21st century scenario forcing increases more for CMIP6 than CMIP5 models

Hege-Beate Fredriksen[1], Christopher J Smith[2], Angshuman Modak[3], and Maria A.A. Rugenstein[4]

[1]Department of Physics and Technology, UiT The Arctic University of Norway
[2]International Institute for Applied Systems Analysis
[3]Department of Meteorology, Stockholm University
[4]Colorado State University

November 22, 2022

Abstract

We present new estimates of the forcing for models participating in Coupled Model Intercomparison Project 6 (CMIP6) by applying the method developed in Fredriksen et al. (2021). Validating our approach, these estimates are overall consistent with the fixed-SST estimates available for a small subset of the models. We estimate forcing for experiments with abrupt changes of CO_2, 1% increase of CO2, historical forcings, and future scenarios. Furthermore, we compare our new estimates to CMIP5 forcing, and demonstrate that CMIP6 forcing is lower than CMIP5 forcing at the end of the historical period, but grows faster than CMIP5 in the future scenarios, ending up at higher levels than CMIP5 at the end of the 21st century. The radiative efficiency of CO_2 has not changed, suggesting that the stronger future increase in CO_2 concentrations in CMIP6 compared to CMIP5 explains the forcing difference.

21st century scenario forcing increases more for CMIP6 than CMIP5 models

Hege-Beate Fredriksen[1], Christopher J. Smith[2,3], Angshuman Modak[4], and Maria Rugenstein[5]

[1]Department of Physics and Technology, UiT the Arctic University of Norway, Tromsø, Norway
[2]Priestley International Center for Climate, University of Leeds, Leeds, UK
[3]International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
[4]Department of Meteorology (MISU), Stockholm University, Stockholm, Sweden
[5]Colorado State University, Fort Collins, USA

Key Points:

- Our method to estimate effective radiative forcing based on common model diagnostics is consistent with fixed-SST estimates
- We present estimates for abrupt CO_2, 1% CO_2, historical, and future scenario experiments
- Forcing estimates for the 21st century grows faster for CMIP6 than for CMIP5 models

Corresponding author: Hege-Beate Fredriksen, `hege-beate.fredriksen@uit.no`

Abstract

We present new estimates of the forcing for models participating in Coupled Model Intercomparison Project 6 (CMIP6) by applying the method developed in Fredriksen et al. (2021). Validating our approach, these estimates are overall consistent with the fixed-SST estimates available for a small subset of the models. We estimate forcing for experiments with abrupt changes of CO_2, 1% increase of CO_2, historical forcings, and future scenarios. Furthermore, we compare our new estimates to CMIP5 forcing, and demonstrate that CMIP6 forcing is lower than CMIP5 forcing at the end of the historical period, but grows faster than CMIP5 in the future scenarios, ending up at higher levels than CMIP5 at the end of the 21st century. The radiative efficiency of CO_2 has not changed, suggesting that the stronger future increase in CO_2 concentrations in CMIP6 compared to CMIP5 explains the forcing difference.

Plain Language Summary

To understand climate model responses, it is useful to separate between the drivers of climate change and their responses. We present new estimates of the drivers, called the effective radiative forcing, for the latest generation of climate models (CMIP6). This estimates the energy input at the top of the atmosphere and is a measure of human and natural influences on climate. Normally this requires additional climate model experiments to make these estimates, but since these have only been run for a few models, we are here aiming to make the best alternative estimates based on existing data, following the method in Fredriksen et al. (2021). We show that our forcing estimates are growing faster during the 21st century for the new CMIP6 models than for the previous generation of models (CMIP5), and suggest this can be attributed to the higher CO_2 concentrations in future scenarios for CMIP6 compared to CMIP5.

1 Introduction

The [effective] radiative forcing (ERF) describes the energy input into the Earth system that drives climate change. As well as a common currency to compare the energetic impacts of different human and natural influences on the climate, it is used to develop scenarios characterising possible futures, for example in representative concentration pathway (RCP) and shared socioeconomic pathway (SSP) scenarios (Moss et al., 2010; O'Neill et al., 2016). However, [effective] radiative forcing is difficult to observe and complex climate models such as the general circulation models (GCMs) developed as part of the Coupled Model Intercomparison Project Phase 6 (CMIP6) are often the best or only way to determine ERF. Accurately quantifying ERF will allow us to attribute cause and effect in climate model behavior and better constrain climate sensitivity.

Unfortunately, only a small number of CMIP6 models—9 out of 51— provided estimates of ERF, for the historical period and one scenario to 2100. These ERF estimates were derived from atmosphere-only runs of CMIP6 models using pre-industrial sea-surface temperatures and sea-ice distributions (Hansen et al., 2005; Forster et al., 2016), known as the fixed-SST method. The experiment is a Tier 2 simulation provided by the Radiative Forcing Model Intercomparison Project (RFMIP) contribution to CMIP6 (Pincus et al., 2016). To obtain estimates of ERF from more models and scenarios, we can use estimates of the climate feedback parameter from each model's abrupt-4xCO2 experiment (a mandatory experiment for all CMIP6 models) obtained from a Gregory regression, and use this to relate outputs of modelled top-of-atmosphere energy imbalance and surface temperature to time-varying ERF (for a full description of this method, see Forster et al. (2013)). However, this method is biased, as it is now well-known that the climate feedback parameter is not constant in time (e.g. Senior & Mitchell, 2000; Winton et al., 2010; Armour, 2017; Rugenstein et al., 2020; Andrews et al., 2022). Comparing the two methods for historical ERF shows that the bias is worse in models that show significant non-stationarity in their climate feedback parameter (Smith & Forster, 2021).

Acknowledging that fixed-SST ERF is not yet widely available from models, we can seek to improve ERF estimated from the abrupt-4xCO2 climate feedback. By calculating ERF assuming a time-scale dependent feedback parameter with three different time scales, we show in Fredriksen et al. (2021) that we can well describe the surface temperature output of the historical and RCP scenarios for the majority of CMIP5 models. In this paper we extend the analysis to CMIP6 models and scenarios, with the added confidence of comparing results with fixed-SST estimates in 10 cases, and also compare ERF in RCP (CMIP5) and SSP (CMIP6) scenarios with the same nominal year-2100 radiative forcing.

2 Data

We study CMIP6 models that have published the four variables *tas* (near-surface air temperature), *rlut*, *rsut* and *rsdt* (top of atmosphere longwave upwelling, shortwave upwelling, and shortwave downwelling radiation respectively) for both the piControl and the abrupt-4xCO2 experiment in March 2022. For these 51 models (listed in Table S1 and S2), we look at all members we could find for the experiments abrupt-2xCO2, abrupt-0p5xCO2, 1pctCO2, historical, hist-GHG, hist-aer, hist-nat, ssp119, ssp126, ssp245, ssp370, ssp585, piClim-4xCO2, piClim-control and piClim-histall. The number of members used for analysis is listed in Tables S1-S4. The piClim-* experiments are atmosphere-only simulations using climatological SSTs and sea ice distributions from the models' pre-industrial climate, and we often refer to these as fixed-SST experiments. For many models, the piClim-histall experiments are extended with the SSP2-4.5 scenario to year 2100. In addition to the 9 publicly available piClim-histall experiments, we have included an experiment done with the model MPI-ESM1-2-LR not available yet through CMIP6. In our Figure 3 we include also estimates presented in Fredriksen et al. (2021) for 21 CMIP5 models. For each variable, we have studied the annual anomalies relative to a linear trend (computed from all piControl years) evaluated in the corresponding period of the control run.

3 Method

The linear energy balance framework describes, to first order, the correspondence between forcing, feedbacks and global mean temperature:

$$N = F + \lambda T \tag{1}$$

where N is the top-of-the-atmosphere (TOA) net radiative downward flux (in W m^{-2}), λ is the climate feedback parameter (in W m^{-2} K^{-1}), T is the surface air temperature change (in K) relative to an unperturbed steady state where $N = F = 0$ and F is the external radiative forcing (in W m^{-2}), for instance due to a change in atmospheric composition. λ is often determined from idealised experiments where the CO_2 concentration is abruptly quadrupled, using the Gregory method (Gregory et al., 2004). Once λ is known, Eq. (1) can be rearranged to determine $F(t)$ from any experiment where the evolution of $T(t)$ and $N(t)$ are known (Forster et al., 2013), here referred to as the 1-λ forcing.

We use here the method described in detail in Fredriksen et al. (2021) to compute what we will refer to as the 3-λ forcing. This method attempts to correct the biases in the 1-λ forcing by using three different λ's instead. We assume the global temperature responds linearly to the forcing, and can be described as

$$T(t) = \sum_{n=1}^{K} T_n(t) = \sum_{n=1}^{K} c_n \exp(-t/\tau_n) * F(t) \tag{2}$$

where the $*$ denotes a convolution, and that N can be decomposed similarly, where different λ's are associated with each component of $T(t)$:

$$N(t) = \sum_{n=1}^{K} N_n(t) = F(t) + \sum_{n=1}^{K} \lambda_n T_n(t) \tag{3}$$

We determine the c_n's and λ_n's using abrupt-4xCO2 experiments as in Fredriksen et al. (2021), except that the method is further developed to use data points from all ensemble members if several are available for a model to better constrain the estimate. Additional members are averaged over when computing the parameters of the temperature response, and treated as extra data points when plotting T vs N to determine the λ_n's. As before, we use 150 years of data for estimation to treat all models equally. Many models have run the experiments for longer than that, and these extra years are included in the figures, allowing us to visually inspect how our fit performs at longer scales.

In Eqs. (2) and (3) we use $K = 4$, but the slowest response is assumed to be so slow, that it can be approximated as a constant heat flux $N_4 = b_4$ going into the deeper oceans without affecting the surface temperature during the first 150 years after quadrupling. Hence, we are in practice studying the 3-time scale responses $T(t) \approx \sum_{n=1}^{3} T_n(t)$ and $N(t) \approx b_4 + \sum_{n=1}^{3} N_n(t)$.

