taken as the fraction of the adjustment from the constant lapse rate to the total tropospheric temperature adjustment. The ~~land surface temperature change and land surface albedo change are added~~surface albedo change is also removed, whereas no cloud adjustment is included justified by cloud adjustments to a large extent depending on heating/cooling in the troposphere (Smith et al., 2018b). This was known as ERF_kernel in Tang et al. (2019).

– *Feedback corrected effective radiative forcing (ERF_λ):* An amount corresponding to the global-average near-surface air temperature (GSAT) warming multiplied by the model's climate feedback parameter from its corresponding CMIP abrupt4xCO2 run is subtracted from the fixed-SST ERF. The same value of λ from abrupt4xCO2 is applied to the GSAT change in all experiments. This method was first investigated by Hansen et al. (2005) and is known as ERF_fSST_ΔTland in Tang et al. (2019).

Table 2 shows the ERF diagnosed from each forcing and each model using the climatological-SST method, and fig. 1 shows the ERF, diagnosed IRF, and adjustments from each RFMIP Tier 1 experiment. Values for the different methods for calculating

Table 2. Effective radiative forcing from each Tier 1 time-slice RFMIP experiment for each model (W m^{-2}). Also shown is the $4\times CO_2$ ERF scaled to 2014 concentrations (as $1.4\times CO_2$) and the residual forcing (anthropogenic − WMGHGs − aerosol − land-use). WMGHGs = well-mixed greenhouse gases. Note that not all models performed all experiments.

#	Model	$4\times CO_2$	$1.4\times CO_2$	WMGHGs	aerosols	land-use	anthropogenic	residual
1	ACCESS-CM2	7.95	1.80	3.04	−1.09		1.90	
2	CanESM5	7.61	1.72	2.87	−0.85	−0.08	2.37	0.43
3	CESM2	8.91	2.02	3.03	−1.37	−0.04	2.05	0.43
4	CNRM-CM6-1	8.00	1.81	2.74	−1.15		1.61	
5	CNRM-ESM2-1	7.93	1.80	2.51	−0.74	−0.07	1.66	−0.04
6	EC-Earth3	8.09	1.83	2.75	−0.80	−0.13	2.09	0.28
7	GFDL-CM4	8.24	1.87	3.13	−0.73	−0.33	2.34	0.27
8	GFDL-ESM4	7.74	1.75	3.23	−0.70	−0.28	2.17	−0.08
9	GISS-E2-1-G p1	~~6.90~~ 7.35	~~1.56~~ 1.67	2.89	−1.32	−0.00	1.93	0.35
10	GISS-E2-1-G p3				−0.93			
11	HadGEM3-GC31-LL	8.09	1.83	3.11	−1.10	−0.11	1.81	−0.08
12	IPSL-CM6A-LR	8.00	1.81	2.82	−0.63	−0.05	2.32	0.18
13	MIROC6	7.32	1.66	2.69	−1.04	−0.03	1.80	0.17
14	MPI-ESM1-2-LR	8.35	1.89	2.69		−0.10	2.13	
15	MRI-ESM2-0	7.65	1.73	3.03	−1.21	−0.17	1.95	0.29
16	NorESM2-LM	8.15	1.85	2.80	−1.21	0.26	2.06	0.20
17	NorESM2-MM	8.38	1.90		−1.26			
18	UKESM1-0-LL	7.94	1.80	~~2.94~~ 2.95	~~−1.13~~ −1.11	~~−0.30~~ −0.18	~~1.71~~ 1.79	0.12
	Mean	~~7.93~~ 7.98	~~1.80~~ 1.81	~~2.87~~ 2.89	~~−1.04~~ −1.01	~~−0.08~~ −0.09	~~1.97~~ 2.00	0.20
	Standard dev.	~~0.47~~ 0.38	~~0.11~~ 0.09	~~0.18~~ 0.19	0.23	~~0.14~~ 0.13	~~0.26~~ 0.23	0.17

forcing are given in Tables S1–S5. Instantaneous forcing (IRF) is calculated as the difference of the ERF and the sum of adjustments, with exception being land-use forcing where IRF is calculated directly from the surface albedo kernel. In keeping with the definitions of ERF and adjustments, IRF is defined at the TOA in this study. Adjustment calculations are explained in detail in section 4.

For ease of comparison we show $1.4\times CO_2$ instead of $4\times CO_2$, with the scaling to present-day concentrations assumed to apply to ERF, IRF and all adjustments proportionally. Figure 1 also shows the ERF_reg (for $4\times CO_2$), ERF_ts, ERF_λ and RF. In general, the methods that correct for land surface temperature change (ERF_ts, ERF_trop and ERF_λ) result in forcings that are slightly stronger than non-corrected ERF, although differences between these methods are comparable to the magnitude of internal year-to-year variability and small compared to the contribution of adjustments. ~~The exception is for ERFfor $1.4\times CO_2$, where there is substantial land-surface warming that contributes to tropospheric warming under the fixed lapse-rate assumption.~~

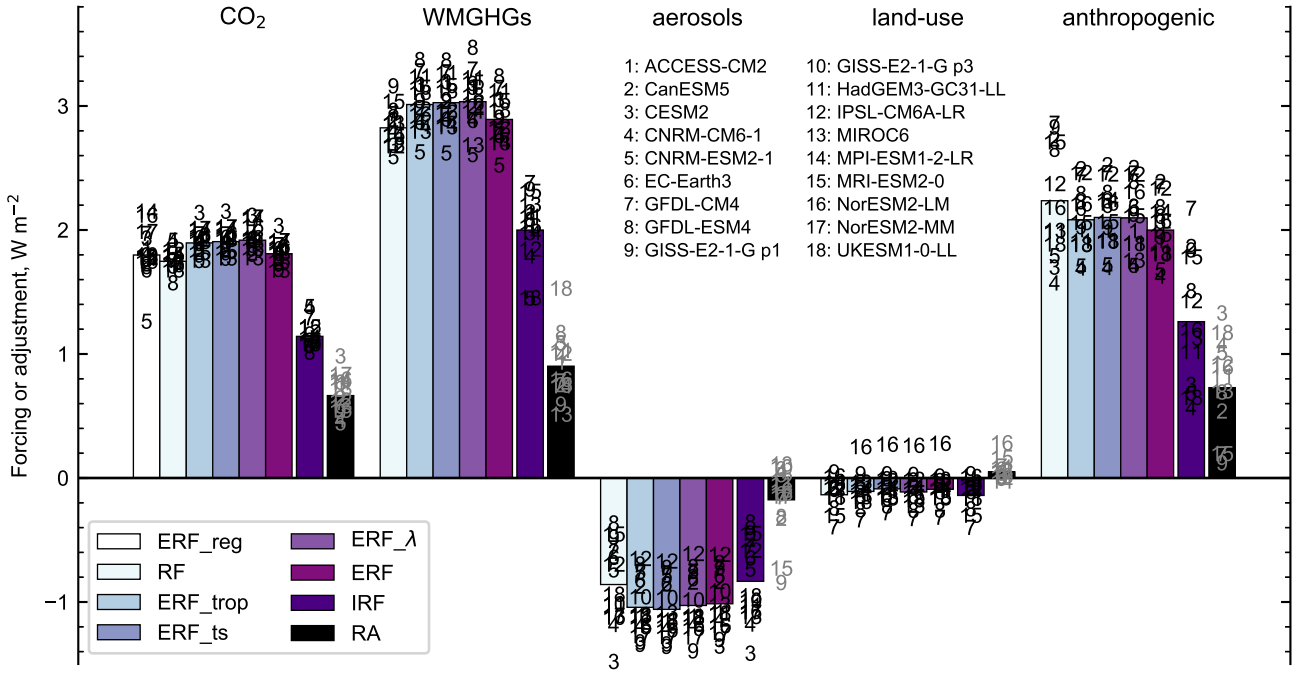

Figure 1. Comparison of radiative forcing (RF, which by definition includes stratospheric temperature adjustment), effective radiative forcing with tropospheric correction (ERF_trop), effective radiative forcing with land-surface kernel correction (ERF_ts), feedback-corrected ERF (ERF_λ), and fixed-SST ERF. For CO_2 forcing, ERF from a Gregory regression (ERF_reg) from each model's corresponding abrupt4xCO2 CMIP simulation is also given. The ERF is compared with the IRF and adjustments (RA) for each of the present-day RFMIP-ERF time slice experiments ($1.4\times CO_2$ is shown instead of $4\times CO_2$ for better comparison with other forcing agents). Individual models are numbered.

For CO_2 ~~in we exclude GISS-E2-1-G (model 6) due to an anomalous tropospheric cooling that biases ERF. For CO_2~~, ERF_reg results in a ~~lower~~ similar mean estimate of ERF ~~than any of~~ to the fixed-SST ~~methods (Forster et al., 2016) , which is possibly due to~~ method. Excluding CNRM-ESM2-1 for reasons described in the next section, the ~~non-linear nature of how climate feedbacks evolve over time in coupled model runs (Armour, 2017) and any change in pattern of SSTs that are not included in the climatological-SST~~ $4\times CO_2$ ERF_reg is 8.09 W m^{-2} compared to 7.99 W m^{-2} for ERF.

4 Forcing adjustments

4.1 Non-cloud adjustments

Adjustments to the radiative forcing describe flux changes resulting from changing atmospheric or surface state, in response to a forcing, but unrelated to the change in globally-averaged surface temperature (thus decoupling them from climate feedbacks, Myhre et al. (2013); Sherwood et al. (2015)). Adjustments to non-cloud changes in this study are calculated using radiative ker-

nels ~~(Chung and Soden, 2015; Vial et al., 2013; Smith et al., 2018b)~~ (Shell et al., 2008; Soden et al., 2008; Block and Mauritsen, 2013; Hu

The difference in an atmospheric state variable x (air temperature, surface temperature, specific humidity or surface albedo) between a forcing perturbation (pert) and piClim-control (base) is multiplied by the kernel K_x to derive the adjustment A_x:

$$A_x = K_x(x_{\text{pert}} - x_{\text{base}}) \tag{2}$$

The radiative kernel describes the change in TOA fluxes for a unit change in state for $x \in \{T, T_s, q, \alpha\}$ where T is atmospheric air temperature, T_s is surface temperature, q is water vapour and α is surface albedo. K_T and K_q are four-dimensional (month, pressure level, latitude, longitude) and K_{T_s} and K_α are three dimensional (month, latitude, longitude). Kernels are produced for both longwave and shortwave radiation changes. Typical unit changes are 1 K for temperature, the change in specific humidity that maintains constant relative humidity for a temperature increase of 1 K for water vapour, and 1% additive for surface albedo. For the division of temperature into stratospheric and tropospheric components, the WMO definition of the lapse-rate tropopause is used from each model's piClim-control run, using geopotential height as an approximation of geometric height on model pressure levels.

The water vapour kernel describes the change in TOA flux for a perturbation that maintains relative humidity for ~~an increase in specific humidity corresponding to~~ a temperature increase of 1 K~~in the kernel climatology~~, the effect being that specific humidity increases. The assumption therefore is that relative humidity is approximately constant ~~in the~~ between perturbation and control runs, which is found to be true in coupled experiments ~~where changes in specific humidity are larger than in this study due to surface-temperature driven feedbacks (Held and Soden, 2000)~~ (Held and Soden, 2000; Held and Shell, 2012) . Note that the difference in states is taken for the logarithm of water vapour concentration in eq. (2). More details on the application of the kernel method can be found in Smith et al. (2018b, Supplementary Material).

In this paper we use radiative kernels derived from the atmospheric component of the HadGEM3-GC31-LL model (HadGEM3-GA7.1), interpolated to the 19 standard CMIP6 pressure levels (Smith et al., 2020). With the exception of stratospheric temperature adjustments to greenhouse-gas forcing, structural differences introduced by using different kernels are well within 0.1 W m^{-2} (Soden et al., 2008; Smith et al., 2018b), and the HadGEM3-GA7.1 kernel is representative of the population of radiative kernels commonly used in the literature for tropospheric and surface adjustments (fig. S1); we use this particular kernel for its improved stratospheric resolution as outlined in Smith et al. (2020).

Stratospheric adjustments to greenhouse-gas driven experiments are expected to equilibriate within a few model months (Sherwood et al., 2015). We find that the time to reach equilibrium varies between models for a 4×CO_2 forcing. Figure 2 shows the time taken for the stratospheric temperature adjustment, and hence stratospheric temperatures, to adjust to a 4×CO_2 forcing. In CNRM-ESM2-1, concentrations of CO_2 are relaxed towards the 4× pre-industrial level below 560 hPa, and allowed to propagate throughout the atmosphere, therefore taking around 15 years to reach an approximate uniform atmospheric concentration. ~~This~~ A similar specification is implemented in the abrupt-4xCO2 run of CNRM-ESM2-1, causing ERF_reg to be biased low (fig. 1). This highlights one advantage of the fixed-SST based methods over the Gregory regression, as these "spin-up" years can simply be discarded with a fixed-SST measure of ERF. The CO_2 treatment in CNRM-ESM2-1 is in contrast to the physical climate model from the same group (CNRM-CM6-1). However, even in some physical models, we find

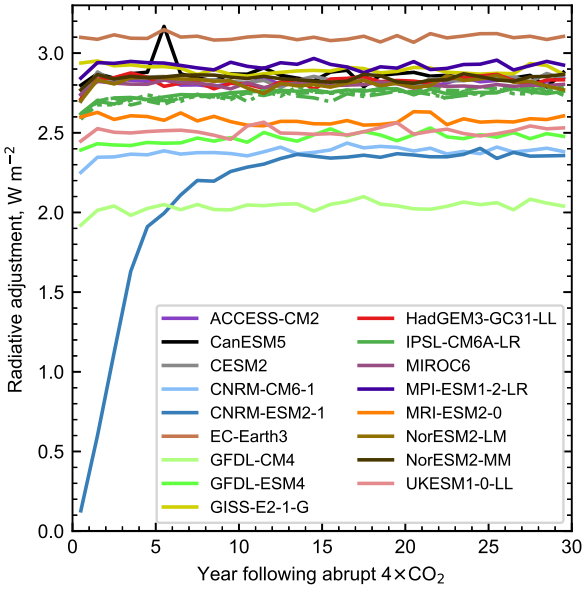

Figure 2. Transient response of the stratospheric temperature adjustment to a 4×CO_2 forcing. The small spike in year 6 in CanESM5 is due to an unseasonably low tropical tropopause in July of year 6, resulting in much of the temperature adjustment at the 100 hPa level to be counted in the stratosphere.

that the time to reach equilibrium varies between models and may be up to 10 years (e.g. in IPSL-CM6A-LR; fig. 2). For this reason, we discard the first few years of model output where the stratosphere is still adjusting to a forcing for the 4×CO_2, well-mixed greenhouse gas (WMGHG) and anthropogenic forcing experiments (table 1). We find this issue is not present in the aerosol or land-use experiments. It is important to emphasise that our stratospheric adjustment is calculated in a different way to the usual RF method which uses an offline radiative transfer method. It may therefore be the case that differences are due to a change in tropopause height in ~~greenhouse-gas -driven~~ greenhouse-gas-driven experiments (Santer et al., 2003).

4.2 Cloud adjustments

The radiative effect of clouds depends on their coverage (both within layer and total), ice water content, liquid water content, droplet effective radius and ice particle habit. Cloud properties vary extensively from model to model, and unlike pressure level diagnostics of temperature and humidity, cloud diagnostics are not output on 19 standard pressure levels in CMIP. A number of different approaches have therefore been used to estimate cloud adjustments, depending on availability of diagnostics and model specific setup, and we can exploit methods originally designed for cloud feedback calculations for calculating adjustments. Where cloud adjustments can be calculated with more than one method, we take the mean of each available method. In some models and experiments, cloud adjustments cannot be calculated and no estimate is made.

4.2.1 ISCCP simulator kernel

The ISCCP simulator ~~cloud kernel (Zelinka et al., 2012) allows diagnosis of flux changes due to clouds from ISCCP simulator diagnostics. The ISCCP simulator algorithm (Klein and Jakob, 1999; Webb et al., 2001) maps model-native clouds into cloud types that would be reported by ISCCP satellites (Rossow et al., 1996) , which is a~~ (Klein and Jakob, 1999; Webb et al., 2001) provides a joint 7×7 histogram of cloud ~~optical thickness~~ visble-wavelength optical depth (τ) and cloud top pressure (CTP). These outputs can be multiplied by the ISCCP simulator kernel (Zelinka et al., 2012) to estimate the impact of cloud changes on top-of-atmosphere fluxes. Ten models included ISCCP simulator diagnostics within their RFMIP output (table 1).

The ISCCP simulator kernel reports all flux changes resulting from clouds. For CO_2, WMGHG and land-use forcings, it is assumed that cloud droplet effective radius does not change ~~as aerosols do not change~~ (except for the land-use experiment in NorESM2-LM as discussed in section 5.4, but this model did not include ISCCP simulator diagnostics), and therefore in these experiments the SW flux changes from the ISCCP simulator kernel are treated as the cloud adjustment. For aerosol and total anthropogenic forcing this is usually not the case as most models include aerosol-radiation interactions (significant in the SW), with ice particle behaviour also changing in the MRI-ESM-2.0, MIROC6 and CESM2 models which affects LW fluxes. NorESM2-LM also includes the effects of mineral dust and BC on heterogeneous ice nucleation (Kirkevåg et al., 2018). Following Boucher et al. (2013) we treat the cloud-albedo response to aerosols as part of the IRF, and the ISCCP simulator kernel is unable to separate this effect from any ~~rapid~~ adjustment. We assume that any LW effect from aerosol-cloud interactions is small except in those models that include aerosol effects on ice clouds.

4.2.2 Approximate partial radiative perturbation with liquid water path adjustment

The approximate partial radiative perturbation (APRP; section 5.3.3) method uses standard climate model diagnostics to estimate the components of SW ERF attributed to cloud fraction change, and all-sky and clear-sky scattering and absorption. With no changes in aerosol forcing, the changes in cloud absorption, cloud scattering and cloud amount calculated from ~~the approximate partial radiative perturbation method (APRP ;)~~ APRP can be taken to be the SW cloud adjustment. We use this estimate for CO_2, WMGHG and land-use forcing.

