Response to Reviewer #1

This paper has two objectives: 1) to present a new set of radiation kernels with high top, and 2) to intercompare this and a few other sets of radiation kernels, specifically with respect to the estimate of the radiative impact of stratospheric temperature adjustment in response to CO2. These are both potentially important contributions to make and warrant the efforts here. While I find the first objective well done and generally welcome a new kernel set to enrich the feedback analysis toolbox, I find the second objective relatively poorly executed. I'd suggest the authors take into consideration of the following comments and questions in revising this paper.

Yi, thank you for the considered review of our paper. The first point about the addition of a new kernel is well taken and we are pleased that this is a useful addition to the existing set of model kernels. We note the deficiencies highlighted in your second point and trust we have addressed these satisfactorily in the response to your comments and in the forthcoming revised paper.

There lacks a solid basis for the "recommendation" of the three kernel datasets concluded by the paper. For making such an important and strong statement as to which kernels are better, a principle (criterion) needs to be explicitly stated and justified for comparing them – this is currently missing in the paper. Note that a high-top kernel does not guarantee a higher accuracy in its assessment of radiative impact because the atmosphere which the kernel is based on can be biased – one should be especially cautious if the atmosphere is from a GCM – or because the radiation code used for the kernel computation is biased against the radiation code used in the target GCM simulation. On the other hand, a lower-top kernel also does not necessary lead to a poorer assessment, as shown by the GFDL kernel included here, due to fortuitous compensation of errors or due to some technical details of radiative transfer. Fore instance, some kernels may have used high-top atmospheric profiles in their computation but then truncated to lower top when applied to computing feedback. Moreover, computing and applying kernels at lower vertical resolution may be less subject to the nonlinear coupling between different vertical layers – one can test the non-radiation closure due to this issue, for example, by comparing the sum of vertical kernels to the true radiation change computed using the same radiation model from a vertically uniform 1-K temperature change. To make a more objective and informative assessment, I suggest adding: 1) the comparisons of a) the global mean radiation change (Ax) due to layers above 1hPa, 10 hPa, tropopause and surface (whole column), respectively, assuming a uniform 1 K change of atmospheric temperature - this would disclose how the different kernels differ with respect to the radiative sensitivity to different portions of the atmosphere and whether there may be compensation of errors from different vertical portions; and b), like a), but using the atmospheric temperature adjustment to CO2 forcing as simulated by one representative GCM or the multiple model mean.

As described further on in the response the "recommendation" is weakened to a suggestion. We have included the comparisons that you suggest as ~~additional figures~~ an additional figure 6.

2) additional kernels, especially those observation-based kernels, such as the kernels of Huang et al. (2017) based on ERA-interim and of Yue et al. (2016) based on satellite. The former one (available from https://huanggroup.wordpress.com/research/) was computed with a high-top atmospheric profile using RRTMG and provides kernel values up to 1hPa, which would provide a good comparison to ECMWF kernel here based on another radiation model (Oslo) – e.g., for assessing radiation code dependency noted above.

Thank you for the suggestion here. The Huang et al. (2017) kernels will be added to the analysis which were missed in the original submission. This kernel has a very strong negative temperature response at the 1 hPa level, in negative excess of -2 W m-2 (100hPa)-1 K-1 (shown in the update to figure 3) which provide large estimates of the stratospheric adjustment (update to figure 5).

The Yue et al. (2016) observational kernels focus on clouds and are more appropriately compared with the ISSCP simulator kernel (Zelinka et al., 2012) rather than our kernels derived from atmospheric state variables. While we use the ISCCP kernel to derive cloud radiative adjustments in the IPSL model from ISSCP simulator diagnostics, a comparison of cloud kernels produced by other groups is beyond the scope of this paper. But we thank you for making us aware of this paper and include references to it as further evidence of the utility of kernel approaches.

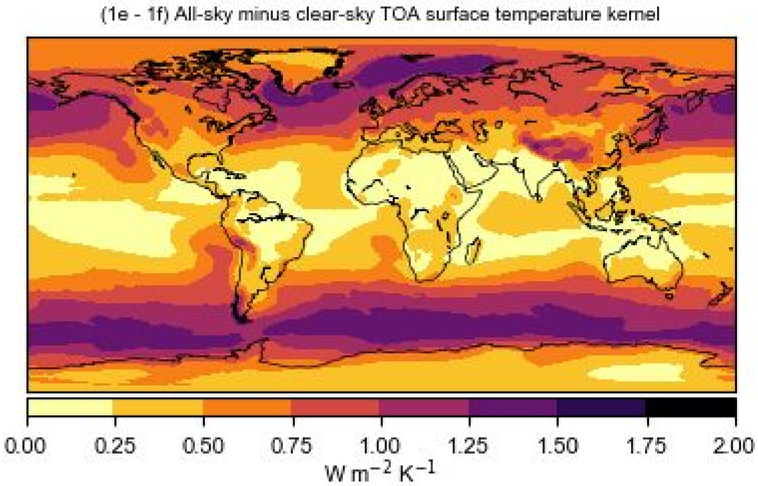
Additional comments: Line 17. It is recommended to include Zhang & Huang (2014) here, as this is one of the earliest that quantified CO2 forcing, including both instantaneous forcing and the adjustment components, using kernels. The quantification of adjustment in multiple models reported by this work would make good comparisons to the results reported here, e.g., Table 2, 3.

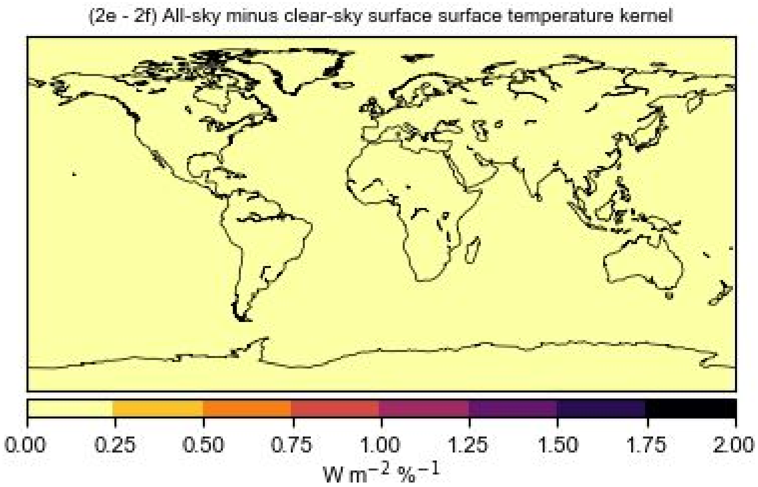
The reference to Zhang and Huang (2014) near line 17 for using kernels to diagnose adjustments has been included, so thank you for reminding us of this study.

It is an excellent suggestion to compare our results to Zhang & Huang (2014). It in fact gives more weight to the claim that stratospheric changes are important. A new table 4 will be included comparing the results and showing that the IPSL-CM6A-LR model with our kernel is outside the 2-sigma range of the 11 models' stratospheric temperature adjustment in Zhang & Huang (2014). IPSL-CM6A-LR is fairly typical of CMIP6 models in terms of ERF and stratospheric adjustment as shown in Smith et al. (2020), and this paper also discusses the fact that ERF is increased in CMIP6 compared to CMIP5 for 4xCO2. Therefore we speculate that an increase in stratospheric temperature adjustment could be partially responsible for an increase in CMIP6 ERF, although we can't prove it without IRF calculations from more models and the fact that most CMIP5 model output does not include the 5 hPa and 1 hPa levels, again preventing a formal comparison.

Line 108. It is not obvious to me that cloud masking has a lesser impact on the surface flux. Please clarify and be more quantitative here.

In figures 1 and 2 the subfigures were actually mislabelled. This statement was meant to refer to (what is erroneously labelled) the difference between 2(c)-2(g) and 1(c)-1(g), which are the surface temperature kernels for all-sky and clear-sky for the surface and TOA. In the attachment to this review the differences are shown. The differences in the TOA kernels are large but for the surface kernels are small. The figure captions and surrounding text will be updated. Thank you for spotting this.





[Figure A: differences between surface temperature kernels for TOA and surface fluxes]

Line 148-150. Can you illustrate the biases of these low-top kernels mentioned here?

To keep the structure of the paper as it is we refer to the following section where we show the decomposition in the IPSL-CM6A-LR model using our kernel but explain the bias here. Using either equation 7 or 8 we get a small positive residual (table 3). Figure 5 implies that the low-top kernels (excluding GFDL) are around 0.5 W m^{-2} or more lower in their stratospheric adjustment than HadGEM3-GA7.1, so that the residuals would be more positive assuming the tropospheric adjustments are similar across kernels. Although we don't compare non-stratospheric adjustments in this paper, I showed previously that kernels agree well for tropospheric adjustments (Smith et al., 2018), as they do for climate feedbacks (e.g. Soden et al., 2008).

The following addition to the manuscript is made around line 148:

> **We show in section 5 that adjustments calculated using the HadGEM3-GA7.1 kernel in the IPSL-CM6A-LR model for a quadrupled CO2 experiment provide small residuals (i.e. the adjustments are appropriately captured), suggesting that assuming there are no compensating errors, low-top kernels would underestimate the stratospheric temperature response and produce larger residuals.**

Line 201. Again, the first that applied this residual method was Zhang&Huang [2014].

Thank you for this suggestion. This reference has been added here.

Line 10, 206. Can't approve such a "recommendation" for the reasons above. And such a recommendation could lead to wrongly denial of the use of the other kernels - both the lower top ones like the GFDL one that can achieve similar quantitative results and those the authors failed to include for comparison here.