In Eq. (2), we note that we can move an arbitrary factor between the c_n and $F(t)$ without changing the temperature response, so different definitions of the forcing can in fact be used in a linear/impulse response model, as long as one is consistent about the definition when applying the model to different experiments. Here we strive to make a forcing definition that does not involve adjustments from surface temperature responses, and is consistent with the fixed-SST forcing estimates. When the parameters c_n and λ_n have been estimated from the abrupt-4xCO2 experiment, we have defined a separation of forcing and response, which we can use to compute $F(t)$ for other experiments by rearranging Eq. (3). Since this equation needs to know the components of $T(t)$, we need to iterate until convergence between (i) performing the convolutions in Eq. (2) to find the components and (ii) computing the forcing which is needed for the convolutions.

The method relies on a linear response model for predicting the temperature components, so a criterion for making good forcing estimates is that the linear model actually predicts the temperature well. In several figures we therefore include the difference between the temperature predicted by the linear model and the temperature output of the GCM. A difference close to zero is considered a necessary, but not sufficient criterion that we have a good forcing estimate. More importantly, we are interested in estimates that are consistent with the fixed-SST forcing (corrected for land temperature responses). Hence, whenever available, our forcing estimates are compared to fixed-SST estimates of the forcing.

As a thought experiment of why we think it is important to have a good curved fit to all points in a Gregory plot to make good forcing estimates for other experiments, we can consider the result of using 1-λ methods for estimating time-varying forcing for abrupt-4xCO2, and test how close this is to a constant. Assuming we have a typical Gregory plot where feedbacks become less negative with time and we make a regression for the first 150 years of data, the time-varying forcing $F(t) = N(t) - \lambda T(t)$ will have higher values in the beginning where the values of $N(t)$ are above the straight line. Similarly, if making a regression for the first 20 years, then the later time period will get stronger forcing estimates. So if these forcing estimation methods cannot reproduce the constant 4xCO_2 forcing, we would expect them to give biased time-varying forcings also for other experiments.

4 Results

For the 18 models where fixed-SST forcing is available for abrupt-4xCO_2, we find a generally good correspondence between our forcing estimates and the fixed-SST forcing (see Figure 1 and estimated parameters in Tables S6 and S7). Since the land temperatures have responded a little in the fixed-SST experiments (Andrews et al., 2021), we can expect these estimates to be comparable to our curve after a few months of response. Several methods exist for adjusting these forcing estimates to isolate the forcing at zero temperature response, and in Figure 1 we include the ERF_trop estimates from Smith et al. (2020). The light blue curves provides some insight into the uncertainties associated with our estimates, and we note that their spread varies substantially between models. We can expect eventual over- or under-estimations of the 4xCO_2 forcings here to follow the transient forcing estimates presented in other figures. Uncertainties

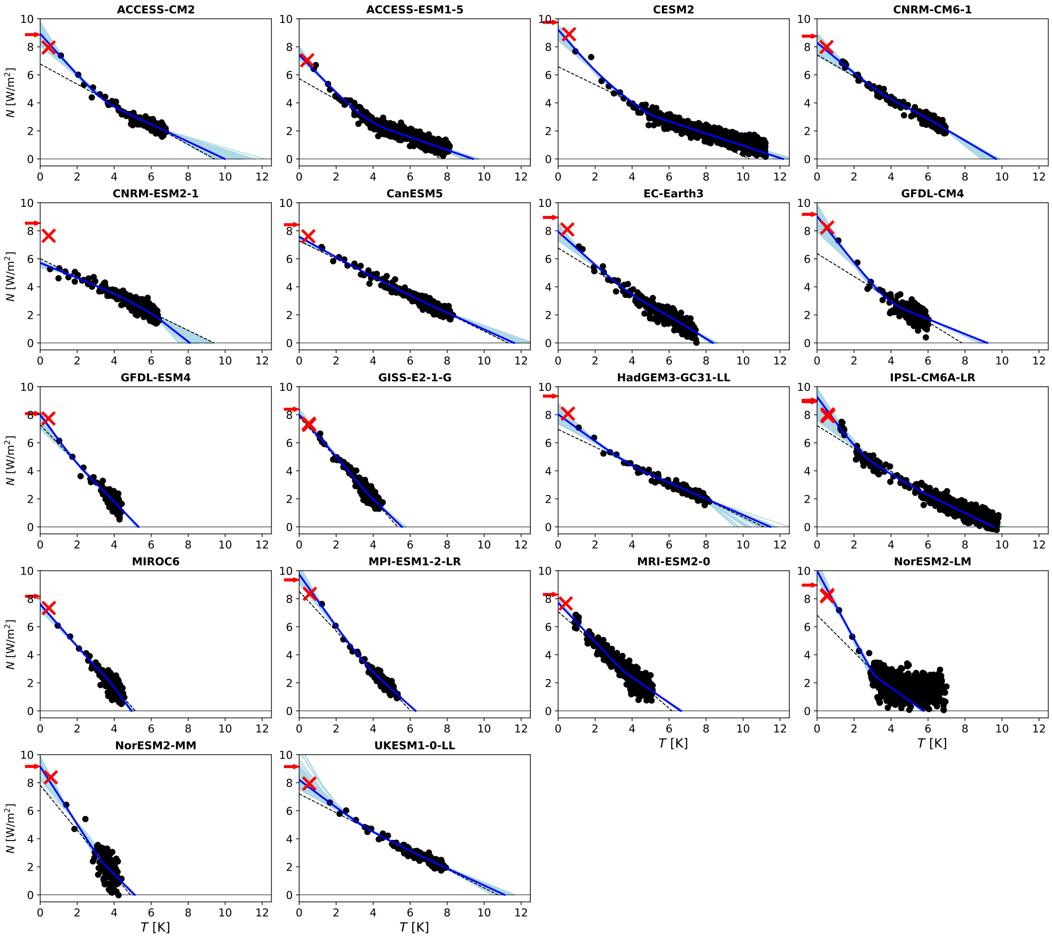

Figure 1. The top-of-the-atmosphere net radiation anomalies (N) versus the temperature anomalies (T) for the 18 models where we know the fixed-SST forcing (plotted as red crosses) for abrupt-4xCO2 simulations. The black dots are annual mean values, and all available members are included. The black dashed fit is the standard 150-year linear regression still used in the Sixth Assessment Report. The light blue curves are fits done to the first 150 years of the response with the 3-λ method for 1000 different random choices of time scales, and the dark blue curves show the best (least-squares) fit for the temperature response, as in Fredriksen et al. (2021). The red arrows show the ERF_trop forcing estimates from Smith et al. (2020).

in the forcing estimates are often larger if the model's surface temperature responds quickly, as there will be fewer points close to the y-axis to constrain the intercept. Similarly, if a model is still far from equilibrium after 150 years we can expect a larger spread in the estimated climate sensitivity (as derived from the intercept with the x-axis). Internal variability in T and N also plays a role in determining the uncertainty of the fit.

One exception from the good correspondence happens for the model CNRM-ESM2-1, because the model did not use an abrupt concentration change of $4\times CO_2$ throughout the whole atmosphere, instead adjusting the surface emissions of CO_2 to maintain $4\times CO_2$ in the lowest atmospheric layer and allowing CO_2 to percolate throughout the atmosphere, which takes around 15 years to reach a uniform $4\times CO_2$ concentration (Smith et al., 2020). This results in regression estimates being biased low compared to the fixed-SST estimates. These "effective" lower estimates actually work well in predicting temperature responses with our linear model, but since the forcing has been specified differently to other CMIP6 models we have not included this model in further analyses.

We show similar figures for the 33 models without fixed-SST forcing estimates in the supporting information (Figures S1 and S2), and for the 12 models with abrupt-2xCO2 and the 9 models with abrupt-0p5xCO2 experiments in Figures S3 and S4, respectively. Our curved fit through the points appears to be a generally better fit than straight lines, so we expect to find reasonable forcing estimates (i.e. relative to fixed-SST forcing estimates) also for these experiments. Forcing estimates for the abrupt-2xCO2 and abrupt-0p5xCO2 experiments are given in Table S8. We find that $4\times CO_2$ forcing is on average 2.11 times stronger than the $2\times CO_2$ forcing, and the absolute value of the $0.5\times CO_2$ forcing is a little weaker than the $2\times CO_2$ forcing for most models, consistent with a radiative forcing depending superlogarithmically on the CO_2 concentration (Etminan et al., 2016). However, the smaller signal-to-noise ratio makes these lower forcing estimates more uncertain. We note this in particular for the 0p5xCO_2 experiments, where a few estimates in the high end of the uncertainty range cause the mean absolute forcing to be slightly stronger than for $2\times CO_2$.

All abrupt CO_2 experiments could in principle have been used for estimating the 3 λ's, but many more models have published 4xCO_2 experiments than abrupt $2\times$ and $0.5\times$ CO_2 experiments. The higher signal-to-noise ratio of the abrupt-4xCO_2 experiments is also an advantage. However, the stronger the response, the stronger the effect of state-dependent feedbacks or other parameters may be, which violates our linear response assumption (Bloch-Johnson et al., 2021). Finding a way to tell how much of the feedback change is due to state-dependence (inconsistent with linear response) or due to pattern effect/time-scale dependence (can be consistent with linear response) will be important in future work.

The historical and SSP2-4.5 3-λ forcing is consistent with the fixed-SST forcing for most of the 10 models where this is available and always better than or as good as the 1-λ forcing estimates (Figure 2). Hence we expect the 3-λ ERF to be a good approximation also for the many models and experiments that lack fixed-SST forcing. The black curves in the right column show that land temperatures have not responded much in these fixed-SST experiments compared to abrupt-4xCO2 experiments, so these forcing estimates probably do not need to be corrected for land responses to the same degree. For models with little curvatures in Figure 1, the 1-λ forcing is as expected very similar to the 3-λ forcing. For IPSL-CM6A-LR the fixed-SST forcing falls in the middle of the 1-λ and 3-λ forcing, suggesting that both the 3-λ and 1-λ forcings are slightly biased in different directions. From Figure 1 we note that our 3-λ IPSL-CM6A-LR forcing estimate is in the higher end of a large uncertainty range.

In addition to the comparison with fixed-SST forcing, the ability to predict the GCM temperature also serves as a measure of how good the forcing estimate is, therefore we have included in the right column of Figure 2 the difference between the temperature predicted from our 3-λ forcing and linear response model and the output of the complex model. For positive differences, the forcing is probably overestimated, and vice versa. Temperature differences are typically within a ± 0.5°C interval, suggesting that our combination of forcing and linear response can generally well describe global mean temperatures. Our smaller temperature emulation error for some models compared to Jackson et al. (2022) (which can e.g. be up to 0.5°C for IPSL-CM6A-LR) is probably explained by our different forcing definition, in particular related to its correction for land temperature responses.