For aerosol forcing, the effect of cloud amount changes calculated by APRP (A_{CLT}) is an adjustment, but the cloud scattering is a combination of radiative forcing due to aerosol-cloud interactions (RFaci), treated as part of the IRF, and adjustments due to cloud liquid water path (LWP) changes (A_{LWP}; Bellouin et al., 2020b). For the LWP adjustment we use a relationship obtained in Gryspeerdt et al. (2019) in which LWP adjustment (W m^{-2}) scales linearly with vertically integrated in-cloud liquid water path (kg m^{-2}):

$$A_{\mathrm{LWP}} = -\frac{1000}{37.6}\left(\frac{\mathrm{clwvi_{pert}} - \mathrm{clivi_{pert}}}{\mathrm{clt_{pert}}/100} - \frac{\mathrm{clwvi_{base}} - \mathrm{clivi_{base}}}{\mathrm{clt_{base}}/100}\right). \quad (3)$$

In eq. (3), clwvi, clivi and clt are the CMIP6 variable labels for total cloud water path, ice water path and total cloud fraction in percent. We then isolate the RFaci as

$$\mathrm{RFaci} = \mathrm{ERFaci} - A_{\mathrm{LWP}} - A_{\mathrm{CLT}}. \quad (4)$$

with ERFaci, the effective radiative forcing due to aerosol-cloud interactions, calculated from APRP (section 5.3.3).

For anthropogenic total forcing, the RFaci calculated in eq. (4) from the aerosol forcing experiment is subtracted from the total derived cloud change under APRP, which includes contributions from greenhouse gases and land use as well as RFaci. For models not including ice cloud nucleation, the LW cloud adjustment for aerosols is estimated from the change in cloud radiative effect (CRE; difference between all-sky and clear-sky fluxes). For other experiments this results in a biased estimate of cloud adjustment due to masking of LW adjustments.

4.2.3 Offline monthly mean partial radiative perturbation

~~Not all models provide ISCCP simulator diagnostics so for LW cloud adjustments we also produce an offline simulation~~ A direct estimate of cloud radiative effect can be obtained by substituting model cloud fields into an offline radiative transfer model. We perform these offline calculations using the SOCRATES radiative transfer code (Edwards and Slingo, 1996). This is produced by substituting fields of 3D cloud fraction, cloud water content and cloud ice content from each model and experiment into a climatology for the year 2014 provided by ERA5 (Copernicus Climate Change Service, 2017). Taking the cloud fields in each experiment minus those from the control gives $A_{\mathrm{LWP}} + A_{\mathrm{CLT}}$ in each model. ~~This offline substitution method is performed for years 16–25 of each model's output.~~ As only monthly mean diagnostics are available from models in general, we only attempt this in the LW which is assumed to be less biased than the SW (Mülmenstädt et al., 2019; Bellouin et al., 2020a). The monthly mean cloud fraction, ice water content and liquid water content variables in all experiments are scaled by a model-dependent factor that ranges between 0.68 and 1.5 to ensure that TOA LW outgoing flux is approximately ~~240~~ 240.2 W m^{-2} in the control experiment, in line with TOA observations (Loeb et al., 2018).

4.2.4 Kernel masking

In the land-use experiment, IRF is directly estimated from the surface albedo kernel such that IRF = A_α. As there are no other unknowns in the kernel decomposition, cloud adjustments can be calculated using the ~~kernel masking method (Soden et al., 2008) , where~~ difference between all-sky and clear-sky fluxes (Soden et al., 2008) , such that

$$A_c = (\mathrm{ERF} - \mathrm{ERF}^{\mathrm{clr}}) - (A_\alpha - A_\alpha^{\mathrm{clr}}) - \sum_{i \in \{\mathrm{T}, \mathrm{T_s}, \mathrm{q}\}} (A_i - A_i^{\mathrm{clr}}) \tag{5}$$

where the clr superscript in eq. (5) refers to fluxes calculated with clear-sky radiative kernels.

5 Multi-model results

Figure 3 shows the contribution to the total adjustment in each experiment from land surface temperature, tropospheric temperature, stratospheric temperature, water vapour, surface albedo and clouds. No corrections for tropospheric or land surface warming as discussed in section 3 have been performed for these results.

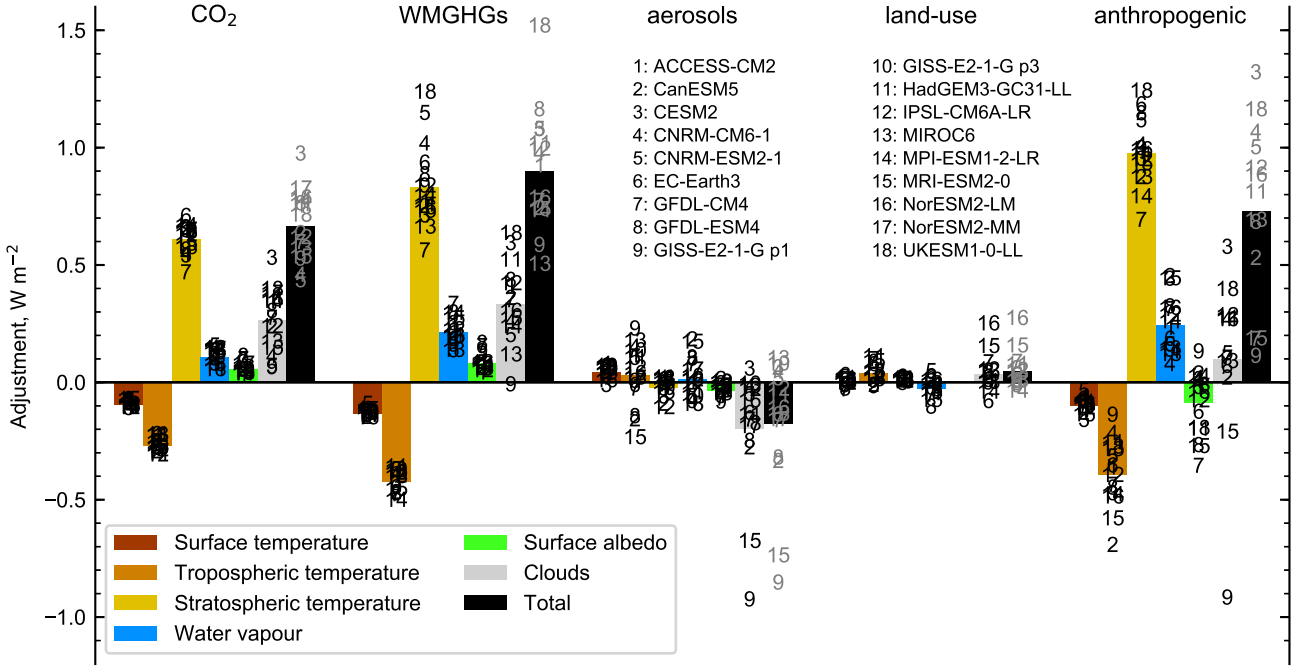

Figure 3. Adjustments broken down by mechanism in each of the present-day RFMIP-ERF time slice experiments. Black/grey ~~circles~~ numbers indicate individual models, coloured bars indicate the multi-model mean.

Figure 4 shows the effect on TOA radiative flux arising from cloud responses from the ISCCP simulator for each experiment from models that provided these diagnostics (table 1). In this figure, histogram boxes not marked with a cross are where 75% or more of the models agree on the sign of the cloud fraction or radiative flux change, following Zelinka et al. (2012).

5.1 Carbon dioxide

The multi-model mean ERF from a quadrupling of CO_2 is ~~7.93~~ 7.98 W m^{-2} (± ~~0.47~~ 0.38 W m^{-2}; all ~~uncertainties~~ ranges given as one standard deviation). ~~This equates to an ERF~~ A point of comparison for ERF is Etminan et al. (2016) , who computed estimates of and parametric fits for radiative forcing accounting for masking by clouds and stratospheric temperature equilibration, using a tropical and mid-latitude profile to represent the global mean. The implied ERF from RFMIP models for 2×CO_2 ~~of 3.79~~ is 3.81 (± ~~0.23~~0.18) W m^{-2} ~~scaled down~~ when scaling down the 4×CO_2 results using the Etminan et al. (2016) formula, comparable to a radiative forcing of 3.80 W m^{-2} for a doubling of CO_2 in Etminan et al. (2016) . Both estimates are slightly higher than the best estimate of 3.71 W m^{-2} from the IPCC's Fifth Assessment Report (AR5; Myhre et al. (2013)). ~~The Etminan et al. (2016) formula produces a value of 3.80 W m^{-2} for a doubling of~~ Scaling down the 4×CO_2 ~~which is very close to our multi-model mean ERF. On this basis~~forcing using Etminan et al. (2016) , our derived multi-model mean for 1.4×CO_2 is ~~1.80~~ 1.81 (± ~~0.11~~0.09) W m^{-2}~~, which is the same result produced from the Etminan et al. (2016) formula. In individual~~

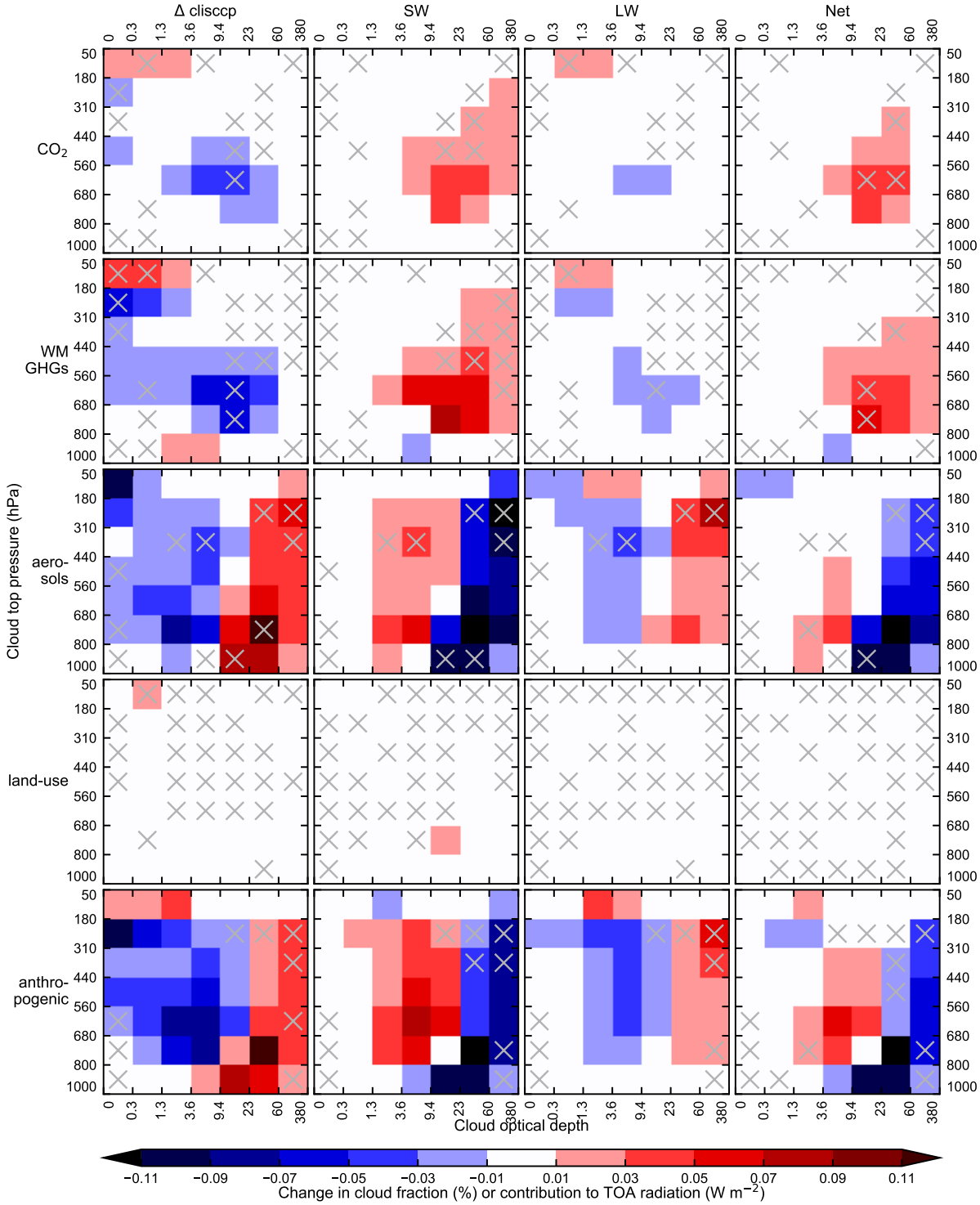

Figure 4. Global mean change in ISCCP-simulated cloud fraction in CTP-τ space (first column) and consequential changes in SW (second column), LW (third column) and net (fourth column) radiation when convoluted with the ISCCP cloud kernel. Grey crosses show where less than 75% of models agree on sign. Figure shows the multi-model mean cloud fraction and radiative effect. For $1.4\times CO_2$ the change in cloud fraction, as well as the radiative fluxes, are scaled down from the $4\times CO_2$ experiment using Etminan et al. (2016).

models which have performed 1.4×CO_2 experiments, our scaled-down 4×CO_2 calculation produces very similar results (table S1). The Etminan et al. (2016) relationship produces a CO_2 RF that increases faster than logarithmically, for which there is a growing body evidence (Hansen et al., 2005; Gregory et al., 2015; Colman and McAvaney, 2009; Jonko et al., 2013; Caballero and Huber, . As shown in fig. 1 and discussed in section 5.1, ERF is approximately equal to RF for CO_2, and we apply the Etminan formula to ERF.

The 4×CO_2 ERF from 13 17 CMIP6 models is larger, but not significantly so (p-value 0.19 0.13 using a Welch's t-test), than the 4×CO_2 ERF from 13 CMIP5 models of 7.53 (±0.89) W m^{-2} (Kamae and Watanabe, 2012). In addition, CMIP6 models are notable for their smaller spread in CO_2 ERF than CMIP5 models (fig. 5). Zelinka et al. (2020) show that ERF_reg150 for 4×CO_2 also increases in CMIP6 compared to CMIP5, and attribute 20% of the increase in multi-model mean equilibrium effective climate sensitivity (ECS) in CMIP6 to this. We note that a long standing problem in GCMs has been on the diversity in the forcing of CO_2 (Soden et al., 2018), which may result both from model broadband radiation parameterisation error in the IRF component (Pincus et al., 2015) and differences in base state climatology between models. The reduction in spread of CO_2 forcing in CMIP6 may be indicative that model radiation parameterisations are improving. In particular, , for example as documented in HadGEM3-GC31-LL , the IRF for 4×CO_2 is improved relative to previous versions of the model when comparing the GCM radiative transfer parameterisation to a narrowband radiative transfer model (Andrews et al., 2019) and UKESM1-0-LL (Andrews et al., 2019) , but could also be from a convergence in model base states, including clouds.

The breakdown of ERF into adjustments is shown in table 3 with the corresponding 4×CO_2 values in table S6. Stratospheric temperature adjustment dominates for CO_2-driven simulations, which is well-known (Smith et al., 2018b; Myhre et al., 2013). Tropospheric adjustments approximately sum to zero, such that the overall adjustment approximately equals the stratospheric adjustment, and RF is a good approximation to ERF (Smith et al., 2018b). Nevertheless, individual tropsopheric adjustments are non-zero and significant. A warming land surface and troposphere leads to a negative adjustment (more outgoing LW radiation to space) that is partially offset by increased tropospheric water vapour (analogous to the water vapour feedback). Cloud adjustments are overall positive, dominated by a reduction in mid-troposphere clouds driven by tropospheric warming, leading to a positive SW radiative effect (fig. 4). The LW effect is small in comparison, so that the SW effect dominates the net cloud adjustment.

The GISS-E2-1-G model shows anomalous tropospheric behaviour compared to the other models, exhibiting a positive tropospheric temperature adjustment, negative water vapour adjustment and negative cloud adjustment. This effect is traced to a slight cooling in the mid-troposphere in this model whereas other models show a distinct warming. This phenomenon causes the ERF definition that relies on tropospheric corrections (ERF) to differ significantly from other methods. Excluding this anomalous model, the largest spread in total adjustments are due to clouds.

The spatial pattern of adjustments is shown in fig. 6. In figs. 6 to 8, 12 and 13, cloud changes are only shown from the ISCCP simulator kernels in subfigures (g–i) and are not the means of all participating models, whereas ERF and non-cloud adjustments in (a–f) are multi-model means. Hatched areas are defined where less than 75% of models agree on the sign of the change. Stratospheric cooling is spatially uniform and results in a robustly positive adjustment of +0.60 +0.61 W m^{-2}, i.e. around one third of the total ERF. Tropospheric temperature adjustments are globally negative and robustover much of the

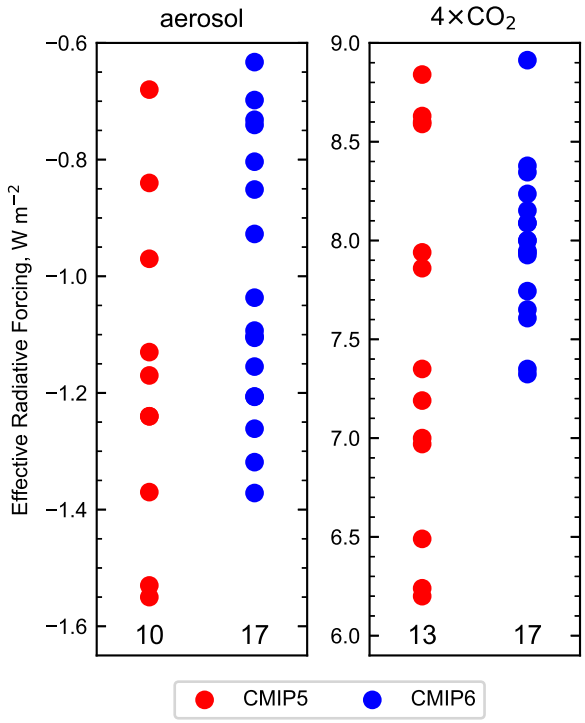

Figure 5. Aerosol and 4×CO_2 effective radiative forcing from CMIP5 sstClim4xCO2 and sstClimAerosol experiments (Kamae and Watanabe, 2012; Zelinka et al., 2014) and CMIP6 RFMIP experiments. Numbers at the bottom of each plot give the number of participating models.

~~Northern Hemisphere~~. Cloud changes show several robust spatial patterns, including positive changes over Eurasia and North America land.

