Changes in IPCC scenario assessment emulators between SR1.5 and AR6 unravelled

Zebedee R.J. Nicholls[1], Malte Alexander Meinshausen[1], Jared Lewis[2], Christopher J Smith[3], Piers Forster[3], Jan Sigurd Fuglestvedt[4], Joeri Rogelj[5], Jarmo Kikstra[6], keywan Riahi[6], and Edward Anthony Byers[7]

[1]University of Melbourne
[2]Climate Resource
[3]University of Leeds
[4]Center for International Climate and Environmental Research - Oslo (CICERO)
[5]Imperial College London
[6]IIASA
[7]International Institute for Applied Systems Analysis

November 28, 2022

Abstract

The IPCC's scientific assessment of the timing of net-zero emissions and 2030 emission reduction targets consistent with limiting warming to 1.5C or 2C rests on large scenario databases. Updates to this assessment, such as between the IPCC's Special Report on Global Warming of 1.5C (SR1.5) of warming and the Sixth Assessment Report (AR6), are the result of intertwined, sometimes opaque, factors. Here we isolate one factor: the Earth System Model emulators used to estimate the global warming implications of scenarios. We show that warming projections using AR6-calibrated emulators are consistent, to within around 0.1C, with projections made by the emulators used in SR1.5. The consistency is due to two almost compensating changes: the increase in assessed historical warming between the IPCC's Fifth Assessment Report (AR5) and AR6, and a reduction in projected warming due to improved agreement between the emulators' response to emissions and the underlying assessment.

Changes in IPCC scenario assessment emulators between SR1.5 and AR6 unravelled

Z. Nicholls[1,2,3], M. Meinshausen[2,3], J. Lewis[1,2,3], C. J. Smith[1,4], P. M. Forster[4], J. Fuglestvedt[5], J. Rogelj[1,6,7], J. S. Kikstra[1,6,7], K. Riahi[1], E. Byers[1]

[1]International Institute for Applied System Analysis, IIASA, Laxenburg, Austria
[2]Climate & Energy College, School of Geography, Earth and Atmospheric Sciences, The University of Melbourne, Parkville, Victoria, Australia
[3]Climate Resource, Northcote, Victoria, Australia
[4]Priestley International Centre for Climate, University of Leeds, Leeds, United Kingdom
[5]CICERO, Oslo, Norway
[6]Centre for Environmental Policy, Imperial College London, London, United Kingdom
[7]Grantham Institute for Climate Change and the Environment, Imperial College London, United Kingdom

Key Points:

- Emulators used in IPCC SR1.5 and AR6 are remarkably consistent, despite their entirely new calibrations
- The consistency is mostly due to two factors: change in assessed historical warming and improvements to emulator calibration methods

Corresponding author: Zebedee Nicholls, `zebedee.nicholls@climate-energy-college.org`

Abstract

The IPCC's scientific assessment of the timing of net-zero emissions and 2030 emission reduction targets consistent with limiting warming to 1.5°C or 2°C rests on large scenario databases. Updates to this assessment, such as between the IPCC's Special Report on Global Warming of 1.5°C (SR1.5) of warming and the Sixth Assessment Report (AR6), are the result of intertwined, sometimes opaque, factors. Here we isolate one factor: the Earth System Model emulators used to estimate the global warming implications of scenarios. We show that warming projections using AR6-calibrated emulators are consistent, to within around 0.1°C, with projections made by the emulators used in SR1.5. The consistency is due to two almost compensating changes: the increase in assessed historical warming between the IPCC's Fifth Assessment Report (AR5) and AR6, and a reduction in projected warming due to improved agreement between the emulators' response to emissions and the underlying assessment.

Plain Language Summary

The IPCC's latest physical science report, the Working Group 1 (WG1) Contribution to the Sixth Assessment Report (AR6), was released in August 2021. That report includes an update to the tools used to project the climate outcome of emission scenarios. Here we apply these newly calibrated tools, called earth system model emulators, to the set of scenarios assessed in the IPCC's Special Report on warming of 1.5°C (SR1.5). We find that two compensating changes lead to a remarkable consistency (peak warming projections within 0.1°C) between the projections made by the emulators used in SR1.5 and their updated, AR6-calibrated descendants. Firstly, updates to the historical warming assessment since the IPCC's 2013 physical science report (AR5) increase future warming projections. However, improved consistency between the emulators and the assessment of the underlying physics, particularly the short-term warming response to emissions, lowers warming projections by an approximately equivalent amount. Our work reinforces the key messages from the IPCC: limiting warming to around 1.5°C is a great and urgent challenge, and it is up to us to decide whether we pull out all the stops to hold temperatures around 1.5°C or whether we sail on by.

1 Introduction

To assess the characteristics of scenarios in line with different levels of global warming, emission scenarios are grouped in distinct categories based on their global-mean temperature outcomes (Rogelj et al., 2011). This practice was followed in both SR1.5 (Rogelj et al., 2018) and the Working Group 3 (WG3) Contribution to AR6. The emissions scenarios are typically generated by Integrated Assessment Models (IAMs, Weyant, 2017), which combine assumptions about future population, economy, climate policy and technology to project internally consistent evolutions of future greenhouse gas and other emissions.

Over 400 scenarios were assessed in SR1.5 (Huppmann et al., 2018), and AR6 WG3 assessed over 1200 (Riahi et al., 2022). During the IPCC drafting process, projections for these scenarios have to be delivered in a matter of weeks, which requires computationally efficient models, also known as Earth System model emulators. These emulators quantify the climate implications of each scenario's emissions, which in turn are used to categorise scenarios according to their global warming outcomes (Riahi et al., 2022).

Before AR5, IAMs self-reported climate outcomes of scenarios. However, climate system representations vary in complexity, sophistication, and accuracy between IAMs (van Vuuren et al., 2011; Harmsen et al., 2015), so comparing self-reported climate outcomes from different IAMs can be complex and inaccurate. To eliminate the unnecessary noise that results from the use of an unwieldy set of poorly calibrated climate models, the WG3 Contribution to AR5 initiated a harmonised approach to the climate assessment of IAM scenarios (Clarke et al., 2014). IAM scenarios were assessed with a single calibrated climate model, also referred to as a climate emulator, in a probabilistic setup (Meinshausen et al., 2009, 2011; Rogelj et al., 2012). The probabilistic calibration aims to make the climate response of the emulator reflect the state of climate science knowledge and its surrounding uncertainties as closely as possible.

IPCC AR5 used the MAGICC6 model to assess the scenarios submitted to the AR5 scenario database as part of the wider assessment process. The 2018 IPCC Special Report on Global Warming of 1.5°C (SR1.5, Forster et al., 2018; Rogelj et al., 2018) used the exact same AR5-setup of MAGICC6, together with a second climate emulator, the SR1.5-setup of FaIR1.3 (Millar et al., 2017; C. J. Smith et al., 2018). At the time of SR1.5, differences in the temperature projections by these emulators remained unexplained and were instead highlighted as a knowledge gap. This affected the accuracy by which the global warming implications of scenarios could be assessed and scenarios could be grouped in 1.5°C compatible or 2°C compatible classes (Rogelj et al., 2018). For consistency with AR5, the AR5-setup of MAGICC6 was used for classification of scenarios in SR1.5 and information from the SR1.5-setup of FaIR 1.3 was used to inform the overall uncertainty assessment (Rogelj et al., 2018).

Scientific efforts and lessons learned since SR1.5 have now closed this knowledge gap. Climate emulator intercomparison exercises have developed protocols to compare and understand differences between emulators and their calibrations (Nicholls & Lewis, 2021; Nicholls et al., 2021). These advances were applied as part of the AR6 physical science assessment (WGI), where a cross-chapter activity calibrated and vetted four emulators using a wide range of assessed climate system characteristics. This activity ensured that the probabilistic parameterisations of the emulators closely matched AR6 findings related to equilibrium climate sensitivity (ECS), transient climate response (TCR), transient climate response to emissions (TCRE), ocean heat uptake, historical temperature observations and the assessed projected global-mean temperatures under various ScenarioMIP scenarios (O'Neill et al., 2016; Tebaldi et al., 2021).

Comparing this set of AR6-calibrated climate emulators with previous setups allows us to explore how advances in our understanding of the physical climate system affect which emissions pathways are consistent with holding warming below 1.5°C compared to preindustrial levels. Given the widespread use of these emulators in the literature, the analysis is also useful for teams who wish to anticipate and under the changes when updating from the AR5- to the AR6-versions of the emulators. Throughout this paper we focus on the difference between the AR5-setup of MAGICC6, which was used for scenario categorisation in SR1.5, and AR6-calibrated MAGICCv7.5.3, which is used for scenario categorisation in AR6 WG3. The differences with the SR1.5-setup of FaIR1.3 and AR6-calibrated FaIRv1.6.2, used for SR1.5 and AR6, respectively, are discussed where appropriate, but are not examined in the same detail.

2 Materials and Methods

We use the 368 scenarios underlying Table 2.4 in SR1.5, a subset of the SR1.5 scenario database's complete set of more than 400 scenarios (Rogelj et al., 2018; Huppmann et al., 2019). We focus on this subset as it formed the basis of many of SR1.5's top-level statements and excludes scenarios that have greenhouse gas emissions that were deemed unrealistic at the time of SR1.5 or bias the full set because of strong similarity (Rogelj et al., 2018). For these 368 scenarios, we reassess their climate outcomes with the newly AR6-calibrated emulators and reapply the scenario classification rules from SR1.5. Any differences can thus be attributed to changes in the calibrated climate emulators and associated changes in our physical science understanding.

We reassess the SR1.5 scenarios with the AR6-calibrated emulators using the WG3 climate assessment pipeline (Kikstra et al., 2022 (in prep.)). The pipeline is built on three key tools: Aneris for harmonising the emissions timeseries to historical emissions (M. J. Gidden et al., 2018; M. Gidden et al., 2022), Silicone for infilling emissions species not natively reported by the IAMs (Lamboll et al., 2020), and OpenSCM-Runner for running the climate models (Nicholls et al., 2020).

The MAGICCv7.5.3 and FaIRv1.6.2 AR6 setups are documented in Forster et al. (2021). For the SR1.5 emulators, we use output from the SR1.5 database (Huppmann et al., 2018) without modification. To run MAGICCv7.5.3 in an AR5-like setup, we use MAGICCv7.5.3's RCMIP Phase 2 HadCRUT4.6.0.0 calibration and the AR5 recent past warming estimate of 0.61°C for 1986-2005 relative to 1850-1900.

3 Results

3.1 Scenario categorisation

We find that the key outputs used for categorisation are broadly consistent between the AR5-setup of MAGICC6 and AR6-calibrated MAGICCv7.5.3 (Figure 1). Differences are limited to 0.7% in the median across all the scenarios (5-95% range across scenarios of -3.5% to 4.9%) for peak 1.5°C exceedance probability, 0.0% (-9.1% to 3.4%) for peak 2.0°C exceedance probability and 0.0% (-11.1% to 2.5%) for 2100 1.5°C exceedance probability (Supplementary Figure S1). In terms of median temperature projections, the median difference across the scenarios is 0.02°C (-0.15°C to 0.06°C) for median peak warming and -0.05°C (-0.16°C to 0.05°C) for median 2100 warming (Supplementary Figures S2 and S3).

These differences are smaller than the usually applied rounding precision of 0.1°C and natural variability. They demonstrate a remarkable consistency between the SR1.5

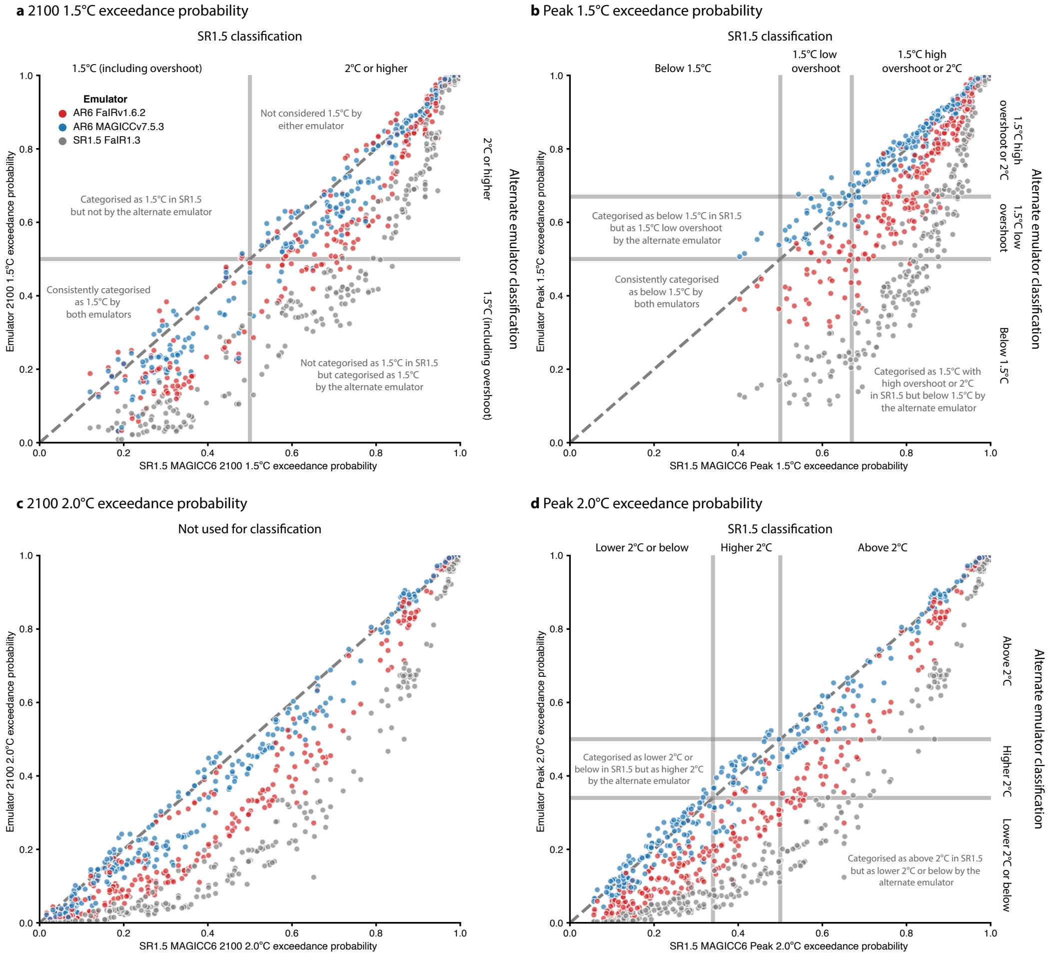

Figure 1. The classification-relevant exceedance probabilities of SR1.5 scenarios are similar when re-assessed with the AR6-calibrated MAGICCv7.5.3, slightly lower with the AR6-calibrated FaIRv1.6.2 and lower in the SR1.5-calibration of FaIR1.3. a) 1.5°C exceedance probabilities in 2100 from AR6-calibrated MAGICCv7.5.3 (blue dots), AR6-calibrated FaIRv1.6.2 (red dots) and SR1.5-calibrated FaIR1.3 (grey dots) compared to the data used for SR1.5 categorisation based on the AR5-setup of MAGICC6. b) As in panel a, but for peak warming. c) As in panel a, but for 2°C warming. d) As in panel a, but for 2°C peak warming. The vertical and horizontal lines delineate the scenario classifications. To aid comparisons, dashed diagonal lines show the 1:1 line (points below the diagonal indicate higher outcomes with the AR5-setup of MAGICC6 than with the other considered emulator setups).

and updated AR6 emulator setups. For example, AR6 reports assessed temperature projections to the nearest tenth of a degree (Lee et al., 2021). The reason for this choice is the scientific uncertainties that must be considered when making long-term projections, such as the historical anthropogenic warming uncertainty of 0.8 - 1.3°C (likely range for 2000-2019 relative to 1850-1900, Eyring et al., 2021), the contribution of internal variability of about 0.15°C for a 20-year average (5-95% range, Lee et al., 2021) or uncertainty in the zero emissions commitment (Jones et al., 2019; MacDougall et al., 2020) of about 15% of total warming (1-sigma Lee et al., 2021). The contribution of internal variability is key to keep in mind: our climate model emulators only model the externally forced warming response, almost entirely human driven with a small (approximately 1%) contribution from the solar cycle, and natural variations around this are not included in the assessment of warming performed here.