Figure 3 shows the multi-model mean 3-λ forcing from all available models for 7 different experiments (left column), and the corresponding global mean temperature difference between the linear responses and the GCMs (right column). The large ensembles of temperature differences show that temperature responses are on average slightly

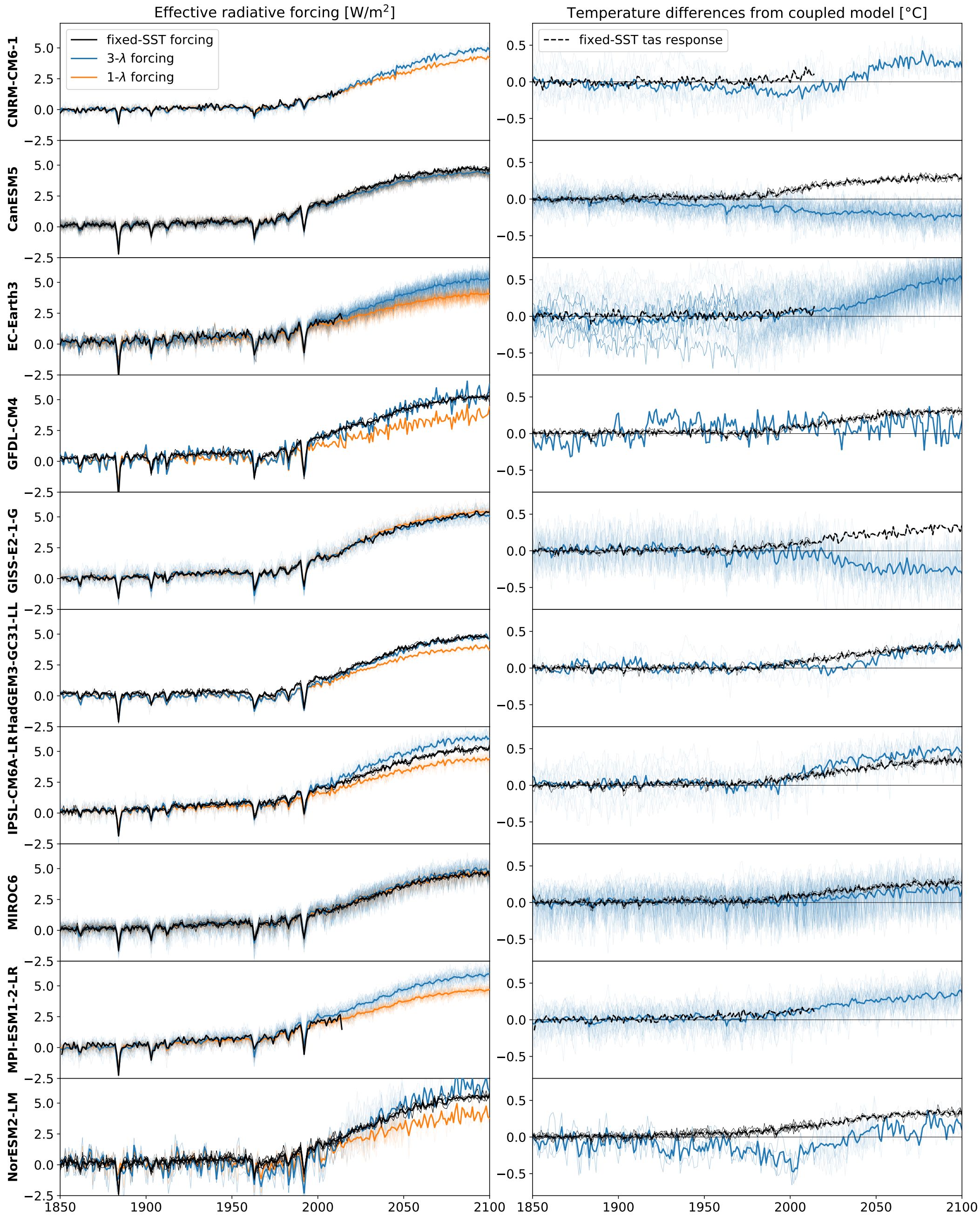

Figure 2. Left column: comparison of 1-λ (orange) and 3-λ (blue) forcing estimates to the fixed-SST (black) forcing estimates, for the models where transient fixed-SST estimates are available (10 models). Some models have run these experiments for the historical period, and some have extended it for the SSP2-4.5 scenario. Thin lines are estimates from single members, and thicker lines are ensemble means. Right column: the surface air temperature differences between the output of the coupled model and the estimated response to the 3-λ forcing (blue). In addition, we have included the change in surface air temperature from the fixed-SST runs, which should give the fixed-SST forcing a tiny negative bias when the temperature response grows (black).

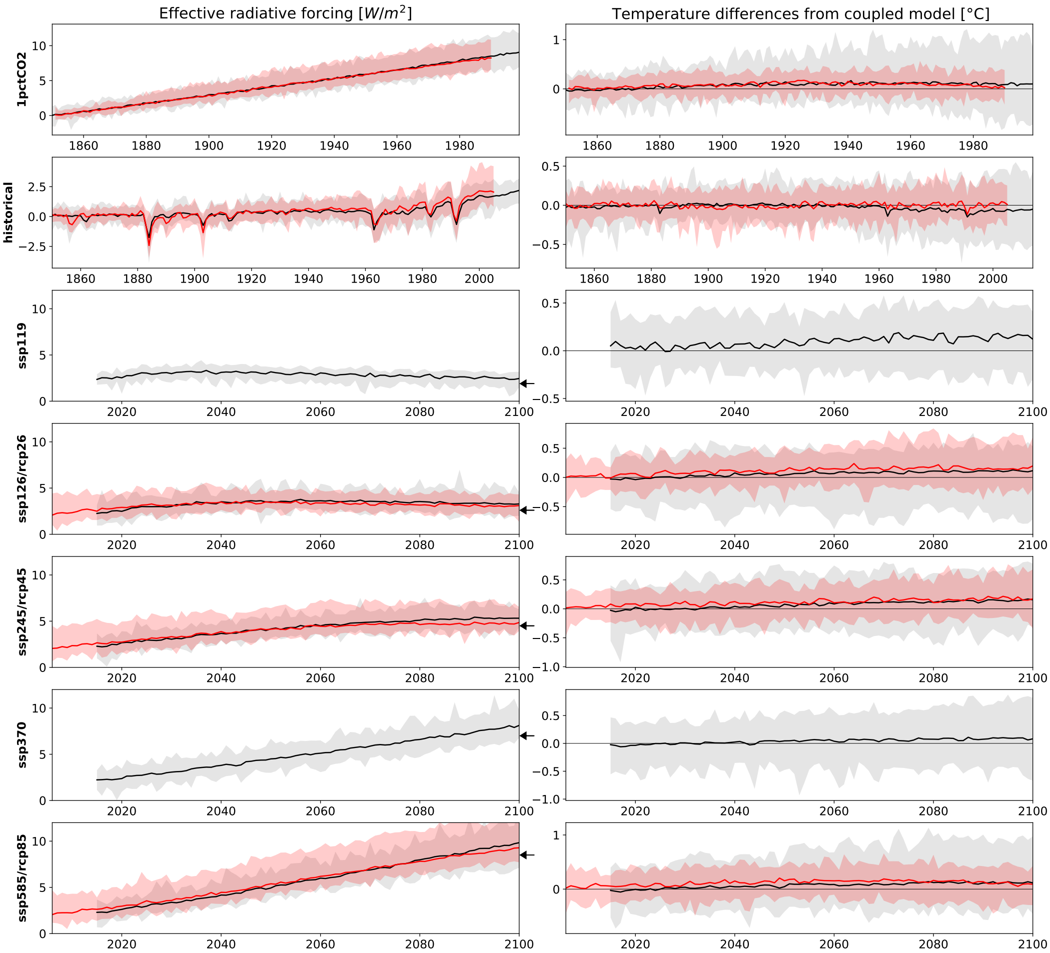

Figure 3. Left column: Effective radiative forcing estimates using the 3-λ method from up to 50 models from CMIP6 (gray/black) and up to 21 models from CMIP5 (red). The shading shows the min and max values of the member mean forcing for each year, and the red/black curves the mean of all models. Each row shows an experiment denoted by the y-axis label. The black arrows denote the forcing levels 1.9, 2.6, 4.5, 7.0 and 8.5 W m^{-2} in year 2100. Right: The difference between the temperature predicted by the linear temperature response and the global mean temperature from the coupled model.

overestimated for the 1pctCO2 and future scenario experiments, and hence the forcings in the left column are probably slightly overestimated too. We hypothesize this could be due to state-dependencies in the feedback parameters. We include also the CMIP5 estimates from Fredriksen et al. (2021) in this figure for comparison, and note that for the 1pctCO2 experiment (top row) the average forcing estimates are remarkably similar for CMIP5 and CMIP6.

For the last decades of the historical experiment, we find the CMIP6 forcing to be lower than the CMIP5 forcing, but maybe slightly underestimated, as evidenced by the comparison between our estimated temperatures and those from the GCMs. For the future comparable scenarios however (SSP1-2.6, SSP2-4.5 and SSP5-8.5 for CMIP6; RCP2.6, RCP4.5 and RCP8.5 for CMIP5), the CMIP6 forcing grows more than the CMIP5 forcing, and ends up at higher values than CMIP5 at the end of the 21st century with no clear difference in the bias in temperatures compared to CMIP5 models. This suggests that CMIP6 ERF is higher than equivalent nominal scenarios in CMIP5. The multi-model mean difference in year 2100 is 0.18, 0.46, and 0.55 Wm^{-2} for RCP2.6/SSP1-2.6, RCP4.5/SSP2-4.5, RCP8.5/SSP5-8.5, respectively.

A closer look at the historical period (Figure 4) shows that our 3-λ total forcing for around 1995 onwards is a little stronger than the 1-λ forcing, as used in Smith and Forster (2021). Studying the components separately, we find that the greenhouse gas forcing becomes more positive, and the aerosol forcing becomes more negative when using the 3-λ method. In general, the different forcing definitions give more different results the stronger the temperature response is.