5.2 Well-mixed greenhouse gases

The ERF from all well-mixed greenhouse gases is evaluated to be ~~2.87 (±0.18~~ 2.89 (±0.19 W m^{-2}) for 1850–2014, implying a contribution of ~~1.07~~ 1.08 (±0.21) W m^{-2} from non-CO_2 WMGHGs (uncertainties in quadrature and this definition excludes changes in ozone). Tier 1 of RFMIP does not contain additional granularity to break down non-CO_2 forcing by species, however dedicated experiments to derive ERF from methane, nitrous oxide and halocarbons separately are part of the protocol for the Aerosol and Chemistry Model Intercomparison Project (AerChemMIP; Thornhill et al., 2020; Collins et al., 2017).

There is also a substantial adjustment arising from WMGHG forcing, and again this is mostly driven by stratospheric cooling implied by the observation that ERF and RF are approximately equal. This confirms PDRMIP model behaviour for CO_2 and CH_4 forcing (Smith et al., 2018b), which found that tropospheric and land adjustments, while individually significant, approximately sum to zero leaving just the stratospheric temperature adjustment. Unlike in Smith et al. (2018b), who found

Table 3. ERF, IRF and adjustments (W m^{-2}) by component from 1.4×CO_2, scaled down from the 4×CO_2 RFMIP experiment. ts=surface temperature, ta_tr=tropospheric temperature, ta_st=stratospheric temperature, hus=water vapour, albedo=surface albedo, cl=clouds.

#	Model	ERF	IRF	Adj.	ts	ta_tr	ta_st	hus	albedo
1	ACCESS-CM2	1.80	1.05	0.75	−0.09	−0.23	0.64	0.07	0.02
2	CanESM5	1.72	1.09	0.63	−0.10	−0.30	0.65	0.10	0.05
3	CESM2	2.02	~~1.04~~ 1.05	~~0.98~~ 0.97	−0.12	−0.29	0.64	0.11	0.09
4	CNRM-CM6-1	1.81	1.36	0.45	−0.10	−0.29	0.54	0.14	0.05
5	CNRM-ESM2-1	1.80	1.37	0.43	−0.08	−0.28	0.53	0.15	0.05
6	EC-Earth3	1.83			−0.09	−0.27	0.70	0.11	0.05
7	GFDL-CM4	1.87	1.28	0.59	−0.09	−0.28	0.46	0.13	0.09
8	GFDL-ESM4	1.75	1.00	0.76	−0.08	−0.26	0.56	0.13	0.11
9	GISS-E2-1-G p1	~~1.56~~ 1.67	~~1.18~~ 1.14	~~0.38~~ 0.52	~~−0.06~~ −0.09	~~0.01~~ −0.23	~~0.61~~ 0.65	~~−0.20~~ 0.07	~~0.06~~ 0.05
11	HadGEM3-GC31-LL	1.83	~~1.09~~ 1.08	0.75	−0.11	−0.23	0.64	0.05	0.03
12	IPSL-CM6A-LR	1.81	1.20	0.61	−0.11	−0.31	0.62	0.14	0.04
13	MIROC6	1.66	1.09	0.57	−0.10	−0.26	0.63	0.07	0.05
14	MPI-ESM1-2-LR	1.89	1.12	0.77	−0.11	−0.31	0.66	0.14	0.05
15	MRI-ESM2-0	1.73	1.20	0.53	−0.08	−0.28	0.58	0.13	0.04
16	NorESM2-LM	1.85	~~1.06~~ 1.07	~~0.79~~ 0.78	−0.11	−0.28	0.64	0.12	0.06
17	NorESM2-MM	1.90	1.08	0.82	−0.11	−0.30	0.64	0.14	0.07
18	UKESM1-0-LL	1.80	~~1.10~~ 1.09	~~0.70~~ 0.71	−0.11	−0.23	0.57	~~0.06~~ 0.05	0.03
	Mean	~~1.80~~ 1.81	~~1.17~~ 1.14	~~0.63~~ 0.66	−0.10	~~−0.25~~ −0.27	~~0.60~~ 0.61	~~0.09~~ 0.11	0.05
	St. dev.	~~0.11~~ 0.09	0.11	~~0.16~~ 0.14	0.01	~~0.08~~ 0.03	0.06	~~0.09~~ 0.03	0.02

that the stratospheric temperature adjustment to methane was approximately zero, we find a larger stratospheric temperature adjustment for WMGHGs compared to CO_2 implying a positive non-CO_2 WMGHG stratospheric adjustment, although this cannot be attributed to individual gases.

The multi-model mean non-CO_2 WMGHG ERF of ~~1.07~~ 1.08 W m^{-2} is close to the 1850–2014 ~~theoretical~~ RF of 1.09 W m^{-2} made up of CH_4 (0.55 W m^{-2}) plus N_2O (0.17 W m^{-2}) from Etminan et al. (2016), plus halocarbons (0.37 W m^{-2}) using relationships from Myhre et al. (2013).

As for CO_2 only forcing, the total adjustment approximately equals the stratospheric temperature adjustment, implying that tropospheric and surface adjustments approximately cancel (table 4) so that the ~~uncertainty~~ spread in their sum is smaller than for each component individually. ~~The GISS-E2-1-G model does not show the anomalous cooling for tropospheric temperature and water vapour but does show a negative cloud adjustment not seen in other models.~~ For the ISCCP-simulator cloud adjustments, a ~~very~~ similar pattern can be seen from all WMGHGs to CO_2-only forcing, with a larger reduction in mid-troposphere cloud fraction leading to a greater positive SW adjustment which dominates the net adjustment.

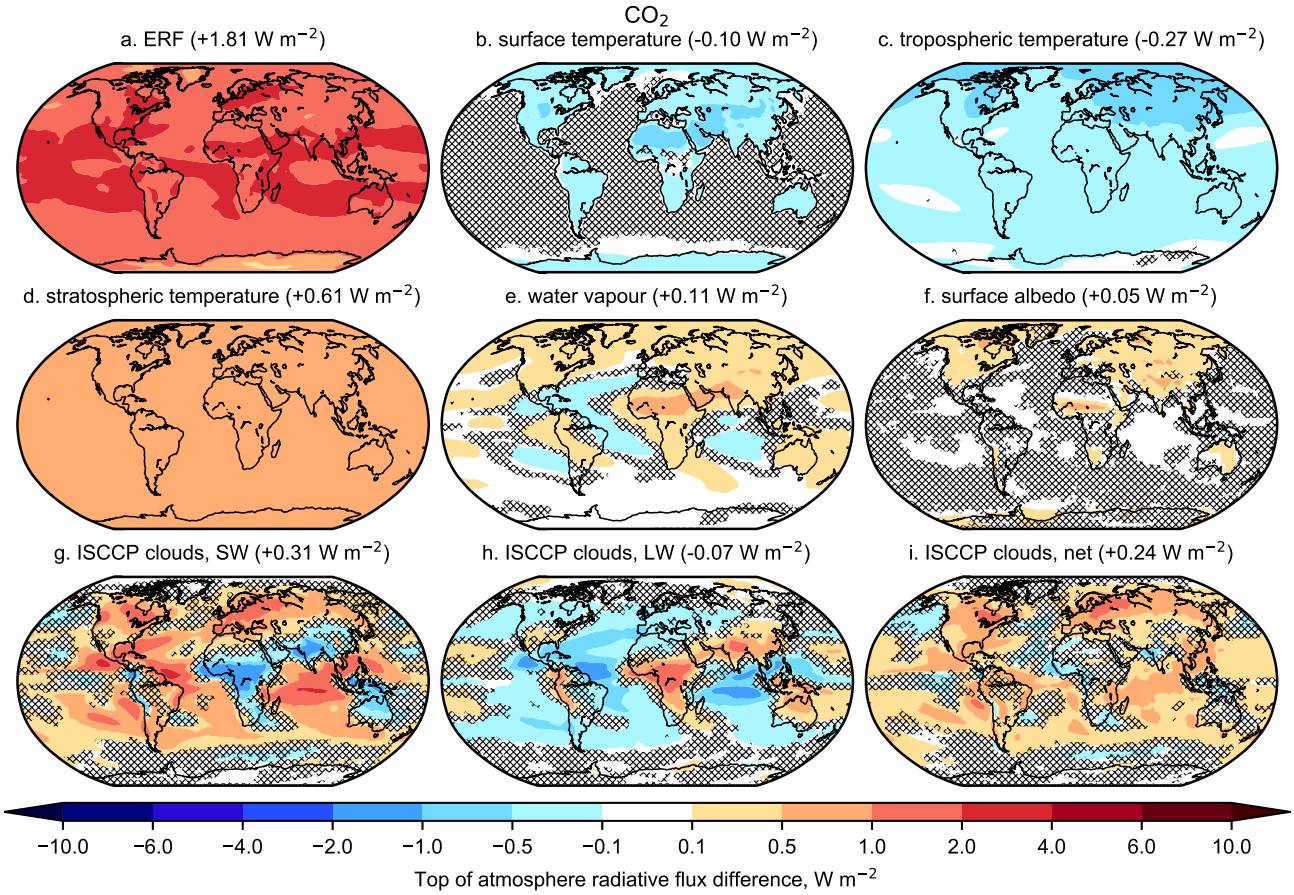

Figure 6. Multi-model mean spatial patterns of (a) effective radiative forcing, (b–f) adjustments and (g–i) cloud contributions to ERF for 1.4×CO_2. Hatched regions are where ~~changes are not significant at~~ less than 75% of models agree on the ~~one standard deviation level~~sign of the change.

The spread in ERF and stratospheric temperature adjustments is larger for WMGHG than for CO_2 forcing alone. One factor may be the inclusion or exclusion of stratospheric chemistry, which affects ozone formation. The effect can be seen by comparing Earth system (ESM) and physical models from the same group: the UKESM1-0-LL ESM (model ~~13~~18) to the HadGEM3-GC31-LL physical model (model ~~7~~11), and CNRM-ESM2-1 (model ~~4~~5) to CNRM-CM6-1 (model ~~3~~4). The physical models show ERFs around 0.2 W m^{-2} greater than the ESMs, a greater IRF, and a smaller stratospheric temperature adjustment. Additionally, for UKESM1-0-LL, large and compensating ERFs from CH_4 (+0.93 W m^{-2}) and halocarbons (−0.33 W m^{-2}), resulting from interactive chemistry, bring the total WMGHG ERF closer to the no-chemistry ERFs total from HadGEM3-GC31-LL (O'Connor et al., 2020).

The spatial patterns are overall similar to the CO_2 experiment (fig. 7) with a larger magnitude. ~~The tropospheric warming is more robust to all-WMGHG forcing, as GISS-E2-1-G conforms to other models and shows a tropospheric warming.~~

Table 4. As for table 3 but for 1850–2014 well-mixed greenhouse gas forcing.

#	Model	ERF	IRF	Adj.	ts	ta_tr	ta_st	hus	albedo	
1	ACCESS-CM2	3.04	2.12	0.92	−0.13	−0.40	0.77	0.24	0.04	0
2	CanESM5	2.87	2.13	0.74	−0.15	−0.47	0.74	0.21	0.05	0
3	CESM2	3.03	~~1.94~~ 1.96	~~1.10~~ 1.07	−0.15	−0.44	0.70	0.25	0.12	~~0.61~~ 0
4	CNRM-CM6-1	2.74	~~1.73~~ 1.77	~~1.01~~ 0.97	−0.15	−0.40	1.01	0.17	0.08	~~0.30~~ 0
5	CNRM-ESM2-1	2.51	~~1.44~~ 1.43	1.07	−0.10	−0.38	1.14	0.13	0.08	~~0.19~~ 0
6	EC-Earth3	2.75			−0.13	−0.45	0.93	0.20	0.06	
7	GFDL-CM4	3.13	~~2.34~~ 2.37	~~0.80~~ 0.77	−0.13	−0.48	0.56	0.33	0.15	~~0.37~~ 0
8	GFDL-ESM4	3.23	2.07	1.16	−0.13	−0.48	0.88	0.29	0.16	0
9	GISS-E2-1-G p1	2.89	~~2.34~~ 2.31	~~0.55~~ 0.58	−0.14	−0.38	0.83	0.15	0.13	~~−0.04~~ −0
11	HadGEM3-GC31-LL	3.11	~~2.11~~ 2.09	~~0.99~~ 1.01	−0.15	−0.36	0.75	0.19	0.06	~~0.50~~ 0
12	IPSL-CM6A-LR	2.82	~~1.81~~ 1.83	~~1.01~~ 0.99	−0.13	−0.39	0.83	0.22	0.05	~~0.44~~ 0
13	MIROC6	2.69	~~2.14~~ 2.19	~~0.54~~ 0.50	−0.13	−0.41	0.66	0.19	0.07	~~0.16~~ 0
14	MPI-ESM1-2-LR	2.69	~~2.00~~ 1.96	~~0.69~~ 0.73	−0.14	−0.51	0.79	0.29	0.06	~~0.19~~ 0
15	MRI-ESM2-0	3.03	2.30	0.73	−0.12	−0.46	0.72	0.27	0.05	0
16	NorESM2-LM	2.80	~~1.93~~ 2.02	~~0.87~~ 0.78	−0.14	−0.39	0.74	0.19	0.08	~~0.38~~ 0
18	UKESM1-0-LL	~~2.94~~ 2.95	1.44	~~1.50~~ 1.51	−0.16	−0.38	1.23	~~0.16~~ 0.13	~~0.05~~ 0.06	~~0.60~~ 0
	Mean	~~2.87~~ 2.89	~~1.97~~ 2.00	~~0.89~~ 0.90	−0.14	−0.42	~~0.82~~ 0.83	~~0.21~~ 0.22	0.08	~~0.33~~ 0
	St. dev.	~~0.18~~ 0.19	~~0.29~~ 0.27	0.25	~~0.02~~ 0.01	~~0.05~~ 0.04	~~0.18~~ 0.17	0.06	~~0.03~~ 0.04	~~0.18~~ 0

5.3 Aerosols

5.3.1 Forcing and adjustments

Present-day aerosol ERF is ~~−1.04~~ −1.01 (±0.23) W m^{-2} from ~~12 models. We exclude MPI-ESM1.2-LR in this estimate, since it uses prescribed aerosol optical properties and an associated effect on clouds from the simple plumes parameterization (MACv2-SP, Fiedler et al., 2017; Stevens et al., 2017) that will form the separate model inter-comparison of RFMIP-SpAer.~~ 17 models. The full range of aerosol ERF estimates for 2014 versus 1850 is −0.63 to −1.37 W m^{-2}. This is a narrower range of ERF than similar experiments performed with CMIP5 models for year 1850 and year 2000 forcings (Zelinka et al., 2014), particularly in relation to the lower (more negative) bound of aerosol forcing. Based on the 2000–1850 estimate of −1.17 (±0.30) W m^{-2} from Zelinka et al. (2014), aerosol forcing in CMIP6 models is less negative than in CMIP5, but this difference again is not significant (p-value ~~0.26~~0.15). Some of this multi-model mean difference is likely due to lower emissions of aerosol precursors in 2014 relative to 2000 along with updated historical estimates for CMIP6 (Hoesly et al., 2018; Lamarque et al., 2010), although it is not clear that this explains the reduction in model spread in CMIP6. It should also be borne in mind that our

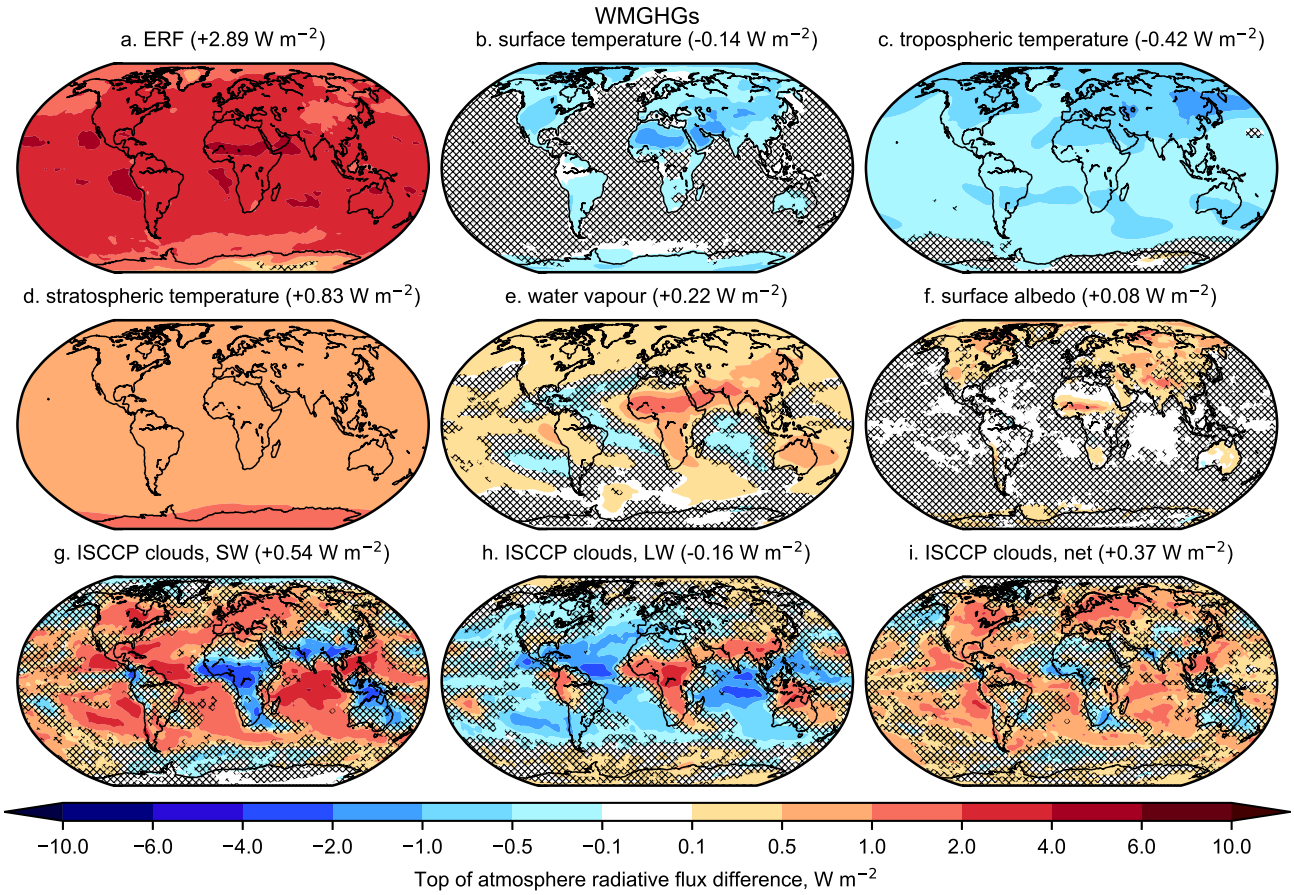

Figure 7. As fig. 6 but for present-day WMGHG forcing.

range does not include the E3SM model, which diagnosed aerosol forcing to be −1.65 W m^{-2} for 2005–2014 from a pair of parallel all-forcing and pre-industrial aerosol forcing atmosphere-only runs (fig. 25 in Golaz et al., 2019) . This highlights the likelihood that the inclusion of more models submitting results to RFMIP would extend the CMIP6 range of aerosol forcing, but the same may also have been true in CMIP5 where only a subset of models performed the sstClimAerosol experiment.