Using the AR5 MAGICC6 setup, 42 scenarios were classified as 1.5°C with no or low overshoot, 36 were classified as 1.5°C with high overshoot and 54 were classified as lower 2°C (Table 1). Using the AR6-calibrated MAGICCv7.5.3 setup, 41 scenarios are classified as 1.5°C with no or low overshoot, 38 are classified as 1.5°C with high overshoot and 64 are classified as lower 2°C.

Using the AR6-calibrated FaIRv1.6.2 and especially FaIR1.3, more scenarios are classified in these low categories due to cooler projections. Specifically, 78 scenarios are assessed as 1.5°C with low or no overshoot with the AR6-calibrated FaIRv1.6.2 emulator (red dots below the 67% exceedance probability line in Figure 1b). The lower projections from AR6-calibrated FaIRv1.6.2 are the result of a slightly lower TCR (Forster et al., 2021; C. Smith et al., 2021) and lower projections of atmospheric CO_2 and CH_4 concentrations (a topic we return to in Section 4.3). At the time of SR1.5, a total of 149 scenarios would have been classified as 1.5°C with low or no overshoot had the SR1.5-setup of FaIR1.3 been chosen for the classification of scenarios (grey dots below the 67% exceedance probability line in Figure 1).

We see the broad consistency between the AR5 MAGICC6 setup's and the AR6-calibrated MAGICCv7.5.3's projections reflected in the similarity of the scenario classification. The only case where this isn't true is if we draw a distinction between 1.5°C no overshoot and 1.5°C low overshoot scenarios (where 5 scenarios are classified as no overshoot with the AR5 MAGICC6 setup while no scenarios are classified as no over-

Table 1. Classification rules for scenarios from the IPCC SR1.5 (only scenarios included in SR1.5 Table 2.4, adapted from Rogelj et al., 2018), classification of scenarios in SR1.5 and classification based on AR6-calibrated emulators.

Class name	**Classification rule (P(1.5°C) is the probability that temperatures exceed 1.5°C)**	**Number of scenarios in SR1.5 Table 2.4**	**Number of scenarios with other SR15 emulator**	**Number of scenarios with AR6-calibrated emulator**	
Emulator		*MAGICC6*	*FaIR1.3*	*MAGICCv7.5.3*	*FaIRv1.6.2*
Below 1.5°C	$0.34 < P(1.5°C) \leq 0.5$	5	127	0	36
1.5°C low-overshoot	$0.5 < P(1.5°C) \leq 0.67$ AND $P(1.5°C \text{ in } 2100) \leq 0.5$	37	22	41	42
1.5°C no and low overshoot	Combination of two categories above i.e., $P(1.5°C) \leq 0.67$ AND $P(1.5°C \text{ in } 2100) \leq 0.5$	42	149	41	78
1.5°C high-overshoot	$0.67 < P(1.5°C)$ AND $P(1.5°C \text{ in } 2100) \leq 0.5$	36	1	38	19
Lower 2°C	$P(2°C) \leq 0.34$ AND $P(1.5°C \text{ in } 2100) > 0.5$	54	76	64	92
Higher 2°C	$0.34 < P(2°C) \leq 0.5$ AND $P(1.5°C \text{ in } 2100) > 0.5$	54	13	52	36
Above 2°C	$P(2°C) > 0.5$	182	128	173	143

shoot using the AR6-calibrated MAGICCv7.5.3, Figure 1). However, following the SR1.5 choice means that scenarios in the '1.5°C with low overshoot' category must have a peak 1.5°C exceedance probability between 50% and 67% (a range of approximately 0.12°C in terms of median warming, Supplementary Figure S4). While across all scenarios the changes of 1.5°C exceedance probabilities are much less than this, the very strong mitigation scenarios discussed feature approximately 10% changes, which is enough to cause them all to change category. The small difference between warming to date and the 1.5°C limit means that the 1.5°C no overshoot and 1.5°C low overshoot categories are very close.

3.2 Temperature threshold crossing times

Alongside the changes in categories, we also consider the change in the point in time when overshoot scenarios cross and return below the 1.5°C threshold (Figure 2). We find that, while scenarios cross the 1.5°C threshold 4 years earlier (in the median) using the AR6-calibrated MAGICCv7.5.3 compared to the AR5-setup of MAGICC6, many scenarios also return below 1.5°C sooner than previously thought. However, there is quite some uncertainty in the change in the year in which temperatures return below 1.5°C, with the median being a 4 year earlier return and a 5-95% range of 19 years earlier to 12 years later. The range reflects the fact that small changes in the rate of cooling lead to large changes in crossing times (a result of the geometry of determining the point at which two nearly parallel lines, the 1.5°C limit and the declining temperatures, cross). In addition, both the uncertainty in the climate system's response to net zero or net negative CO_2 emissions and the wide range of non-CO_2 emissions pathways (specifically after net zero CO_2) in the SR1.5 database contribute to the uncertainty as to when exactly temperature will return back below the 1.5°C limit if temporarily overshot. This uncertainty and the ill-defined geometrical nature of estimating the time of returning below a temperature threshold after an overshoot suggests that this characteristic can be more robustly described by the decade of peak warming and the decadal rate of temperature reduction thereafter, be it zero or negative (Rogelj et al., 2019).

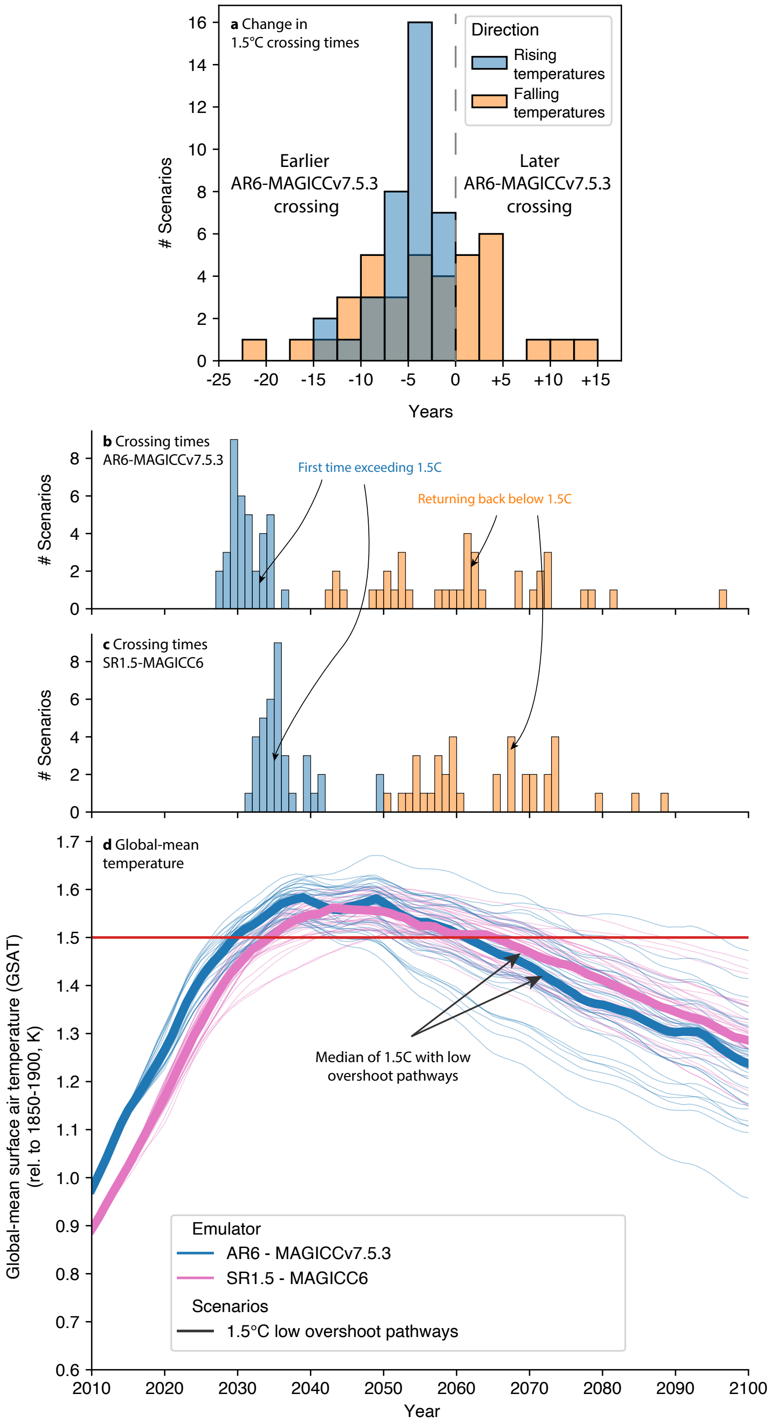

Figure 2. Change in time at which 1.5°CC warming is first crossed and then returned below in scenarios which were classified as 1.5°CC with low overshoot in SR1.5. a) Crossing times based on the AR6-calibrated MAGICCv7.5.3 relative to the crossing times based on the SR1.5 data (AR5-setup of MAGICC6). b) Crossing times based on the AR6-calibrated MAGICCv7.5.3. c) Crossing times based on the SR1.5 data (AR5-setup of MAGICC6). d) Timeseries of temperature evolution in the considered pathways.

4 Discussion

4.1 Causes of categorisation changes

We find two key causes for changes in the IPCC categorisation: changes in the historical temperature assessment and other changes in the physical science assessment, which includes the ability of calibrated emulators to reflect that science. The upwards revision of the historical warming in AR6 meant that the best-estimate for 1986-2005 relative to 1850-1900 was 0.69°C, compared to 0.61°C in AR5 (Gulev et al., 2021). Similarly, for 2003-2012 relative to 1850-1900, AR6's best-estimate warming was 0.90°C, compared to 0.78°C in AR5. These increases are 0.1°C, or around 15% in terms of 1.5°C exceedance probabilities (Supplementary Figure S4).

To disentangle the multiple updates between the AR5 setup of MAGICC6 and AR6-calibrated MAGICCv7.5.3 – apart from historical temperatures – we first compare results using the AR5 setup of MAGICC6 and the MAGICCv7.5.3 calibration presented in RCMIP Phase 2 (Nicholls et al., 2021). The latter is calibrated to HadCRUT.4.6.0.0 (Morice et al., 2012) and literature published before AR6, hence is a rough approximation of how a MAGICCv7.5.3 calibration to AR5 would perform. The RCMIP Phase 2 calibration of MAGICCv7.5.3 projects median peak warming that is 0.13°C less (5-95% range across scenarios of 0.25°C less to 0.06°C less) than the AR5 setup of MAGICC6 (Figure 3 and Supplementary Figures S5 and S6). In other words, updating from MAGICC6's AR5-setup to a setup more directly calibrated to AR5 would likely cause a drop in projections. The major driver for this change is the different historical warming estimate, with other effects playing only a minor role (Supplementary Text S1).

Next, we consider the overall change i.e., the difference in warming projections by the AR5-setup of MAGICC6 and the AR6-calibration of MAGICCv7.5.3 (Supplementary Figure S7). The difference can arise from changes in any of the steps (specifically parameterisations thereof) along the cause-effect chain from emissions to atmospheric concentrations to effective radiative forcing to warming. We firstly observe that the AR5-setup of MAGICC6 generally has lower effective radiative forcing than the AR6-calibration of MAGICCv7.5.3 (Supplementary Figure S8, with a breakdown of the contribution of different climate forcers discussed in Supplementary Text S2). Therefore, differences in the parameterisations that link emissions and effective radiative forcing are not the reason for higher warming projections when using the MAGICC6 AR5-setup.