The small underestimation of the CMIP6 linear temperature responses for the historical period seems to stem mainly from the response to aerosol forcing in some of the models (not shown). One reason for this could be a similar (possible state-dependence) effect for this negative forcing as we have with the small overestimation for the positive forcing. Or it could be that the more spatially inhomogeneous aerosol forcing triggers more localised responses (for instance land temperatures may respond much faster and stronger than ocean temperatures), resulting in a global mean response to aerosols that differ from the global mean response to CO_2. This may lead to small errors in the linear response assumption, which could cause small errors in the global mean forcing es-

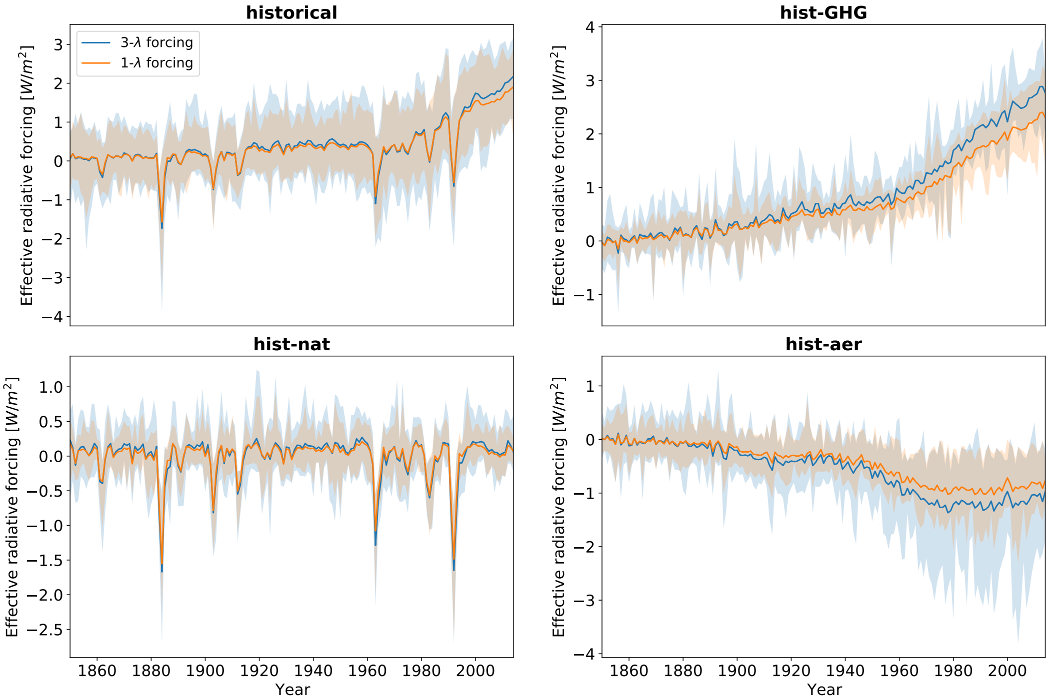

Figure 4. The effective radiative forcing for the historical, hist-GHG, hist-nat and hist-aer experiments, computed using both the 1-λ (orange) and the 3-λ method (blue). The shading shows the min and max values of the model ensemble, and the solid curves the model means.

timates too. Hence we cannot rule out that efficacy factors (Hansen et al., 2005) different from 1 could be needed for this forcing.

5 Discussion

Future temperature projections from CMIP6 show stronger warming than the corresponding projections from CMIP5 (Tebaldi et al., 2021). The CMIP5 RCP scenarios have a different composition of greenhouse gases, aerosols and other forcers than the CMIP6 SSP scenarios, but they are designed such that they should reach approximately the same forcing levels by the end of the 21st century (Gidden et al., 2019). However, Wyser et al. (2020) shows that at least half of the temperature increase from CMIP5 to CMIP6 for the model EC-Earth3-Veg is due to the increase in the ERF, and in particular the greenhouse gas concentrations. Chapter 4 of the IPCC's Sixth Assessment Report Working Group 1 (Lee et al., 2021) shows that ERF is substantially higher for CMIP6 SSPs that are nominally the same forcing as CMIP5 RCPs (SSP1-2.6 versus RCP2.6 for example), and comes to a similar conclusion, namely that the increase in forcing contributes to about half of the temperature increase in CMIP6 models compared to CMIP5 models with the other half attributed to the increase in climate sensitivity.

Our results confirm that the ERF is indeed increasing more for CMIP6 than for CMIP5 during the 21st century. From the 1pctCO2 experiment we note that given the same increase in CO_2 concentrations, the ERF is similar in CMIP5 and CMIP6 models, so the radiative efficiency of CO_2 has not changed between the model generations. This suggests that the higher year-2100 CO_2 concentrations in CMIP6 compared to similar CMIP5 scenarios (Meinshausen et al., 2020) explains the increase in ERF.

Despite the higher temperature increase in the future scenarios, the historical temperature increase is less in CMIP6 than in CMIP5 (Flynn & Mauritsen, 2020; Smith & Forster, 2021). Smith and Forster (2021) explain this as a combination of stronger feedbacks and lower historical forcing, from both aerosols and greenhouse gases (GHGs). Our new 3-λ historical forcing estimates diverge from the 1-λ forcing used in Smith and Forster (2021) only after 1995, and is hence not changing their conclusions. However, our finding that similar CO_2 concentrations lead to similar forcing, should also imply that the similar global annual mean historical CO_2 concentrations (Meinshausen et al., 2017) lead to similar CO_2 forcing in CMIP5 and CMIP6, though differences in the latitudinal and seasonal variations in concentrations or the effective forcing from other greenhouse gases may also affect the resulting GHG forcing. The similar CO_2 forcing may suggest that the stronger aerosol forcing plays a larger role in explaining the lower historical forcing in CMIP6.

The higher temperature increase for CMIP6 during the 21st century is also partly explained by the increase in climate sensitivity, as evidenced by less negative global feedbacks, increased equilibrium climate sensitivity (ECS) and transient climate response (TCR) (Flynn & Mauritsen, 2020; Zelinka et al., 2020). The climate sensitivity does not tell the full story of how much warming we can expect at all times, and an increased sensitivity may not necessarily contribute to increased temperature responses during the historical period. In addition to forcing differences, the pattern of warming also matters, as some regions more effectively radiate out excess energy than others, hence modulating the effective global climate sensitivity with time (the *pattern effect*). During the late 20th century, the warming pattern in the tropical Equatorial Pacific has led to lower estimates of the effective climate sensitivity, (e.g. Zhou et al., 2016; Andrews et al., 2018, 2022). This pattern is not expected to persist in the future, likely causing the effective climate sensitivity to increase in the near future.

Normalizing the abrupt-4xCO2 responses by our estimated forcing yields also a measure of the model sensitivity per unit forcing, which may be helpful for understanding if temperature responses are stronger because of a high climate sensitivity or a higher forcing. We find that CMIP6 models have an average response slightly stronger than CMIP5 models, and a larger model spread (Figure S5). Hence our results confirm that CMIP6 models are overall more sensitive, but the relative role of the climate sensitivity for explaining higher temperature responses is highly model dependent.

6 Open research

The original CMIP6 datasets are available at `https://esgf-node.llnl.gov/projects/cmip6/`. Processed data and code will be permanently stored in zenodo (link will be made and inserted here when paper is accepted. In the meantime, data and code can be accessed through github: `https://github.com/Hegebf/CMIP6-forcing`, which was made public just before submission)

Acknowledgments

We thank the participants in the Tri-MIPathlon-3 discussion. C.S. was supported by a NERC-IIASA collaborative research fellowship (NE/T009381/1). A.M. received funding from the European Research Council grant 770765. M.R. was further funded by the National Aeronautics and Space Administration under Grant No. 80NSSC21K1042. 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.

References

Andrews, T., Bodas-Salcedo, A., Gregory, J. M., Dong, Y., Armour, K. C., Paynter, D., ... Liu, C. (2022). On the effect of historical SST patterns on radiative feedback. *Journal of Geophysical Research: Atmospheres*, e2022JD036675. https://doi.org/10.1029/2022JD036675

Andrews, T., Gregory, J. M., Paynter, D., Silvers, L. G., Zhou, C., Mauritsen, T., ... Titchner, H. (2018). Accounting for Changing Temperature Patterns Increases Historical Estimates of Climate Sensitivity. *Geophysical Research Letters*, *45*(16), 8490-8499. https://doi.org/10.1029/2018GL078887

Andrews, T., Smith, C. J., Myhre, G., Forster, P. M., Chadwick, R., & Ackerley, D. (2021). Effective Radiative Forcing in a GCM With Fixed Surface Temperatures. *Journal of Geophysical Research: Atmospheres*, *126*, e2020JD033880. https://doi.org/10.1029/2020JD033880

Armour, K. C. (2017). Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks. *Nature Climate Change*, *7*, 331 – 335. https://doi.org/10.1038/nclimate3278

Bloch-Johnson, J., Rugenstein, M., Stolpe, M. B., Rohrschneider, T., Zheng, Y., & Gregory, J. M. (2021). Climate Sensitivity Increases Under Higher CO_2 Levels Due to Feedback Temperature Dependence. *Geophysical Research Letters*, *48*, e2020GL089074. https://doi.org/10.1029/2020GL089074

Etminan, M., Myhre, G., Highwood, E. J., & Shine, K. P. (2016). Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing. *Geophysical Research Letters*, *43*(24), 12,614–12,623. https://doi.org/10.1002/2016GL071930

Flynn, C. M., & Mauritsen, T. (2020). On the climate sensitivity and historical warming evolution in recent coupled model ensembles. *Atmospheric Chemistry and Physics*, *20*(13), 7829–7842. https://doi.org/10.5194/acp-20-7829-2020

Forster, P. M., Andrews, T., Good, P., Gregory, J. M., Jackson, L. S., & Zelinka, M. (2013). Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *Journal of Geophysical Research*, *118*, 1139–1150. https://doi.org/10.1002/jgrd.50174

Forster, P. M., Richardson, T., Maycock, A. C., Smith, C. J., Samset, B. H., Myhre, G., ... Schulz, M. (2016). Recommendations for diagnosing effective radiative forcing from climate models for CMIP6. *Journal of Geophysical Research: Atmospheres*, *121*(20), 12,460–12,475. https://doi.org/10.1002/2016JD025320