Atmospheric adjustments are small in magnitude in the aerosol forcing experiment, but large enough such that there is a noticeable difference between ERF and RF (fig. 1; Table S3). The small non-cloud adjustments in most models shows that the aerosol forcing is dominated by scattering aerosols (sulfate, organics, and for a limited number of models, nitrates) rather than black carbon (Smith et al., 2018b). Additionally, in two of the four models that provide the single-forcing BC experiment in AerChemMIP (CNRM-ESM2-1 and UKESM1-0-LL) the overall adjustment is small (Thornhill et al., 2020), in contrast to findings in PDRMIP models (Smith et al., 2018b). In MRI-ESM2-0 (model ~~11~~15) there are strong tropospheric temperature and cloud changes to black carbon forcing resulting in a negative adjustment overall (Thornhill et al., 2020).

For aerosol forcing, the aerosol-cloud interactions dominate, with ~~a close to uniform increase in optical thickness~~ an increase in cloud optical depth at all cloud heights. As cloud droplet effective radius decreases, cloud albedo, and hence optical ~~thickness~~depth, increases. This also implies that absorbing aerosols play only a minor role in most models, as BC induces strong adjustments that cause a general increase in cloud height in PDRMIP models from an increasing tropospheric stability ~~(Smith et al., 2018b; Stjern et al., 2017) for which there~~ (Smith et al. (2018b); Stjern et al. (2017) ; fig. S2). There is no evidence of this in the RFMIP aerosol forcing experiment, although some models do also include aerosol-cloud interactions from BC, and the effect may be due to the BC forcing being a smaller fraction of the total aerosol forcing than sulfate (Thornhill et al., 2020) . Figure S2 shows ISCCP simulator results for the five PDRMIP experiments from the CMIP5-era HadGEM2-ES model, where it can be seen that the aerosol forcing experiment is qualitatively more similar to the $5\times SO_4$ forcing experiment than the $10\times$BC experiment in PDRMIP. The increase in cloud albedo leads to a strong negative SW radiative effect that is partially compensated by LW effects (note that the ISCCP simulator kernel does not distinguish RFaci from adjustments).

Unlike for WMGHGs, aerosol forcing adjustments are dominated by cloud effects with only small non-cloud components (table 5). For aerosol forcing, all model years are used, as the stratospheric temperature adjustment is negligible. The spread in values of cloud adjustments is large, and spans positive and negative values. This reconfirms that atmospheric processes in response to aerosol forcing remains one of the largest uncertainties in climate models. There is also a spread in tropospheric temperature and water vapour adjustments with multi-model means near zero, suggesting that some models respond to aerosols with substantial atmospheric warming or cooling.

For many regions, particularly Southern Asia and the Eastern Pacific, the aerosol ERF is driven by large and negative cloud changes (fig. 8). The small adjustment overall and increase in cloud optical ~~thickness~~ depth for all ISCCP cloud categories suggests this is driven by an increase in cloud condensation nuclei leading to a more negative RFaci. There are some regions such as the Sahara in which a positive ERF arises and not easily explained by any adjustment component. This may be a reduction in mineral dust loading and increase in BC loading, leading to a positive forcing (e.g. as seen in NorESM2-LM, fig. S3).

The total derived cloud adjustment for aerosols is -0.20 W m^{-2}, derived of -0.04 W m^{-2} from SW cloud liquid water path adjustment, ~~-0.15~~ -0.13 W m^{-2} from SW cloud fraction change, and ~~-0.02~~ -0.03 W m^{-2} from cloud changes in the LW (table ~~S2~~S7).

5.3.2 Relationship to climate sensitivity

The increase in the upper bound, and in the overall spread, of ~~equilibrium climate sensitivity (ECS)~~ ECS in the CMIP6 model population compared to CMIP5 is well-documented ~~(?Zelinka et al., 2020)~~ (Forster et al., 2020; Zelinka et al., 2020) . Figure 9 shows the relationships between ECS and transient climate response (TCR) and aerosol ERF in CMIP6, taking ECS and TCR from each model's abrupt4xCO2 and 1pctCO2 CMIP runs respectively. There are weak and non-significant positive correlations between ECS and aerosol forcing (~~$r = 0.23$~~$r = 0.12$) and between TCR and aerosol forcing (~~$r = 0.30$~~$r = 0.26$). This suggests that, as a population, models with high sensitivity are not tuning present-day aerosol forcing to be strong in

Table 5. As for table 3 but for 1850–2014 aerosol forcing.

#	Model	ERF	IRF	Adj.	ts	ta_tr	ta_st	hus	
1	ACCESS-CM2	−1.09			0.07	0.11	0.01	−0.00	
2	CanESM5	−0.85	−0.51	−0.34	0.02	−0.16	−0.10	0.18	
3	CESM2	−1.37	~~−1.41~~ −1.43	~~0.04~~ 0.06	−0.02	−0.00	−0.07	0.10	
4	CNRM-CM6-1	−1.15	~~−1.22~~ −1.19	~~0.06~~ 0.04	0.07	0.15	−0.00	−0.08	
5	CNRM-ESM2-1	−0.74	~~−0.78~~ −0.75	~~0.04~~ 0.01	0.06	0.10	−0.01	−0.06	
6	EC-Earth3	−0.80	−0.66	−0.14	0.06	−0.02	0.01	−0.02	
7	GFDL-CM4	−0.73	−0.56	−0.17	0.05	−0.03	−0.02	0.07	
8	GFDL-ESM4	−0.70	−0.37	−0.33	0.04	−0.15	−0.01	0.10	
9	GISS-E2-1-G p1	−1.32	~~−0.45~~ −0.46	~~−0.87~~ −0.86	0.06	0.22	−0.03	−0.09	
10	GISS-E2-1-G p3	−0.93	−1.00	0.07	0.05	0.13	−0.03	−0.06	
11	HadGEM3-GC31-LL	−1.10	~~−1.03~~ −1.04	~~−0.07~~ −0.06	0.05	0.10	0.01	−0.05	
12	IPSL-CM6A-LR	−0.63	−0.60	−0.03	0.05	0.06	−0.11	0.01	
13	MIROC6	−1.04	~~−1.21~~ −1.13	~~0.15~~ 0.10	0.06	0.17	0.01	−0.10	
15	MRI-ESM2-0	−1.21	~~−0.48~~ −0.46	~~−0.72~~ −0.74	0.04	−0.24	−0.00	0.17	
16	NorESM2-LM	−1.21	~~−1.13~~ −1.09	~~−0.08~~ −0.11	0.00	0.03	−0.05	0.02	
17	NorESM2-MM	−1.26	−1.10	−0.16	0.03	0.01	−0.02	0.05	
18	UKESM1-0-LL	~~−1.13~~ −1.11	~~−1.00~~ −0.97	~~−0.13~~ −0.14	0.06	~~0.02~~ 0.01	~~0.00~~ 0.01	0.01	
	Mean	~~−1.04~~ −1.01	~~−0.87~~ −0.83	−0.18	~~0.04~~ 0.05	0.03	~~−0.03~~ −0.02	0.01	~~−0.04~~
	St. dev.	0.23	~~0.33~~ 0.31	~~0.30~~ 0.27	~~0.03~~ 0.02	~~0.13~~ 0.12	0.04	~~0.09~~ 0.08	

order to reproduce observed warming[1]: it would be expected that these correlations would be negative if this was the case (Smith et al., 2018a) . In CMIP5 models, aerosol forcing was stronger in models with higher ECS and TCR, but not significantly so (Forster et al., 2013)~~.~~ , although significance emerges if one considers only models which include an aerosol indirect effect (Chylek et al., 2016) . In CMIP3 there was a strong and significant negative correlation between climate sensitivity and aerosol forcing (Kiehl, 2007) . It may be the case that aerosol forcing over ~~the whole historical transient aerosol forcing~~ some historical periods is stronger in CMIP6 than in CMIP5, as despite higher climate sensitivity, CMIP6 models warm less than CMIP5 models and observations up until 2000 (Flynn and Mauritsen, 2020).

5.3.3 Decomposition of aerosol forcing into aerosol-radiation and aerosol-cloud effects

The approximate partial radiative perturbation (APRP) method (Taylor et al., 2007) can be used to decompose shortwave (SW) aerosol forcing into aerosol-radiation interactions (ERFari), aerosol-cloud interactions (ERFaci), and the surface albedo

[1] MPI-ESM1-2 (Mauritsen et al., 2019) is the only documented exception. MIROC6 (Tatebe et al., 2019) did tune the aerosol forcing to better correspond to the AR5 best estimate but explicitly did not tune for surface temperature.

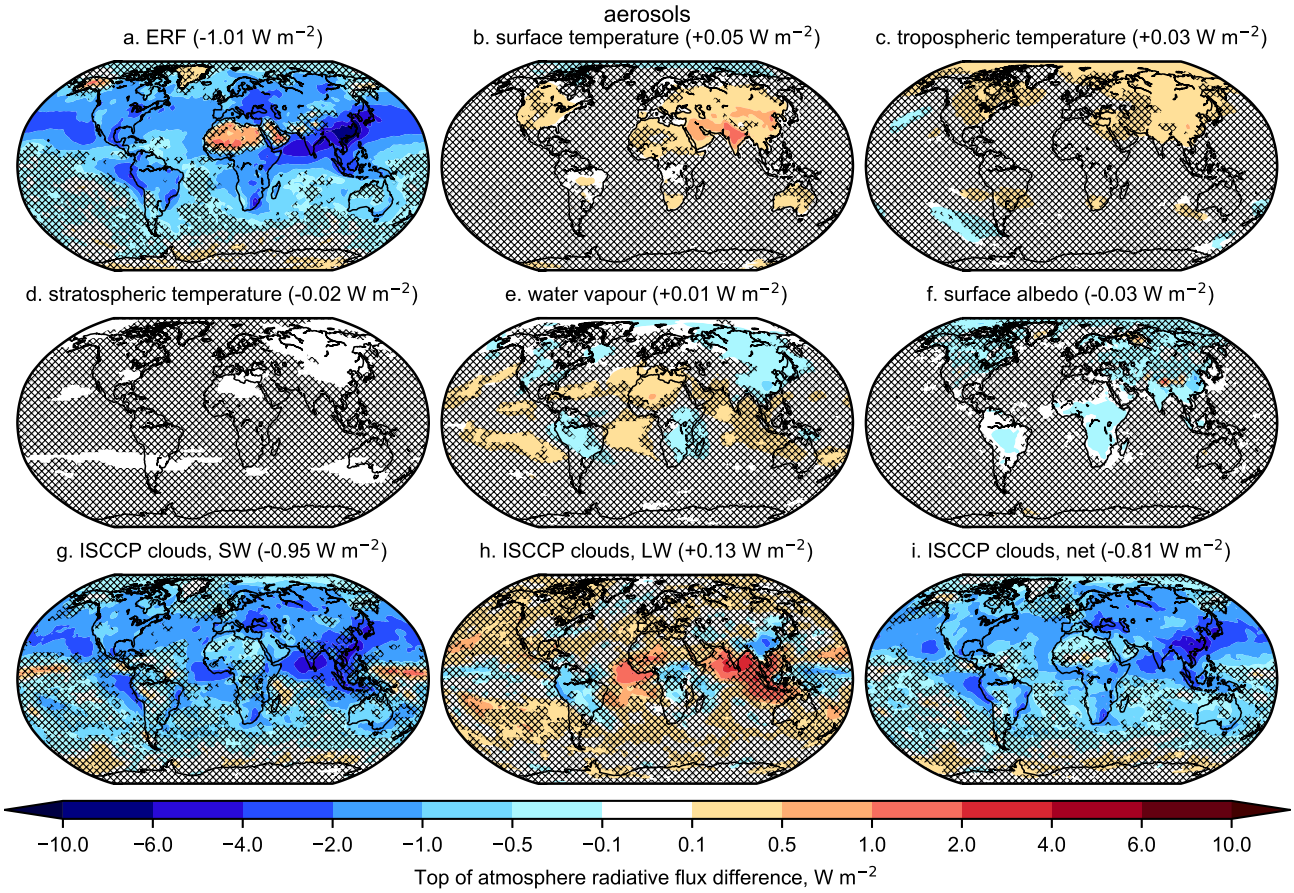

Figure 8. As fig. 6 but for present-day aerosol forcing.

adjustment (Zelinka et al., 2014). In section 5.3.4 we compare other methods to estimate ERFari and ERFaci. ERFari is the component of aerosol forcing that arises from the direct radiative effect of aerosol absorption and scattering (RFari) plus any adjustments (formerly known as the semi-direct effect) arising from perturbations in tropospheric heating rates, humidity, and their consequential effects on where clouds form (Boucher et al., 2013). These adjustments tend to be strong for black carbon but weak for scattering aerosol (Smith et al., 2018b; Stjern et al., 2017). ERFaci is composed of any changes in cloud albedo resulting from aerosols acting as cloud condensation nuclei and changing cloud droplet effective radius (RFaci, formerly the first indirect or Twomey effect, Twomey (1977)), plus adjustments relating to cloud lifetime and precipitation efficiency that changes liquid water path and cloud fraction (formerly second indirect or Albrecht effect, Albrecht (1989)). RFaci tends to be strong for sulfate aerosol, but several models also include cloud interactions to other aerosol species, and four models (CESM2, MIROC6, MRI-ESM2-0 and NorESM2-LM) include aerosol interaction on ice clouds. The direct plus Twomey effects (RFari+aci) are treated as the IRF component of aerosol forcing, with the remaining components of ERFari+aci as adjustments (Boucher et al., 2013).

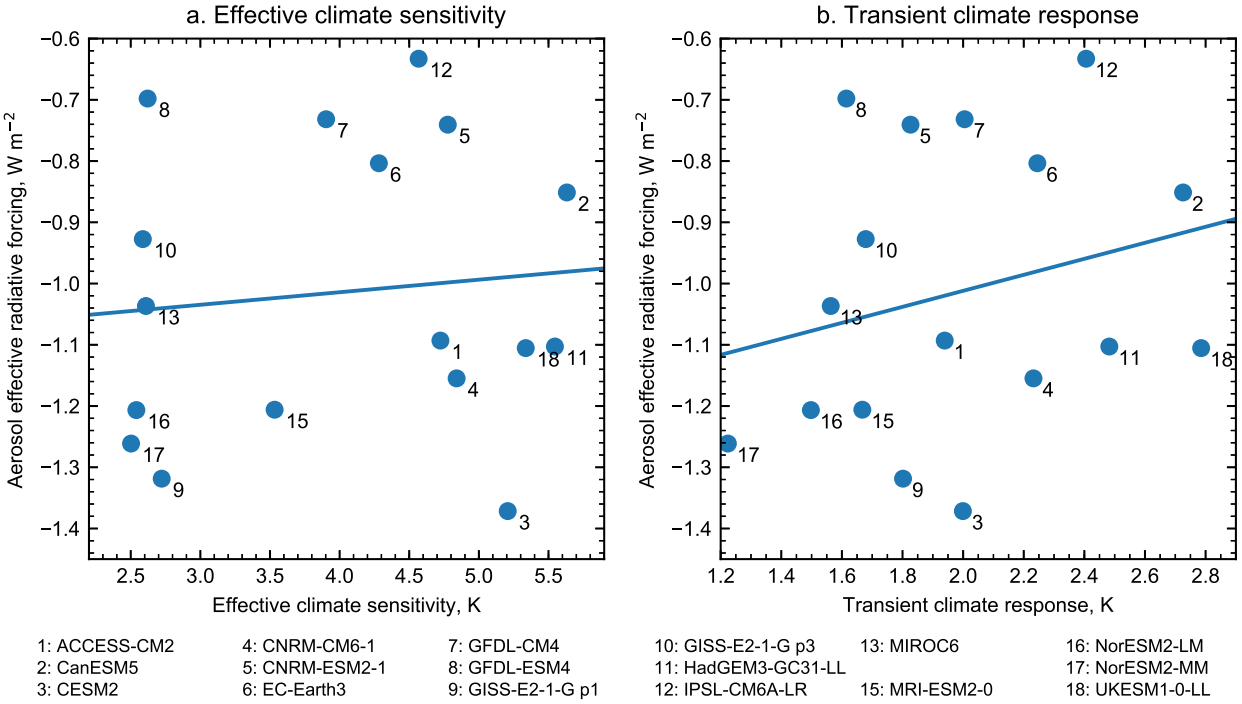

Figure 9. Relationship between (a) ECS and (b) TCR and aerosol ERF in the CMIP6 model ensemble. MPI-ESM1.2-LR (model ~~10~~14) did not produce the ~~aer, abrupt4xCO2 or 1pctCO2 experiments~~piClim-aer experiment.

There is no ~~equalivent~~ equivalent longwave (LW) method to APRP, so we take the approach of Zelinka et al. (2014) and use the cloud radiative effect to decompose LW ERF into ERFari and ERFaci. The advantages of these techniques are that they only require standard CMIP output, and all participating models can provide estimates. Results are displayed in table 6 and shown in fig. 10. In table S8 the equivalent SW ERFari for clear sky conditions are shown.