We are inclined to agree that for a data description paper recommendation may be a bit strong and have changed the sentence near the end:

> **We suggest that radiative kernels with a higher stratospheric resolution and model top are better able to fully capture stratospheric adjustments to CO2 forcing in general, and generate smaller residuals. This effect has become more prominent with the additional 5 hPa and 1 hPa model levels archived as standard in processed CMIP6 model output compared to CMIP5.**

and in the abstract:

> **We show in the IPSL-CM6A-LR model where a full set of climate diagnostics are available that the HadGEM3-GA7.1 kernel exhibits linear behaviour and the residual error term is small, and from a survey of kernels available in the literature that in general low-top radiative kernels underestimate the stratospheric temperature response.**

References

Huang, Y., Y. Xia and X. Tan, (2017), On the pattern of CO_2 radiative forcing and poleward energy transport, J. Geophys. Res.- Atmosphere, 122, 10,578–10,593, doi: 10.1002/2017JD027221.

Kernel data: https://huanggroup.wordpress.com/research/

Yue, Q., Kahn, B. H., Fetzer, E. J., Schreier, M., Wong, S., Chen, X., & Huang, X. (2016). ObservationâA˘ Rbased longwave cloud radiative ker- ˘ nels derived from the AâA˘ RTrain. Journal of Climate, 29(6), 2023–2040. ˘ https://doi.org/10.1175/JCLIâA˘ RDâ ˘ A˘ R15â ˘ A˘ R0257.1 ˘

Zhang, M. and Y. Huang (2014), Radiative forcing of quadrupling CO_2, J. of Climate, 27, 2496–2508. doi: http://dx.doi.org/10.1175/JCLI-D-13-00535.1

Response to Reviewer #2

This manuscript documents a new set of “radiative kernels,” calculations of the sensitivity of radiative flux to atmospheric state, developed from a current-generation climate model with a domain reaching well into the stratosphere. The construction of the kernels was motivated by a desire to understand the fast response of stratospheric temperature to changes in carbon dioxide concentration and the authors demonstrate the value added by the new kernels. The construction of the kernels is described and their accuracy and generality assessed.

The data are well worth publishing. They require substantial computational resources to produce, extend the vertical domain in an almost-unique way, and use an accurate radiative transfer code. The free availability of the data has been verified. The manuscript is effective at documenting how the kernels are produced, providing enough details for readers to understand and potentially replicate the steps. It is also effective at motivating why this implementation is useful, noting that the diagnosis of the fast climate response (the adjustment) to increased concentrations of carbon dioxide depends importantly on having a deep vertical domain. Beyond a few small points of expression noted below the manuscript could be most improved by more context for the uninitiated and a more general treatment of some ideas.

Thank you for your positive overall comments and we are pleased that you agree that the data and description paper is worth publishing in a form close to present.

The introduction, which introduces the concept of and motivation for a radiative kernel, may be more general than is needed for the present manuscript. The generality makes it open to objections as to how the ideas are expressed. The general idea of a kernel is the ability to compute flux perturbations from state perturbations. As originally implemented by Soden, Shell, and others, these were restricted to specific characteristics of atmospheric state (air and surface temperature, water vapor, surface albedo, and excluding clouds) based partly on prioritization and partly on based on the availability of data. The does not establish a “standard” (line 47) nor does it exclude in principle other variables from being relevant (line 28). Readers may also wonder how the general material on the use of kernels (lines 39-54) is directly relevant to the construction of the present kernels.

Thank you for these suggestions. In our opinion the introduction is not overly long at present so we are inclined to keep lines 39-54 in the paper as general background. We note that others have taken a different approach and have assumed more background knowledge (e.g. Prendergrass et al 2018). Readers familiar with radiative kernels could easily skip over the introduction. Those not familiar may welcome it, and as the kernel method is used increasingly outside the climate feedback community in which it was developed, it may make the paper more self-contained.

Line 47 has been changed to “Cloud adjustments and feedbacks cannot be determined directly using atmospheric state kernels”. The previous wording could be taken to imply that cloud adjustments/feedbacks could not be calculated at all using only atmospheric kernels, which was not the intention.

Line 28: after this sentence, included “Although other (non-cloud) variables may also be relevant, the majority of adjustments are expected to be captured under this framework (Vial et al., 2013).”

The authors might revisit the introduction and focus it more tightly on the subject of the manuscript. This might include a not too profound explanation of how kernels can be used to diagnose both feedbacks and adjustments, and and explanation as to why yet another set of kernels might be desirable (i.e. the material that begins section 4). Care should be taken not to confuse routine practice with standardization.

To improve the motivation, the first paragraph from section 4 has been assimilated into the introduction. Following the previous response, we believe the present level of background is appropriate.

Section 2:

It would be worth noting explicitly that these kernels rely on two almost distinct aspect of HadGEM: the radiation code SOCRATES run at low spectral resolution, and the climatology of atmospheric state including clouds, even if experience shows relatively weak dependence on the background state.

This section has been updated to include a sentence:

> **The kernel is therefore dependent on two aspects of the HadGEM3-GA7.1 model: the pre-industrial background climatology (including clouds), and the broadband version of the radiation code.**

The authors might also explain some of their choices and any expected impacts. These might include the choice to develop kernels for pre-industrial conditions, the relatively highly-resolved vertical structure and coarse horizontal resolution of the simulations, and the high time resolution.

In the updated section 2 we have added a few sentences at opportune points, justifying some of the methodological choices. More explanation for each is given below.

The pre-industrial conditions were chosen for the kernel base state as these kernels were designed to be used first and foremost with the RFMIP piClim-X forcing experiments (X = 4xCO2, present-day GHG, present-day aerosols, present-day land-use and present-day total anthropogenic), covering both negative and positive forcing relative to the pre-industrial. The base climate of the target models is pre-industrial except for the perturbed component(s) and differences are taken with respect to piClim-control (an atmosphere-only pre-industrial control run). Therefore, a pre-industrial climatology for the kernel is appropriate. Other choices were of course possible such as linearising around present-day conditions.

The relatively high (for a GCM) vertical resolution is the default configuration of the HadGEM3-GA7.1 model, and its importance for stratospheric adjustment is already stated.

The horizontal resolution (1.875° x 1.25°) is the lowest resolution used for HadGEM3 and UKESM1 in CMIP6. Many other CMIP6 models that are ultimately the target of the radiative kernels (excluding HighResMIP) are on similar resolutions to this. We are not aware of an exhaustive list, but Table 1 in Smith et al. (2020) lists those used in RFMIP, so it is not apparent that an increase in resolution in the base climate would improve accuracy after the kernels are re-gridded to the resolution of target climate models. But the main reason why a finer resolution wasn't considered is computing time. The "MM" resolution of HadGEM3 has an atmospheric grid of 0.83° x 0.56° with 5 times as many grid points, so running the offline radiation model would have taken 5 times longer. To do this in a reasonable amount of time would require a reduction in the vertical or time resolution (or use of HPC). The high vertical resolution of these kernels are one of their strengths, and it would be difficult to reduce the number of temporal radiation calls as explained below.

A two-hour time step was used as instantaneous rather than time-mean climate output is needed to run the offline radiative transfer code (a discussion of why time-mean output does not work well is in Bellouin et al, 2020; section 4). Two hours was considered fine enough to avoid biases by undersampling the diurnal cycle of

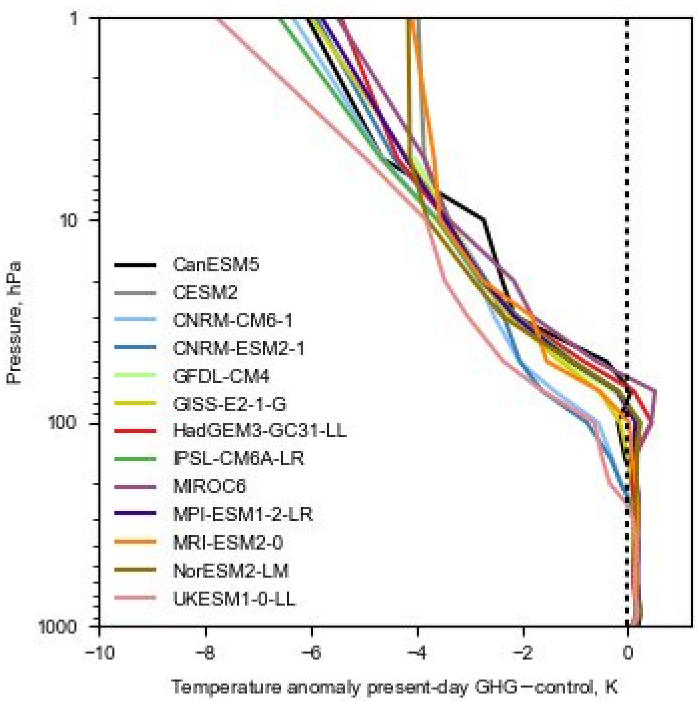
temperature, humidity and clouds, and shortwave solar geometry. On the last point, using longer timesteps like 6 hours between shortwave calls leads to some longitudes receiving substantially more incoming solar radiation than others over the course of the year. This effect could be achieved alternatively by using a timestep that does not divide 24 (e.g. 22, 23, 25 or 26 hours) and running several years of the climate model to sample diurnal, seasonal and interannual variability, such as is sometimes done in PRP calculations.

Section 4 illustrates the added value of the new kernels quite nicely. The use case is important but a little narrow. Is the value also added for other greenhouse gas forcings?

Thank you for your positive comments here and agree it is beneficial to demonstrate why the kernels are useful. In response to an earlier comment we moved the first paragraph to the introduction.

It is indeed useful for other GHG forcings. We generalise from CO2 to GHGs in a few places in Section 4, although the focus of the results will still be on CO2 as we have the double-call results from IPSL-CM6A-LR for this experiment.

We didn't re-run every kernel calculation using the piClim-ghg experiment in this paper. However, in supplementary figure 1 in Smith et al. (2020) we compare four of the kernels for all RFMIP experiments, showing that HadGEM3, ECMWF-Oslo and GFDL kernels show larger stratospheric adjustments for piClim-ghg than the CCSM kernel which is in accordance with the 4xCO2 results in this paper. In attachment to this response we show the temperature profile for the GHG experiment, which moves in the same direction as for 4xCO2 but with a lower magnitude.



[Figure B: stratospheric temperature differences from piClim-ghg minus piClim-control for 13 models in RFMIP].