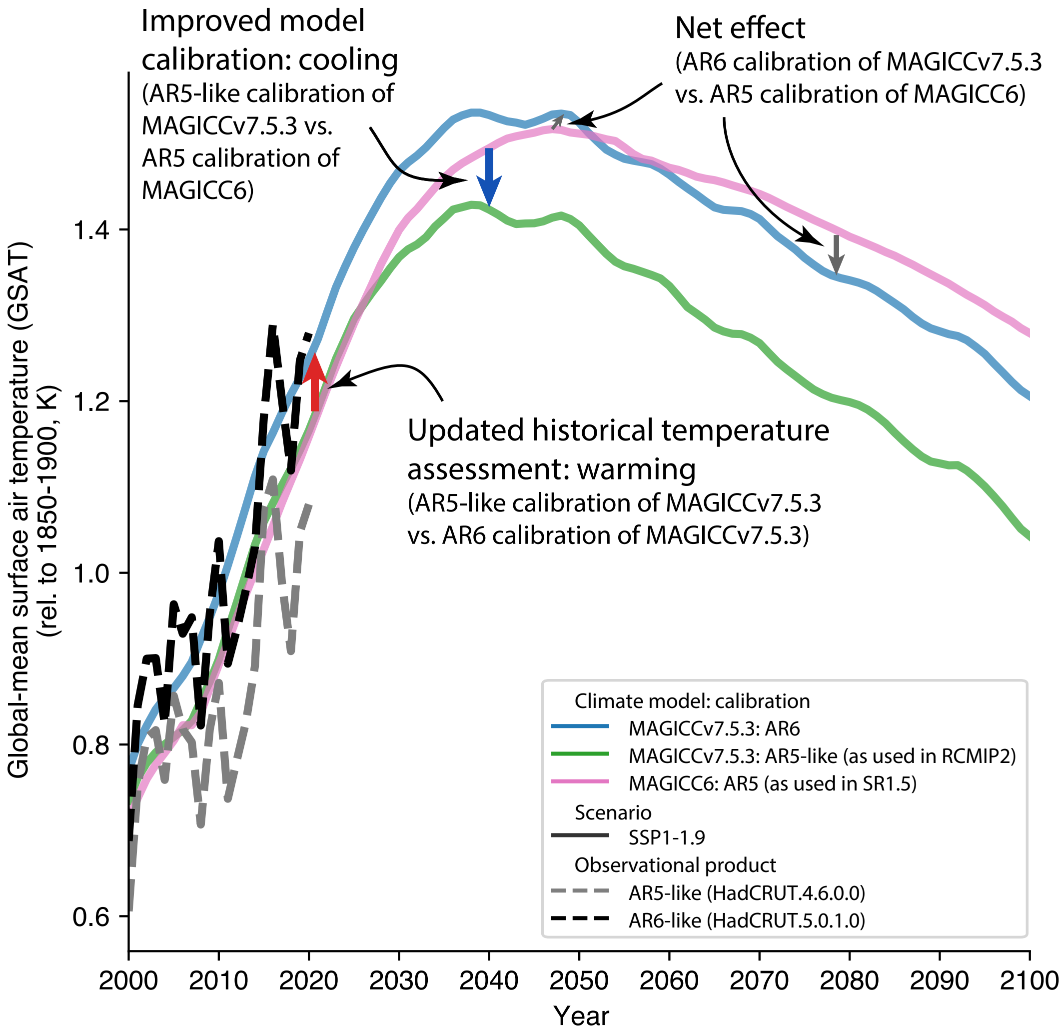

Figure 3. Contributions to changes in temperature projections, illustrated using the SSP1-1.9 scenario. We compare the AR5-setup of MAGICC6 as used in SR1.5 (pink line), MAGICCv.7.5.3 as calibrated in RCMIP Phase 2 (green line) and MAGICCv7.5.3 as calibrated in AR6 (blue line). For comparison, we also plot HadCRUT4.6.0.0 (grey dashed line) and HadCRUT5.0.1.0 (black dashed line). HadCRUT4.6.0.0 is used as a proxy for the AR5 historical temperature assessment (which the AR5-setup of MAGICC6 and MAGICCv.7.5.3 as used in RCMIP Phase 2 are calibrated to) while HadCRUT5.0.1.0 is used as a proxy for the AR6 historical temperature assessment (which MAGICCv7.5.3 as used in AR6 is calibrated to).

Given that effective radiative forcings do not explain the change, we instead focus on the parameterisation linking effective radiative forcing and warming. A key measure of this is the transient climate response (TCR). In MAGICC, TCR is not a model parameter, but an emergent property that is influenced by multiple parameters that control ocean heat uptake and climate feedbacks. In AR5, the assessment was a likely range from 1 to 2.5°C (with no explicit central assessment) while in AR6 the range slightly narrowed to 1.4 to 2.2°C with a central assessment of 1.8°C. As the AR6-calibrated MAGICCv7.5.3 matches the AR6 TCR assessment well (see AR6 WG1 Cross-Chapter Box 7.1, Table 2, Forster et al., 2021), we conclude that the calibration of MAGICC6 used in SR1.5 had a TCR which was higher than assessed ranges available at the time (as also suggested by Leach et al., 2018).

The overall change in projections between AR6-calibrated MAGICCv7.5.3 and the AR5-setup of MAGICC6 includes both the warming from changes in the IPCC assessment of historically observed warming and the cooling from other forcing and feedback related changes, which manifest in a lower TCR in the AR6-calibrated MAGICCv7.5.3 version compared to the AR5-setup of MAGICC6. The two contributions (historical warming and other effects) approximately cancel, leading to changes in exceedance probabilities of around 10% as discussed previously.

4.2 Implications for mitigation

The relatively small differences in climate projections lead to small changes in key mitigation milestones describing scenario categories, such as net zero CO_2 years (Figure 4) or 2030 emissions reductions. Using the AR5-setup of MAGICC6, no and low overshoot 1.5°C scenarios had a net zero CO_2 year of 2050 (2038 to 2061 5-95% range). In contrast, the AR6-calibrated MAGICCv7.5.3 has a net zero CO_2 year of 2050 (2038 to 2075) and the AR6-calibrated FaIRv1.6.2 has a net zero CO_2 year of 2052 (2042 to 2070).

The importance of these changes for policy and economic transition is a separate question, but they may not be seen as zero in all contexts (e.g., the difference in the 95^{th} percentile is 14 years). These differences in mitigation milestones arise even though climate science has remained remarkably consistent (differences of 0.05°C in the median). A key point from SR1.5 remains relevant, "because of numerous geophysical uncertainties and model dependencies [...] absolute temperature characteristics of the various path-

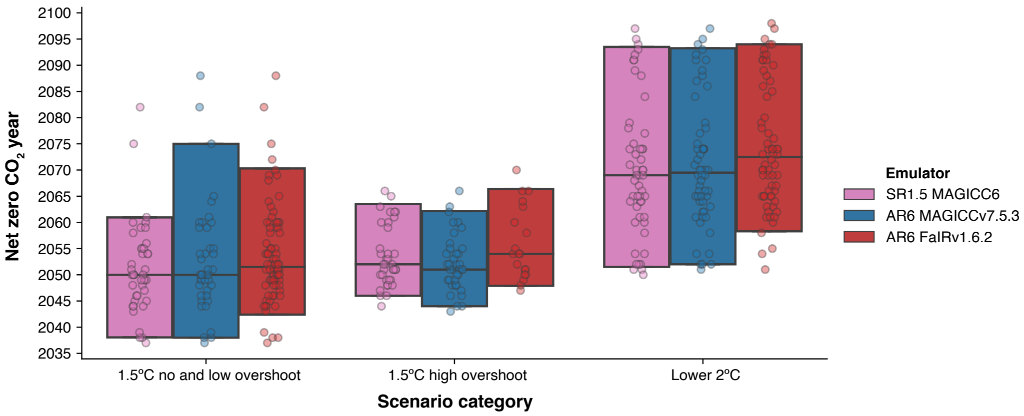

Figure 4. Sensitivity of net zero CO_2 year in different categories to emulator choice. For each category (x-axis), we show the distribution (black line shows median, box shows 5-95% range and dots show individual scenarios) of net zero CO_2 year based on either the SR1.5 classification emulator (AR5-setup of MAGICC6), the AR6-calibrated MAGICCv7.5.3 or the AR6-calibrated FaIRv1.6.2. For the number of scenarios in each distribution, see Table 1.

way categories are more difficult to distinguish than relative features" (Rogelj et al., 2018). The fact that our classifications rely on absolute temperatures, in which we have lower confidence, raises the question of whether there are ways to analyse mitigation pathways that rely on the relative differences where we have more confidence.

Another point which is not always immediately obvious is that the connection between changes in physical climate assessment and emissions milestones for scenario categories is not one-to-one. For example, the net zero CO_2 years of 1.5°C with low and high overshoot scenarios are similar despite their (by definition) different climate outcomes (Figure fig:mitigation-metric-changes). The key reason is that the SR1.5 scenario database can be described as an ensemble of opportunity (Tebaldi & Knutti, 2007; Rogelj et al., 2011; Huppmann et al., 2018) and is not a systematic sample of the underlying scenario space (Fujimori et al., 2019).

4.3 Emissions-driven uncertainty

The MAGICC and FaIR emulators show improved agreement in AR6 compared to SR1.5. This is particularly so in experiments where concentrations of greenhouse gases are prescribed to the models, where the emulators' median warming projections agree to within 0.05°C under the SSP1-1.9 and SSP1-2.6 scenarios (Forster et al., 2021; C. Smith et al., 2021). These concentration-driven experiments are directly comparable to both the WG1 temperature assessment (Gulev et al., 2021; Eyring et al., 2021) and CMIP ScenarioMIP (Eyring et al., 2016; O'Neill et al., 2016) experiments, both of which are based on large scientific efforts.

However, the agreement between emulators is reduced once we consider experiments where emissions of greenhouse gases are prescribed to the models, rather than concentrations. The switch to emissions-driven experiments introduces uncertainty in greenhouse gas cycles, particularly the carbon and methane cycles (Forster et al., 2021). Another key uncertainty in these emissions-driven experiments is the zero emissions commitment, which has a range of -0.34°C to 0.28°C (for the change in temperature 50 years after CO_2 emissions compatible with warming of around 2°C cease) across Earth System Models (Lee et al., 2021), and was assessed by AR6 be centred around zero and likely (with greater than 66% probability) fall in the ±0.3°C range. In their AR6-calibrations, MAGICCv7.5.3 projects higher CO_2 and methane concentrations than FaIRv1.6.2 (Supplementary Figure S9). Unfortunately, a lack of validation data for emissions-driven experiments, particularly in scenarios where emissions are falling or net negative, restricts our ability to derive robust conclusions about which one of the two projections are more likely. The AR6-calibrated FaIRv1.6.2's airborne fraction is slightly closer to Earth System Model (ESM) experiments (Forster et al., 2021), although this is based on idealised rather than scenario-based experiments. There are also few ESM experiments to compare with the methane projections and none which are directly comparable.

These carbon and methane cycle differences are part of the reason for differences in the AR6-calibrated MAGICCv7.5.3 and FaIRv1.6.2 models' temperature projections (Supplementary Figures S10 and S11). Improvements in reduced complexity carbon and methane cycle representations and their evaluation is a clear area for future research. Nonetheless, the difference in model projections of order 0.1°C is a reasonable representation of our current emissions-driven uncertainty. It is also worth noting the progress seen since SR1.5, where emulator disagreement was around 0.3°C in the median and largely unexplained.

5 Conclusions

When applied to the SR1.5 scenarios database, the projections from the AR6-calibrated emulators are remarkably close to their predecessors used in SR1.5. From a climate model emulator perspective, the key insights from SR1.5 remain valid and policies enacted based on the key insights from SR1.5 are supported by the latest scientific evidence. For example, reducing CO_2 emissions by 50% by 2030 and reaching net zero CO_2 emissions around 2050 will – from a geophysical perspective – more likely than not limit peak warming to around 1.5°C (i.e., with greater than 50% likelihood). Updates to the design of scenarios (Rogelj et al., 2019; Riahi et al., 2021) with stronger reductions early on and slower approaches towards net-zero might add further insights into how near-term action can help push back net zero years, but they do not change the validity of a 2050 net-zero CO_2 year as a guide to mitigation action in the next one or two decades given current emission trends.

Our best projection remains that the world is going to see 1.5°C warming by the early 2030s (averaged over a 20-year period and acknowledging that individual years will exceed 1.5°C beforehand due to natural variability). Thus, while decisive mitigation efforts this decade will be crucial in determining whether we shoot beyond 1.5°C, adaptation actions will have to be taken on the basis of a minimal warming level around 1.5°C.

Assuming we do reach net zero and then achive net negative CO_2 emissions, the response of the Earth System thereafter is uncertain (Jones et al., 2019; MacDougall et al., 2020; Lee et al., 2021). Despite this uncertainty, there is robust evidence that every tonne of CO_2 matters and every avoided emission lowers the risk of climate damage (Canadell et al., 2021). Our results reinforce this and other key messages that have been delivered by the IPCC for many years. On the other hand, the lack of sufficient action and global emissions reductions is irrefutably pushing the Paris Agreement goals out of reach and putting our global society at risk.

6 Open Research

The code and data used to produce the plots is preserved at 10.5281/zenodo.6584386 and developed openly at https://gitlab.com/magicc/nicholls-et-al-2022-emulator-changes.

Acknowledgments

Z.N. and J.L. acknowledge the European Union's Horizon 2020 research and innovation program ESM2025 – Earth System Models for the Future (grant agreement no. 101003536). J.S.K. was supported by the UK Natural Environment Research Council under grant agreement NE/S007415/1. J.S.K., E.B. and K.R. acknowledge funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 821471 (ENGAGE).

References

Canadell, J. G., Monteiro, P. M. S., Costa, M. H., Cotrim da Cunha, L., Cox, P., Eliseev, A. V., ... Zickfeld, K. (2021). Global carbon and other biogeochemical cycles and feedbacks [Book Section]. In V. Masson-Delmotte et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (chap. 5). Cambridge University Press. Retrieved from https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Chapter_05.pdf

Clarke, L., Jiang, K., Akimoto, K., Babiker, M., Blanford, G., Fisher-Vanden, K., ... al, E. (2014). Assessing transformation pathways [Book Section]. In O. Edenhofer et al. (Eds.), *Climate change 2014: Mitigation of climate change. contribution of working group iii to the fifth assessment report of the intergovernmental panel on climate change* (p. 413–510). Cambridge University Press.

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 [Journal Article]. *Geophysical Research Letters*, *43*(24), 12,614-12,623. Retrieved from http://dx.doi.org/10.1002/2016gl071930 doi: 10.1002/2016gl071930

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization [Journal Article]. *Geoscientific Model Development (Online)*, *9*(LLNL-JRNL-736881).