The total ERFari+aci from the APRP method is ~~−1.05~~ −1.04 (± ~~0.22~~0.20) W m^{-2}, agreeing very well with the ERF estimate of ~~−1.04~~ −1.01 (±0.23) W m^{-2}. ERFari+aci is approximately ~~20~~22% from ERFari and ~~80~~78% from ERFaci, and is comprised of a SW contribution of ~~−1.33~~ −1.26 W m^{-2} offset by a LW contribution of ~~+0.28~~ +0.23 W m^{-2}. The model spread in both the SW and LW individual components is larger than for the net forcing. This is driven by the four models that include ice cloud interactions that show positive LW ERFaci offset by strong negative SW ERFaci. MRI-ESM2.0 in particular has a very large positive LW ERFaci of +1.47 W m^{-2}, which comes from ice cloud nucleation by black carbon aerosols with temperature below −38°C in high-level clouds in the tropics (Oshima et al., in prep.) . For the SW component the ERFari/ERFaci split is ~~25% to 75~~approximately 28% to 72%.

Multi-model mean SW ERFari is ~~−0.33~~ −0.35 W m^{-2}, comprised of an absorption of ~~+0.26~~ +0.28 W m^{-2} offset by scattering of ~~−0.59~~ −0.63 W m^{-2}. The SW ERFaci is ~~−0.99~~ −0.91 W m^{-2}, made up of scattering (~~−0.83~~ −0.77 W m^{-2}), absorption (−0.01 W m^{-2}) and cloud fraction change (~~−0.15~~ −0.13 W m^{-2}).

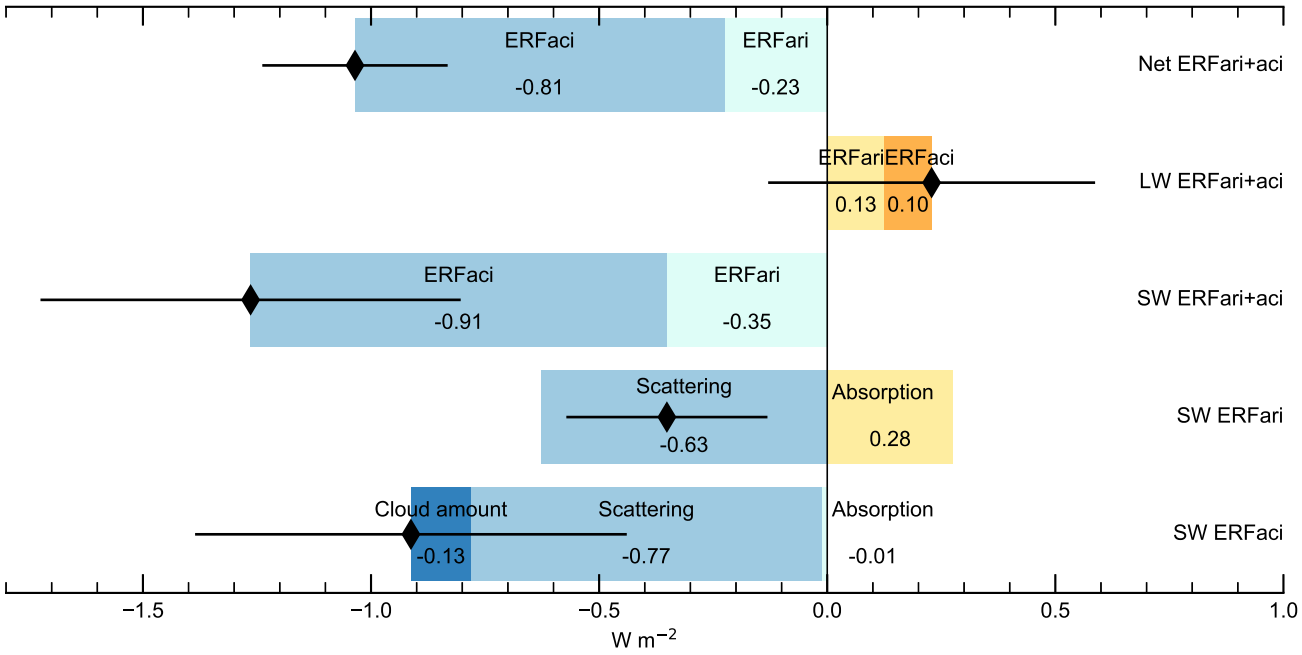

Figure 10. Components of the aerosol forcing diagnosed from the Approximate Partial Radiative Perturbation (for SW aerosol components) and from the cloud radiative effect (for LW components). Black diamonds represent multi-model means, black bars show one standard deviation.

5.3.4 Comparison of ERFari and ERFaci methods

~~Six~~ Eight models also archived radiation diagnostics from aerosol-free radiation calls (the double call method) as recommended by Ghan (2013) ~~. The aerosol forcing is estimated as~~

$$\mathrm{ERFariSW} = -\Delta\mathrm{rsut} - (-\Delta\mathrm{rsutaf})$$
$$\mathrm{ERFaciSW} = -\Delta\mathrm{rsutaf} - (-\Delta\mathrm{rsutcsaf})$$
$$\mathrm{ERFariLW} = -\Delta\mathrm{rlut} - (-\Delta\mathrm{rlutaf})$$
$$\mathrm{ERFaciLW} = -\Delta\mathrm{rlutaf} - (-\Delta\mathrm{rlutcsaf})$$

~~with variable names corresponding to CMIP6 TOA upwelling radiation diagnostics (r<l,s>utcsaf; l=longwave, s=shortwave, cs = clear-sky, af=aerosol-free) and Δ representing the present-day minus pre-industrial difference.~~

which allows separation into ERFari and ERFaci. This can be compared with the APRP estimates in the SW and cloud radiative effect for the LW. Figure 11 shows different methods of estimating ERFaci and ERFari from the aerosol forcing

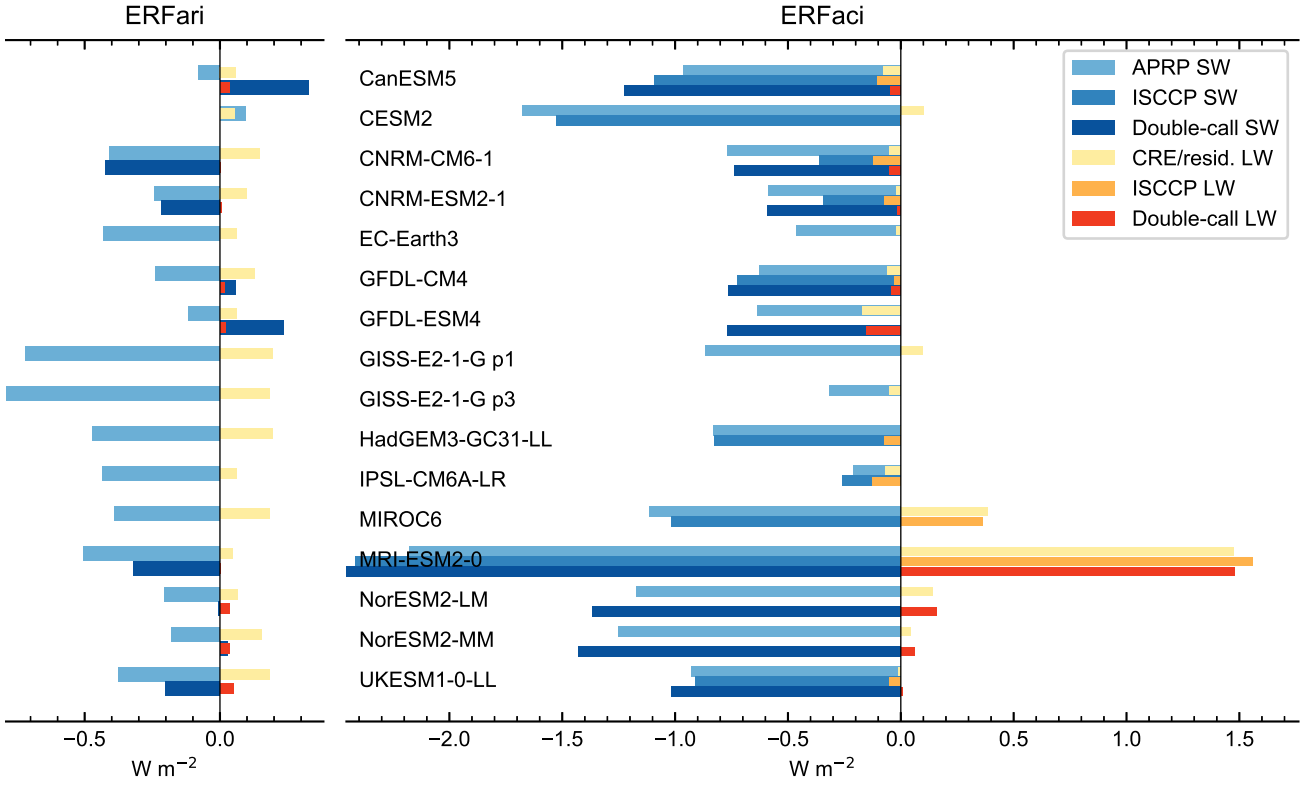

Figure 11. Comparison of methods to estimate ERFaci and ERFari from the aerosol experiment. CRE/resid. is the LW cloud radiative effect for ERFaci and the difference of LW ERF and CRE for ERFari. Not all methods are available in all models.

experiment. For ERFaci in both the SW and LW, different methods provide similar estimates. For ERFari, the APRP and double-call methods sometimes disagree ~~, however the ERFari is smaller than ERFaci . Additionally, it can also be noted that the LWERFari+aci from the~~ on sign for SW forcing, but this component is relatively small compared to the SW ERFaci where estimates are generally more consistent between APRP and the double call. Similarly in the LW, the CRE and double call methods produce similar results for ERFaci with larger relative differences for the smaller ERFari component. The double-call method ~~does not always equal the total ERF and may introduce a residual. For the CRE residual method the implied ERFari +aci equals the true ERF by definition, but some non-cloud rapid adjustments may be aliased into the total~~is considered to be quite reliable, but the good agreement between the APRP or CRE and double call methods suggest that these simpler tools are useful tool to diagnose ERFari and ERFaci from climate models, which is advantageous due to there being no requirement for specialised model diagnostics.

5.4 Land-use change

Land use ERF is small and not significant at ~~−0.08 (±0.14~~−0.09 (±0.13) W m^{-2}. Forcing and adjustments are difficult to distinguish from zero and it is unlikely that this forcing played a large role historically for global mean impacts. In ~~11 of 12 models (land cover change is not provided in CNRM-CM6.1, so this experimentwas not run)~~13 of the 14 models that ran this experiment, land-use ERF is negative, and the multi-model mean and standard deviation is affected by a relatively large positive forcing in the NorESM2-LM model~~, whereas all other models show a negative ERF~~. In fig. S3g we show that this is due to cloud adjustments in this model. This is a consequence of interactive isoprene and monoterpene specified from the land surface changes, causing a reduction in organic matter, reducing cloud condensation nuclei and increasing SW cloud adjustment (unlike for the aerosol forcing experiment, the Twomey effect in response to a land use forcing is treated as an adjustment and not a forcing, because anthropogenic aerosol emissions are not perturbed). In other models, where the ERF is small and negative, it should also be borne in mind that internal variability may make it more difficult to isolate the forcing signal from the noise in free-running simulations (Forster et al., 2016), although the multi-model mean is likely to be more robust than individual model results. This experiment was partly motivated by a large land use forcing of -0.4 W m^{-2} in the CMIP5 HadGEM2-ES model (Andrews et al., 2017)~~; the successor UKESM1-0-LL model also shows a relatively strong land-use forcingof −0.30~~, which showed a large change in regional dust loading that contributed to this forcing. Our multi-model mean ERF of -0.09 W m^{-2} ~~, although the physical model from the same group (HadGEM3-GC31-LL) does not show such a strong land use forcing (~~(-0.12 W m^{-2} if NorESM2-LM is excluded) agrees well with an observational-constrained analysis from CMIP5 models of -0.11 W m^{-2} (Lejeune et al., 2020) , and is within the likely range of the AR5 assessment of -0.15 (-0.05 to -0.25) W m^{-2}~~)~~.

The radiative forcing from land use change is driven by the resulting change in surface albedo. For example, deforestation for agricultural use converts relatively dark forest cover to brighter cropland, exerting a negative forcing (Betts, 2000, 2001). The surface albedo kernel-derived flux change is taken to be the IRF. It is not a perfect measure as it includes changes in snow and ice cover over land, and any biophysical response, as both changes in land surface temperatures and surface properties can affect snow cover. However, the land surface temperature change is very small in the land-use experiment, evidenced by the small land surface temperature adjustment in fig. 3. In Fig. S4 we show changes in aerosol optical depth at 550 nm for models that provided this diagnostic. There is diversity in the model aerosol loadings to land-use forcing that does not appear to explain the diversity in land-use ERF between models. In particular, CanESM5 has a strong aerosol optical depth increase to land-use change, but a relatively weak ERF of -0.08 W m^{-2}. Changes in surface properties such as how snow cover settles over different land types and the biophysical response is not easy to discern from model output. Again, all available model years are used because stratospheric temperature adjustment does not play a large role.

The spatial pattern of land-use forcing and adjustments (fig. 12) is generally not significant in many parts of the world due to the small size of the forcing. The exception to this is water vapour and SW cloud adjustments over the Amazon; deforestation from pre-industrial to present-day is likely to have reduced evapotranspiration from vegetation, reducing tropospheric humidity

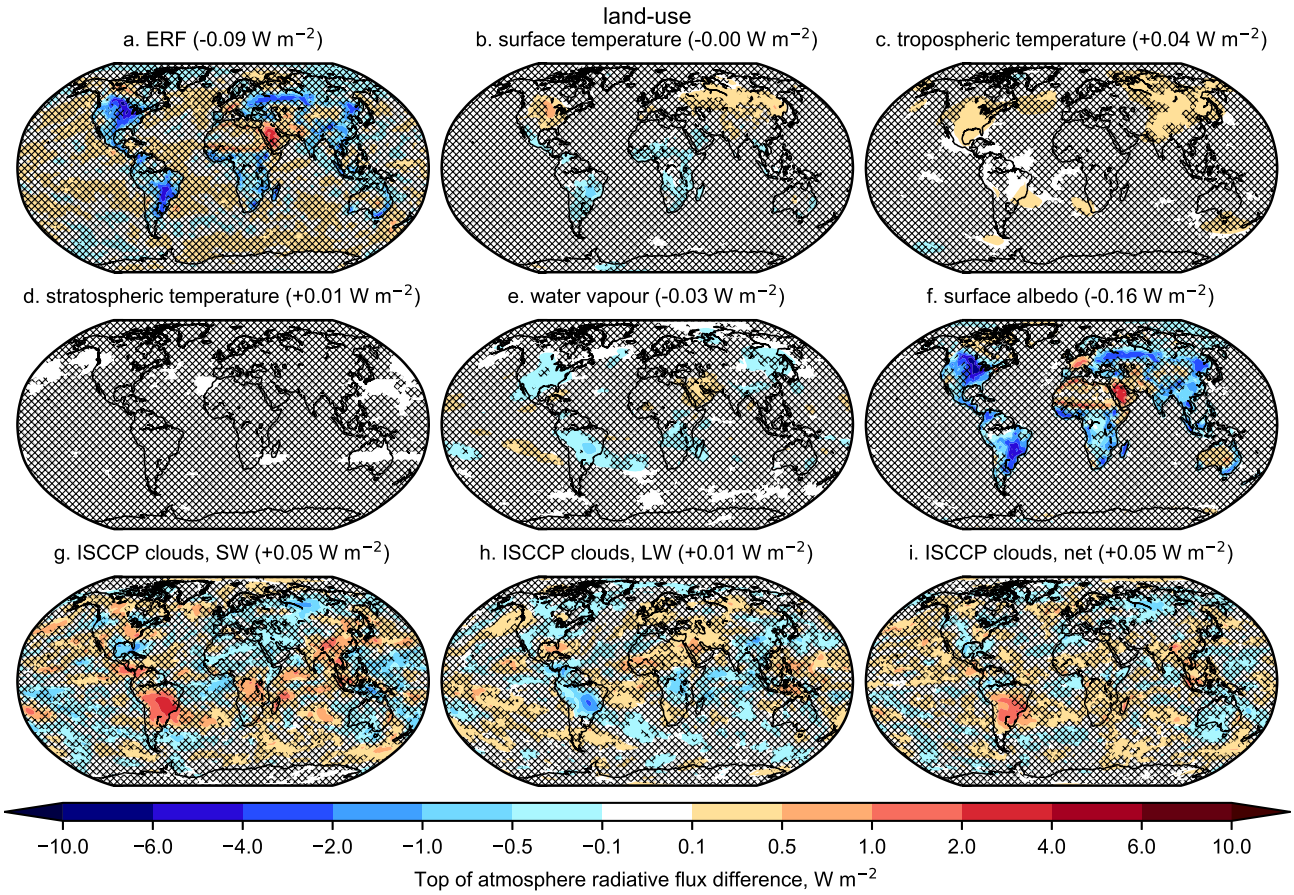

Figure 12. As fig. 6 but for present-day land use forcing. Subplot (f) represents the IRF in this experiment.

and low level cloud cover. These spatial patterns are also coincident with a decrease in organic carbon loading in NorESM2-LM (fig. S3).

5.5 Anthropogenic total

The total anthropogenic ERF for 1850–2014 stands at ~~1.97~~ 2.00 (± ~~0.26~~0.23) W m^{-2}~~(12 models; MPI-ESM2-0 did not run this experiment; CNRM-CM6-1 is included but did not include effects of land use change)~~. Inter-model spread is larger, both in relative and absolute terms, in the total anthropogenic forcing than it is for any of its individual components, suggesting that individual models respond very differently to the same combinations of forcing. In the absence of non-linearities between forcing components, the residual ERF of ~~+0.17~~ +0.21 W m^{-2} from the land-use, aerosol and WMGHG components compared to the total anthropogenic would mostly be comprised of ozone forcing, although the sum of individual forcings does not necessarily equal the total forcing in some models (Thornhill et al., 2020; O'Connor et al., 2020). As for the aerosol forcing experiment, there is no significant correlation between total anthropogenic forcing and ECS or TCR.

The total anthropogenic forcing shows the offsetting influences of the greenhouse gas and aerosol forcing components on the ERF, IRF and adjustments. The total anthropogenic ISCCP-simulator cloud changes are also a combination of the WMGHG and aerosol contributions, with the net effect being dominated by aerosol. For non-cloud adjustments the combination of strong positive adjustments from the greenhouse gas forcing with a small negative adjustment from the aerosol forcing results in an adjustment that is of comparable magnitude to the IRF. The exception is the GISS-E2-1-G model r1i1p1f1 variant that has a very strong negative cloud adjustment, driven by a large increase in cloud fraction in the aerosol experiment (table 6). Additionally, stratospheric adjustment is stronger for total anthropogenic forcing than for WMGHGs alone, suggesting a role for ~~tropospheric~~ ozone forcing in contributing to this adjustment.