Also note that in figure 4 we have removed the missing top level from GFDL-CM4, which was previously displayed as zero rather than missing.

Section 5 is the least organized and clear and the section seems to assume a lot of background knowledge. The point of the section is to demonstrate the accuracy and applicability of the kernels. The narrative should be constructed to as to make this goal clear, explain how accuracy and applicability can be assessed, and finally to demonstrate the results.

Thank you for this suggestion. It has been re-written to introduce the aim of the exercise (the discussion on IRF appears too soon and is given too much weight) before showing the results. Following suggestions from reviewer #1 we have also included a comparison of the results here to a previous study from CMIP5. To address the assumption of background knowledge on interpretation of section 5, we note that the introduction section may come in useful here to the uninitiated reader, which is a reason why we feel it is appropriate to keep most of it.

Lines 2-3: “the utility of radiative kernels. . . is most appropriate” The last word isn’t quite right. Utility can be greater or less but not appropriate.

Thanks for pointing out the confusing wording. Changed “most appropriate” to “greatest”.

Line 23: Kernels represent derivatives of flux with respect to state, not differential equations

A slight terminological liberty taken on our part. We’ll keep the notation of eq. (1) as it is used by others (e.g. Shell et al., (2008) where they have used (F-Q) in place of R, and Huang et al. (2017)) and provides a nice concise representation but explain that it isn’t strictly a differential equation.

Line 32: climate model (not mode)

Typo corrected - thank you

Line 84: the equation should have units

Updated to confirm that 10000 and p_thick are both in units of Pa here.

Line 113: The sudden appearance of PDRMIP may confuse the uninitiated

PDRMIP (Precipitation Driver and Response Model Intercomparison Project) acronym is now introduced, which was an oversight in the first submission.

Line 125: the limitations of low-topped kernels are presumably independent of whether the state comes from a “climate model” or any other source

Agreed: revised this sentence to be more general: “For kernels built from underlying atmospheric profiles where the top of the profile is not sufficiently high or with too coarse a resolution in the stratosphere, this additional upper stratospheric cooling is missed.”

Line 161: cars break down - what is meant here is “decomposition” or similar

As part of the re-write of section 5, this sentence will be revised.

Line 163: “ways of calculating the residual can be obtained” is a confusing phrasing

Agree this is not meaningful: changed to “Two different ways of calculating the residual exist.”

The HadGEM3-GA7.1 radiative kernel: the importance of a well-resolved stratosphere

Christopher J. Smith[1,2], Ryan J. Kramer[3,4], and Adriana Sima[5]

[1]School of Earth and Environment, University of Leeds, Leeds LS2 9JT, United Kingdom
[2]International Institute for Applied Systems Analysis (IIASA), A-2361 Laxenburg, 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]LMD/IPSL, Sorbonne Université, ENS, PSL Université, École polytechnique, Institut Polytechnique de Paris, CNRS, Paris France

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

Abstract. We present top-of-atmosphere and surface radiative kernels based on the atmospheric component (GA7.1) of the HadGEM3 general circulation model developed by the UK Met Office. We show that the utility of radiative kernels for forcing adjustments in idealised CO_2 perturbation experiments is ~~most appropriate~~ greatest where there is sufficiently high resolution in the stratosphere in both the target climate model and the radiative kernel. This is because stratospheric cooling to a CO_2 perturbation continues to increase with height, and low-resolution or low-top kernels or climate model output are unable to fully resolve the full stratospheric temperature adjustment. In the sixth phase of the Coupled Model Intercomparison Project (CMIP6), standard atmospheric model data is available up to 1 hPa on 19 pressure levels, which is a substantial advantage compared to CMIP5. We show in the IPSL-CM6A-LR model where a full set of climate diagnostics are available that the HadGEM3-GA7.1 kernel exhibits linear behaviour and the residual error term is small~~. From~~ , and from a survey of kernels available in the literature ~~we recommend three kernels for adjustment calculations to CO_2 and well-mixed greenhouse gas perturbations based on their stratospheric resolution:~~ that in general low-top radiative kernels underestimate the stratospheric temperature response. The HadGEM3-GA7.1 ~~, ECMWF-Oslo, and ECHAM6.~~radiative kernels are available at https://doi.org/10.5281/zenodo.3594673 (Smith, 2019) .

1 Introduction

Radiative kernels describe how a small change in an atmospheric state variable affects the Earth's energy balance (Soden et al., 2008; Shell et al., 2008). They allow an analysis of climate feedbacks (Shell et al., 2008; Soden et al., 2008; Sanderson and Shell, 2012; Jonko et al., 2012; Block and Mauritsen, 2013; Huang, 2013) or forcing adjustments ~~(Vial et al., 2013; Chung and Soden, 2015~~ standardised climate model diagnostics such as those from the Coupled Model Intercomparison Projects (CMIPs). The use of radiative kernels is efficient, removing the need for time- and memory-consuming calculations of climate feedbacks online through partial radiative perturbation calculations (Wetherald and Manabe, 1988) or offline using a standalone version of the model radiative transfer code (Colman and McAvaney, 2011).

A radiative kernel K_X is in effect a four-dimensional (time, height, latitude, longitude) array ~~of partial differential equations~~ describing how radiation fluxes R change with an atmospheric state variable X

$$K_X(t,z,y,x) = \left.\frac{\partial R}{\partial X}\right|_{(t,z,y,x)}. \tag{1}$$

Although strictly not a partial differential equation, this statement provides a concise written form and is used by others (e.g. Shell et al. (2008); Huang et al. (2017)). R may be upwelling, downwelling or net, shortwave or longwave, radiation changes at any atmospheric level. Most commonly net top-of-atmosphere (TOA), surface and tropopause-level fluxes are of greatest interest. X here represents atmospheric temperature (T_a), surface (skin) temperature (T_s), water vapour (q) and surface albedo (α). Although other (non-cloud) variables may also be relevant, the majority of adjustments are expected to be captured under this framework (Vial et al., 2013) . For determining adjustments to a radiative forcing A_X, the kernel K_X is multiplied by the change in atmospheric state variable ΔX between two integrations of a climate model such that

$$A_X = K_X \Delta X. \tag{2}$$

ΔX is calculated as the difference of two atmosphere-only climate ~~mode~~ model integrations using climatological sea-surface temperatures and sea ice distributions, one of which is driven by a forcing perturbation (e.g. a quadrupling of CO_2) and the other a control. For temperature and albedo the adjustment is linear with ΔX, and logarithmic for water vapour (Sanderson and Shell (2012) and Smith et al. (2018, Supplementary Material) describe how the adjustment to water vapour is applied in practice). For determining climate feedbacks λ_X, the perturbation is normalised by the change in global mean near-surface air temperature T such that

$$\lambda_X = K_X \frac{\partial X}{\partial T}. \tag{3}$$

The individual contributions from each feedback component λ_X contribute the total climate feedback $\lambda = \lambda_{T_a} + \lambda_{T_s} + \lambda_q + \lambda_\alpha + \lambda_c$ where c represents cloud feedback in the forcing-feedback representation of the Earth's energy budget $\Delta N = F - \lambda \Delta T$. Here, ΔN is the Earth's energy imbalance and F the effective radiative forcing. Likewise, the effective radiative forcing can be decomposed into

$$F = F_i + A_{T_a} + A_{T_s} + A_q + A_\alpha + A_c \tag{4}$$

with F_i being the instantaneous radiative forcing (IRF).

Usage of radiative kernels assumes that radiative perturbations change linearly with changes in atmospheric state. Where perturbations are small, linearity is an appropriate assumption both for feedbacks (Jonko et al., 2012) and adjustments (Smith et al., 2018).

Cloud adjustments and feedbacks cannot be determined ~~using standard~~ directly using atmospheric state kernels. They may be diagnosed using the cloud kernel based on ISCCP simulator diagnostics (Zelinka et al., 2012) or from the residual of all-sky and clear-sky radiative kernels (Soden et al., 2008; Shell et al., 2008). For adjustments this calculation is

$$A_c = (F - F^{\mathrm{clr}}) - (F_i - F_i^{\mathrm{clr}}) - \sum_{X \in \{T_a, T_s, q, \alpha\}} (A_X - A_X^{\mathrm{clr}}) \tag{5}$$

where the clr superscript represents fluxes calculated in the absence of clouds. In eq. (5), the instantaneous radiative forcing must be known or estimated. This method is commonly used, requiring the production of all-sky and clear-sky kernel sets to calculate all-sky and clear-sky adjustments.

This paper introduces the top-of-atmosphere and surface radiative kernels from the HadGEM3-GA7.1 model. HadGEM3-GA7.1 has 85 vertical levels up to 85 km vertical height (about 0.005 hPa). The construction of the HadGEM3-GA7.1 kernel was motivated by the observation that adjustments to a doubling of CO_2 in models participating in the Precipitation Driver and Response Model Intercomparison Project (PDRMIP; Myhre et al. (2017)) were around 0.3 W m^{-2} larger using the ECMWF-Oslo kernel (Myhre et al., 2018) than other kernels used in the same study (Smith et al., 2018, Supplementary Figure 3) . The ECMWF-Oslo kernel was built from ECMWF-Interim reanalysis data (Dee et al., 2011) which has 60 vertical levels up to 0.1 hPa. Most other kernels used in Smith et al. (2018) were derived from climate models with a lower model top, and it is possible that low-top kernels were underestimating the stratospheric adjustment in Smith et al. (2018) .

For stratospheric temperature adjustments we compare the ~~radiative~~ HadGEM3-GA7.1 kernel to other kernels in the literature using available $4\times CO_2$ results from climate models contributing to the Radiative Forcing Model Intercomparison Project (RFMIP). We find that ~~only~~ in general, kernels based upon climate models with a high stratospheric resolution ~~can adequately resolve the~~ are better at resolving the stratospheric adjustment to a CO_2 forcing.