Eyring, V., Gillett, N. P., Achuta Rao, K. M., Barimalala, R., Barreiro Parrillo, M., Bellouin, N., ... Sun, Y. (2021). Human influence on the climate system

[Book Section]. In V. Masson-Delmotte et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (chap. 3). Cambridge University Press. Retrieved from https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Chapter_03.pdf

Forster, P., Huppmann, D., Kriegler, E., Mundaca, L., Smith, C., Rogelj, J., & Séférian, R. (2018). Mitigation pathways compatible with 1.5°c in the context of sustainable development supplementary material [Book Section]. In *Global Warming of 1.5 °C: an IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* (p. 2SM1-2SM50). IPCC/WMO. Retrieved from http://www.ipcc.ch/report/sr15/

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J. L., Frame, D., ... Zhang, H. (2021). The earth's energy budget, climate feedbacks, and climate sensitivity [Book Section]. In V. Masson-Delmotte et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (chap. 7). Cambridge University Press. Retrieved from https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Chapter_07.pdf

Fujimori, S., Rogelj, J., Krey, V., & Riahi, K. (2019). A new generation of emissions scenarios should cover blind spots in the carbon budget space [Journal Article]. *Nature Climate Change*, *9*(11), 798-800. Retrieved from https://doi.org/10.1038/s41558-019-0611-9 doi: 10.1038/s41558-019-0611-9

Gidden, M., Hörsch, J., Nicholls, Z., & Bot, S. (2022, May). *iiasa/aneris: v0.3.1* [Software]. Zenodo. Retrieved from https://doi.org/10.5281/zenodo.6545476 doi: 10.5281/zenodo.6545476

Gidden, M. J., Fujimori, S., van den Berg, M., Klein, D., Smith, S. J., van Vuuren, D. P., & Riahi, K. (2018). A methodology and implementation of automated emissions harmonization for use in integrated assessment models [Journal Article]. *Environmental Modelling & Software*, *105*, 187-200.

Retrieved from https://www.sciencedirect.com/science/article/pii/S1364815217307867 doi: https://doi.org/10.1016/j.envsoft.2018.04.002

Gulev, S. K., Thorne, P. W., Ahn, J., Dentener, F. J., Domingues, C. M., Gerland, S., ... Vose, R. S. (2021). Changing state of the climate system [Book Section]. In V. Masson-Delmotte et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (chap. 2). Cambridge University Press. Retrieved from https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Chapter_02.pdf

Harmsen, M. J. H. M., van Vuuren, D. P., van den Berg, M., Hof, A. F., Hope, C., Krey, V., ... Schaeffer, M. (2015). How well do integrated assessment models represent non-co2 radiative forcing? [Journal Article]. *Climatic Change*, *133*(4), 565-582. Retrieved from http://dx.doi.org/10.1007/s10584-015-1485-0 doi: 10.1007/s10584-015-1485-0

Huppmann, D., Kriegler, E., Krey, V., Riahi, K., Rogelj, J., Calvin, K., ... Zhang, R. (2019). *IAMC 1.5 °C Scenario Explorer and Data hosted by IIASA (release 2.0* [Generic]. Integrated Assessment Modeling Consortium & International Institute for Applied Systems Analysis. Retrieved from https://data.ene.iiasa.ac.at/iamc-1.5c-explorer doi: 10.5281/zenodo.3363345

Huppmann, D., Rogelj, J., Kriegler, E., Krey, V., & Riahi, K. (2018). A new scenario resource for integrated 1.5 °c research [Journal Article]. *Nature Climate Change*, *8*(12), 1027–1030. Retrieved from http://dx.doi.org/10.1038/s41558-018-0317-4 doi: 10.1038/s41558-018-0317-4

Jones, C. D., Frölicher, T. L., Koven, C., MacDougall, A. H., Matthews, H. D., Zickfeld, K., ... Burger, F. A. (2019). The zero emissions commitment model intercomparison project (zecmip) contribution to c4mip: quantifying committed climate changes following zero carbon emissions [Journal Article]. *Geoscientific Model Development*, *12*(10), 4375-4385. Retrieved from https://www.geosci-model-dev.net/12/4375/2019/http://dx.doi.org/10.5194/gmd-12-4375-2019 doi: 10.5194/gmd-12-4375-2019

Kikstra, J., Nicholls, Z., et al. (2022 (in prep.)). The ipcc sixth assessment report wgiii climate assessment of mitigation pathways: from emissions to global temperatures [Journal Article].

Lamboll, R. D., Nicholls, Z. R. J., Kikstra, J. S., Meinshausen, M., & Rogelj, J. (2020). Silicone v1.0.0: an open-source python package for inferring missing emissions data for climate change research [Journal Article]. *Geoscientific Model Development*, *13*(11), 5259-5275. Retrieved from https://gmd.copernicus.org/articles/13/5259/2020/http://dx.doi.org/10.5194/gmd-13-5259-2020 doi: 10.5194/gmd-13-5259-2020

Leach, N. J., Millar, R. J., Haustein, K., Jenkins, S., Graham, E., & Allen, M. R. (2018). Current level and rate of warming determine emissions budgets under ambitious mitigation [Journal Article]. *Nature Geoscience*, *11*(8), 574.

Lee, J. Y., Marotzke, J., Bala, G., Cao, L., Corti, S., Dunne, J. P., … Zhou, T. (2021). Future global climate: Scenario-based projections and near-term information [Book Section]. In V. Masson-Delmotte et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (chap. 4). Cambridge University Press. Retrieved from https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Chapter_04.pdf

MacDougall, A. H., Frölicher, T. L., Jones, C. D., Rogelj, J., Matthews, H. D., Zickfeld, K., … Ziehn, T. (2020). Is there warming in the pipeline? a multi-model analysis of the zero emissions commitment from co2 [Journal Article]. *Biogeosciences*, *17*(11), 2987-3016. Retrieved from https://bg.copernicus.org/articles/17/2987/2020/http://dx.doi.org/10.5194/bg-17-2987-2020 doi: 10.5194/bg-17-2987-2020

Meinshausen, M., Meinshausen, N., Hare, W., Raper, S. C. B., Frieler, K., Knutti, R., … Allen, M. R. (2009). Greenhouse-gas emission targets for limiting global warming to 2°c [Journal Article]. *Nature*, *458*(7242), 1158-1162. Retrieved from https://doi.org/10.1038/nature08017http://dx.doi.org/10.1038/nature08017 doi: 10.1038/nature08017

Meinshausen, M., Raper, S. C. B., & Wigley, T. M. L. (2011). Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, magicc6 - part 1: Model description and calibration [Journal Article]. *Atmospheric Chemistry and Physics*, *11*(4), 1417-1456. Retrieved from http://dx.doi.org/10.5194/acp-11-1417-2011 doi: 10.5194/acp-11-1417-2011

Millar, R. J., Nicholls, Z. R., Friedlingstein, P., & Allen, M. R. (2017). A modified impulse-response representation of the global near-surface air temperature and atmospheric concentration response to carbon dioxide emissions [Journal Article]. *Atmospheric Chemistry and Physics*, *17*(11), 7213-7228. Retrieved from https://www.atmos-chem-phys.net/17/7213/2017/http://dx.doi.org/10.5194/acp-17-7213-2017 doi: 10.5194/acp-17-7213-2017

Morice, C. P., Kennedy, J. J., Rayner, N. A., & Jones, P. D. (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The hadcrut4 data set [Journal Article]. *Journal of Geophysical Research: Atmospheres*, *117*(D8). Retrieved from http://dx.doi.org/10.1029/2011jd017187 doi: 10.1029/2011jd017187

Myhre, G., Shindell, D., Bréon, F. M., Collins, W., Fuglestvedt, J., Huang, J., ... Zhang, H. (2013). Anthropogenic and natural radiative forcing [Book Section]. In T. F. Stocker et al. (Eds.), *Climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change* (p. 659–740). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Retrieved from http://www.climatechange2013.orghttp://dx.doi.org/10.1017/CBO9781107415324.018 doi: 10.1017/CBO9781107415324.018

Nicholls, Z., & Lewis, J. (2021, 2021/3). *Reduced complexity model intercomparison project (rcmip) protocol* [Dataset]. Zenodo. Retrieved from https://doi.org/10.5281/zenodo.4589756http://dx.doi.org/10.5281/zenodo.4589756 doi: 10.5281/zenodo.4589756

Nicholls, Z., Lewis, J., Smith, C. J., Kikstra, J., Gieseke, R., & Willner, S. (2020). *OpenSCM-Runner: Thin wrapper to run simple climate models (emissions driven runs only)* [Software]. GitHub. Retrieved from https://github.com/openscm/openscm-runner

Nicholls, Z., Meinshausen, M., Lewis, J., Corradi, M. R., Dorheim, K., Gasser, T., ... Woodard, D. L. (2021). Reduced complexity model intercomparison project phase 2: Synthesizing earth system knowledge for probabilistic climate projections [Journal Article]. *Earth's Future*, *9*(6), e2020EF001900. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020EF001900 doi: https://doi.org/10.1029/2020EF001900

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 [Journal Article]. *Geoscientific Model Development*, *9*(9), 3461-3482. Retrieved from https://gmd.copernicus.org/articles/9/3461/2016/http://dx.doi.org/10.5194/gmd-9-3461-2016 doi: 10.5194/gmd-9-3461-2016

Riahi, K., Bertram, C., Huppmann, D., Rogelj, J., Bosetti, V., Cabardos, A.-M., … Zakeri, B. (2021). Cost and attainability of meeting stringent climate targets without overshoot [Journal Article]. *Nature Climate Change*, *11*(12), 1063-1069. Retrieved from https://doi.org/10.1038/s41558-021-01215-2 doi: 10.1038/s41558-021-01215-2

Riahi, K., Schaeffer, R., Arango, J., Calvin, K., Guivarch, C., Hasegawa, T., … Van Vuuren, D. P. (2022). Mitigation pathways compatible with long-term goals [Book Section]. In P. R. Shukla et al. (Eds.), *Ipcc, 2022: Climate change 2022: Mitigation of climate change. contribution of working group iii to the sixth assessment report of the intergovernmental panel on climate change* (chap. 3). Cambridge, UK and New York, NY, USA: Cambridge University Press. Retrieved from http://dx.doi.org/10.1017/9781009157926.005 doi: 10.1017/9781009157926.005

Rogelj, J., Hare, W., Lowe, J., Van Vuuren, D. P., Riahi, K., Matthews, B., … Meinshausen, M. (2011). Emission pathways consistent with a 2 °c global temperature limit [Journal Article]. *Nature Climate Change*, *1*(8), 413-418.

Rogelj, J., Huppmann, D., Krey, V., Riahi, K., Clarke, L., Gidden, M., … Meinshausen, M. (2019). A new scenario logic for the paris agreement long-term temperature goal [Journal Article]. *Nature*, *573*(7774), 357-363. Retrieved from http://dx.doi.org/10.1038/s41586-019-1541-4 doi: 10.1038/s41586-019-1541-4

Rogelj, J., Meinshausen, M., & Knutti, R. (2012). Global warming under old and new scenarios using ipcc climate sensitivity range estimates [Journal Article]. *Nature Climate Change*, *2*(4), 248-253. Retrieved from https://doi.org/10.1038/nclimate1385http://dx.doi.org/10.1038/nclimate1385 doi: 10.1038/nclimate1385

Rogelj, J., Shindell, D., Jiang, K., Fifita, S., Forster, P., Ginzburg, V., … Vilariño, M. V. (2018). Mitigation pathways compatible with 1.5°c in the context of sustainable development [Book Section]. In V. Masson-Delmotte et al. (Eds.), *Global Warming of 1.5°C: An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty.* Geneva, Switzerland: World Meteorological Organization. Retrieved from https://www.ipcc.ch/sr15/

Smith, C., Nicholls, Z. R. J., Armour, K., Collins, W., Forster, P., Meinshausen, M., ... Watanabe, M. (2021). The earth's energy budget, climate feedbacks, and climate sensitivity supplementary material [Book Section]. In V. Masson-Delmotte et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (chap. 7.SM). Cambridge University Press. Retrieved from https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Chapter_07_Supplementary_Material.pdf

Smith, C. J., Forster, P. M., Allen, M., Leach, N., Millar, R. J., Passerello, G. A., & Regayre, L. A. (2018). Fair v1.3: a simple emissions-based impulse response and carbon cycle model [Journal Article]. *Geoscientific Model Development*, *11*(6), 2273-2297. Retrieved from http://dx.doi.org/10.5194/gmd-11-2273-2018 doi: 10.5194/gmd-11-2273-2018

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 [Journal Article]. *Atmospheric Chemistry and Physics*, *20*(16), 9591-9618. Retrieved from https://acp.copernicus.org/articles/20/9591/2020/http://dx.doi.org/10.5194/acp-20-9591-2020 doi: 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 [Journal Article]. *Earth System Dynamics*, *12*(1), 253-293. Retrieved from https://esd.copernicus.org/articles/12/253/2021/http://dx.doi.org/10.5194/esd-12-253-2021 doi: 10.5194/esd-12-253-2021

Tebaldi, C., & Knutti, R. (2007). The use of the multi-model ensemble in probabilistic climate projections [Journal Article]. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *365*(1857), 2053-2075. Retrieved from https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2007.2076 doi: doi:10.1098/rsta.2007.2076

van Vuuren, D. P., Lowe, J., Stehfest, E., Gohar, L., Hof, A. F., Hope, C., … Plattner, G.-K. (2011). How well do integrated assessment models simulate climate change? [Journal Article]. *Climatic Change*, *104*(2), 255-285. Retrieved from http://dx.doi.org/10.1007/s10584-009-9764-2 doi: 10.1007/s10584-009-9764-2

Weyant, J. (2017). Some contributions of integrated assessment models of global climate change [Journal Article]. *Review of Environmental Economics and Policy*, *11*(1), 115–137. Retrieved from http://dx.doi.org/10.1093/reep/rew018 doi: 10.1093/reep/rew018

Supporting Information for 'Changes in IPCC scenario assessment emulators between SR1.5 and AR6 unravelled'

Z. Nicholls[1,2,3], M. Meinshausen[2,3], J. Lewis[1,2,3], C. J. Smith[1,4], P. M. Forster[4], J. Fuglestvedt[5], J. Rogelj[1,6,7], J. S. Kikstra[1,6,7], K. Riahi[1], E. Byers[1]

[1]International Institute for Applied System Analysis, IIASA, Laxenburg, Austria

[2]Climate & Energy College, School of Geography, Earth and Atmospheric Sciences, The University of Melbourne, Parkville, Victoria, Australia

[3]Climate Resource, Northcote, Victoria, Australia

[4]Priestley International Centre for Climate, University of Leeds, Leeds, United Kingdom

[5]CICERO, Oslo, Norway

[6]Centre for Environmental Policy, Imperial College London, London, United Kingdom

[7]Grantham Institute for Climate Change and the Environment, Imperial College London, United Kingdom

Contents of this file

Corresponding author: Z. Nicholls, Climate & Energy College, School of Geography, Earth and Atmospheric Sciences, The University of Melbourne, Parkville, Victoria, Australia (zebedee.nicholls@climate-energy-college.org)

1. Text S1: Changes in effective radiative forcing by species between MAGICC6 and the AR6-calibration of MAGICCv7.5.3

2. Text S2: Reasons for the higher warming projections in MAGICC6 compared to the projections from MAGICCv7.5.3 when it is calibrated to HadCRUT4.6.0.0, a rough proxy for the AR5 assessment

3. Figure S1: Difference between the emulator output underlying the SR1.5 scenario categorisation and alternate emulator output for exceedance probabilities of 1.5°C and 2°C warming for 2100 and peak

4. Figure S2: Projected median and upper 67% warming projections for SR1.5 scenarios are between MAGICC6, AR6-calibrated MAGICCv7.5.3, AR6-calibrated FaIR1.6 and the SR1.5 FaIR 1.3 version.