The pattern of anthropogenic forcing is spatially inhomogeneous; positive where aerosol forcing is weak, and negative where localised aerosol-cloud effects dominate (fig. 13). The influence of WMGHG forcing on temperature and water vapour adjustments, and of aerosol forcing on the cloud response, is evident. For all forcings, but particularly for land-use, aerosol and total anthropogenic, many of the forcing and adjustment terms do not show robust signals regionally. This indicates that adjustments are best considered as global-mean quantities that affect the globally-resolved forcing-feedback framework (eq. (1)).

6 Conclusions

Effective radiative forcing is the driving process behind long-term changes in global mean surface temperature. As ERF is now preferred to RF, climate models are the best tools we have to determine the heating impacts of various species on the Earth atmosphere system.

From CMIP5 to CMIP6, both CO_2 and aerosol forcing has become more consistent across the population of ~~models.~~ participating models. This has helped to address a concern from CMIP5: that forcing was poorly characterised in CMIP5 models and inconsistently determined (Stouffer et al., 2017) . Multi-model mean CO_2 and all-WMGHG ERF estimates agree very well with RF estimates from Etminan et al. (2016) using a line-by-line radiative transfer model (sections 5.1 and 5.2). A comprehensive review of aerosol forcing placed the 16–84% uncertainty range in present-day aerosol ERF at -1.60 to -0.65 W m^{-2} (Bellouin et al., 2020b). Results from CMIP6 models show a ~~smaller range of -1.27 to -0.81~~ relatively tight spread of -1.37 to -0.63 W m^{-2} ~~(one standard deviation, equivalent to 16–84% range for a Gaussian distribution), though~~ for the full range. Although 17 models is a reasonable sample size of the CMIP6 population, more models may submit forcing results to CMIP6 that would widen this range ~~. Present-day aerosol forcing does not explain the increase in the range of ECS in~~ (and indeed, we would encourage modelling groups to do so). One example is E3SM which did not perform the RFMIP aerosol forcing experiment but where it would be likely that the 1850–2014 aerosol forcing would be more negative than -1.37 W m^{-2} (fig. 25 in Golaz et al., 2019) . While the increase in $4\times CO_2$ forcing compared to CMIP5 may explain some of the increase in climate sensitivity in CMIP6 models ~~, particularly~~ (Zelinka et al., 2020) , the model range of present-day aerosol forcing does not, particularly for the upper bound.

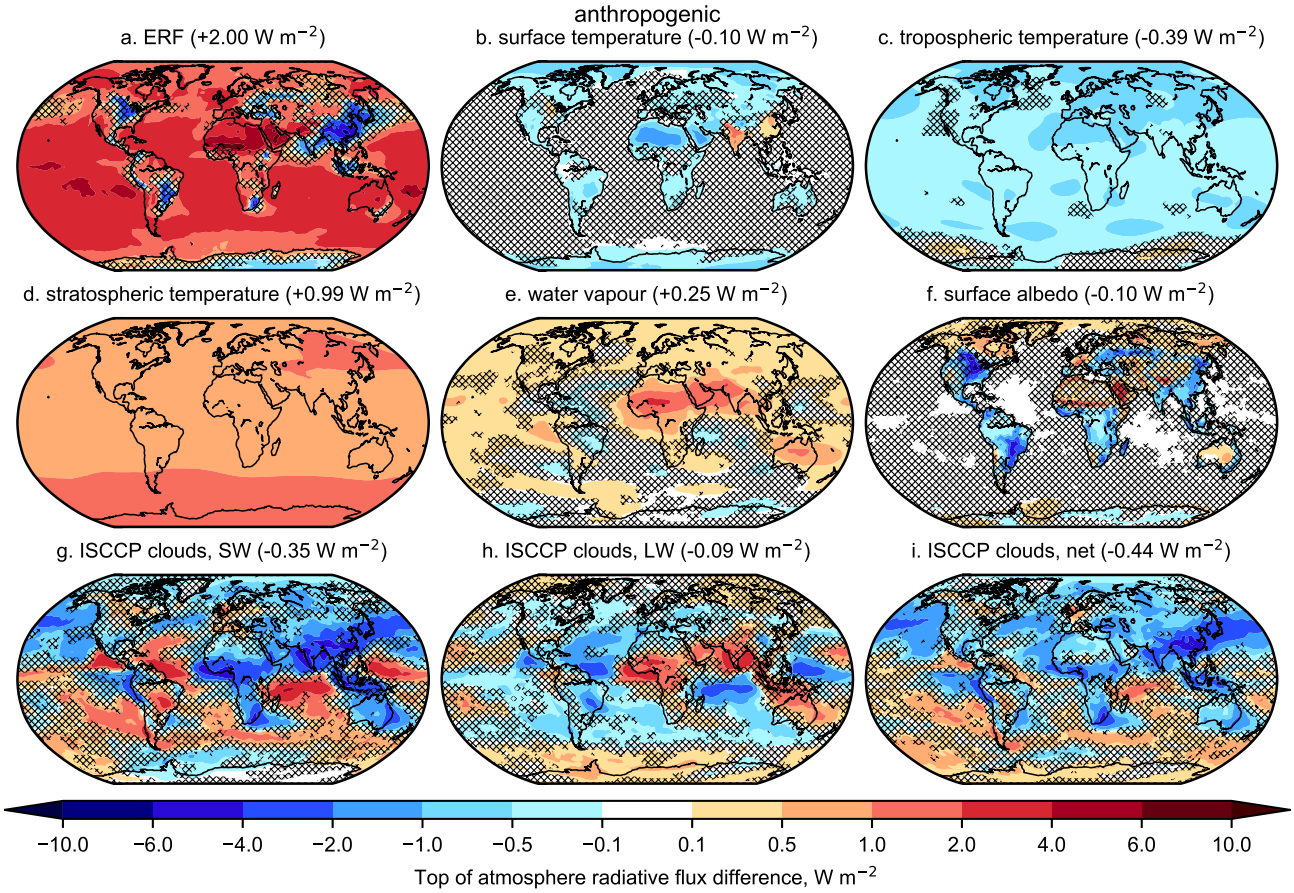

Figure 13. As fig. 6 but for present-day anthropogenic forcing.

We determine a multi-model mean anthropogenic ERF of ~~1.97 (±0.26~~2.00 (±0.23) W m^{-2} for 1850–2014. This is less than the anthropogenic ERF in AR5 for 1850–2011 of 2.24 W m^{-2} (~~although with large uncertainty~~ Myhre et al. (2013) ; although this figure has a wide uncertainty range), and extrapolating trends forward would suggest an anthropogenic ERF of around 2.4 W m^{-2} from AR5 for 1850–2014. The two main reasons for this difference are a stronger negative aerosol forcing in CMIP6 compared to the AR5 assessment (~~−1.04~~ −1.01 W m^{-2} in CMIP6 for 1850–2014 versus −0.72 W m^{-2} in AR5 for 1850–2011), and a weaker ozone forcing (~~+0.17~~ +0.21 W m^{-2} versus +0.31 W m^{-2}), if residual anthropogenic forcing is attributed to ozone.

Forcing adjustments produce insight into the atmospheric mechanisms that contribute to ERF. Warming of the troposphere results in a negative adjustment due to the increase in outgoing LW radiation, and increasing water vapour counteracts this effect partially by its role as a greenhouse gas. ~~With one exception, all~~ All models agree on tropospheric warming and moistening for WMGHG and ~~all-anthropogenic~~ all anthropogenic forcing. These tropospheric adjustments are small for aerosol forcing but models do not agree on the sign of the change. ~~Cloud adjustments remain the largest uncertainty~~ The instantaneous radiative

forcing and cloud adjustments are generally the largest sources of inter-model spread in the forcing component in climate models~~, as well as in the feedback and climate response (Zelinka et al., 2020) ,~~ . Since IRF is not directly calculated in this study, some of this spread may be from residuals in the kernel decomposition and the true spread in IRF may be smaller than reported here. One strand of RFMIP will include benchmarking of GCM radiative transfer against line-by-line codes. Radiative transfer is a well-grounded theoretical problem where the diversity in line-by-line codes is small (Pincus et al., 2015) , so this component of inter-model diversity has a measurable yardstick for improvement. Cloud responses are more difficult to constrain and exhibit a wide range of behaviour to both greenhouse gas and aerosol forcing. ~~Constraining cloud responses to both forcingand climate response should therefore continue to be a priority~~However, progress is beginning to be made. For greenhouse gas forcing, techniques from the climate feedback literature that have observational parallels, such as analysing cloud-controlling factors (Klein et al., 2017) , can be applied to adjustments. Use of the ISCCP simulator diagnostics with the ISCCP cloud kernel, another method conceptualised by climate feedback investigations (Zelinka et al., 2012) , allows cloud adjustments to be calculated directly facilitating better inter-model comparison. For aerosol forcing, observational methods exist to determine RFari and RFaci using satellite and reanalysis data (Bellouin et al., 2013, 2020a) . Ultimately, reducing uncertainty in effective radiative forcing will reduce uncertainty in climate projections due to the central role of forcing in driving the Earth's global mean temperature response.

Data availability. RFMIP model data used in this study is freely available from the CMIP6 repository on the Earth System Grid Foundation nodes (https://esgf-node.llnl.gov/search/cmip6/) with the exception of ACCESS-CM2 which is currently unpublished. The HadGEM3-GA7.1 kernels are available at https://doi.org/10.5281/zenodo.3594673.

Author contributions. C.J.S. co-ordinated the project, analysed the data and led the writing of the manuscript. R.J.K. produced adjustment calculations using radiative kernels. G.M. provided adjustment calculations using offline radiation simulations. K.A. provided model results of effective radiative forcing and effective climate sensitivity. W.C. provided analysis of different definitions of effective radiative forcing. R.P. and P.M.F. oversaw the RFMIP-ERF project in coordination with the World Climate Research Programme under the CMIP6 protocol. All other authors ran models and published results on the Earth System Grid Foundation, without which this work would not have been possible. All authors contributed to the writing and review.

Competing interests. The authors declare no competing interests.

Disclaimer. TEXT

Acknowledgements. We thank Ed Gryspeerdt, Mark Zelinka, Roland Séférian and the participants and hosts of the Tri-MIP-athlon consortium meetings in Reading 2018 and Princeton 2019 for fruitful discussions. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF.

C.J.S. was supported by a NERC/IIASA Collaborative Research Fellowship (NE/T009381/1). C.J.S. and P.M.F. were supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 820829 (CONSTRAIN project). R.J.K. is supported by an appointment to the NASA Postdoctoral Program at NASA Goddard Space Flight Center. ~~RP~~ R.P. was supported by the US Department of Energy's Office of Biological and Environmental Research under grant 7457436 from Lawrence Berkeley National Lab. T.A., F.M.O'C., E.R., and A.W. were supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra and the Newton Fund through the Met Office Climate Science for Service Partnership Brazil (CSSP Brazil). F.M.O'C. also acknowledges support from the EU Horizon 2020 Research Programme CRESCENDO project, grant agreement number 641816. A.K. and D.O. were supported by the Research Council of Norway (grant nos. 229771, 285003, and 285013), by Notur/NorStore (NN2345K and NS2345K), and through EU H2020 grant no. 280060. H.S. was supported by TOUGOU (MEXT, Japan). The MIROC6 simulations were performed using the Earth Simulator at JAMSTEC and the NEC SX at NIES. The CMIP6 project at IPSL used the HPC resources of TGCC under the allocations 2017-R0040110492 and 2018-R0040110492 (project gencmip6) provided by GENCI (Grand Équipement National de Calcul Intensif). The CESM project is supported primarily by the National Science Foundation. Part of this material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977. The simulations with EC-Earth3 were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at the National Supercomputer Centre (NSC) partially funded by the Swedish Research Council through grant agreement no. 2016-07213. M.D. and C.M. were supported by funding from the Earth Systems and Climate Change Hub of the Australian Government's National Environmental Science Program. ACCESS modelling was undertaken with the assistance of resources from the National Computational Infrastructure (NCI Australia), an NCRIS enabled capability supported by the Australian Government.

References

Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, Science, 245, 1227–1231, 1989.

Allen, R., Amiri-Farahani, A., Lamarque, J.-F., Smith, C., Shindell, D., Hassan, T., and Chung, C.: npj Climate and Atmospheric Science, 2, https://doi.org/10.1038/s41612-019-0073-9, 2019.

Andrews, T., Gregory, J. M., Webb, M. J., and Taylor, K. E.: Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models, Geophys. Res. Lett., 39, L09 712, https://doi.org/10.1029/2012GL051607, 2012.

Andrews, T., Betts, R. A., Booth, B. B. B., Jones, C. D., and Jones, G. S.: Effective radiative forcing from historical land use change, Clim. Dynam., 48, 3489–3505, https://doi.org/10.1007/s00382-016-3280-7, 2017.

Andrews, T., Andrews, M. B., Bodas-Salcedo, A., Jones, G. S., Kuhlbrodt, T., Manners, J., Menary, M. B., Ridley, J., Ringer, M. A., Sellar, A. A., Senior, C. A., and Tang, Y.: Forcings, Feedbacks, and Climate Sensitivity in HadGEM3-GC3.1 and UKESM1, J. Adv. Model. Earth Sy., 11, https://doi.org/10.1029/2019MS001866, 2019.

Armour, K.: Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks, Nat. Clim. Change, 7, 331–335, https://doi.org/10.1038/nclimate3278, 2017.

Bellouin, N., Quaas, J., Morcrette, J.-J., and Boucher, O.: Estimates of aerosol radiative forcing from the MACC re-analysis, Atmospheric Chemistry and Physics, 13, 2045–2062, https://doi.org/10.5194/acp-13-2045-2013, https://www.atmos-chem-phys.net/13/2045/2013/, 2013.

Bellouin, N., Davies, W., Shine, K., Quaas, J., Muelmenstaedt, J., Block, K., Forster, P., Smith, C., Lee, L., Regayre, L., Brasseur, G., Sudarchikova, N., Bouarar, I., Boucher, O., and Myhre, G.: Radiative forcing of climate change from the Copernicus reanalysis of atmospheric composition, Earth Syst. Sci. Data, https://doi.org/10.5194/essd-2019-251, 2020a.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K., Christensen, M., Daniau, A.-L., Dufresne, J.-L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B.: Bounding global aerosol radiative forcing of climate change, Rev. Geophys., https://doi.org/10.1029/2019RG000660, 2020b.

Betts, R.: Offset of the potential carbon sink from boreal forestation by decreases in surface albedo, Nature, 408, 187–190, https://doi.org/10.1038/35041545, 2000.

Betts, R. A.: Biogeophysical impacts of land use on present-day climate: near-surface temperature change and radiative forcing, Atmos. Sci. Lett., 2, 39–51, https://doi.org/10.1006/asle.2001.0037, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1006/asle.2001.0037, 2001.

Bi, D., Dix, M., Marsland, S., O'Farrell, S., Sullivan, A., Bodman, R., Law, R., Harman, I., Srbinovsky, J., Rashid, H., Dobrohotoff, P., Mackallah, C., Yan, H., Hirst, A., Savita, A., Dias, F. B., Woodhouse, M., Fiedler, R., and Heerdegen, A.: Configuration and spin-up of ACCESS-CM2, the new generation Australian Community Climate and Earth System Simulator Coupled Model, Journal of Southern Hemisphere Earth Systems Science, submitted.

Block, K. and Mauritsen, T.: Forcing and feedback in the MPI-ESM-LR coupled model under abruptly quadrupled CO_2, J. Adv. Model. Earth Sy., pp. 696–691, https://doi.org/10.1002/jame.20041, 2013.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S., Sherwood, S., Stevens, B., and Zhang, X.: 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, edited by Stocker,

T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., pp. 571–658, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, L., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traoré, A.-K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and evaluation of the IPSL-CM6A-LR climate model, J. Adv. Model. Earth Sy., submitted.

Caballero, R. and Huber, M.: State-dependent climate sensitivity in past warm climates and its implications for future climate projections, P. Natl. Acad. Sci. USA, 110, 14 162–14 167, https://doi.org/10.1073/pnas.1303365110, 2013.

Chung, E.-S. and Soden, B. J.: An assessment of methods for computing radiative forcing in climate models, Env. Res. Lett., 10, 074 004, https://doi.org/10.1088/1748-9326/10/7/074004, https://doi.org/10.1088%2F1748-9326%2F10%2F7%2F074004, 2015.

Chylek, P., Vogelsang, T. J., Klett, J. D., Hengartner, N., Higdon, D., Lesins, G., and Dubey, M. K.: Indirect Aerosol Effect Increases CMIP5 Models' Projected Arctic Warming, Journal of Climate, 29, 1417–1428, https://doi.org/10.1175/JCLI-D-15-0362.1, 2016.

Collins, W. J., Lamarque, J.-F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., Shindell, D., and Smith, S. J.: AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6, Geosci. Model Dev., 10, 585–607, https://doi.org/10.5194/gmd-10-585-2017, https://www.geosci-model-dev.net/10/585/2017/, 2017.

Colman, R. and McAvaney, B.: Climate feedbacks under a very broad range of forcing, Geophysical Research Letters, 36, https://doi.org/10.1029/2008GL036268, 2009.

Copernicus Climate Change Service: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, https://cds.climate.copernicus.eu/cdsapp#!/home, accessed 28 October 2019, 2017.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), J. Adv. Model. Earth Sy., 12, e2019MS001 916, https://doi.org/10.1029/2019MS001916, 2020.

Dunne et al.: The GFDL Earth System Model version 4.1 (GFDL-ESM4.1): Model description and simulation characteristics, J. Adv. Model. Earth Sy., in prep.

Edwards, J. M. and Slingo, A.: Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model, Q. J. Roy. Meteor. Soc., 122, 689–719, https://doi.org/10.1002/qj.49712253107, 1996.

Etminan, M., Myhre, G., Highwood, E. J., and Shine, K. P.: Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing, Geophys. Res. Lett., 43, 12,614–12,623, https://doi.org/10.1002/2016GL071930, 2016GL071930, 2016.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, https://www.geosci-model-dev.net/9/1937/2016/, 2016.