2 ~~Methods~~Construction of the HadGEM3-GA7.1 radiative kernel

One year of a pre-industrial, atmosphere-only (i.e. with climatological sea-surface temperatures and sea ice distributions) integration of the HadGEM3-GA7.1 general circulation model (Williams et al., 2018; Mulcahy et al., 2018) was run. HadGEM3-GA7.1 is the atmospheric component of the HadGEM3-GC3.1 physcial model and UKESM1.0 Earth System model that represents the UK research community's contribution to CMIP6. The model was run at LL (N96) resolution with a latitude-longitude grid of 1.25° by 1.875° and 85 vertical levels extending up to 85 km (approximately 0.005 hPa) and a native model timestep of 20 minutes. A pre-industrial base climatology was chosen as the first identified use case for this kernel set was the RFMIP-ERF Tier 1 single-forcing experiments (Pincus et al., 2016) which compare individual perturbations from a pre-industrial control baseline, described in Smith et al. (2020) .

Model diagnostics of air temperature, specific humidity, surface (skin) temperature, surface albedo (ratio of broadband upwelling to downwelling shortwave surface radiation), model level pressure, surface pressure, cloud fraction, cloud water content, cloud ice content, effective (time-averaged) solar zenith angle and gridbox daylight fraction every two model hours were saved. 2-hourly sampling is considered to give an appropriate representation of the diurnal cycle of clouds and reduce biases due to variations in solar geometry with longitude compared to longer timesteps, while keeping computational demands to a minimum. These model outputs were transplanted into an offline version of the SOCRATES radiative transfer code (version

17.03; Manners et al. (2015); Edwards and Slingo (1996)) and top-of-atmosphere and surface radiative fluxes calculated for each two-hour timestep in both the shortwave and longwave spectra, for all-sky and clear-sky. SOCRATES is a broadband radiation code that uses 6 bands in the shortwave and 9 bands in the longwave and is the same radiation scheme used in the ~~online version of~~ HadGEM3-GC3.1 and UKESM1.0 climate models. Aerosols were neglected and greenhouse gases, including the prescribed CMIP6 monthly climatology in ozone concentrations, were set to their pre-industrial (1850) values. Following the protocol for RFMIP (Pincus et al., 2016), sea-surface temperatures and sea-ice distributions from 50 years of the HadGEM3-GC3.1 coupled model were used to build the climatology (Andrews et al., 2019). The kernel is therefore dependent on two aspects of the HadGEM3-GA7.1 model: the pre-industrial background climatology (including clouds), and the broadband version of the radiation code.

To build the kernel, each vertical level of the model on each 2-hour timestep was perturbed separately, firstly by 1 K for air temperature, and secondly by a perturbation in specific humidity that maintains relative humidity for an increase in 1 K (without actually changing the layer temperature). The surface temperature and surface albedo were also perturbed by 1 K and 1% (additive) individually each timestep. For each perturbation, surface and TOA fluxes are again saved for clear-sky and all-sky in the shortwave (SW) and longwave (LW), and the difference compared to the control simulation gives the radiative kernel for each model level or surface. Building the kernels took in total approximately three months of computing time on 24 processors on the University of Leeds "cluj" Linux cluster. Running the base HadGEM3 model at higher (MM) resolution, with five times as many grid points as LL, was not determined to be necessary, as kernels are usually ultimately regridded to the resolutions of other CMIP6 models which are approximately as coarse as LL (e.g. Table 1 in Smith et al. (2020)).

Following this, the air temperature and water vapour kernel outputs were normalised by multiplying by ~~$10000/p_{\mathrm{thick}}$~~ $10000\ \mathrm{Pa}/p_{\mathrm{thick}}$ where p_{thick} [Pa] is the thickness of each level in pressure co-ordinates. This allows the 85-level native model kernel to be reduced down to the 19-level standard CMIP6 pressure levels by providing a weighted average contribution to each pressure level. The kernels are further averaged by month. In the 19-level format they can be used with standard "Amon" model output from any CMIP6 model, which is one of the key advantages of radiative kernels.

3 Kernel ~~results~~profiles

3.1 Top-of-atmosphere kernels

Figure 1 shows the TOA radiative kernels for HadGEM3-GA7.1 for clear-sky and all-sky. The air temperature, all-sky kernel (fig. 1a) shows a peak in cooling in the tropical upper troposphere, showing the importance of this region for changes in radiative balance. There are also substantial contributions to the TOA radiation balance from the lower troposphere in the mid-latitudes. For clear-sky (fig. 1b) there is more latitude-height homogeneity in the troposphere, showing the impact of removing clouds. A key feature of the air temperature kernels is the increasing strength of the LW outgoing radiation with increasing stratospheric height. The temperature kernel is negative throughout the atmosphere, in keeping with the fact that an increase in temperature results in additional Planck emission of LW radiation to space.

Water vapour kernels (fig. 1c,d) also show a peak in the upper tropical troposphere, which is opposite in sign to the negative temperature adjustment owing to the fact that water vapour is a significant greenhouse gas. In contrast to the temperature kernel, the water vapour kernel is very insensitive in the dry upper stratosphere.

The impact of cloud masking is more easily seen for the surface temperature kernels (fig. 1e,f) and surface albedo kernels (fig. 1g,h).

3.2 Surface kernels

Surface kernels are most useful for determining precipitation adjustments (Myhre et al., 2018) and feedbacks (Previdi, 2010), where the precipitation adjustment is proportional to the atmospheric absorption, calculated as the difference in TOA and surface adjustments. Figure 2 shows the surface radiative kernels for HadGEM3-GA7.1 for clear-sky and all-sky. Both the air temperature (fig. 2a,b) and water vapour (fig. 2c,d) kernels are more sensitive for perturbations close to the surface than higher in the atmosphere (note non-linear colour scales). Cloud masking for the surface temperature kernel has less of an effect for surface fluxes than for TOA fluxes (fig. 2e,f), whereas the surface albedo kernel shows quite a similar spatial pattern (fig. 2g,h) to its the TOA counterpart.

4 Comparison to other kernels for stratospheric temperature

~~The construction of the HadGEM3-GA7.1 kernel was motivated by the observation that adjustments to a doubling of CO_2 in PDRMIP models (Myhre et al., 2017) was around 0.3 W m^{-2} larger using the ECMWF-Oslo kernel (Myhre et al., 2018) than other kernels used in the same study (Smith et al., 2018, Supplementary Figure 3) . The ECMWF-Oslo kernel was built from ECMWF-Interim reanalysis data (Dee et al., 2011) which has 60 vertical levels up to 0.1 hPa. In contrast, most other kernels used in Smith et al. (2018) had a model top much lower.~~

Figure 3 shows the air temperature kernel for the stratosphere and upper troposphere for a selection of kernels available in the literature (table 1). In all cases, radiative kernels have been interpolated from their native vertical resolution (except for CCSM4, which is available only on the standard 17 CMIP5 pressure levels) to the 19 CMIP6 pressure levels for consistency with CMIP6 model output. For our calculations of stratospheric temperature adjustment, where kernels do not extend up to the 1 hPa top level of CMIP6 model output, kernels have been extended upwards using the value from the highest level where data does exist, but in fig. 3 missing data has been masked out. This extending upwards of the top level has been applied previously in adjustment calculations where the top level of the climate model is higher than the top level of the kernel (e.g. in Smith et al. (2018)). However, extending the top level of a radiative kernel upwards cannot make up for the fact that more radiation is emitted to space from the upper stratosphere for each additional K of temperature change. For kernels built from ~~climate models with a low model top or a coarse model~~ underlying atmospheric profiles where the top of the profile is not sufficiently high or with too coarse a resolution in the stratosphere, this additional upper stratospheric cooling is missed. In fig. 3, it can be seen that the kernels based on a high-top atmospheric model with a high number of native model levels—ECHAM6, ECMWF-

Table 1. Radiative kernels considered in this study.

Base model	Native model vertical levels	Top level (hPa)	3rd level (hPa)	Reference
BMRC	17	8.75	53.63	Soden et al. (2008)
CCSM4	17	10	30	Shell et al. (2008)
CESM	30	3.64	14.36	Pendergrass et al. (2018)
ECHAM5	19	10	50.39	Previdi (2010)
ECHAM6	47	0.0099	0.11	Block and Mauritsen (2013)
ECMWF-RRTMG	24	1	6	Huang et al. (2017)
ECMWF-Oslo	60	0.11	0.5	Myhre et al. (2018)
GFDL	25	3.32	53.63	Soden et al. (2008)
HadGEM2	38	2.99	13.02	Smith et al. (2018)
HadGEM3-GA7.1	85	0.005	0.03	this study

Oslo, ECMWF-RRTM and HadGEM3-GA7.1—have a marked increase in both the magnitude and the rate of negative LW outgoing flux at the 5 hPa and 1 hPa levels.

The consequences for a CO_2-induced greenhouse-gas-induced stratospheric cooling are such that the additional stratospheric adjustment from greater cooling high in the stratosphere is not accounted for with either kernels or models that are truncated too low. Figure 4 shows the atmospheric temperature anomalies simulated in atmosphere-only simulations from CMIP6 models participating in RFMIP-ERF Tier 1 experiments (Pincus et al., 2016) for a 30-year time slice simulation where CO_2 concentrations are quadrupled relative to a pre-industrial control (piClim-4xCO2). Stratospheric cooling continues to increase above 5 hPa in 12 out of the 13 10 hPa in all models where data is available, with the only exception being GFDL-CM4 which has a layer of missing data at 1 hPa. A similar pattern of stratospheric cooling occurs in the piClim-ghg experiment which evaluates forcing from present-day greenhouse gases (not shown). Standard CMIP6 diagnostics call for model output on 19 pressure levels: 1000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, 10, 5 and 1 hPa, whereas in CMIP5 the standard set of 17 pressure levels did not include 5 and 1 hPa. Therefore, CMIP5 models were missing important additional stratospheric cooling where kernels were used for adjustment calculations.