Fredriksen, H.-B., Rugenstein, M., & Graversen, R. (2021). Estimating Radiative Forcing With a Nonconstant Feedback Parameter and Linear Response. *Journal of Geophysical Research: Atmospheres*, *126*(24), e2020JD034145. https://doi.org/10.1029/2020JD034145

Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., ... Takahashi, K. (2019). Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century. *Geoscientific Model Development*, *12*(4), 1443–1475. https://doi.org/10.5194/gmd-12-1443-2019

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., ... Williams, K. D. (2004). A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, *31*, L03205. https://doi.org/10.1029/2003GL018747

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., ... Zhang, S. (2005). Efficacy of climate forcings. *Journal of Geophysical Research: Atmospheres*, *110*(D18). https://doi.org/10.1029/2005JD005776

Jackson, L. S., Maycock, A. C., Andrews, T., Fredriksen, H.-B., Smith, C. J., & Forster, P. M. (2022). Errors in Simple Climate Model Emulations of Past and Future Global Temperature Change. *Geophysical Research Letters*, *49*(15), e2022GL098808. https://doi.org/10.1029/2022GL098808

Lee, J.-Y., Marotzke, J., Bala, G., Cao, L., Corti, S., Dunne, J., ... Zhou, T. (2021). *Future Global Climate: Scenario-Based Projections and NearTerm Information. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)].* (Tech. Rep.). https://doi.org/10.1017/9781009157896.006

Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., ... Wang, R. H. J. (2020). The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. *Geoscientific Model Development*, *13*(8), 3571–3605. https://doi.org/10.5194/gmd-13-3571-2020

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., ... Weiss, R. (2017). Historical greenhouse gas concentrations for climate modelling (CMIP6). *Geoscientific Model Development*, *10*(5), 2057–2116. https://doi.org/10.5194/gmd-10-2057-2017

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., Van Vuuren, D. P., ... Wilbanks, T. J. (2010). The next generation of scenarios for climate change research and assessment. *Nature*, *463*(7282), 747–756. https://doi.org/10.1038/nature08823

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., ... Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

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

Rugenstein, M., Bloch-Johnson, J., Gregory, J., Andrews, T., Mauritsen, T., Li, C., ... Knutti, R. (2020). Equilibrium Climate Sensitivity Estimated by Equilibrating Climate Models. *Geophysical Research Letters*, *47*(4), e2019GL083898. https://doi.org/10.1029/2019GL083898

Senior, C. A., & Mitchell, J. F. B. (2000). The time-dependence of climate sensitivity. *Geophysical Research Letters*, *27*(17), 2685-2688. https://doi.org/10.1029/2000GL011373

Smith, C. J., & Forster, P. M. (2021). Suppressed late-20th Century warming in CMIP6 models explained by forcing and feedbacks. *Geophysical Research Letters*, *48*, e2021GL094948. https://doi.org/10.1029/2021GL094948

Smith, C. J., Kramer, R. J., Myhre, G., Alterskjær, K., Collins, W., Sima, A., … Forster, P. M. (2020). Effective radiative forcing and adjustments in CMIP6 models. *Atmospheric Chemistry and Physics*, *20*(16), 9591–9618. https://doi.org/10.5194/acp-20-9591-2020

Tebaldi, C., Debeire, K., Eyring, V., Fischer, E., Fyfe, J., Friedlingstein, P., … Ziehn, T. (2021). Climate model projections from the Scenario Model Intercomparison Project (ScenarioMIP) of CMIP6. *Earth System Dynamics*, *12*(1), 253–293. https://doi.org/10.5194/esd-12-253-2021

Winton, M., Takahashi, K., & Held, I. M. (2010). Importance of Ocean Heat Uptake Efficacy to Transient Climate Change. *Journal of Climate*, *23*(9), 2333-2344. https://doi.org/10.1175/2009JCLI3139.1

Wyser, K., Kjellström, E., Koenigk, T., Martins, H., & Döscher, R. (2020). Warmer climate projections in EC-Earth3-Veg: the role of changes in the greenhouse gas concentrations from CMIP5 to CMIP6. *Environmental Research Letters*, *15*(5), 054020. https://doi.org/10.1088/1748-9326/ab81c2

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., … Taylor, K. E. (2020). Causes of Higher Climate Sensitivity in CMIP6 Models. *Geophysical Research Letters*, *47*(1), e2019GL085782. https://doi.org/10.1029/2019GL085782

Zhou, C., Zelinka, M. D., & Klein, S. A. (2016). Impact of decadal cloud variations on the Earth's energy budget. *Nature Geoscience*, *9*(12), 871–874. https://doi.org/10.1038/ngeo2828

Supporting Information for "21st century scenario forcing increases more for CMIP6 than CMIP5 models"

Hege-Beate Fredriksen[1], Christopher J. Smith[2,3], Angshuman Modak[4], and

Maria Rugenstein[5]

[1]Department of Physics and Technology, UiT the Arctic University of Norway, Tromsø, Norway

[2]Priestley International Center for Climate, University of Leeds, Leeds, UK

[3]International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

[4]Department of Meteorology (MISU), Stockholm University, Stockholm, Sweden

[5]Colorado State University, Fort Collins, USA

Contents of this file:

1. Figures S1 to S5
2. Table S1-S8

Introduction

We present tables with an overview of the data included in this study, and estimated parameters from the abrupt $4xCO_2$ experiments, in addition to forcing estimates from $2xCO_2$ and $0p5xCO_2$ experiments.

We show also results from abrupt-4xCO_2 experiments similar to those in Figure 1 in the main manuscript, but where no fixed-SST estimates are available. And similar plots for abrupt-2xCO_2 and abrupt-0p5xCO_2 experiments.

We note also that for many models the branch information found in the metadata contains errors. We have done our best in trying to correct obvious mistakes, but cannot rule out more branch time errors, which could have small impacts on the computed anomalies.

	abrupt-4xCO2	abrupt-2xCO2	abrupt-0p5xCO2
ACCESS-CM2	1	-	-
ACCESS-ESM1-5	2	-	-
AWI-CM-1-1-MR	1	-	-
BCC-CSM2-MR	1	-	-
BCC-ESM1	1	-	-
CAMS-CSM1-0	2	-	-
CESM2	1	1	1
CESM2-FV2	1	-	-
CESM2-WACCM	1	-	-
CESM2-WACCM-FV2	1	-	-
CMCC-CM2-SR5	1	-	-
CMCC-ESM2	1	-	-
CNRM-CM6-1	6	1	1
CNRM-CM6-1-HR	1	-	-
CNRM-ESM2-1	3	-	-
CanESM5	2	1	1
E3SM-1-0	1	-	-
EC-Earth3	2	-	-
EC-Earth3-AerChem	1	-	-
EC-Earth3-CC	1	-	-
EC-Earth3-Veg	1	-	-
FGOALS-f3-L	3	-	-
FGOALS-g3	1	-	-

Table S1. Number of abrupt-4xCO2, abrupt-2xCO2 and abrupt-0p5xCO2 members used in this study, part I.

	abrupt-4xCO2	abrupt-2xCO2	abrupt-0p5xCO2
GFDL-CM4	1	-	-
GFDL-ESM4	1	-	-
GISS-E2-1-G	4	4	1
GISS-E2-1-H	3	2	-
GISS-E2-2-G	1	1	-
GISS-E2-2-H	1	1	-
HadGEM3-GC31-LL	1	1	1
HadGEM3-GC31-MM	1	-	-
ICON-ESM-LR	1	-	-
IITM-ESM	1	-	-
INM-CM4-8	1	-	-
INM-CM5-0	1	-	-
IPSL-CM5A2-INCA	1	-	-
IPSL-CM6A-LR	12	1	1
KIOST-ESM	1	-	-
MIROC-ES2L	1	-	-
MIROC6	1	1	1
MPI-ESM-1-2-HAM	1	-	-
MPI-ESM1-2-HR	1	-	-
MPI-ESM1-2-LR	1	-	-
MRI-ESM2-0	14	1	1
NESM3	1	-	-
NorCPM1	1	-	-
NorESM2-LM	1	-	-
NorESM2-MM	1	-	-
SAM0-UNICON	1	-	-
TaiESM1	1	1	1
UKESM1-0-LL	1	-	-
Number of models	51	12	9

Table S2. Number of abrupt-4xCO2, abrupt-2xCO2 and abrupt-0p5xCO2 members used in this study, part II.

Table S3. Number of members where we have computed transient forcing time series, part I.

	1pctCO2	historical	hist-nat	hist-GHG	hist-aer	ssp119	ssp126	ssp245	ssp370	ssp585
ACCESS-CM2	1	5	3	3	3	-	5	5	5	5
ACCESS-ESM1-5	1	40	3	3	3	-	40	40	40	40
AWI-CM-1-1-MR	1	5	-	-	-	-	1	1	5	1
BCC-CSM2-MR	1	3	3	3	3	-	1	1	1	1
BCC-ESM1	1	3	-	-	-	-	-	-	3	-
CAMS-CSM1-0	2	3	-	-	-	2	2	2	2	2
CESM2	1	11	3	3	2	-	3	3	3	3
CESM2-FV2	1	3	-	-	-	-	-	-	-	-
CESM2-WACCM	1	3	-	-	-	-	1	5	3	5
CESM2-WACCM-FV2	1	3	-	-	-	-	-	-	-	-
CMCC-CM2-SR5	1	7	-	-	-	-	1	1	1	1
CMCC-ESM2	1	1	-	-	-	-	1	1	1	1
CNRM-CM6-1	1	29	10	10	10	-	6	10	6	6
CNRM-CM6-1-HR	1	1	-	-	-	-	1	1	1	1
CNRM-ESM2-1	10	11	-	-	-	5	5	10	5	5
CanESM5	6	65	50	50	30	50	50	50	50	50
E3SM-1-0	1	5	-	-	-	-	-	-	-	-
EC-Earth3	1	72	-	-	-	51	57	71	57	57
EC-Earth3-AerChem	1	2	-	-	-	-	-	-	2	-
EC-Earth3-CC	1	1	-	-	-	-	-	1	-	1
EC-Earth3-Veg	1	9	-	-	-	3	7	8	6	8
FGOALS-f3-L	3	3	-	-	-	-	1	1	1	1
FGOALS-g3	3	6	3	3	3	1	4	4	5	4

Table S4. Number of members where we have computed transient forcing time series, part II.