5. Figure S3: Difference between the emulator output underlying the SR1.5 scenario categorisation and alternate emulator output for median and 67^{th}-percentile projections for 2100 and peak temperatures

6. Figure S4: Relationship between exceedance probabilities and temperature projections across multiple emulators

7. Figure S5: Difference between the emulator output underlying the SR1.5 scenario categorisation and the RCMIP Phase 2 calibration of MAGICCv7.5.3 for median and 67^{th}-percentile projections for 2100 and peak temperatures

8. Figure S6: Difference between the emulator output underlying the SR1.5 scenario categorisation and the RCMIP Phase 2 calibration of MAGICCv7.5.3 for median and

67$^{\text{th}}$-percentile projections for 2100 and peak temperatures shown as histogram rather than scatter plot

9. Figure S7-S8, S12-S15: Comparison of output from the emulator underlying the SR1.5 scenario categorisation (MAGICC6) and the AR6-calibration of MAGICCv7.5.3

10. Figure S9: CO_2 and CH_4 concentration projections from the AR6-calibrated MAGICCv7.5.3 and FaIRv1.6.2

11. Figure S10: Difference between the AR6-calibrated MAGICCv7.5.3 and AR6-calibrated FaIRv1.6.2 emulators for median and 67$^{\text{th}}$-percentile projections for 2100 and peak temperatures

12. Figure S11: Difference between the AR6-calibrated MAGICCv7.5.3 and AR6-calibrated FaIRv1.6.2 emulators for median and 67$^{\text{th}}$-percentile projections for 2100 and peak temperatures shown as histogram rather than scatter plot

13. Figure S12-S17: Comparison of output from the emulator underlying the SR1.5 scenario categorisation (MAGICC6) and the RCMIP Phase 2 calibration of MAGICCv7.5.3

14. Figure S18-S21: Comparison of more output from the emulator underlying the SR1.5 scenario categorisation (MAGICC6) and the AR6-calibration of MAGICCv7.5.3

15. Table S1: Emulator model versions, their key characteristics and the main purpose for their use in this study

Text S1.

Here we discuss the reasons for the higher warming projections in MAGICC6 compared to the projections from MAGICCv7.5.3 when it is calibrated to HadCRUT4.6.0.0 (as used in RCMIP Phase 2), a rough proxy for the AR5 assessment (Supplementary Figure

S12). MAGICC6 generally has lower anthropogenic effective radiative forcing than MAGICCv7.5.3 in its RCMIP Phase 2 configuration (Supplementary Figure S13). Therefore, differences in the parameterisations that link emissions and effective radiative forcing are not the reason for MAGICC6's higher warming projections and we hence conclude that differences in model calibration, particularly transient climate response (TCR), explain the difference instead.

We next discuss some of the changes in effective radiative forcing of different species and why these changes are approximately zero. As the RCMIP Phase 2 calibration of MAGICCv7.5.3 reflects literature which is more recent than AR5, this comparison provides an insight into some, but not all, the changes in effective radiative forcing between AR5 and AR6.

CO_2 effective radiative forcing is again relatively unchanged between the two emulators (Supplementary Figure S14), reflecting the close agreement between the CO_2 ERF estimates used for the RCMIP Phase 2 tuning (Smith et al., 2020) and the AR5 assessment (Myhre et al., 2013). The differences in CO_2 effective radiative forcing are again largely driven by differences in the carbon cycle response (Supplementary Figure S15), with a similar range of differences on either side of the median.

Like for the AR6-calibration of MAGICCv7.5.3 compared to MAGICC6, aerosol effective radiative forcing is more sensitive to changes in aerosol emissions (Supplementary Figure S16). These changes arise because the ERF used in RCMIP Phase 2 (Smith et al., 2020) assesses aerosol forcing for 1750-2014 to be -1.01 W m^{-2}, an approximately 10% increase on AR5's 1750-2011 assessment (Myhre et al., 2013) of -0.9 W m^{-2}.

Finally, methane effective radiative forcing increases (Supplementary Figure S17). The increase follows the upwards revision presented in (Etminan et al., 2016, without the inclusion of rapid adjustments, which is not included in the RCMIP Phase 2 calibration of MAGICCv7.5.3).

Taking all the changes together, we see that CO_2 effective radiative forcing is largely unchanged, aerosol effective radiative forcing is initially more negative before being more positive and methane effective radiative forcing is more positive. In sum, the combination of these and other changes leads to an increase of 0.0 - 0.2 W m^{-2} in anthropogenic effective radiative forcing. We note again that these changes are scenario dependent (Supplementary Figure S9) and bespoke analysis is required to understand each specific scenario's drivers.

Text S2.

Here we consider the changes in effective radiative forcing of different species and why these changes cancel out to approximately zero change (with a slight skew towards an increase) in total anthropogenic effective radiative forcing between MAGICC6 and the AR6-calibration of MAGICCv7.5.3.

CO_2 effective radiative forcing has a median change of approximately zero between MAGICC6 and AR6-calibrated MAGICCv7.5.3 (Supplementary Figure S18). This almost zero change is the result of a slight increase in the assessment of CO_2 effective radiative forcing (for a given atmospheric CO_2 concentration) and a decrease in projected atmospheric CO_2 concentrations (Supplementary Figure S19). It is also worth noting that, while the median is unchanged, there is a range of differences due to differing climate-carbon feedbacks (particularly those driven by the different temperature projections).

Aerosol effective radiative forcing is more sensitive to changes in aerosol emissions. As a result, the negative aerosol effective radiative forcing is stronger today but reduced in the future as aerosol emissions drop (differences are taken as the AR6-calibration of MAGICCv7.5.3 minus the MAGICC6 output, Supplementary Figure S20). These changes arise because the AR6 assessment of ERF (Forster et al., 2021) assesses aerosol forcing for 2005-2014 relative to 1750 to be -1.3 W m^{-2}, compared to AR5's assessment (Myhre et al., 2013) for 2011 relative to 1750 of -0.9 W m^{-2}.

Finally, methane effective radiative forcing increases in the short-term before returning to a median difference of zero with a range of -0.075 W m^{-2} to 0.1 W m^{-2} (Supplementary Figure S21). The increase follows the upwards revision presented in Etminan et al. (2016), tempered by AR6's inclusion of rapid adjustments in the assessment of methane effective radiative forcing in line with Smith et al. (2020). The increased sensitivity to methane emissions (via methane concentrations) also leads to a more pronounced reduction in methane effective radiative forcing after its peak.

In summary, CO_2 effective radiative forcing has a median change of around zero with a slight skew towards a decrease, aerosol effective radiative forcing is initially more negative before being more positive and methane effective radiative forcing is initially more positive before being around zero or slightly more negative. In general, the combination of these and other changes lead to the increase in anthropogenic effective radiative forcing. However, the scenario dependence of these changes should also be noted as individual scenarios might not match well with the overall trends discussed here (Supplementary Figure S8).

References

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 [Journal Article]. *Geophysical Research Letters*, *43*(24), 12,614-12,623. Retrieved from http://dx.doi.org/10.1002/2016gl071930 doi: 10.1002/2016gl071930

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J. L., Frame, D., ... Zhang, H. (2021). The earth's energy budget, climate feedbacks, and climate sensitivity [Book Section]. In V. Masson-Delmotte et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (chap. 7). Cambridge University Press. Retrieved from https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Chapter_07.pdf

Myhre, G., Shindell, D., Bréon, F. M., Collins, W., Fuglestvedt, J., Huang, J., ... Zhang, H. (2013). Anthropogenic and natural radiative forcing [Book Section]. In T. F. Stocker et al. (Eds.), *Climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change* (p. 659–740). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Retrieved from http://www.climatechange2013.orghttp://dx.doi.org/10.1017/

CBO9781107415324.018 doi: 10.1017/CBO9781107415324.018

Nicholls, Z., Meinshausen, M., Lewis, J., Corradi, M. R., Dorheim, K., Gasser, T., ... Woodard, D. L. (2021). Reduced complexity model intercomparison project phase 2: Synthesizing earth system knowledge for probabilistic climate projections [Journal Article]. *Earth's Future*, *9*(6), e2020EF001900. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020EF001900 doi: https://doi.org/10.1029/2020EF001900

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 [Journal Article]. *Atmospheric Chemistry and Physics*, *20*(16), 9591-9618. Retrieved from https://acp.copernicus.org/articles/20/9591/2020/http://dx.doi.org/10.5194/acp-20-9591-2020 doi: 10.5194/acp-20-9591-2020

do not specify file extension

Table S1: Emulator model versions, their key characteristics and the main purpose for their use in this study.

Model version	**Key characteristic**	**Purpose in this study**
SR1.5 MAGICC6	MAGICC6 with its calibration and input data as used to categorise SR1.5 scenarios for the IPCC SR1.5 report	Our baseline reference in this study against which we express most model variations
RCMIP Phase 2 MAGICCv7.5.3 (also called "AR5-like" MAGICCv7.5.3)	MAGICCv7.5.3 as used in RCMIP2, calibrated to match scientific understanding approximately in line with knowledge at the time of SR1.5 (for full details, see Nicholls et al. (2021)). This calibration approximately reflects the science of AR5 and SR1.5, but not the science of AR6. The 'RCMIP Phase 2 MAGICCv7.5.3' results provide an indication of what SR1.5 MAGICC6 could have been with updated calibration efforts that would better match the physical science at the time of SR1.5. The key difference to SR1.5 MAGICC6 is a slightly lower TCR - more in line with the AR5 TCR assessment (which has been retained for AR6).	Provides a MAGICC calibration that approximates the projections might have been made in SR1.5, had it been re-calibrated at the time.
AR6-calibrated MAGICCv7.5.3 (also called simply 'MAGICCv7.5.3')	MAGICCv7.5.3 calibrated to the updated science of AR6 (see Cross-chapter Box 7.1 of Forster et al. (2021))	The reference used for scenario categorisation in AR6. A comparison with 'SR1.5 MAGICC6' indicates the full difference in the climate assessment between SR1.5 and AR6. A comparison with 'RCMIP Phase 2 MAGICCv7.5.3' approximates the effect of the physical science update from AR5 to AR6.

Continued on next page

Table S1 – *Continued from previous page*

Model version	**Key characteristic**	**Purpose in this study**
SR1.5 FaIR1.3	The model used for sensitivity exploration (beyond MAGICC6) as part of the IPCC SR1.5 assessment. The differences between SR1.5 MAGICC6 and SR1.5 FaIR 1.3 were largely unexplained at the time.	Showing, for historical context, the relatively large difference that existed between two key emulators (SR1.5 MAGICC6 and SR1.5 FaIR 1.3) at the time of the IPCC SR1.5.
AR6-calibrated FaIRv1.6.2	FaIRv1.6.2 calibrated to the updated science of AR6 (see Cross-chapter Box 7.1 of Forster et al. (2021))	The difference between AR6 MAGICCv7.5.3 and AR6 FaIRv1.6.2 highlights the uncertainty that arises due to emulator representation of AR6 science. Both AR6 MAGICCv7.5.3 and AR6 FaIRv1.6.2 were assessed as representing WG1 well, so we cannot say with confidence which one is more correct. The difference is much reduced compared to the difference between SR1.5 MAGICC6 and SR1.5 FaIR1.3 (where, had time allowed, deeper investigation may have revealed which model was more in line with the science available at the time).

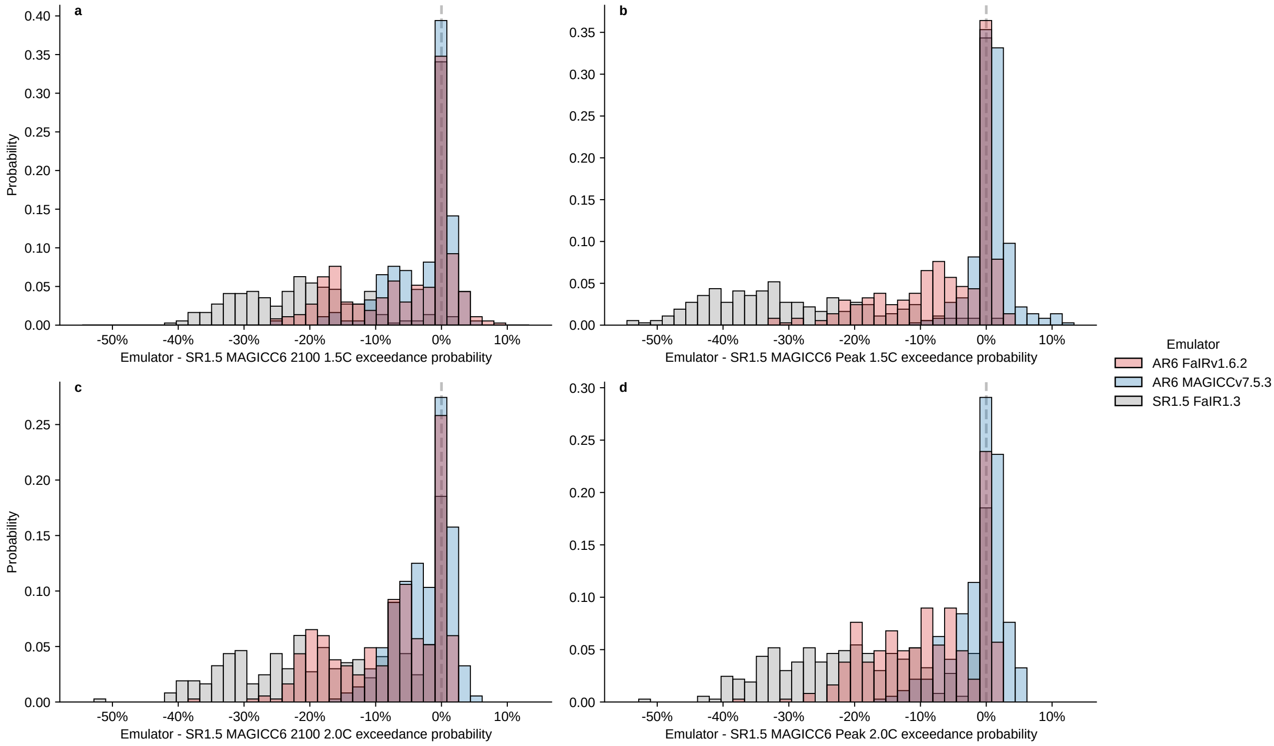

Figure S1. Difference between the emulator output underlying the SR1.5 scenario categorisation and alternate emulator output for exceedance probabilities of 1.5°C (top row) and 2°C warming (bottom row) for 2100 (left column) and peak (i.e. maximum between 2010 and 2100, right column). The data is the same as Figure 1 but has been plotted as a histogram of differences rather than a scatter plot. Scenarios with exceedance probabilities of close to 100% cannot have large changes in exceedance probability (because exceedance probability is capped at 100%), hence the large number of results with no change (zero is marked by the dashed grey line). The difference from the AR6-calibrated MAGICCv7.5.3 emulator is shown in dark blue, from the the AR6-calibrated FaIRv1.6.2 emulator is shown in red and from the alternate SR1.5 emulator FaIR1.3 is shown in grey.