The truncation of stratospheric height in "low top" radiative kernels (all but ECHAM6, ECMWF-Oslo and HadGEM3-GA7.1 as seen in those with a top level lower in altitude than 1 hPa) has substantial consequences for adjustments to a CO_2 greenhouse gas forcing. Figure 5 shows the stratospheric temperature adjustment to for $4\times CO_2$ in the 13 models contributing to RFMIP. A simplified tropopause definition is used here, borrowed from Soden et al. (2008), of a linear in latitude ramp from 100 hPa at the equator to 300 hPa at the poles. There is a spread of around 1 W m^{-2} in calculated stratospheric temperature adjustment for each model using the full range of kernels, which is about 13% of the effective radiative forcing (ERF) for a quadrupling of CO_2 from these models (Smith et al., 2020). It can be seen in fig. 5 that the kernel estimates for stratospheric adjustment to CO_2 forcing are clustered into two groups and one high outlier for most models. The one model where kernel estimates are not clearly separated into high and low clusters is GFDL-CM4 model, for which data is missing at 1 hPa. The "low-top" radiative kernels, with the exception of GFDL and ECHAM5, produce substantially lower estimates of the stratospheric temperature adjustment than the 'high-top" kernels (HadGEM3-GA7.1, ECHAM6 and ECMWF-Oslo). The ~~GFDL kernel~~ high outlier, ECMWF-RRTM, has a large flux change at the 1 hPa level. We show in section 5 that adjustments calculated using the HadGEM3-GA7.1 kernel in the IPSL-CM6A-LR model for a quadrupled CO_2 experiment provide small residuals (i.e. the adjustments are appropriately captured), suggesting that assuming there are no compensating errors, low-top kernels would underestimate the stratospheric temperature response and produce larger residuals.

To investigate the source of the differences between kernels in more detail, fig. 6 shows the vertically integrated kernels from 1 hPa, 10 hPa, the full stratosphere (Soden et al. (2008) definition) and the full atmosphere. As previously, kernels that do not include the uppermost layers have been extended based on the highest layer for which data is reported. Figure 6a gives the temperature adjustment from a uniform 1 K increase throughout the atmosphere. Here, there is little variation between kernels, and much of the total adjustment comes from the troposphere where the bulk of the atmosphere is present. Notably, the whole atmosphere adjustment to a 1 K temperature increase is of the order -3 W m^{-2}, approximating (if slightly underestimating) the Planck feedback (-3.2 W m^{-2} K^{-1}).

Figure 6b shows the layer contributions when each kernel is applied to the atmospheric profile from the IPSL-CM6A-LR model for piClim-4xCO2. In contrast to the isothermal case, the top 1 hPa, top 10 hPa and stratospheric temperature adjustments show substantial diversity between kernels. The large contributions from these layers follow from fig. 4 where the 1 and 10 hPa layers are of the order 10 K and 20 K cooler in the 4$\times CO_2$ run than in the control. The HadGEM3-GA7.1, GFDL, ECMWF-Oslo, ECMWF-RRTM and ECHAM6 kernels have a larger adjustment from the top 10 hPa than the other kernels. Despite being nominally a "low-top" kernel, the GFDL kernel has a similar magnitude and gradient of cooling between 10 and 5 hPa as the high-top kernels ~~, and~~ (fig. 3) and produces a similar result of total stratospheric adjustment to HadGEM3-GA7.1. The ECHAM5 kernel has more cooling around the 100 hPa level than any other kernel ~~. These reasons may explain why the stratospheric adjustments estimated from this kernel are more in line with the three high-top kernels. The one model where kernel estimates are not clearly separated into high and low clusters is again the GFDL-CM4 model, for which the missing data at 1 hPa impacts adjustment estimates from different kernels~~ (fig. 3) and also has stratospheric adjustments that are similar to HadGEM3-GA7.1.

The differences between models for the top 10 hPa propagate through to the tropopause level and whole atmosphere. Differences in tropospheric adjustments between kernels (whole atmosphere minus stratosphere) are small, showing that choice of atmospheric base state and radiative transfer is not critical for tropospheric temperature adjustments. This analysis gives further confidence that the choice of radiative kernel is not that important in climate feedback studies (assuming state changes are sufficiently small to remain linear; (Jonko et al., 2012)) as stratospheric temperature differences are small when differencing coupled atmosphere-ocean simulations (Chung and Soden, 2015b) and kernels show similar behaviour in the troposphere.

5 Accuracy Linearity of the HadGEM3-GA7.1 kernel

Where "double calls" or other methods of determining the IRF are not obtained directly from climate output, the IRF is estimated as a residual of the ERF and all adjustments. If the cloud adjustments are known (e.g. from the ISCCP simulator kernel, Zelinka et al. (2012)), it follows from that

$$F_i = F - A_{T_a} - A_{T_s} - A_q - A_\alpha - A_c.$$

This section shows the linear behaviour of the HadGEM3-GA7.1 kernel used with IPSL-CM6A-LR climate model output. IPSL-CM6A-LR is chosen as all required diagnostics are available in this model to evaluate linearity, including estimates of the IRF from double calls, and ISCCP simulator diagnostics for clouds. This model is also representative of the RFMIP population, with 4×CO_2 ERF and adjustments close to the multi-model average (Smith et al., 2020) .

The IRF is an important and useful concept in itself, as although it is not the best predictor of long-term near-surface global mean temperature changes from a forcing (Hansen et al., 2005) , it can be used to benchmark the performance of radiative transfer parameterisation in climate models (Pincus et al., 2015; Soden et al., 2018) .

These breakdowns of ERF into IRF and adjustments using kernels depend on the kernels being able to perform this decomposition linearly. The *residual*Linearity is evaluated by the size of the residual term, ϵ, describes any error term resulting from a non-linear decomposition. Two different ways of calculating the residual can be obtained. If the all-sky IRF (e. g. from double call) and which is any TOA flux changes not explained by instantaneous radiative forcing or kernel-calculated adjustments. A guideline of linearity for the kernel method is that the residual should be within 10% of the ERF. We take two different approaches to calculate the residual. The first assumes that we have knowledge of the cloud adjustment term A_c (e.g. from the ISCCP simulator kernel)are both known, then the or NASA A-Train kernels convoluted with ISSCP simulated cloud output from climate models (Zelinka et al., 2012; Yue et al., 2016)), and knowledge of the IRF (F_i) from a double-call in the online model. This is rare in practice, as double-calls and ISCCP cloud diagnostics are not routinely archived on ESGF (although model participation is improving for ISCCP clouds). From this "perfect information" method we can calcaulate an all-sky residual ϵ_{all} is as

$$\epsilon_{\text{all}} = F - F_i - A_{T_a} - A_{T_s} - A_q - A_\alpha - A_c. \qquad (6)$$

If the

The second method does not require knowledge of the cloud adjustments, but does require knowledge of the clear-sky IRFis known, , and in this case the clear-sky residual term ϵ_{clr} can be calculated as

$$\epsilon_{\text{clr}} = F^{\text{clr}} - F_i^{\text{clr}} - A_{T_a}^{\text{clr}} - A_{T_s}^{\text{clr}} - A_q^{\text{clr}} - A_\alpha^{\text{clr}}. \qquad (7)$$

In practice, the kernel method is assumed to perform sufficiently well for ϵ_{clr} being within 10% of the ERF. In This clear-sky residual definition is more common in the literature (Smith et al., 2018; Soden et al., 2008; Vial et al., 2013) , although often the clear-sky IRF is estimated rather than calculated directly. However, in some circumstances, IRF is clear-sky and all-sky IRF

Table 2. IPSL-CM6A-LR double call results for $4\times CO_2$ experiments. IRFs are given in W m^{-2}.

Base climatology	Second call	IRF LW	IRF SW	IRF Net	IRF LW CS	IRF SW CS	IRF Net CS
pre-industrial	$4\times CO_2$	3.66	0.83	4.49	5.02	0.46	5.48
$4\times CO_2$	pre-industrial	4.94	0.81	5.75	6.26	0.46	6.72
Mean	Mean	4.30	0.82	5.12	5.64	0.46	6.10

Table 3. IPSL-CM6A-LR forcing and adjustments for the $4\times CO_2$ experiment using the HadGEM3-GA7.1 kernel. Fluxes are given in W m^{-2}. A_{T_a} strat. and A_{T_a} trop. are stratospheric and tropospheric temperature adjustments.

	ERF	IRF	A_{T_a} strat.	A_{T_a} trop.	A_{T_s}	A_q	A_α	A_c (eq. (5))	A_c (ISCCP kernel)	ϵ_{clr}	ϵ_{all}
LW	5.33	4.31	2.74	−1.38	−0.49	0.52		−0.66	−0.75	0.28	0.38
SW	2.68	0.82				0.11	0.18	1.60	1.81	−0.02	−0.23
Net	8.01	5.12	2.74	−1.38	−0.49	0.63	0.18	0.94	1.06	0.26	0.15

are known to be identically zero (e.g. in the LW spectrum to a change in the solar constant; Smith et al. (2018))~~and~~. In these cases, eq. (5) can be used with $F_i = F_i^{clr} = 0$ to determine ~~cloud adjustments~~A_c which is then plugged into eq. (6) to calculate ϵ_{all}.

~~We can test the linear separation in the~~ IPSL-CM6A-LR ~~model, where IRF were archived using a double call. For~~ used two sets of double calls. In the RFMIP piClim-4xCO2 experiment (30-year time-slice atmosphere-only run with quadrupled CO_2) the second radiation call ~~used~~ saw a pre-industrial CO_2 concentration~~, and in~~. In the piClim-control experiment (pre-industrial atmosphere only run) the second radiation call saw $4\times CO_2$. The ~~IRF estimated shows a substantial dependence on the base climatology, with the~~ resulting IRF depends on the direction of the double call and is related to the underlying $4\times CO_2$ or pre-industrial climatology. The $4\times CO_2$ climate and pre-industrial second radiation call ~~showing LW fluxes more than 1.2~~ exhibits an IRF that is 1.26 W m^{-2} greater than the pre-industrial climate with $4\times CO_2$ second radiation call (table 2). We take the mean of the two simulations to the be the IRF.