	1pctCO2	historical	hist-nat	hist-GHG	hist-aer	ssp119	ssp126	ssp245	ssp370	ssp585
GFDL-CM4	1	1	3	-	-	-	-	1	-	1
GFDL-ESM4	1	1	3	1	1	1	1	1	1	1
GISS-E2-1-G	5	46	20	10	15	7	12	31	27	11
GISS-E2-1-H	1	25	-	-	-	2	5	10	6	5
GISS-E2-2-G	1	-	-	-	-	-	-	-	-	-
GISS-E2-2-H	1	5	-	-	-	-	-	-	-	-
HadGEM3-GC31-LL	4	5	10	5	5	-	1	5	-	4
HadGEM3-GC31-MM	1	4	-	-	-	-	1	-	-	4
ICON-ESM-LR	1	5	-	-	-	-	-	-	-	-
IITM-ESM	1	1	-	-	-	-	1	1	1	1
INM-CM4-8	1	1	-	-	-	-	1	1	1	1
INM-CM5-0	1	10	-	-	-	-	1	1	5	1
IPSL-CM5A2-INCA	1	1	-	-	-	-	1	-	1	-
IPSL-CM6A-LR	1	33	10	10	10	6	6	11	11	7
KIOST-ESM	1	1	-	-	-	-	1	1	-	1
MIROC-ES2L	1	26	-	-	-	10	10	25	10	10
MIROC6	1	50	50	3	10	1	50	50	3	50
MPI-ESM-1-2-HAM	1	3	-	-	-	-	-	-	3	-
MPI-ESM1-2-HR	1	10	-	-	-	-	2	2	10	2
MPI-ESM1-2-LR	1	30	-	-	-	30	30	30	30	30
MRI-ESM2-0	2	12	5	5	5	5	5	10	5	6
NESM3	1	5	-	-	-	-	2	2	-	2
NorCPM1	1	30	-	-	-	-	-	-	-	-
NorESM2-LM	1	3	3	3	3	-	1	13	3	1
NorESM2-MM	1	3	-	-	-	-	1	2	1	1
SAM0-UNICON	1	1	-	-	-	-	-	-	-	-
TaiESM1	1	2	-	-	-	-	1	1	1	1
UKESM1-0-LL	4	19	-	-	-	5	16	17	16	5
Number of models	51	50	15	14	14	15	38	38	37	39

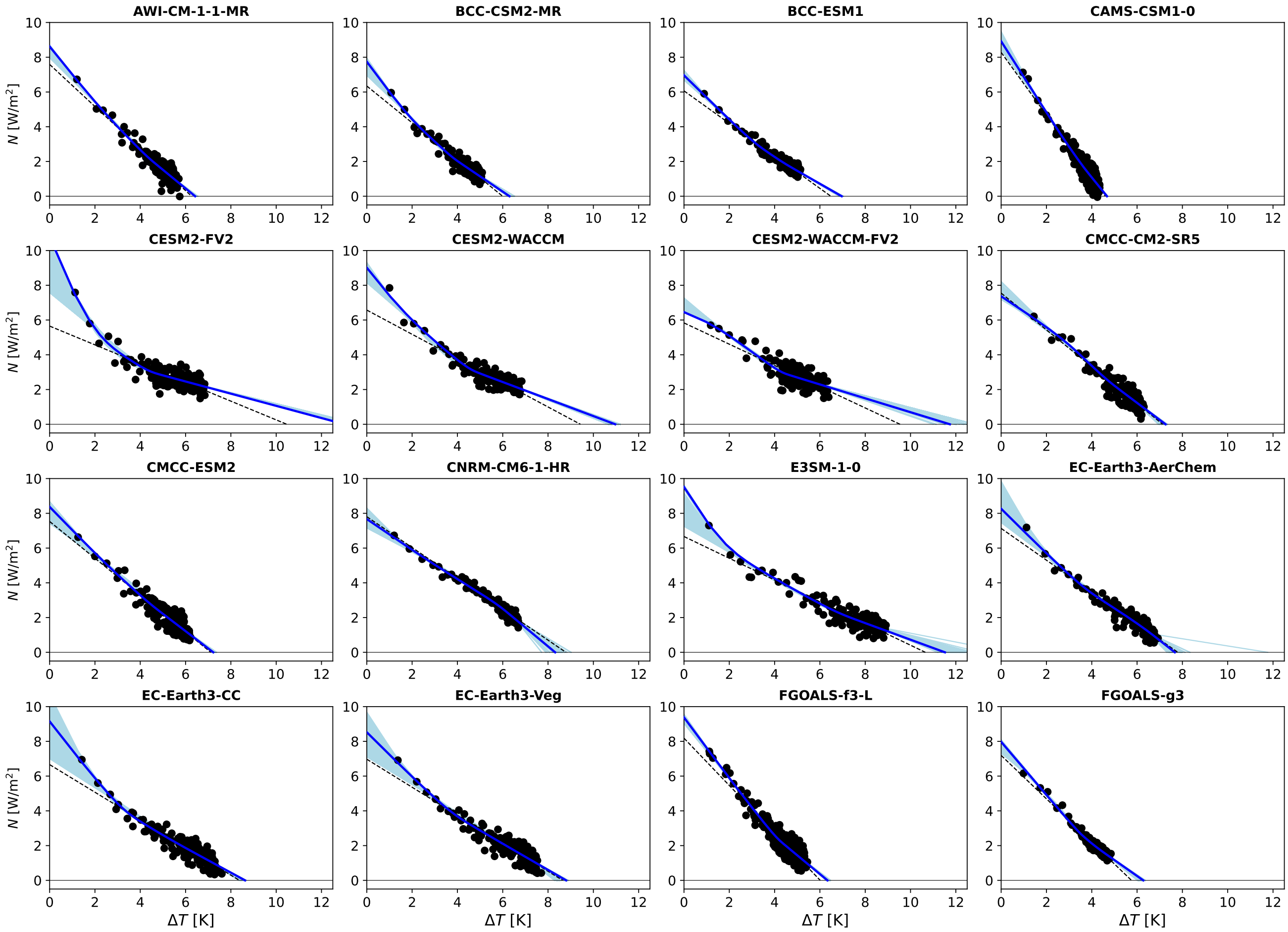

Figure S1. As Figure 1, but for models without fixed-SST forcing. Part I.

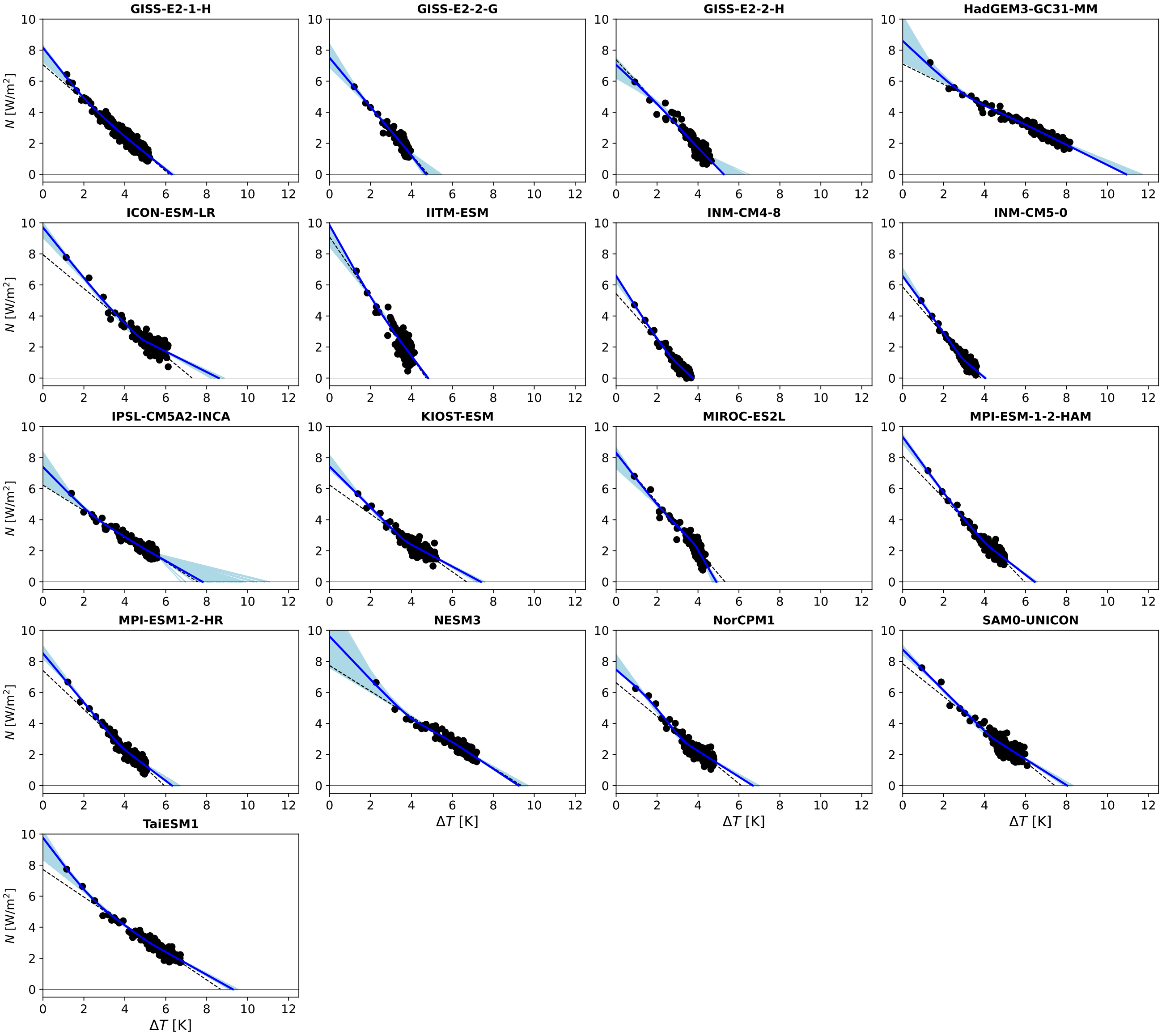

Figure S2. As Figure 1, but for models without fixed-SST forcing. Part II.