Fiedler, S., Stevens, B., and Mauritsen, T.: On the sensitivity of anthropogenic aerosol forcing to model-internal variability and parameterizing a Twomey effect, J. Adv. Model. Earth Sy., 9, 1325–1341, https://doi.org/10.1002/2017MS000932, 2017.

Flynn, C. M. and Mauritsen, T.: On the Climate Sensitivity and Historical Warming Evolution in Recent Coupled Model Ensembles, Atmospheric Chemistry and Physics, https://doi.org/10.5194/acp-2019-1175, 2020.

Forster, P., Richardson, T., Maycock, A., Smith, C., Samset, B., Myhre, G., Andrews, T., Pincus, R., and Schulz, M.: Recommendations for diagnosing effective radiative forcing from climate models from CMIP6, J. Geophys. Res., 121, 12 460–12 475, https://doi.org/10.1002/2016JD025320, 2016.

Forster, P., Maycock, A., McKenna, C., and Smith, C.: Latest climate models confirm need for urgent mitigation, Nat. Clim. Change, 10, 7–10, 2020.

Forster, P. M., Andrews, T., Good, P., Gregory, J. M., Jackson, L. S., and Zelinka, M.: 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, https://doi.org/10.1002/jgrd.50174, 2013.

Forster, P. M. d. F. and Shine, K. P.: Radiative forcing and temperature trends from stratospheric ozone changes, J. Geophys. Res.-Atmos., 102, 10 841–10 855, https://doi.org/10.1029/96JD03510, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD03510, 1997.

Ghan, S.: Technical Note: Estimating aerosol effects on cloud radiative forcing, Atmos. Chem. Phys., 13, 9971–9974, https://doi.org/10.5194/acp-13-9971-2013, http://www.atmos-chem-phys.net/13/9971/2013/, 2013.

Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., van Vuuren, D. P., van den Berg, M., Feng, L., Klein, D., Calvin, K., Doelman, J. C., Frank, S., Fricko, O., Harmsen, M., Hasegawa, T., Havlik, P., Hilaire, J., Hoesly, R., Horing, J., Popp, A., Stehfest, E., and Takahashi, K.: Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century, Geosci. Model Dev., 12, 1443–1475, https://doi.org/10.5194/gmd-12-1443-2019, https://www.geosci-model-dev.net/12/1443/2019/, 2019.

Golaz, J.-C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., Abeshu, G., Anantharaj, V., Asay-Davis, X. S., Bader, D. C., Baldwin, S. A., Bisht, G., Bogenschutz, P. A., Branstetter, M., Brunke, M. A., Brus, S. R., Burrows, S. M., Cameron-Smith, P. J., Donahue, A. S., Deakin, M., Easter, R. C., Evans, K. J., Feng, Y., Flanner, M., Foucar, J. G., Fyke, J. G., Griffin, B. M., Hannay, C., Harrop, B. E., Hoffman, M. J., Hunke, E. C., Jacob, R. L., Jacobsen, D. W., Jeffery, N., Jones, P. W., Keen, N. D., Klein, S. A., Larson, V. E., Leung, L. R., Li, H.-Y., Lin, W., Lipscomb, W. H., Ma, P.-L., Mahajan, S., Maltrud, M. E., Mametjanov, A., McClean, J. L., McCoy, R. B., Neale, R. B., Price, S. F., Qian, Y., Rasch, P. J., Reeves Eyre, J. E. J., Riley, W. J., Ringler, T. D., Roberts, A. F., Roesler, E. L., Salinger, A. G., Shaheen, Z., Shi, X., Singh, B., Tang, J., Taylor, M. A., Thornton, P. E., Turner, A. K., Veneziani, M., Wan, H., Wang, H., Wang, S., Williams, D. N., Wolfram, P. J., Worley, P. H., Xie, S., Yang, Y., Yoon, J.-H., Zelinka, M. D., Zender, C. S., Zeng, X., Zhang, C., Zhang, K., Zhang, Y., Zheng, X., Zhou, T., and Zhu, Q.: The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution, Journal of Advances in Modeling Earth Systems, 11, 2089–2129, https://doi.org/10.1029/2018MS001603, 2019.

Gregory, J., Ingram, W., Palmer, M., Jones, G., Stott, P., Thorpe, R., Lowe, J., Johns, T., and Williams, K.: A new method for diagnosing radiative forcing and climate sensitivity, Geophys. Res. Lett., 31, https://doi.org/10.1029/2003GL018747, 2004.

Gregory, J. M., Andrews, T., and Good, P.: The inconstancy of the transient climate response parameter under increasing CO2, Philos. T. R. Soc. A, 373, https://doi.org/10.1098/rsta.2014.0417, 2015.

Gryspeerdt, E., Mülmenstädt, J., Gettelman, A., Malavelle, F. F., Morrison, H., Neubauer, D., Partridge, D. G., Stier, P., Takemura, T., Wang, H., Wang, M., and Zhang, K.: Surprising similarities in model and observational aerosol radiative forcing estimates, Atmospheric Chemistry and Physics Discussions, 2019, 1–18, https://doi.org/10.5194/acp-2019-533, https://www.atmos-chem-phys-discuss.net/acp-2019-533/, 2019.

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., Russell, G., Aleinov, I., Bauer, M., Bauer, S., Bell, N., Cairns, B., Canuto, V., Chandler, M., Cheng, Y., Del Genio, A., Faluvegi, G., Fleming, E., Friend, A., Hall, T., Jackman, C., Kelley, M., Kiang, N., Koch, D., Lean, J., Lerner, J., Lo, K., Menon, S., Miller, R., Minnis, P., Novakov, T., Oinas, V., Perlwitz, J., Perlwitz, J., Rind, D., Romanou, A., Shindell, D., Stone, P., Sun, S., Tausnev, N., Thresher, D., Wielicki, B., Wong, T., Yao, M., and Zhang, S.: Efficacy of climate forcings, J. Geophys. Res.-Atmos., 110, https://doi.org/10.1029/2005JD005776, d18104, 2005.

Haustein, K., Allen, M., Forster, P., Otto, F., Mitchell, D., Matthews, H., and Frame, D.: A real-time Global Warming Index, Scientific Reports, 7, 15 417, 2017.

Held, I. M. and Shell, K. M.: Using Relative Humidity as a State Variable in Climate Feedback Analysis, Journal of Climate, 25, 2578–2582, https://doi.org/10.1175/JCLI-D-11-00721.1, 2012.

Held, I. M. and Soden, B. J.: Water vapor feedback and global warming, Annu. Rev. Energ. Env., 25, 441–475, https://doi.org/10.1146/annurev.energy.25.1.441, 2000.

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., Shevliakova, E., Winton, M., Zhao, M., Bushuk, M., Wittenberg, A. T., Wyman, B., Xiang, B., Zhang, R., Anderson, W., Balaji, V., Donner, L., Dunne, K., Durachta, J., Gauthier, P. P. G., Ginoux, P., Golaz, J.-C., Griffies, S. M., Hallberg, R., Harris, L., Harrison, M., Hurlin, W., John, J., Lin, P., Lin, S.-J., Malyshev, S., Menzel, R., Milly, P. C. D., Ming, Y., Naik, V., Paynter, D., Paulot, F., Rammaswamy, V., Reichl, B., Robinson, T., Rosati, A., Seman, C., Silvers, L. G., Underwood, S., and Zadeh, N.: Structure and Performance of GFDL's CM4.0 Climate Model, J. Adv. Model. Earth Sy., https://doi.org/10.1029/2019MS001829, 2019.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J.-I., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geosci. Model Dev., 11, 369–408, https://doi.org/10.5194/gmd-11-369-2018, 2018.

Huang, Y.: On the longwave climate feedbacks, J. Climate, 26, 7603–7610, https://doi.org/10.1175/JCLI-D-13-00025.1, 2013.

Johnson, B. T., Haywood, J. M., and Hawcroft, M. K.: Are Changes in Atmospheric Circulation Important for Black Carbon Aerosol Impacts on Clouds, Precipitation, and Radiation?, J. Geophys. Res.-Atmos., 124, 7930–7950, https://doi.org/10.1029/2019JD030568, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD030568, 2019.

Jonko, A. K., Shell, K. M., Sanderson, B. M., and Danabasoglu, G.: Climate Feedbacks in CCSM3 under Changing CO2 Forcing. Part II: Variation of Climate Feedbacks and Sensitivity with Forcing, J. Climate, 26, 2784–2795, https://doi.org/10.1175/JCLI-D-12-00479.1, 2013.

Kamae, Y. and Watanabe, M.: On the robustness of tropospheric adjustment in CMIP5 models, Geophysical Research Letters, 39, https://doi.org/10.1029/2012GL054275, 2012.

Kelley, M., Schmidt, G. A., Nazarenko, L., Miller, R. L., Bauer, S. E., Ruedy, R., Russell, G. L., Aleinov, I., Bauer, M., Bleck, R., Canuto, V., Cesana, G., Cheng, Y., Clune, T. L., Cook, B., Cruz, C. A., Genio, A. D. D., Elsaesser, G. S., Faluvegi, G., Kiang, N. Y., Kim, D., Lacis, A. A., Leboissetier, A., LeGrande, A. N., Lo, K. K., Marshall, J. C., McDermid, S., Matthews, E. E., Mezuman, K., Murray, L. T., Oinas, V., Orbe, C., Garcia-Pando, C. P., Perlwitz, J. P., Puma, M. J., Rind, D., Romanou, A., Shindell, D. T., Sun, S., Tausnev, N.,

Tsigaridis, K., Tselioudis, G., Weng, E., Wu, J., and Yao, M.-S.: GISS-E2.1: Configurations and Climatology, J. Adv. Model. Earth Sy., https://pubs.giss.nasa.gov/docs/tbp/sub_Kelley_ke01200v.pdf, submitted.

Kiehl, J. T.: Twentieth century climate model response and climate sensitivity, Geophysical Research Letters, 34, https://doi.org/10.1029/2007GL031383, 2007.

Kirkevåg, A., Grini, A., Olivié, D., Seland, Ø., Alterskjær, K., Hummel, M., Karset, I. H. H., Lewinschal, A., Liu, X., Makkonen, R., Bethke, I., Griesfeller, J., Schulz, M., and Iversen, T.: A production-tagged aerosol module for Earth system models, OsloAero5.3 – extensions and updates for CAM5.3-Oslo, Geosci. Model Dev., 11, 3945–3982, https://doi.org/10.5194/gmd-11-3945-2018, https://www.geosci-model-dev.net/11/3945/2018/, 2018.

Klein, S. and Jakob, C.: Validation and Sensitivities of Frontal Clouds Simulated by the ECMWF Model, Mon. Weather Rev., 127, 2514–2531, https://doi.org/10.1175/1520-0493(1999)127<2514:VASOFC>2.0.CO;2, 1999.

Klein, S., Hall, A., Norris, J., and Pincus, R.: Low-Cloud Feedbacks from Cloud-Controlling Factors: A Review., Surveys in Geophysics, 38, 2017.

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, Atmos. Chem. and Phys., 10, 7017–7039, https://doi.org/10.5194/acp-10-7017-2010, https://www.atmos-chem-phys.net/10/7017/2010/, 2010.

Lejeune, Q., Davin, E. L., Duveiller, G., Crezee, B., Meier, R., Cescatti, A., and Seneviratne, S. I.: Biases in the albedo sensitivity to deforestation in CMIP5 models and their impacts on the associated historical Radiative Forcing, Earth System Dynamics Discussions, 2020, 1–34, https://doi.org/10.5194/esd-2019-94, 2020.

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., Liang, L., Mitrescu, C., Rose, F. G., and Kato, S.: Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product, J. Climate, 31, 895–918, https://doi.org/10.1175/JCLI-D-17-0208.1, 2018.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2, J. Adv. Model. Earth Sy., 11, 998–1038, https://doi.org/10.1029/2018MS001400, 2019.

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J., and Weiss, R.: Historical greenhouse gas concentrations for climate modelling (CMIP6), Geosci. Model Dev., 10, 2057–2116, https://doi.org/10.5194/gmd-10-2057-2017, https://www.geosci-model-dev.net/10/2057/2017/, 2017.

Mülmenstädt, J., Gryspeerdt, E., Salzmann, M., Ma, P.-L., Dipu, S., and Quaas, J.: Separating radiative forcing by aerosol–cloud interactions and fast cloud adjustments in the ECHAM-HAMMOZ aerosol–climate model using the method of partial radiative perturbations, Atmospheric Chemistry and Physics Discussions, 2019, 1–20, https://doi.org/10.5194/acp-2018-1304, https://www.atmos-chem-phys-discuss.net/acp-2018-1304/, 2019.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: 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, edited by Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., pp. 659–740, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Myhre, G., Forster, P., Samset, B., Hodnebrog, Ø., Sillmann, J., Aalbergsjø, S., Andrews, T., Boucher, O., Faluvegi, G., Fläschner, D., , Iversen, T., Kasoar, M., Kharin, V., Kirkevå g, A., Lamarque, J.-F., Olivié, D., Richardson, T. B., Shindell, D., Shine, K. P., Stjern, C. W., Takemura, T., Voulgarakis, A., and Zwiers, F.: PDRMIP: A precipitation driver and response model intercomparison project—Protocol and preliminary results, B. Am. Meteorol. Soc., 98, 1185–1198, 2017.

O'Connor, F. M., Abraham, N. L., Dalvi, M., Folberth, G., Griffiths, P., Hardacre, C., Johnson, B. T., Kahana, R., Keeble, J., Kim, B., Morgenstern, O., Mulcahy, J. P., Richardson, M. G., Robertson, E., Seo, J., Shim, S., Teixeira, J. C., Turnock, S., Williams, J., Wiltshire, A., and Zeng, G.: Assessment of pre-industrial to present-day anthropogenic climate forcing in UKESM1, Atmospheric Chemistry and Physics Discussions, 2020, 1–49, https://doi.org/10.5194/acp-2019-1152, 2020.

Oshima, N., Yukimoto, S., Deushi, M., Koshiro, T., Kawai, H., Tanaka, T. Y., and Yoshida, K.: Global and Arctic Radiative Forcing of Anthropogenic Gases and Aerosols in MRI-ESM2.0, in prep.

Pendergrass, A. G., Conley, A., and Vitt, F. M.: Surface and top-of-atmosphere radiative feedback kernels for CESM-CAM5, Earth System Science Data, 10, 317–324, https://doi.org/10.5194/essd-10-317-2018, 2018.

Pincus, R., Mlawer, E. J., Oreopoulos, L., Ackerman, A. S., Baek, S., Brath, M., Buehler, S. A., Cady-Pereira, K. E., Cole, J. N. S., Dufresne, J.-L., Kelley, M., Li, J., Manners, J., Paynter, D. J., Roehrig, R., Sekiguchi, M., and Schwarzkopf, D. M.: Radiative flux and forcing parameterization error in aerosol-free clear skies, Geophys. Res. Lett., 42, 5485–5492, https://doi.org/10.1002/2015GL064291, 2015.

Pincus, R., Forster, P. M., and Stevens, B.: The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6, Geosci. Model Dev., 9, 3447–3460, https://doi.org/10.5194/gmd-9-3447-2016, 2016.

Richardson, T., Forster, P., Maycock, A., Smith, C., Wood, T., Andrews, T., Boucher, O., Faluvegi, G., Fläschner, D., Hodnebrog, O., Kasoar, M., Kirkevag, A., Lamarque, J.-F., Mülmenstädt, J., Olivié, D., Samset, B., Shawki, D., Shindell, D., Takemura, T., and Voulgarakis, A.: Efficacy of climate forcings in PDRMIP models, J. Geophys. Res., 124, https://doi.org/https://doi.org/10.1029/2019JD030581, 2019.

Rossow, W. B., Walker, A. W., Beuschel, D. E., and Roiter, M. D.: International Satellite Cloud Climatology Project (ISCCP) Documentation of New Cloud Datasets, 1996.

Santer, B. D., Wehner, M. F., Wigley, T. M. L., Sausen, R., Meehl, G. A., Taylor, K. E., Ammann, C., Arblaster, J., Washington, W. M., Boyle, J. S., and Brüggemann, W.: Contributions of Anthropogenic and Natural Forcing to Recent Tropopause Height Changes, Science, 301, 479–483, https://doi.org/10.1126/science.1084123, 2003.

Schmidt, G. A., Kelley, M., Nazarenko, L., Ruedy, R., Russell, G. L., Aleinov, I., Bauer, M., Bauer, S. E., Bhat, M. K., Bleck, R., Canuto, V., Chen, Y.-H., Cheng, Y., Clune, T. L., Del Genio, A., de Fainchtein, R., Faluvegi, G., Hansen, J. E., Healy, R. J., Kiang, N. Y., Koch, D., Lacis, A. A., LeGrande, A. N., Lerner, J., Lo, K. K., Matthews, E. E., Menon, S., Miller, R. L., Oinas, V., Oloso, A. O., Perlwitz, J. P., Puma, M. J., Putman, W. M., Rind, D., Romanou, A., Sato, M., Shindell, D. T., Sun, S., Syed, R. A., Tausnev, N., Tsigaridis, K., Unger,

N., Voulgarakis, A., Yao, M.-S., and Zhang, J.: Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive, J. Adv. Model. Earth Syst., 6, 141–184, https://doi.org/10.1002/2013MS000265, 2014.

Seland, Ø., Bentsen, M., Seland Graff, L., Olivié, D., Toniazzo, T., Gjermundsen, A., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., Schwinger, J., Tjiputra, J., Schancke Aas, K., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Hafsahl Karset, I. H., Landgren, O., Liakka, J., Onsum Moseid, K., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iverson, T., and Schulz, M.: The Norwegian Earth System Model, NorESM2 – Evaluation of theCMIP6 DECK and historical simulations, Geoscientific Model Development Discussions, 2020, 1–68, https://doi.org/10.5194/gmd-2019-378, 2020.

Sellar, A. A., Jones, C. G., Mulcahy, J., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., de Mora, L., Kuhlbrodt, T., Rumbold, S., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and evaluation of the UK Earth System Model, J. Adv. Model. Earth Sy., https://doi.org/10.1029/2019MS001739, 2019.

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., Decharme, B., Delire, C., Berthet, S., Chevallier, M., Sénési, S., Franchisteguy, L., Vial, J., Mallet, M., Joetzjer, E., Geoffroy, O., Guérémy, J.-F., Moine, M.-P., Msadek, R., Ribes, A., Rocher, M., Roehrig, R., Salas-y Mélia, D., Sanchez, E., Terray, L., Valcke, S., Waldman, R., Aumont, O., Bopp, L., Deshayes, J., Éthé, C., and Madec, G.: Evaluation of CNRM Earth-System model, CNRM-ESM 2-1: role of Earth system processes in present-day and future climate, J. Adv. Model. Earth Sy., https://doi.org/10.1029/2019MS001791, 2019.