Table 3 shows ERF, IRF, adjustments and residuals using the HadGEM3-GA7.1 radiative kernel with the IPSL-CM6A-LR model output ~~. ISCCP simulator diagnostics are also available for this model. We therefore obtain estimates of SW and LW~~ and cloud adjustments from the ISCCP simulator kernel~~and use these estimates along with the IRF to estimate~~. We use ISCCP kernel values in the calculation of ϵ_{all} ~~, alongside the~~ and also compare the ISCCP values to the cloud-masking estimate of cloud adjustment from eq. (5). For LW forcing the residuals are 0.28 W m^{-2} for ϵ_{clr} and 0.38 W m^{-2} for ϵ_{all}. Residuals are present possibly due to a slight breakdown in the linearity assumption for a forcing as large as $4\times CO_2$ (Jonko et al., 2012), however, the residuals are comfortably within the 10% linearity guideline. SW residuals are also within 10% of the ERF, with ϵ_{clr} being particularly small. For the net fluxes, forcings add but residuals partly cancel, such that ϵ_{clr} and ϵ_{all} are 3.2% and 1.9% of the ERF respectively.

Table 4. IPSL-CM6A-LR forcing and adjustments for the 4×CO_2 experiment using the HadGEM3-GA7.1 kernel and ISCCP cloud kernel (Zelinka et al., 2012) compared to the multi-model mean and standard deviation (1σ) in Zhang and Huang (2014) from 11 CMIP5 models. Fluxes are given in W m^{-2}. *Starred values are outside the 2σ range from Zhang and Huang (2014) .

Forcing or adjustment	IPSL-CM6A-LR	Zhang and Huang (2014)
ERF	8.01	7.18 (± 0.72)
IRF	5.12	5.41 (± 0.46)
Stratospheric temperature	*2.74	1.86 (± 0.36)
Tropospheric + surface temperature	−1.87	−1.66 (± 0.21)
Water vapour (LW)	0.63	0.42 (± 0.12)
Clouds (LW)	−0.75	−0.40 (± 0.50)
Total SW	1.89	1.55 (± 0.83)

~~The~~ Our results in table 3 can be compared with the results of Zhang and Huang (2014) for 11 CMIP5 models. The instantaneous forcing and tropospheric adjustments from IPSL-CM6A-LR with the HadGEM3-GA7.1 kernel are in line with the CMIP5 forcing and adjustments except for the stratospheric temperature adjustment~~is the only adjustment~~ , which is outside the 2σ range from CMIP5 models (table 4). As discussed in Smith et al. (2020) , the 4×CO_2 ERF in available CMIP6 models (7.98 W m^{-2}) is (non-significantly) greater than in corresponding CMIP5 sstClim4xCO2 experiments (7.53 W m^{-2}), which is also the case for IPSL-CM6A-LR in our comparison to Zhang and Huang (2014) . IPSL-CM6A-LR is near the centre of the CMIP6 range for stratospheric temperature adjustment (Smith et al., 2020) so is a representative model of this ensemble. This could suggest stratospheric temperature adjustment increase as one driver of the increase in ERF between CMIP5 and CMIP6 models, although as most CMIP5 output is only on 17 model levels up to 10 hPa a formal comparison is difficult.

The stratospheric temperature adjustment is the only adjustment estimate that varies significantly between radiative kernels (Smith et al., 2018). If a ~~low-top kernel was used to estimate~~ kernel that did not resolve stratospheric temperature adjustment well was used instead of HadGEM3-GA7.1, this adjustment (A_{T_a} strat. in table 3~~, this adjustment~~) would be smaller, and the overall residuals for LW and net ~~forcings~~ responses to 4×CO_2 larger. From fig. 5 it can be seen that some kernels produce a stratospheric temperature adjustment around 0.7 W m^{-2} lower than the HadGEM3-GA7.1 kernel, ~~leading~~ which would lead to residuals of the order 1 W m^{-2} using these kernels, or more than 10% of the ERF.

6 Conclusions

This paper serves two purposes—it introduces the radiative kernel based on the high-top HadGEM3-GA7.1 general circulation model, and it compares estimates of the the stratospheric temperature adjustment obtained with a variety of different radiative kernels for quadrupled CO_2 experiments. The HadGEM3-GA7.1 kernel is the first to our knowledge that has been produced using a CMIP6 era model, with a focus on the 19 pressure level diagnostics available in CMIP6 output, although other kernels in the literature have used high-top atmospheric profiles (Huang et al., 2017; Myhre et al., 2018; Block and Mauritsen, 2013) . Radiative kernels are produced for both top-of-atmosphere and surface fluxes and are available on the native 85-level hybrid height grid in addition to the 19 CMIP6 pressure levels.

We show that there is a significant diversity, of about 1 W m^{-2} or 13% of the ERF for a quadrupling of CO_2, for estimates of stratospheric temperature adjustments to CO_2 depending on the radiative kernel used to derive the estimate. As tropospheric and land surface adjustments vary little between kernels to a variety of different forcing agents (Smith et al., 2018, 2020), these differences in stratospheric temperature adjustments lead to differing estimates of the total adjustment, and also of the IRF if it is calculated as a residual (Chung and Soden, 2015b, a; Soden et al., 2018). Climate feedbacks are little affected by the choice of kernel, due to the fact that stratospheric temperatures readjust quickly to an imposed forcing in coupled model simulations (Chung and Soden, 2015b).

While only one model (IPSL-CM6A-LR) archived IRF from a double call and a rigorous multi-model test is not possible, we show that the HadGEM3-GA7.1 kernel diagnoses IRF and adjustments with a small residual owing to the increased stratospheric resolution available compared to many CMIP3- and CMIP5-era kernels. We recommend that stratospheric temperature adjustments are calculated using our kernel, or the ECHAM6 (Block and Mauritsen, 2013) or ECMWF-Oslo kernels (Myhre et al., 2018) suggest that radiative kernels with a higher stratospheric resolution and model top are better able to fully capture stratospheric adjustments to CO_2 forcing in general, and generate smaller residuals. This effect has become more prominent with the additional 5 hPa and 1 hPa model levels archived as standard in processed CMIP6 model output compared to CMIP5. Archiving instantaneous radiative forcing from more models would be beneficial to further test the linearity assumption of the radiative kernel method.

Data availability. The HadGEM3-GA7.1 radiative kernels are available at https://doi.org/10.5281/zenodo.3594673 (Smith, 2019) .

Author contributions. C.J.S. produced the HadGEM3-GA7.1 radiative kernel and led the writing of the manuscript. R.J.K. provided calculations of stratospheric adjustment to 4×CO_2 for all kernels considered in this paper. A.S. provided double call results from the IPSL-CM6A-LR model.

Competing interests. The authors declare no competing interests.

Acknowledgements. C.J.S. was supported by a NERC/IIASA Collaborative Research Fellowship (NE/T009381/1) and 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. This work used the ARCHER UK National Supercomputing Service (http://www.archer.ac.uk).

References

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.

Block, K. and Mauritsen, T.: Forcing and feedback in the MPI-ESM-LR coupled model under abruptly quadrupled CO2, J. Adv. Model. Earth Sy., 5, 676–691, https://doi.org/10.1002/jame.20041, 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, 2015a.

Chung, E.-S. and Soden, B. J.: An Assessment of Direct Radiative Forcing, Radiative Adjustments, and Radiative Feedbacks in Coupled Ocean–Atmosphere Models, J. Climate, 28, 4152–4170, https://doi.org/10.1175/JCLI-D-14-00436.1, 2015b.

Colman, R. A. and McAvaney, B. J.: On tropospheric adjustment to forcing and climate feedbacks, Clim. Dynam., 36, 1649, https://doi.org/10.1007/s00382-011-1067-4, 2011.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

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

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.

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

Huang, Y., Xia, Y., and Tan, X.: On the pattern of CO2 radiative forcing and poleward energy transport, Journal of Geophysical Research: Atmospheres, 122, 10,578–10,593, https://doi.org/10.1002/2017JD027221, 2017.

Jonko, A. K., Shell, K. M., Sanderson, B. M., and Danabasoglu, G.: Climate Feedbacks in CCSM3 under Changing CO2 Forcing. Part I: Adapting the Linear Radiative Kernel Technique to Feedback Calculations for a Broad Range of Forcings, J. Climate, 25, 5260–5272, https://doi.org/10.1175/JCLI-D-11-00524.1, 2012.

Manners, J., Edwards, J. M., Hill, P., and Thelen, J.: SOCRATES (Suite Of Community RAdiative Transfer codes based on Edwards and Slingo) technical guide, Tech. rep., Met Office, UK, 2015.

Mulcahy, J. P., Jones, C., Sellar, A., Johnson, B., Boutle, I. A., Jones, A., Andrews, T., Rumbold, S. T., Mollard, J., Bellouin, N., Johnson, C. E., Williams, K. D., Grosvenor, D. P., and McCoy, D. T.: Improved Aerosol Processes and Effective Radiative Forcing in HadGEM3 and UKESM1, J. Adv. Model. Earth Sy., 10, 2786–2805, https://doi.org/10.1029/2018MS001464, 2018.

Myhre, G., Forster, P., Samset, B., Hodnebrog, Sillmann, J., Aalbergsjø, S., Andrews, T., Boucher, O., Faluvegi, G., Fläschner, D., Iversen, T., Kasoar, M., Kharin, V., Kirkevag, A., Lamarque, J., Olivié, D., Richardson, T., Shindell, D., Shine, K., Stjern, C., 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, https://doi.org/10.1175/BAMS-D-16-0019.1, 2017.

Myhre, G., Kramer, R. J., Smith, C. J., Hodnebrog, O., Forster, P., Soden, B. J., Samset, B. H., Stjern, C. W., Andrews, T., Boucher, O., Faluvegi, G., Fläschner, D., Kasoar, M., Kirkevåg, A., Lamarque, J.-F., Olivié, D., Richardson, T., Shindell, D., Stier, P., Takemura, T., Voulgarakis, A., and Watson-Parris, D.: Quantifying the Importance of Rapid Adjustments for Global Precipitation Changes, Geophys. Res. Lett., 45, 11,399–11,405, https://doi.org/10.1029/2018GL079474, 2018.