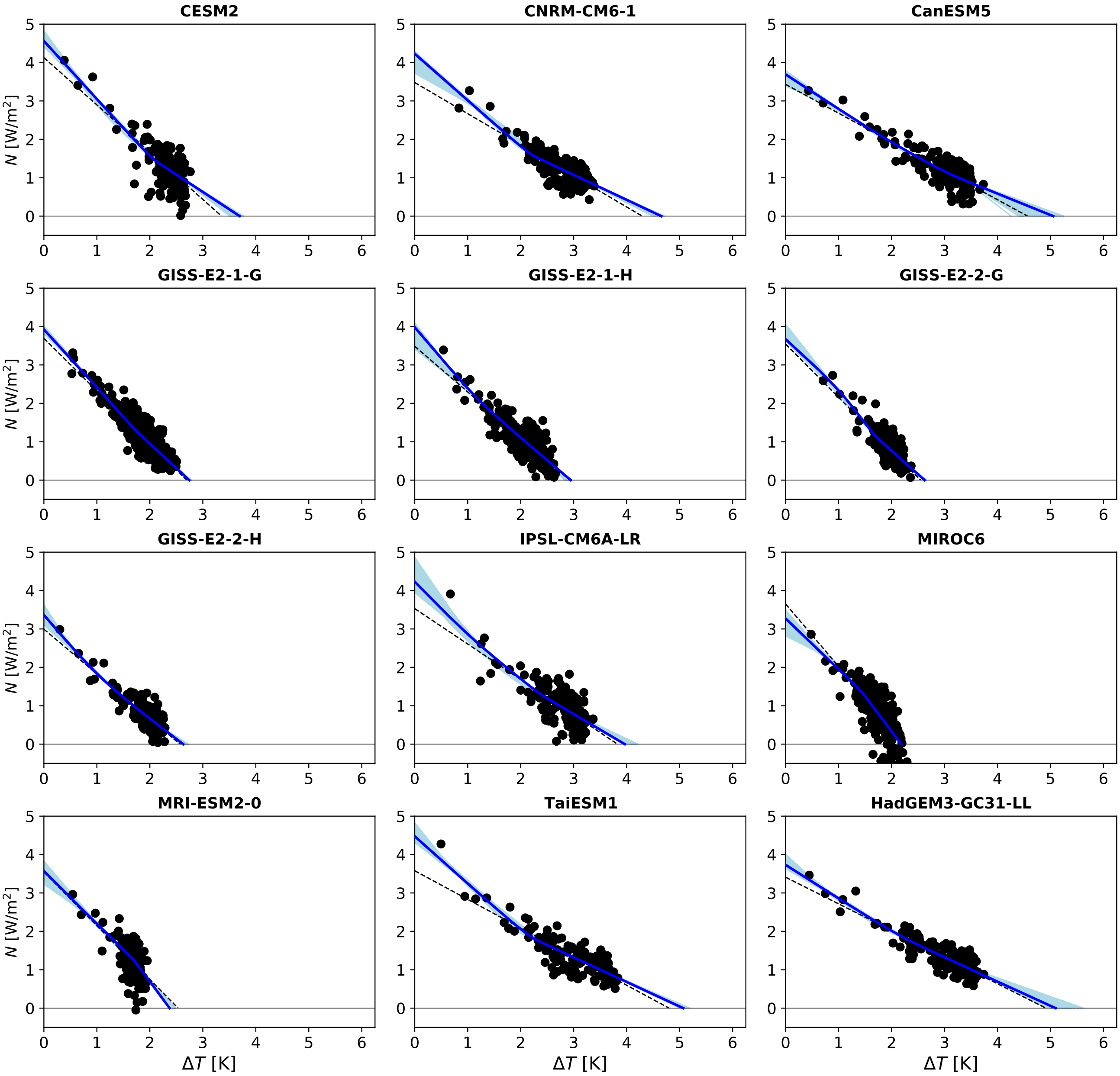

Figure S3. Similar figure as above, but using the abrupt-2xCO2 experiment.

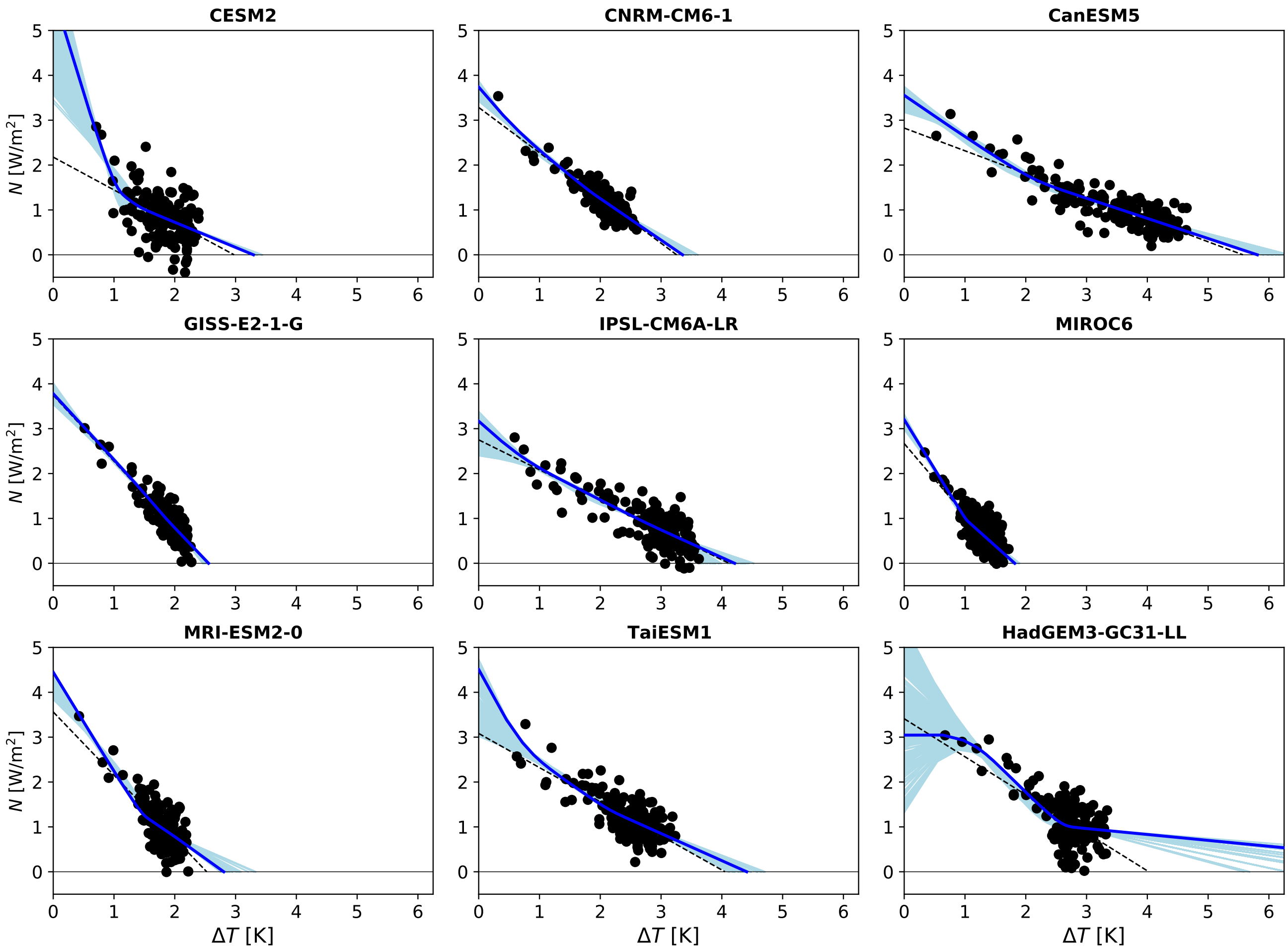

Figure S4. Similar figure as above, but using the abrupt-0p5xCO2 experiment. This experiment has negative responses, but the signs are flipped when performing the estimation and plotting.