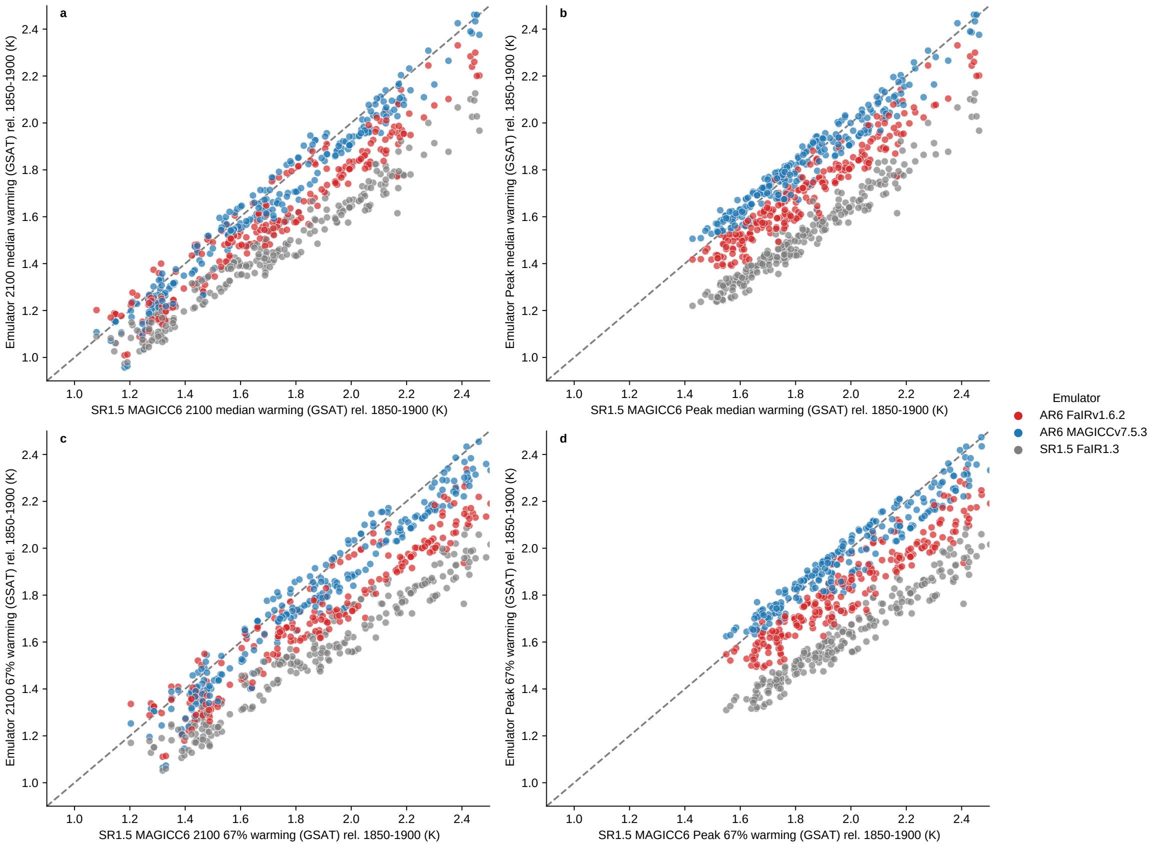

Figure S2. Projected median and upper 67% warming projections for SR1.5 scenarios are very similar between MAGICC6 and the AR6-calibrated MAGICCv7.5.3 (blue dots), slightly higher compared to AR6-calibrated FaIR1.6 (red dots) or the SR1.5 FaIR 1.3 version (grey dots). Emulator output underlying the SR1.5 scenario categorisation (x-axes) and the AR6-calibrated emulators (y-axes) for median (top row) and 67$^{\text{th}}$-percentile (bottom row) projections for 2100 (left column) and peak (right column) temperatures. Scenarios on the diagonal (grey dashed line) did not change temperature projections. Scenarios that exhibit lower global-mean temperatures according to the AR6-calibrated emulators are shown below the diagonal. The AR6-calibrated emulator MAGICCv7.5.3 is shown in dark blue, the AR6-calibrated emulator FaIRv1.6.2 is shown in red is and the grey dots indicate the emulator which was considered as an alternative in SR1.5, namely FaIR1.3.

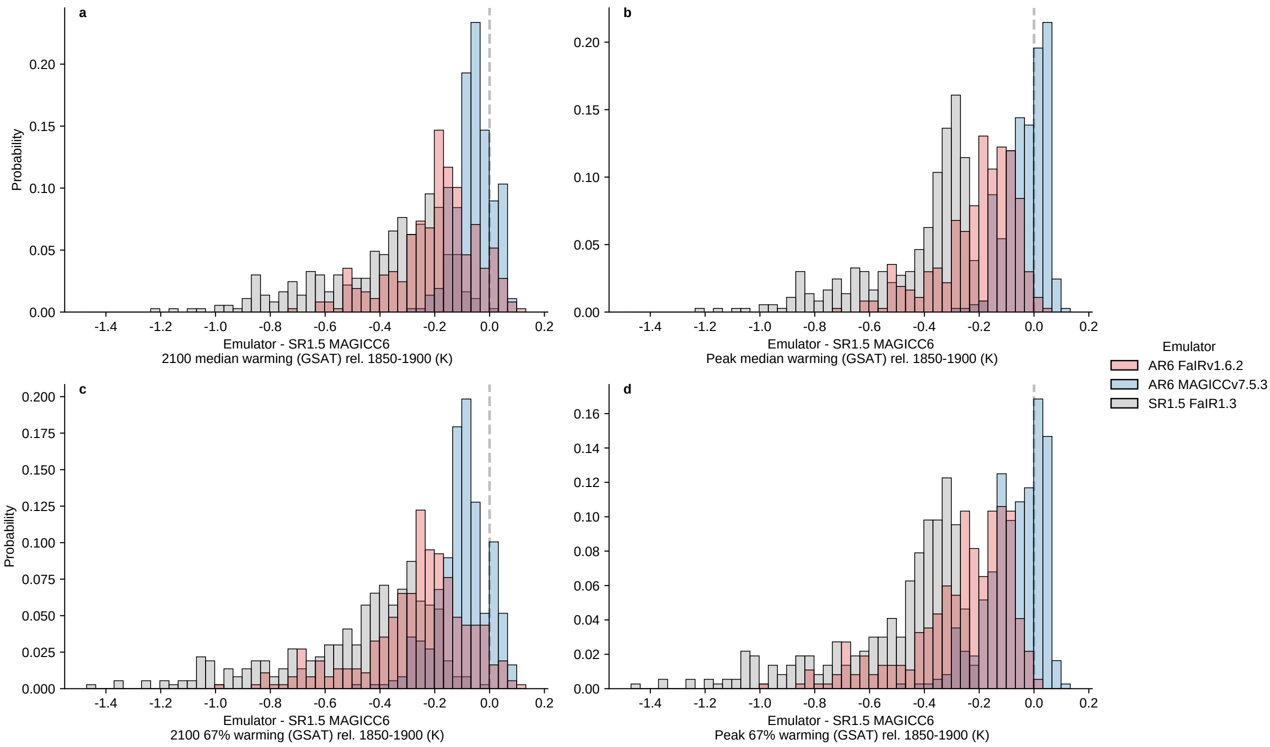

Figure S3. Difference between the emulator output underlying the SR1.5 scenario categorisation and alternate emulator output for median (top row) and 67th-percentile (bottom row) projections for 2100 (left column) and peak (right column) temperatures. The data is the same as Supplementary Figure S2 but has been plotted as a histogram rather than a scatter plot. The difference from the AR6-calibrated MAGICCv7.5.3 emulator is shown in dark blue, from the the AR6-calibrated FaIRv1.6.2 emulator is shown in red and from the alternate SR1.5 emulator FaIR1.3 is shown in grey.

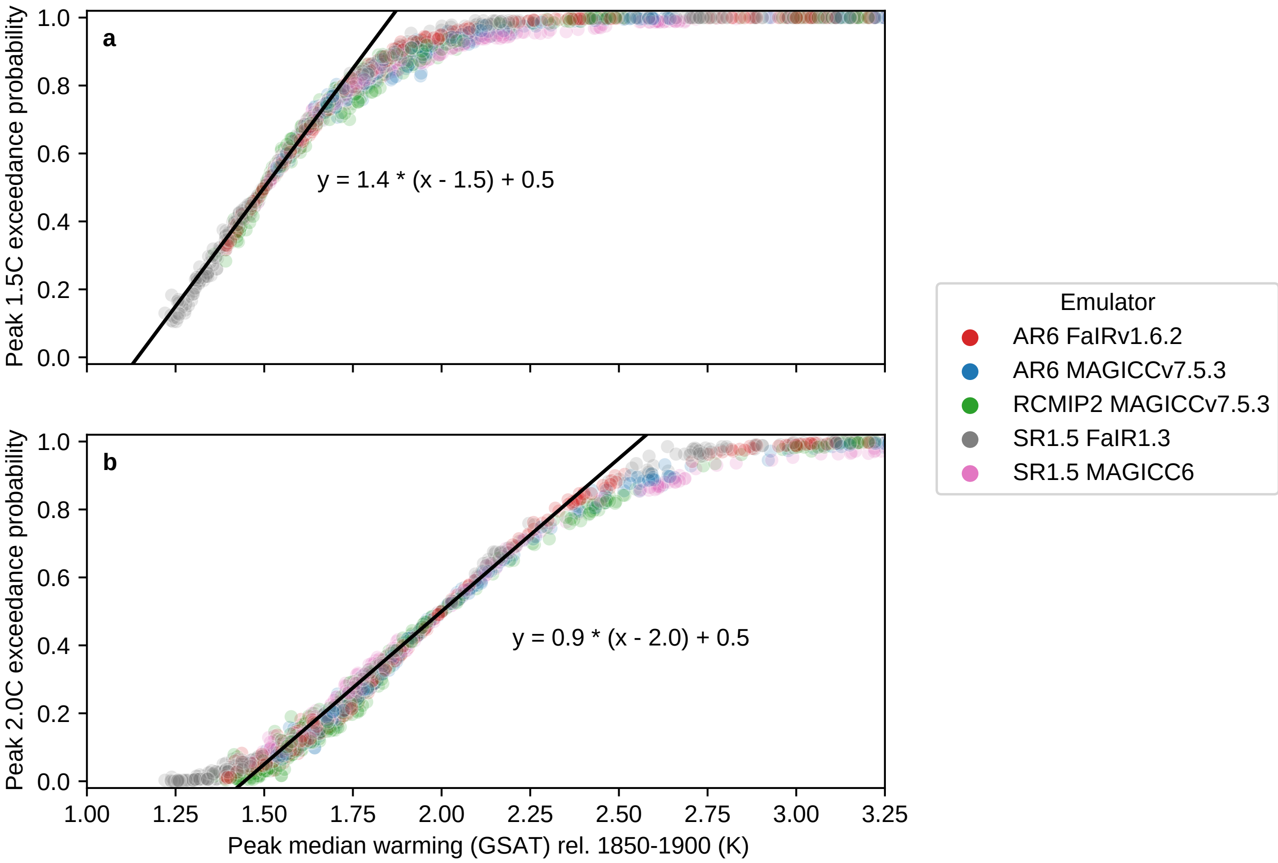

Figure S4. Relationship between exceedance probabilities and temperature projections across multiple emulators. By definition, when the median temperature is 1.5°C (2.0°C), the 1.5°C (2.0°C) exceedance probability is 50%. For every 0.01°C increase in median temperature, 1.5°C exceedance probability increases by around 1.4% and 2.0°C exceedance probability increases by around 0.9%. The difference in gradient is because the uncertainty increases as the median temperature projection increases i.e. we have wider distributions once we get to around 2.0°C warming. The relationship is remarkably consistent across the emulators, with some small variations that become more noticeable as we get into the tails of the distributions i.e. as the median temperature moves away from the exceedance threshold of interest.