Pendergrass, A. G., Conley, A., and Vitt, F. M.: Surface and top-of-atmosphere radiative feedback kernels for CESM-CAM5, Earth Syst. Sci. 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, https://www.geosci-model-dev.net/9/3447/2016/, 2016.

Previdi, M.: Radiative feedbacks on global precipitation, Env. Res. Lett., 5, 025 211, https://doi.org/10.1088/1748-9326/5/2/025211, 2010.

Sanderson, B. M. and Shell, K. M.: Model-Specific Radiative Kernels for Calculating Cloud and Noncloud Climate Feedbacks, J. Climate, 25, 7607–7624, https://doi.org/10.1175/JCLI-D-11-00726.1, 2012.

Shell, K. M., Kiehl, J. T., and Shields, C. A.: 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.

Smith, C.: HadGEM3-GA7.1 radiative kernels, https://doi.org/10.5281/zenodo.3594673, 2019.

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, 2018.

Smith, C. J., Kramer, R. J., Myhre, G., Alterskjær, K., Collins, W., Sima, A., Boucher, O., Dufresne, J.-L., Nabat, P., Michou, M., Yukimoto, S., Cole, J., Paynter, D., Shiogama, H., O'Connor, F. M., Robertson, E., Wiltshire, A., Andrews, T., Hannay, C., Miller, R., Nazarenko, L., Kirkevåg, A., Olivié, D., Fiedler, S., Pincus, R., and Forster, P. M.: Effective radiative forcing and adjustments in CMIP6 models, Atmospheric Chemistry and Physics Discussions, 2020, 1–37, https://doi.org/10.5194/acp-2019-1212, 2020.

Soden, B. J., Held, I. M., Colman, R., Shell, K. M., Kiehl, J. T., and Shields, C. A.: 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.

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.

Wetherald, R. T. and Manabe, S.: Cloud Feedback Processes in a General Circulation Model, J. Atmos. Sci., 45, 1397–1416, https://doi.org/10.1175/1520-0469(1988)045<1397:CFPIAG>2.0.CO;2, 1988.

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.

Yue, Q., Kahn, B. H., Fetzer, E. J., Schreier, M., Wong, S., Chen, X., and Huang, X.: Observation-Based Longwave Cloud Radiative Kernels Derived from the A-Train, Journal of Climate, 29, 2023–2040, https://doi.org/10.1175/JCLI-D-15-0257.1, 2016.

Zelinka, M. D., Klein, S. A., and Hartmann, D. L.: 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.

Zhang, M. and Huang, Y.: Radiative Forcing of Quadrupling CO2, Journal of Climate, 27, 2496–2508, https://doi.org/10.1175/JCLI-D-13-00535.1, 2014.

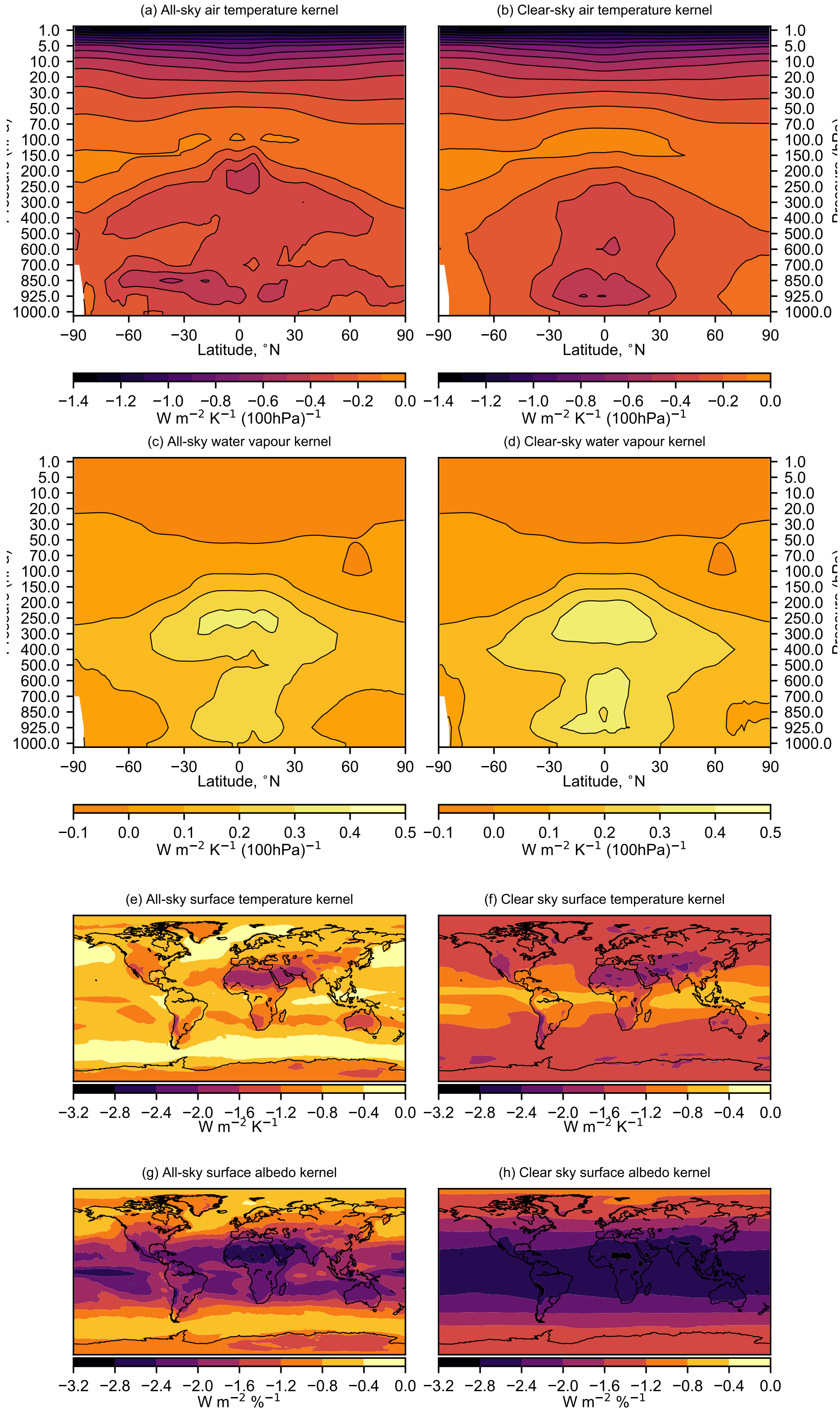


Figure 1. Top-of-atmosphere radiative kernels from HadGEM3-GA7.1.

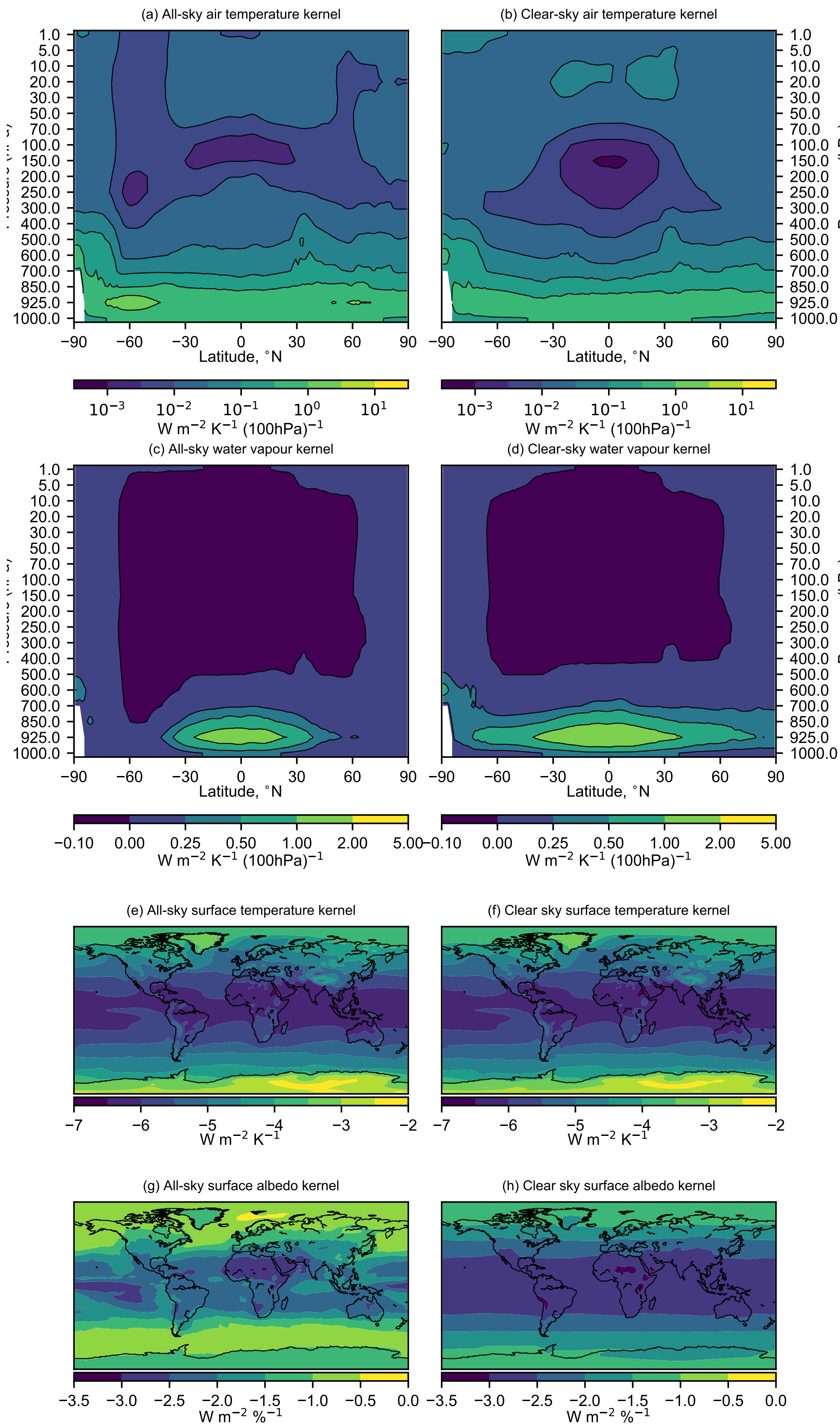


Figure 2. Surface radiative kernels from HadGEM3-GA7.1.