Shell, K., Kiehl, J., and Shields, C.: Using the Radiative Kernel Technique to Calculate Climate Feedbacks in NCAR's Community Atmospheric Model, J. Climate, 21, 2269–2282, https://doi.org/10.1175/2007JCLI2044.1, 2008.

Sherwood, S. C., Bony, S., Boucher, O., Bretherton, C., Forster, P. M., Gregory, J. M., and Stevens, B.: Adjustments in the Forcing-Feedback Framework for Understanding Climate Change, B. Am. Meteorol. Soc., 96, 217–228, https://doi.org/10.1175/BAMS-D-13-00167.1, 2015.

Shine, K. P., Cook, J., Highwood, E. J., and Joshi, M. M.: An alternative to radiative forcing for estimating the relative importance of climate change mechanisms, Geophys. Res. Lett., 30, https://doi.org/10.1029/2003GL018141, 2047, 2003.

Smith, C. J., Forster, P. M., Allen, M., Leach, N., Millar, R. J., Passerello, G. A., and Regayre, L. A.: FAIR v1.3: a simple emissions-based impulse response and carbon cycle model, Geosci. Model Dev., 11, 2273–2297, https://doi.org/10.5194/gmd-11-2273-2018, https://www.geosci-model-dev.net/11/2273/2018/, 2018a.

Smith, C. J., Kramer, R. J., Myhre, G., Forster, P. M., Soden, B. J., Andrews, T., Boucher, O., Faluvegi, G., Fläschner, D., Hodnebrog, O., Kasoar, M., Kharin, V., Kirkevåg, A., Lamarque, J.-F., Mülmenstädt, J., Olivié, D., Richardson, T., Samset, B. H., Shindell, D., Stier, P., Takemura, T., Voulgarakis, A., and Watson-Parris, D.: Understanding Rapid Adjustments to Diverse Forcing Agents, Geophys. Res. Lett., 45, 12,023–12,031, https://doi.org/10.1029/2018GL079826, 2018b.

Smith, C. J., Kramer, R. J., and Sima, A.: The HadGEM3-GA7.1 radiative kernel: the importance of a well-resolved stratosphere, Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2019-254, 2020.

Soden, B., Held, I., Colman, R., Shell, K., Kiehl, J., and Shields, C.: Quantifying Climate Feedbacks Using Radiative Kernels, J. Climate, 21, 3504–3520, https://doi.org/10.1175/2007JCLI2110.1, 2008.

Soden, B. J., Collins, W. D., and Feldman, D. R.: Reducing uncertainties in climate models, Science, 361, 326–327, https://doi.org/10.1126/science.aau1864, 2018.

Stevens, B., Fiedler, S., Kinne, S., Peters, K., Rast, S., Müsse, J., Smith, S. J., and Mauritsen, T.: MACv2-SP: a parameterization of anthropogenic aerosol optical properties and an associated Twomey effect for use in CMIP6, Geosci. Model Dev., 10, 433–452, https://doi.org/10.5194/gmd-10-433-2017, https://www.geosci-model-dev.net/10/433/2017/, 2017.

Stjern, C. W., Samset, B. H., Myhre, G., Forster, P. M., Hodnebrog, Ø., Andrews, T., Boucher, O., Faluvegi, G., Iversen, T., Kasoar, M., Kharin, V., Kirkevåg, A., Lamarque, J.-F., Olivié, D., Richardson, T., Shawki, D., Shindell, D., Smith, C. J., Takemura, T., and Voulgarakis, A.: Rapid Adjustments Cause Weak Surface Temperature Response to Increased Black Carbon Concentrations, J. Geophys. Res.-Atmos., 122, 11,462–11,481, https://doi.org/10.1002/2017JD027326, 2017.

Stouffer, R. J., Eyring, V., Meehl, G. A., Bony, S., Senior, C., Stevens, B., and Taylor, K. E.: CMIP5 Scientific Gaps and Recommendations for CMIP6, Bulletin of the American Meteorological Society, 98, 95–105, https://doi.org/10.1175/BAMS-D-15-00013.1, https://doi.org/10.1175/BAMS-D-15-00013.1, 2017.

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geosci. Model Dev., 12, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019, https://www.geosci-model-dev.net/12/4823/2019/, 2019.

Tang, T., Shindell, D., Faluvegi, G., Myhre, G., Olivié, D., Voulgarakis, A., Kasoar, M., Andrews, T., Boucher, O., Forster, P., Hodnebrog, O., Iversen, T., Kirkevåg, A., Lamarque, J.-F., Richardson, T., Samset, B., Stjern, C., Takemura, T., and Smith, C.: Comparison of Effective Radiative Forcing Calculations Using Multiple Methods, Drivers, and Models, J. Geophys. Res.-Atmos., 124, 4382–4394, https://doi.org/10.1029/2018JD030188, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD030188, 2019.

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M., Saito, F., Chikira, M., Watanabe, S., Mori, M., Hirota, N., Kawatani, Y., Mochizuki, T., Yoshimura, K., Takata, K., O'ishi, R., Yamazaki, D., Suzuki, T., Kurogi, M., Kataoka, T., Watanabe, M., and Kimoto, M.: Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6, Geosci. Model Dev., 12, 2727–2765, https://doi.org/10.5194/gmd-12-2727-2019, https://www.geosci-model-dev.net/12/2727/2019/, 2019.

Taylor, K., Crucifix, M., Braconnot, P., Hewitt, C., Doutriaux, C., Broccoli, A., Mitchell, J., and Webb, M.: Estimating shortwave radiative forcing and response in climate models, J. Climate, 20, 2530–2543, 2007.

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, Atmospheric Chemistry and Physics Discussions, 2020, 1–29, https://doi.org/10.5194/acp-2019-1205, 2020.

Tokarska, K. B., Gillett, N. P., Arora, V. K., Lee, W. G., and Zickfeld, K.: The influence of non-CO_2 forcings on cumulative carbon emissions budgets, Env. Res. Lett., 13, 034 039, https://doi.org/10.1088/1748-9326/aaafdd, 2018.

Twomey, S.: The influence of pollution on the shortwave albedo of clouds, J. Atmos. Sci., 34, 1149–1152, 1977.

Vial, J., Dufresne, J.-L., and Bony, S.: On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates, Clim. Dynam., 41, 3339–3362, https://doi.org/10.1007/s00382-013-1725-9, 2013.

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., Colin, J., Guérémy, J.-F., Michou, M., Moine, M.-P., Nabat, P., Roehrig, R., Salas y Mélia, D., Séférian, R., Valcke, S., Beau, I., Belamari, S., Berthet, S., Cassou, C., Cattiaux, J., Deshayes, J., Douville, H., Ethé, C., Franchistéguy, L., Geoffroy, O., Lévy, C., Madec, G., Meurdesoif, Y., Msadek, R., Ribes, A., Sanchez-Gomez, E.,

Terray, L., and Waldman, R.: Evaluation of CMIP6 DECK Experiments With CNRM-CM6-1, J. Adv. Model. Earth Sy., 11, 2177–2213, https://doi.org/10.1029/2019MS001683, 2019.

Webb, M., Senior, C., Bony, S., and Morcrette, J.-J.: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models, Clim. Dynam., 17, 905–922, https://doi.org/10.1007/s003820100157, 2001.

Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., Davis, P., Graham, T., Hewitt, H. T., Hill, R., Hyder, P., Ineson, S., Johns, T. C., Keen, A. B., Lee, R. W., Megann, A., Milton, S. F., Rae, J. G. L., Roberts, M. J., Scaife, A. A., Schiemann, R., Storkey, D., Thorpe, L., Watterson, I. G., Walters, D. N., West, A., Wood, R. A., Woollings, T., and Xavier, P. K.: The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) Configurations, J. Adv. Model. Earth Sy., 10, 357–380, https://doi.org/10.1002/2017MS001115, 2018.

Wyser, K., van Noije, T., Yang, S., von Hardenberg, J., O'Donnell, D., and Döscher, R.: On the increased climate sensitivity in the EC-Earth model from CMIP5 to CMIP6, Geoscientific Model Development Discussions, 2019, 1–13, https://doi.org/10.5194/gmd-2019-282, https://www.geosci-model-dev-discuss.net/gmd-2019-282/, 2019.

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yabu, S., Yoshimura, H., Shindo, E., Mizuta, R., Obata, A., Adachi, Y., and Ishii, M.: The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component, J. Meteorol. Soc. Jpn. Ser. II, 97, 931–965, https://doi.org/10.2151/jmsj.2019-051, 2019.

Zelinka, M., Klein, S., and Hartmann, D.: Computing and Partitioning Cloud Feedbacks Using Cloud Property Histograms. Part I: Cloud Radiative Kernels, J. Climate, 25, 3715–3735, https://doi.org/10.1175/JCLI-D-11-00248.1, 2012.

Zelinka, M. D., Andrews, T., Forster, P. M., and Taylor, K. E.: Quantifying components of aerosol-cloud-radiation interactions in climate models, J. Geophys. Res.-Atmos., 119, 7599–7615, https://doi.org/10.1002/2014JD021710, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JD021710, 2014.

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., and Taylor, K. E.: Causes of Higher Climate Sensitivity in CMIP6 Models, Geophysical Research Letters, 47, e2019GL085 782, https://doi.org/10.1029/2019GL085782, 2020.

Table 6. Contribution of the components of ~~1850–2014~~ effective radiative forcing from ~~aerosols~~the present-day aerosol time-slice RFMIP experiment.~~ARI: ERF due to aerosol-radiation interactions; ACI: ERF due to aerosol-cloud interactions; scat=scattering; abs=absorption; amt=cloud amount.~~

	SW								
	ARI			ACI					
Model	scat	abs	sum	scat	abs	amt	sum	ARI+ACI	A
ACCESS-CM2	−0.79	0.30	−0.48	−0.80	−0.01	−0.12	−0.93	−1.42	0.
CanESM5	−0.60	0.52	−0.08	−0.88	0.06	−0.14	−0.96	−1.04	0.
CESM2	−0.26	0.35	0.09	−1.70	0.03	−0.01	−1.68	−1.58	0.
CNRM-CM6-1	−0.61	0.20	−0.41	−0.77	−0.05	0.05	−0.77	~~−1.18~~ −1.17	0.
CNRM-ESM2-1	−0.42	0.18	−0.24	−0.52	−0.04	−0.03	−0.59	−0.83	0.
EC-Earth3	−0.75	0.32	−0.43	−0.34	−0.04	−0.08	−0.46	−0.89	0.
GFDL-CM4	−0.65	0.41	−0.24	−0.54	0.00	−0.09	−0.62	−0.86	0.
GFDL-ESM4	−0.65	0.53	−0.12	−0.59	0.01	−0.06	−0.63	−0.75	0.
GISS-E2-1-G p1	−0.91	0.19	−0.72	0.06	0.01	−0.94	−0.87	~~−1.59~~ −1.58	0.
GISS-E2-1-G p3	−0.97	0.18	−0.79	−0.25	−0.01	−0.06	−0.32	−1.11	0.
HadGEM3-GC31-LL	−0.77	0.30	−0.47	−0.75	−0.01	−0.07	−0.83	−1.30	0.
IPSL-CM6A-LR	−0.60	0.17	−0.43	−0.26	−0.01	0.06	~~−0.22~~ −0.21	−0.65	0.
MIROC6	−0.48	0.09	−0.39	−1.04	−0.06	−0.01	−1.12	−1.51	0.
MRI-ESM2-0	−0.70	0.19	−0.51	−1.71	−0.09	−0.38	−2.18	~~−2.69~~ −2.68	0.
NorESM2-LM	−0.42	0.21	−0.21	−1.07	−0.00	−0.10	−1.17	−1.38	0.
NorESM2-MM	−0.40	0.22	−0.18	−1.10	0.02	−0.17	−1.25	−1.43	0.
UKESM1-0-LL	~~−0.72~~ −0.71	0.34	~~−0.39~~ −0.37	−0.82	~~−0.00~~ −0.01	−0.10	~~−0.92~~ −0.93	~~−1.31~~ −1.30	~~0.19~~ 0.
Mean	~~−0.59~~ −0.63	~~0.26~~ 0.28	~~−0.33~~ −0.35	~~−0.83~~ −0.77	−0.01	~~−0.15~~ −0.13	~~−0.99~~ −0.91	~~−1.33~~ −1.26	~~0.12~~ 0.
St. dev.	~~0.17~~ 0.18	0.12	~~0.20~~ 0.22	~~0.49~~ 0.46	0.04	~~0.26~~ 0.22	~~0.49~~ 0.47	~~0.50~~ 0.46	0.

Table 7. As for table 3 but for 1850–2014 land-use forcing.

#	Model	ERF	IRF	Adj.	ts	ta_tr	ta_st	hus	
2	CanESM5	−0.08	−0.10	0.03	0.02	−0.02	0.01	−0.00	
3	CESM2	−0.04	~~−0.08~~ −0.09	0.05	−0.01	0.03	0.00	0.01	~~0.~~
5	CNRM-ESM2-1	−0.07	~~−0.08~~ −0.09	~~0.02~~ 0.03	0.01	−0.02	−0.01	0.04	
6	EC-Earth3	−0.13	−0.13	−0.00	0.02	0.07	0.02	−0.03	
7	GFDL-CM4	−0.33	~~−0.42~~ −0.41	~~0.09~~ 0.08	−0.04	0.09	0.00	−0.06	
8	GFDL-ESM4	−0.28	−0.27	−0.01	−0.03	0.08	0.01	−0.11	
9	GISS-E2-1-G p1	−0.00	0.02	−0.02	−0.02	−0.02	0.01	0.02	
11	HadGEM3-GC31-LL	−0.11	~~−0.16~~ −0.18	~~0.06~~ 0.07	0.01	0.10	0.01	−0.05	~~−0.0~~
12	IPSL-CM6A-LR	−0.05	~~−0.11~~ −0.09	~~0.07~~ 0.05	−0.01	0.02	0.00	−0.01	~~0.~~
13	MIROC6	−0.03	~~−0.10~~ −0.07	~~0.08~~ 0.04	−0.01	0.04	0.00	−0.04	~~0.~~
14	MPI-ESM1-2-LR	−0.10	~~−0.01~~ −0.06	~~−0.09~~ −0.04	−0.01	0.01	0.01	−0.01	~~−0.1~~
15	MRI-ESM2-0	−0.17	~~−0.33~~ −0.32	~~0.16~~ 0.15	0.00	0.08	−0.00	−0.08	~~0.~~
16	NorESM2-LM	0.26	~~−0.01~~ −0.00	0.27	0.01	0.01	0.00	0.00	~~0.~~
18	UKESM1-0-LL	~~−0.30~~ −0.18	~~−0.28~~ −0.18	~~−0.01~~ 0.00	~~0.02~~ 0.00	~~0.08~~ 0.04	~~−0.01~~ 0.01	~~−0.04~~ −0.05	~~−0.0~~
	Mean	~~−0.08~~ −0.09	−0.14	~~0.06~~ 0.05	−0.00	~~0.03~~ 0.04	~~0.00~~ 0.01	~~−0.02~~ −0.03	
	St. dev.	~~0.14~~ 0.13	~~0.09~~ 0.12	0.08	0.02	0.04	0.01	~~0.03~~ 0.04	~~0.~~

Table 8. As for table 3 but for 1850–2014 anthropogenic forcing.

#	Model	ERF	IRF	Adj.	ts	ta_tr	ta_st	hus	albedo
1	ACCESS-CM2	1.90			−0.08	−0.37	0.97	0.23	−0.00
2	CanESM5	2.37	~~1.82~~ 1.85	~~0.55~~ 0.52	−0.14	−0.70	0.87	0.45	0.03
3	CESM2	2.05	~~0.68~~ 0.74	~~1.38~~ 1.31	−0.17	−0.48	0.95	0.44	0.01
4	CNRM-CM6-1	1.61	0.55	1.06	−0.07	−0.22	1.01	0.07	0.01
5	CNRM-ESM2-1	1.66	~~0.66~~ 0.67	~~1.01~~ 1.00	−0.04	−0.32	1.11	0.16	−0.03
6	EC-Earth3	2.09			−0.09	−0.36	1.18	0.18	−0.13
7	GFDL-CM4	2.34	~~2.13~~ 2.16	~~0.21~~ 0.18	−0.11	−0.44	0.69	0.31	−0.36
8	GFDL-ESM4	2.17	1.49	0.68	−0.12	−0.45	1.14	0.32	−0.27
9	GISS-E2-1-G p1	1.93	~~1.79~~ 1.82	~~0.14~~ 0.11	−0.08	−0.14	0.99	0.13	0.13
11	HadGEM3-GC31-LL	1.81	~~0.98~~ 1.00	~~0.83~~ 0.81	−0.09	−0.26	0.97	0.13	−0.20
12	IPSL-CM6A-LR	2.32	~~1.36~~ 1.41	~~0.96~~ 0.91	−0.11	−0.40	0.94	0.28	−0.08
13	MIROC6	1.80	~~1.05~~ 1.11	~~0.74~~ 0.69	−0.10	−0.29	0.87	0.14	−0.02
14	MPI-ESM1-2-LR	2.13			−0.10	−0.48	0.79	0.25	0.02
15	MRI-ESM2-0	1.95	~~1.72~~ 1.77	~~0.23~~ 0.18	−0.10	−0.59	0.93	0.44	−0.28
16	NorESM2-LM	2.06	~~1.09~~ 1.18	~~0.97~~ 0.88	−0.14	−0.49	0.99	0.31	−0.05
18	UKESM1-0-LL	~~1.71~~ 1.79	~~0.55~~ 0.63	1.16	~~−0.05~~ −0.10	~~−0.25~~ −0.28	~~1.22~~ 1.24	~~0.12~~ 0.11	~~−0.28~~ −0.20
	Mean	~~1.97~~ 2.00	~~1.20~~ 1.26	~~0.77~~ 0.73	−0.10	~~−0.38~~ −0.39	~~0.96~~ 0.98	0.25	−0.09
	St. dev.	~~0.26~~ 0.23	~~0.53~~ 0.51	~~0.39~~ 0.37	~~0.04~~ 0.03	~~0.16~~ 0.14	0.14	0.12	0.13