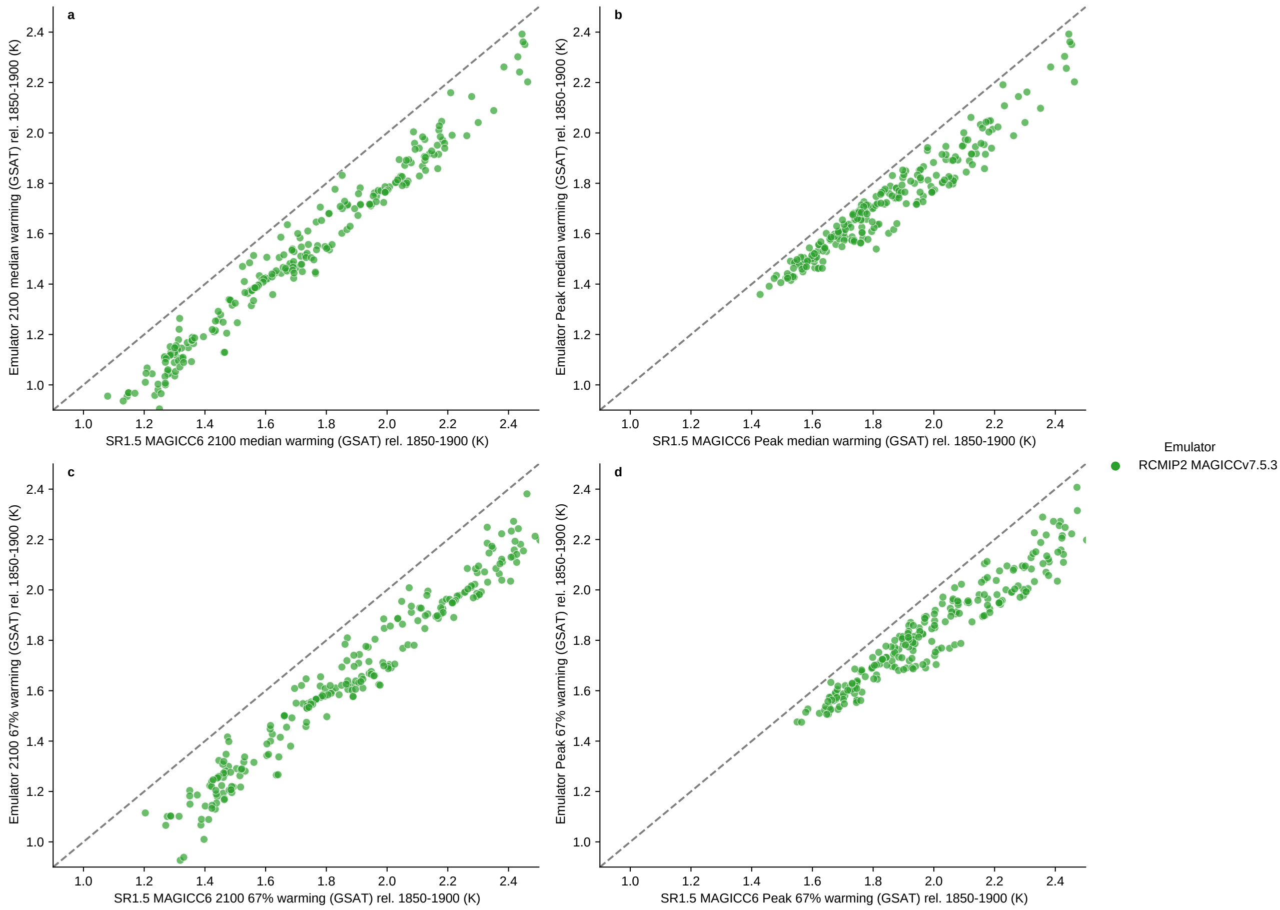

Figure S5. Emulator output underlying the SR1.5 scenario categorisation (x-axes) and the RCMIP Phase 2 calibration of MAGICCv7.5.3 (y-axes) for median (top row) and 67$^{\text{th}}$-percentile (bottom row) projections for 2100 (left column) and peak (right column) temperatures. If temperature projections hadn't changed for a scenario, that scenario would be shown on the diagonal (grey dashed line). Scenarios that exhibit lower global-mean temperatures according to the RCMIP Phase 2 calibration of MAGICCv7.5.3 are shown below the diagonal. The difference between the RCMIP Phase 2 calibrated MAGICCv7.5.3 and MAGICC6 provides an approximate quantification of the change in projections which results from improved emulator calibrations. The figure is the same as Supplementary Figure S2 but for a different combination of models.

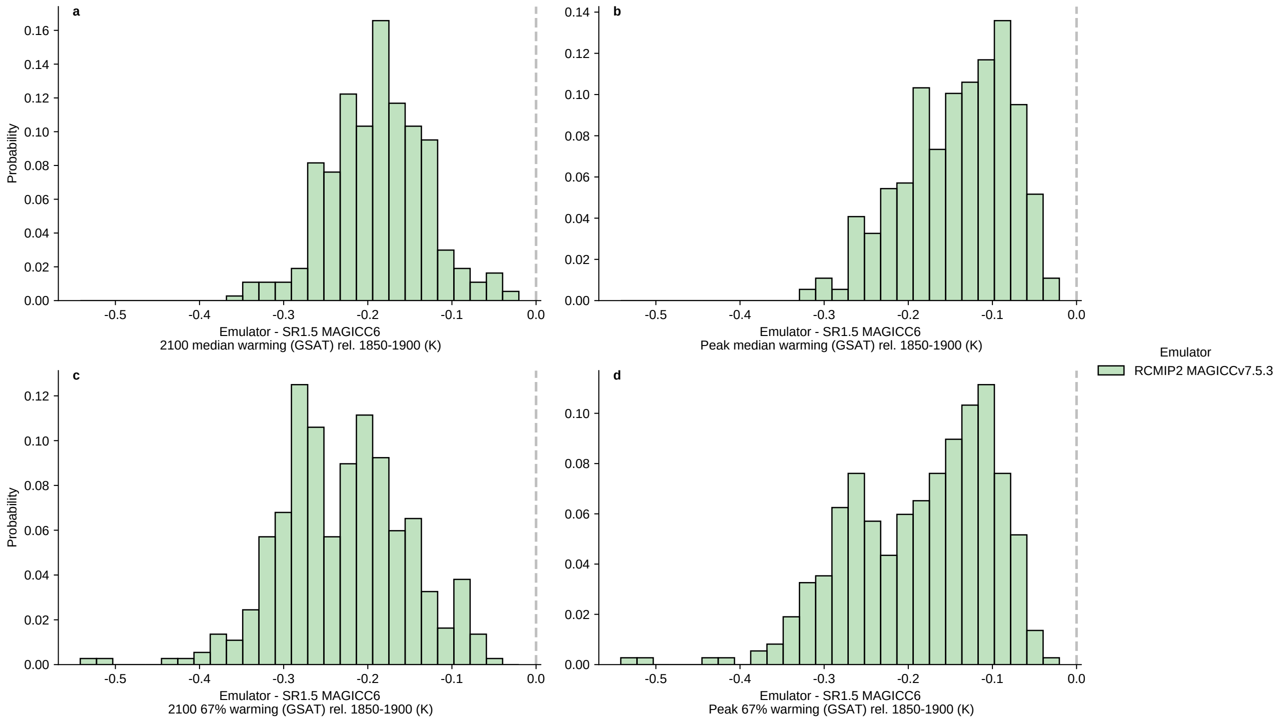

Figure S6. Difference between the emulator output underlying the SR1.5 scenario categorisation and the RCMIP Phase 2 calibration of MAGICCv7.5.3 for median (top row) and 67th-percentile (bottom row) projections for 2100 (left column) and peak (right column) temperatures. The data is the same as Supplementary Figure S5 but has been plotted as a histogram rather than a scatter plot.

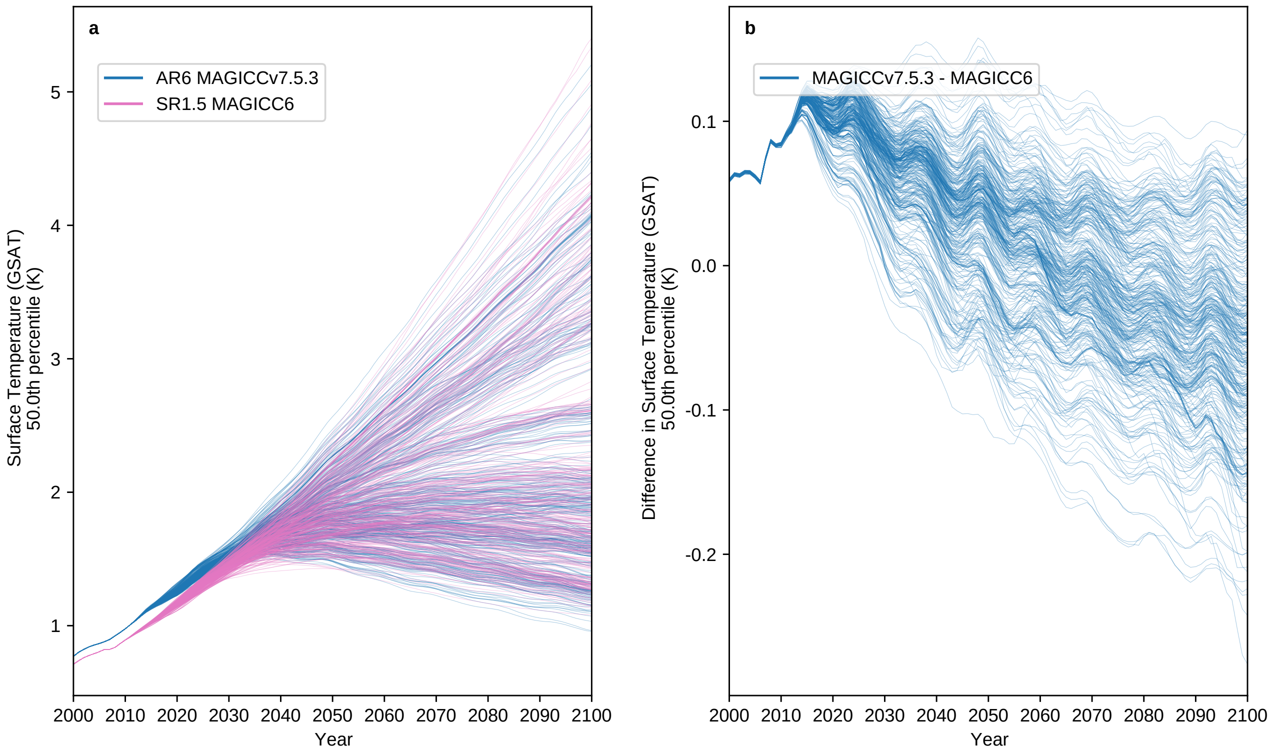

Figure S7. Comparison of output from the emulator underlying the SR1.5 scenario categorisation (MAGICC6) and the AR6-calibration of MAGICCv7.5.3. a) Timeseries of median surface air temperature (GSAT). b) Difference (MAGICCv7.5.3 - MAGICC6) in median surface air temperature (GSAT).

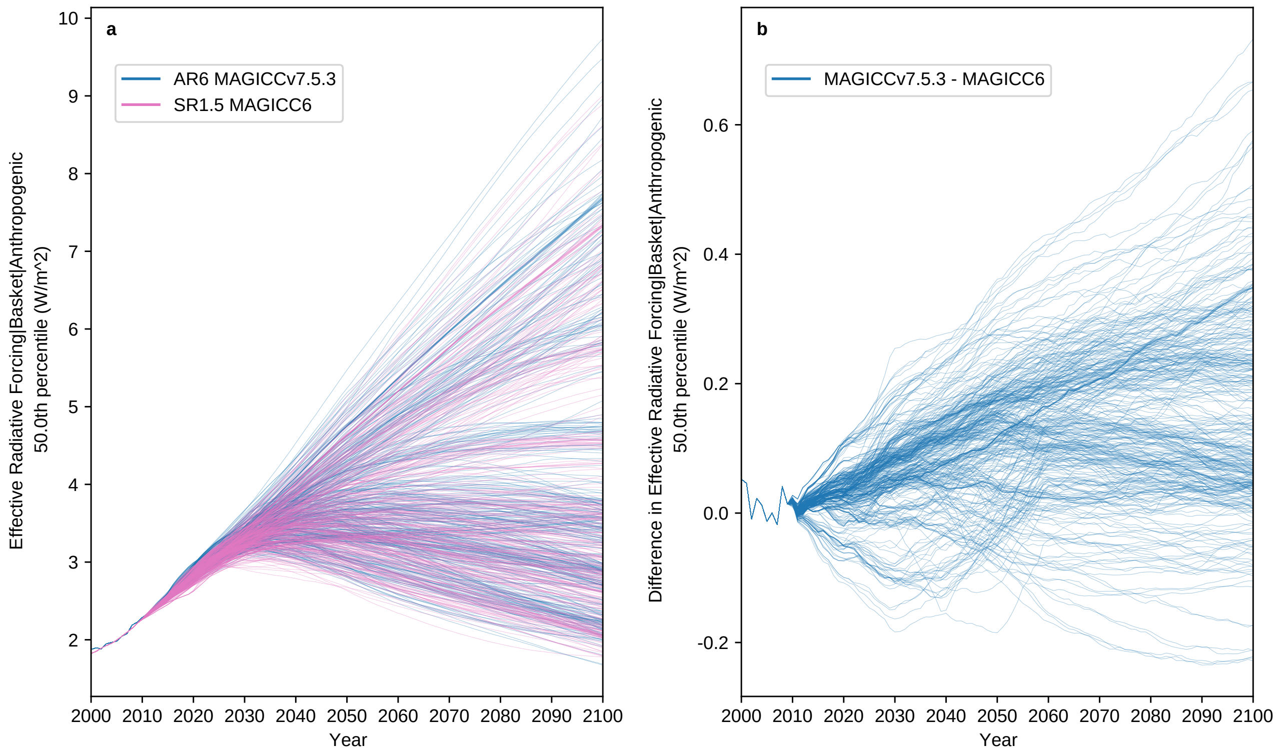

Figure S8. Comparison of output from the emulator underlying the SR1.5 scenario categorisation (MAGICC6) and the AR6-calibration of MAGICCv7.5.3. a) Timeseries of anthropogenic effective radiative forcing. b) Difference (AR6-calibrated MAGICCv7.5.3 - MAGICC6) in anthropogenic effective radiative forcing.