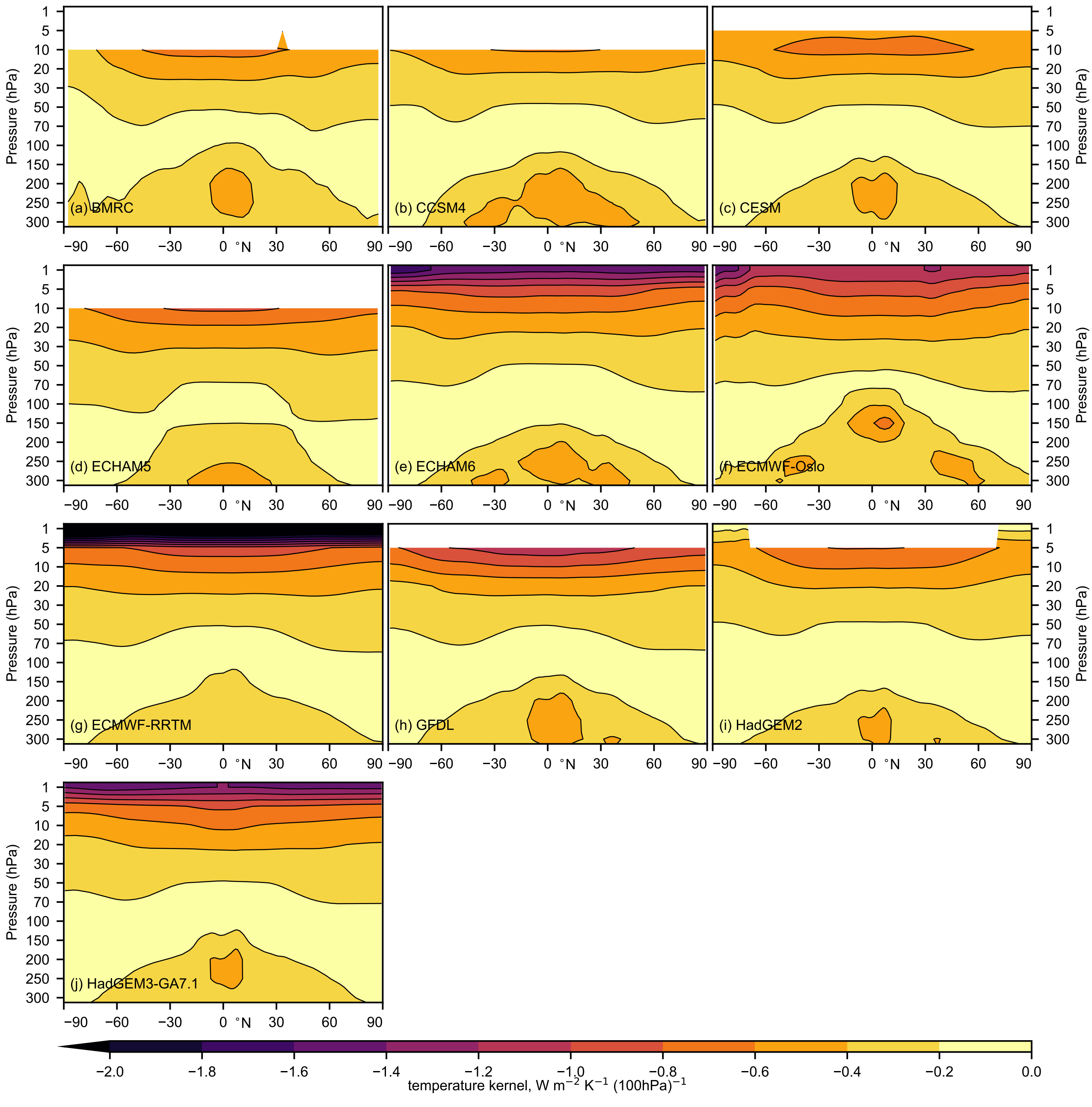


Figure 3. Air temperature radiative kernels available in the literature, truncated at 300 hPa to show the stratospheric temperature contribution. Blank areas are above the top level of the kernel.

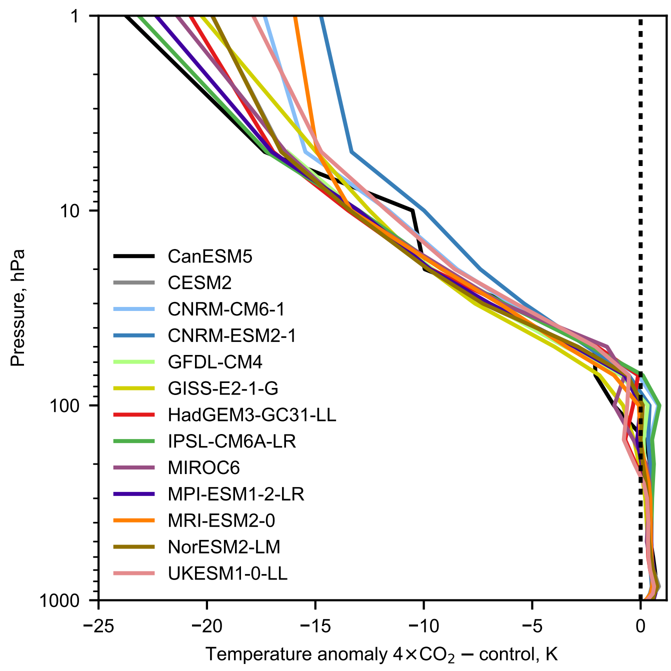


Figure 4. Atmospheric temperature differences for RFMIP models for the piClim-4xCO2 experiment minus piClim-control.

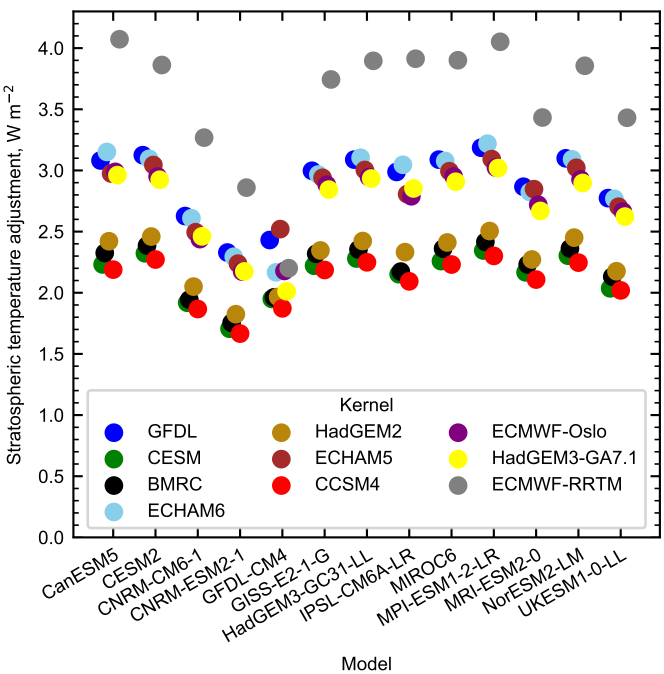


Figure 5. Stratospheric temperature adjustments calculated from RFMIP piClim-4xCO2 experiments using all kernels available in this study.

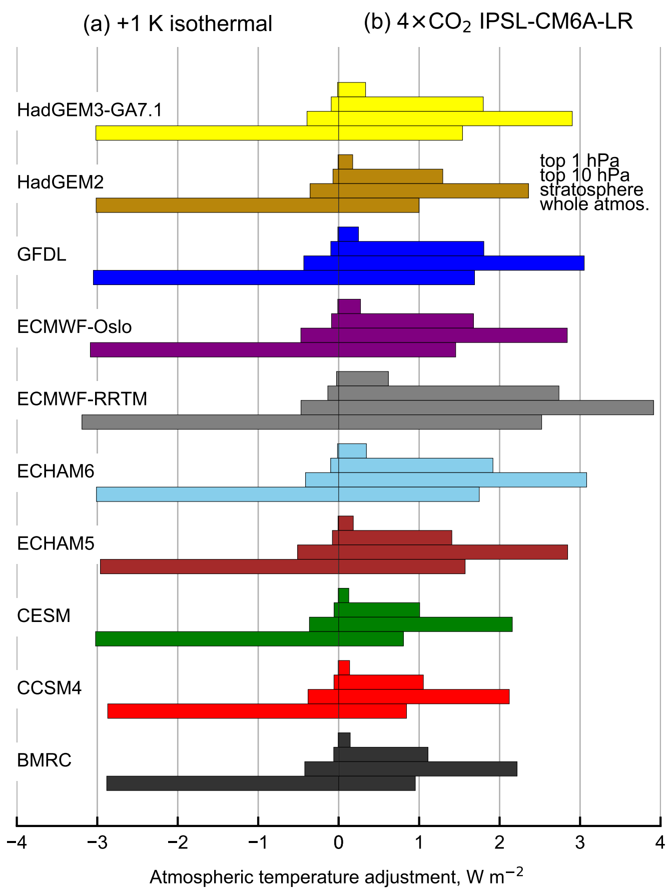


Figure 6. ~~Stratospheric~~ Layer contributions to the TOA temperature ~~adjustments calculated from RFMIP piClim-4xCO2 experiments using all kernels available~~ adjustment in each kernel considered in this study. From top to bottom the four bars show 1 hPa to TOA, 10 hPa to TOA, tropopause to TOA (i.e. stratospheric adjustment) and surface to TOA. (a) The effect of a uniform 1 K increase in atmospheric temperature on TOA fluxes, as diagnosed directly from the radiative kernel. (b) The kernels convoluted with atmospheric profiles from the IPSL-CM6A-LR model under an atmosphere-only quadrupled CO_2 run (RFMIP piClim-4xCO2 minus piClim-control).