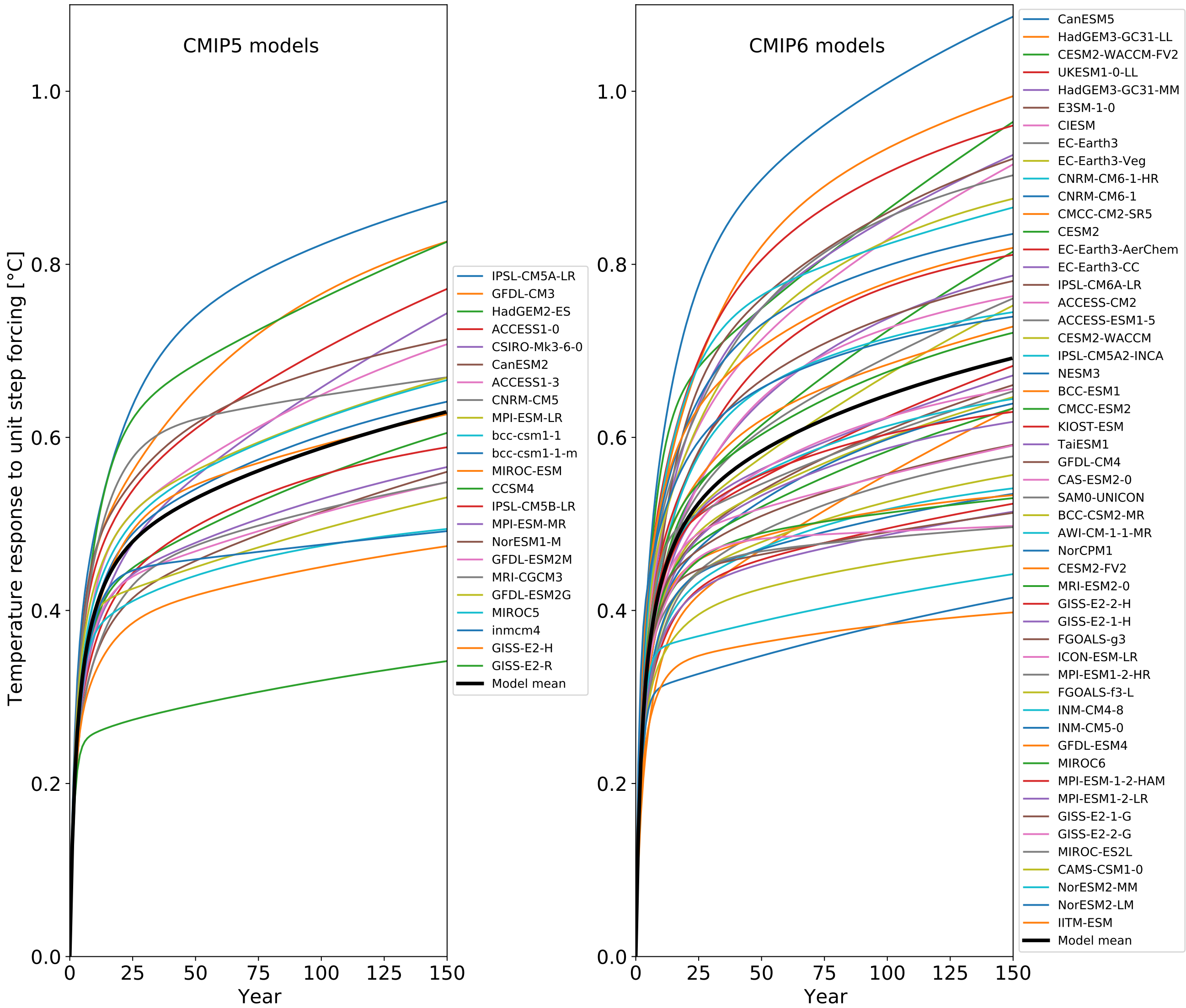

Figure S5. Temperature responses to unit-step forcing. Legend is sorted by the response in year 150. If the temperature responses are strong (weak) it may be because they compensate for weak (strong) forcing estimates.

	piClim-4xCO2	piClim-histall
ACCESS-CM2	1	-
ACCESS-ESM1-5	1	-
CESM2	1	-
CNRM-CM6-1	1	1
CNRM-ESM2-1	1	-
CanESM5	1	3
EC-Earth3	1	1
GFDL-CM4	1	3
GFDL-ESM4	1	-
GISS-E2-1-G	2	3
HadGEM3-GC31-LL	1	3
IPSL-CM6A-LR	5	3
MIROC6	1	3
MPI-ESM1-2-LR	1	1
MRI-ESM2-0	1	-
NorESM2-LM	2	6
NorESM2-MM	1	-
UKESM1-0-LL	1	-
Number of models	18	10

Table S5. Number of members used in this study to compute fixed-SST forcing.

	τ_1	τ_2	τ_3	$-\lambda_1$	$-\lambda_2$	$-\lambda_3$	b_4	$F4x$	$T4x$
ACCESS-CM2	2.57	25.28	139.03	1.58	0.60	0.62	1.63	8.93	9.98
ACCESS-ESM1-5	2.63	14.26	324.14	1.41	1.02	0.46	0.37	7.42	9.40
AWI-CM-1-1-MR	1.17	5.99	87.70	1.72	1.38	1.07	0.58	8.62	6.42
BCC-CSM2-MR	1.11	7.18	184.44	1.93	1.25	0.88	0.11	7.74	6.30
BCC-ESM1	2.72	19.25	337.09	1.32	0.94	0.75	0.02	6.96	6.96
CAMS-CSM1-0	1.87	10.32	131.09	2.07	1.91	1.61	0.09	8.92	4.67
CAS-ESM2-0	1.43	8.93	81.05	2.06	1.07	0.85	0.87	8.81	7.11
CESM2	1.38	6.93	291.55	1.80	1.11	0.44	0.06	9.22	12.15
CESM2-FV2	1.15	6.14	381.16	4.43	0.93	0.35	0.12	10.78	13.02
CESM2-WACCM	1.13	5.75	287.90	2.11	1.17	0.49	0.20	9.01	10.97
CESM2-WACCM-FV2	1.03	5.54	427.24	-0.00	1.00	0.40	0.10	6.45	11.74
CIESM	2.80	16.64	257.81	1.05	0.91	0.51	0.34	9.12	12.63
CMCC-CM2-SR5	1.40	8.28	80.47	0.79	1.33	0.98	0.90	7.37	7.26
CMCC-ESM2	1.25	6.40	105.71	1.39	1.21	1.00	0.59	8.37	7.22
CNRM-CM6-1	1.78	14.26	96.73	1.15	0.66	0.78	1.88	8.29	9.67
CNRM-CM6-1-HR	1.76	12.73	274.71	0.92	0.79	1.09	0.06	7.67	8.31
CNRM-ESM2-1	4.14	40.76	93.08	0.52	0.65	1.04	1.77	5.71	8.09
CanESM5	2.19	13.37	179.44	0.74	0.68	0.60	1.01	7.57	11.62
E3SM-1-0	1.06	9.80	114.58	2.86	0.79	0.48	0.83	9.51	11.52
EC-Earth3	2.66	26.33	107.27	1.24	0.72	0.81	0.52	7.94	8.38
EC-Earth3-AerChem	2.84	28.13	103.24	1.34	0.78	1.03	0.59	8.26	7.67
EC-Earth3-CC	1.71	12.46	80.03	1.78	0.85	0.71	0.67	9.15	8.63
EC-Earth3-Veg	2.27	20.38	88.11	1.35	0.77	0.75	0.63	8.52	8.80
FGOALS-f3-L	1.15	6.13	122.70	1.78	1.71	1.08	0.61	9.38	6.33
FGOALS-g3	2.82	15.44	154.88	1.51	1.50	0.97	0.81	7.92	6.17

Table S6. Parameters estimated from abrupt-4xCO2 experiments, part I.

	τ_1	τ_2	τ_3	$-\lambda_1$	$-\lambda_2$	$-\lambda_3$	b_4	$F4x$	$T4x$
GFDL-CM4	1.96	10.47	175.61	1.72	1.15	0.53	1.02	9.03	9.20
GFDL-ESM4	1.14	5.75	82.88	2.05	1.30	1.36	1.26	8.01	5.32
GISS-E2-1-G	1.16	5.95	341.47	1.45	1.64	1.27	0.44	8.01	5.55
GISS-E2-1-H	1.04	7.26	84.97	1.79	1.23	1.06	1.04	8.16	6.29
GISS-E2-2-G	1.42	10.42	431.60	1.60	1.53	1.68	1.13	7.51	4.72
GISS-E2-2-H	1.33	7.51	83.07	1.16	1.47	1.36	0.85	7.06	5.26
HadGEM3-GC31-LL	2.97	26.13	357.17	1.03	0.57	0.58	0.50	8.03	11.47
HadGEM3-GC31-MM	1.53	15.15	239.48	1.35	0.60	0.65	0.18	8.59	10.92
ICON-ESM-LR	1.31	7.15	397.08	1.73	1.38	0.66	0.29	9.71	8.59
IITM-ESM	1.05	5.37	126.92	2.49	1.92	1.76	1.13	9.85	4.82
INM-CM4-8	1.17	6.53	80.86	2.12	1.89	1.21	0.01	6.58	3.75
INM-CM5-0	1.06	6.38	161.14	1.78	1.92	1.12	0.03	6.56	4.03
IPSL-CM5A2-INCA	1.37	13.05	83.13	1.43	0.84	0.75	1.59	7.39	7.80
IPSL-CM6A-LR	1.08	10.73	93.40	2.07	0.80	0.64	1.17	9.28	9.54
KIOST-ESM	1.05	5.75	285.83	1.30	1.37	0.71	0.02	7.43	7.39
MIROC-ES2L	1.59	9.32	209.35	1.81	1.29	2.29	0.99	8.32	4.89
MIROC6	1.96	13.05	297.24	1.53	1.31	1.77	0.39	7.61	4.94
MPI-ESM-1-2-HAM	1.45	9.67	245.61	1.85	1.59	1.02	0.42	9.33	6.45
MPI-ESM1-2-HR	2.12	10.75	98.71	1.57	1.95	0.99	1.00	8.52	6.31
MPI-ESM1-2-LR	1.87	10.23	257.28	1.92	1.47	1.19	0.10	9.74	6.30
MRI-ESM2-0	1.00	5.26	190.72	1.41	1.41	0.91	0.46	7.71	6.64
NESM3	1.18	11.56	80.78	1.46	0.67	0.87	1.64	9.61	9.24
NorCPM1	1.00	5.02	89.10	0.77	1.83	0.83	1.30	7.46	6.68
NorESM2-LM	1.95	18.67	280.80	2.45	nan	0.94	0.05	9.98	5.76
NorESM2-MM	2.13	12.92	255.69	2.06	nan	1.33	0.03	9.15	5.10
SAM0-UNICON	3.80	26.08	278.66	1.33	nan	0.83	0.12	8.76	8.04
TaiESM1	1.29	8.44	291.73	2.04	0.99	0.73	0.12	9.76	9.28
UKESM1-0-LL	2.03	15.28	96.48	1.04	0.69	0.61	1.64	8.20	11.10
Model mean	1.74	12.20	192.98	1.61	1.17	0.93	0.65	8.36	7.86

Table S7. Parameters estimated from abrupt-4xCO2 experiments, part II.

	F2x	F0p5x	F4x	F4x/F2x	F4x/F0p5x
CESM2	4.56	-5.82	9.22	2.02	-1.58
CNRM-CM6-1	4.23	-3.74	8.29	1.96	-2.22
CanESM5	3.69	-3.55	7.57	2.05	-2.13
GISS-E2-1-G	3.92	-3.78	8.01	2.04	-2.12
GISS-E2-1-H	3.98	nan	8.16	2.05	nan
GISS-E2-2-G	3.67	nan	7.51	2.05	nan
GISS-E2-2-H	3.36	nan	7.06	2.10	nan
HadGEM3-GC31-LL	3.73	-3.04	8.03	2.15	-2.64
IPSL-CM6A-LR	4.23	-3.17	9.28	2.19	-2.93
MIROC6	3.27	-3.20	7.61	2.32	-2.38
MRI-ESM2-0	3.57	-4.45	7.71	2.16	-1.73
TaiESM1	4.48	-4.51	9.76	2.18	-2.17
Model mean	3.89	-3.92	8.18	2.11	-2.21

Table S8. Forcing estimates for 2x, 0.5x, 4x CO_2, and forcing ratios.