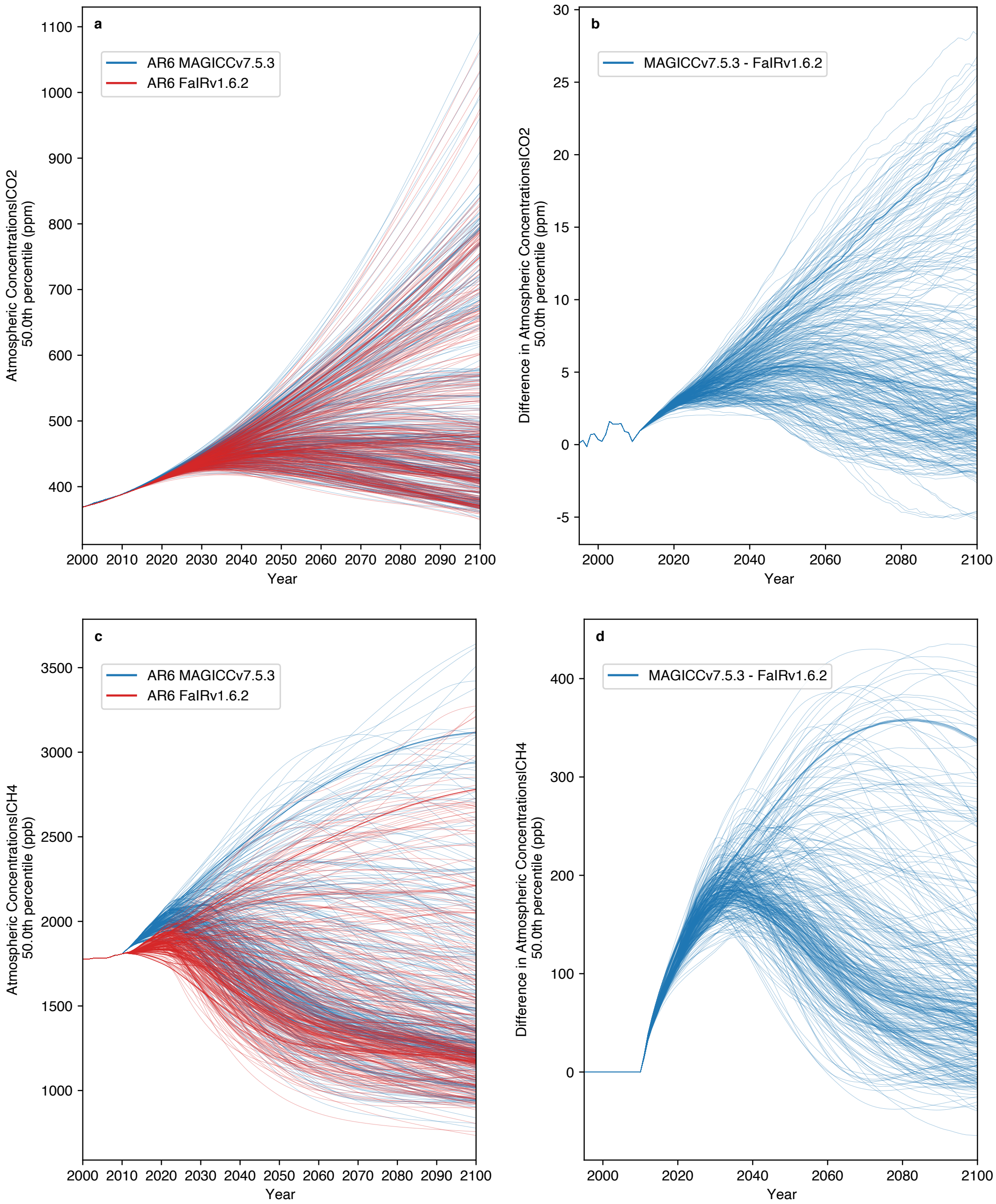

Figure S9. CO_2 (top row) and CH_4 (bottom row) concentration projections from the AR6-calibrated MAGICCv7.5.3 and FaIRv1.6.2. In the first column we show absolute projections from each emulator and in the second column we show the difference between the two emulators. MAGICCv7.5.3 projects higher concentrations (in general) than FaIRv1.6.2 which is part of the reason for differences in emissions-driven runs like those performed in WG3.

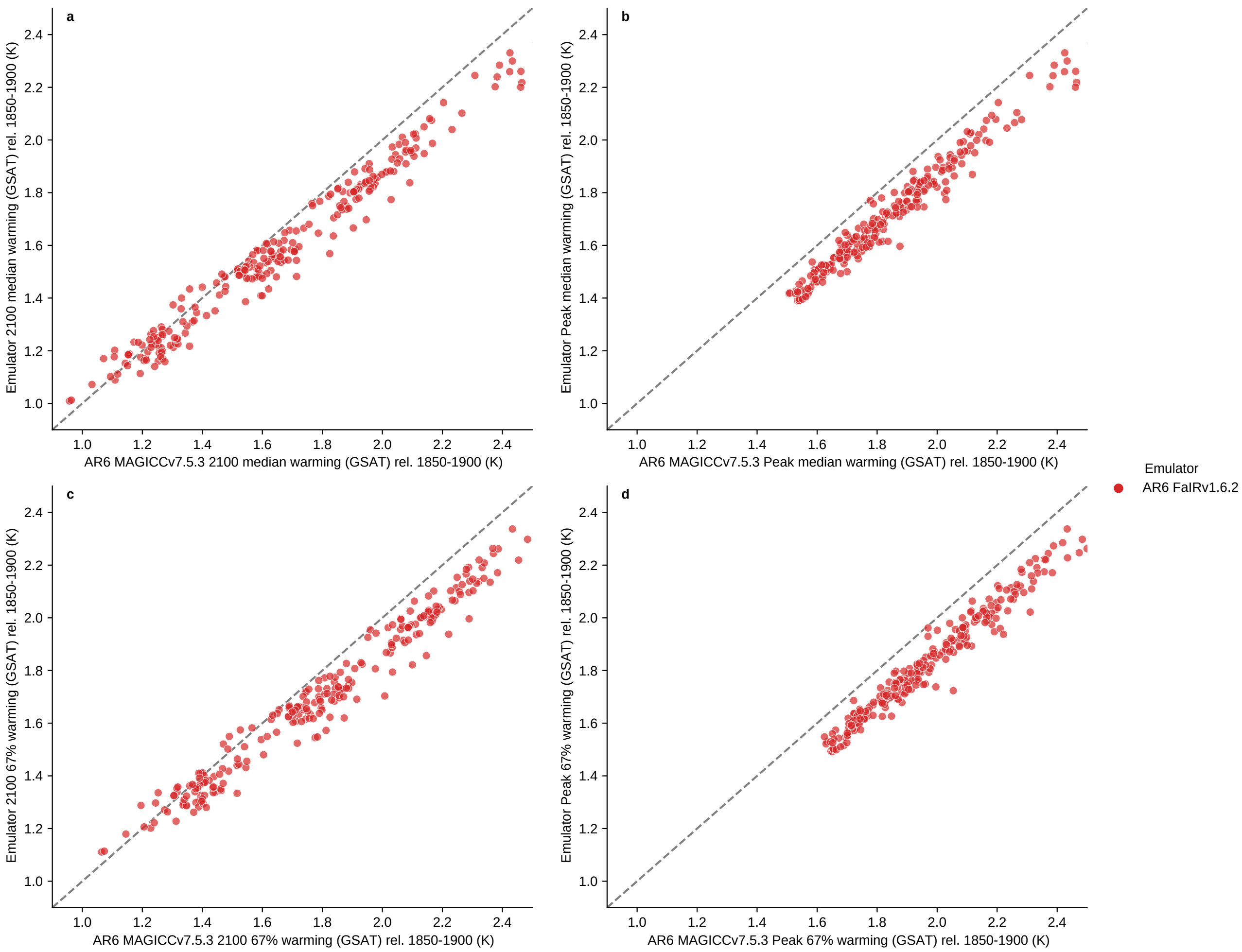

Figure S10. As in Supplementary Figure S5 but for the AR6-calibration of MAGICCv7.5.3 (x-axes) and the AR6-calibration of FaIRv1.6.2.

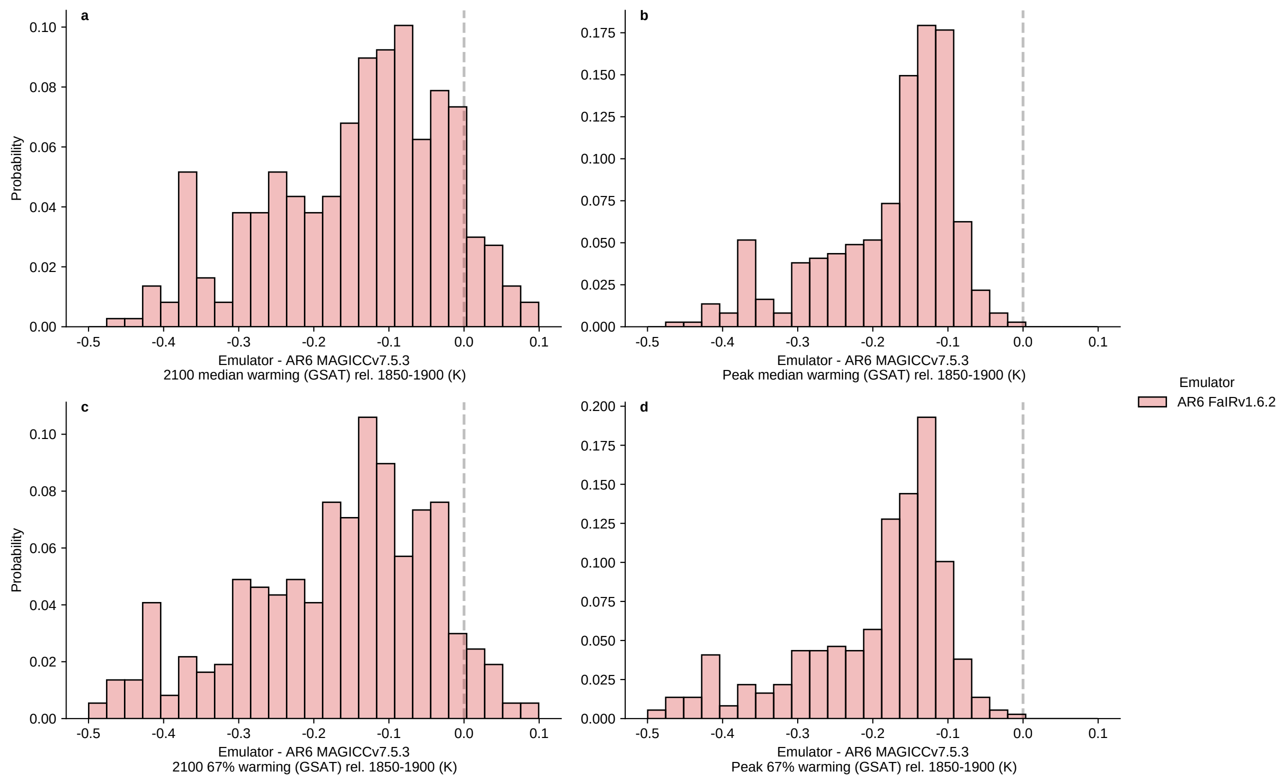

Figure S11. As in Supplementary Figure S6 but for the AR6-calibration of MAGICCv7.5.3 (x-axes) and the AR6-calibration of FaIRv1.6.2. The data is the same as Supplementary Figure S10 but has been plotted as a histogram rather than a scatter plot.

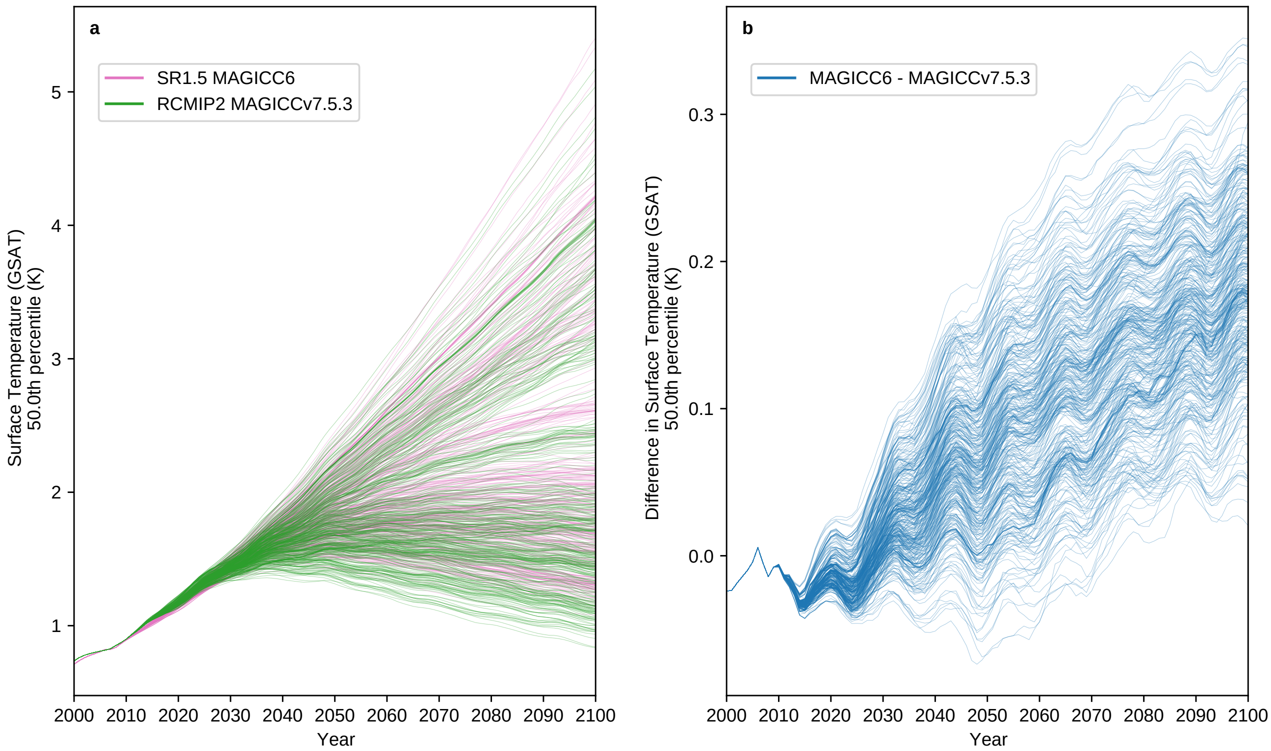

Figure S12. Comparison of output from the emulator underlying the SR1.5 scenario categorisation (MAGICC6) and the RCMIP Phase 2 calibration of MAGICCv7.5.3. a) time series of median surface air temperature (GSAT). b) Difference (MAGICCv7.5.3 - MAGICC6) in median surface air temperature (GSAT).

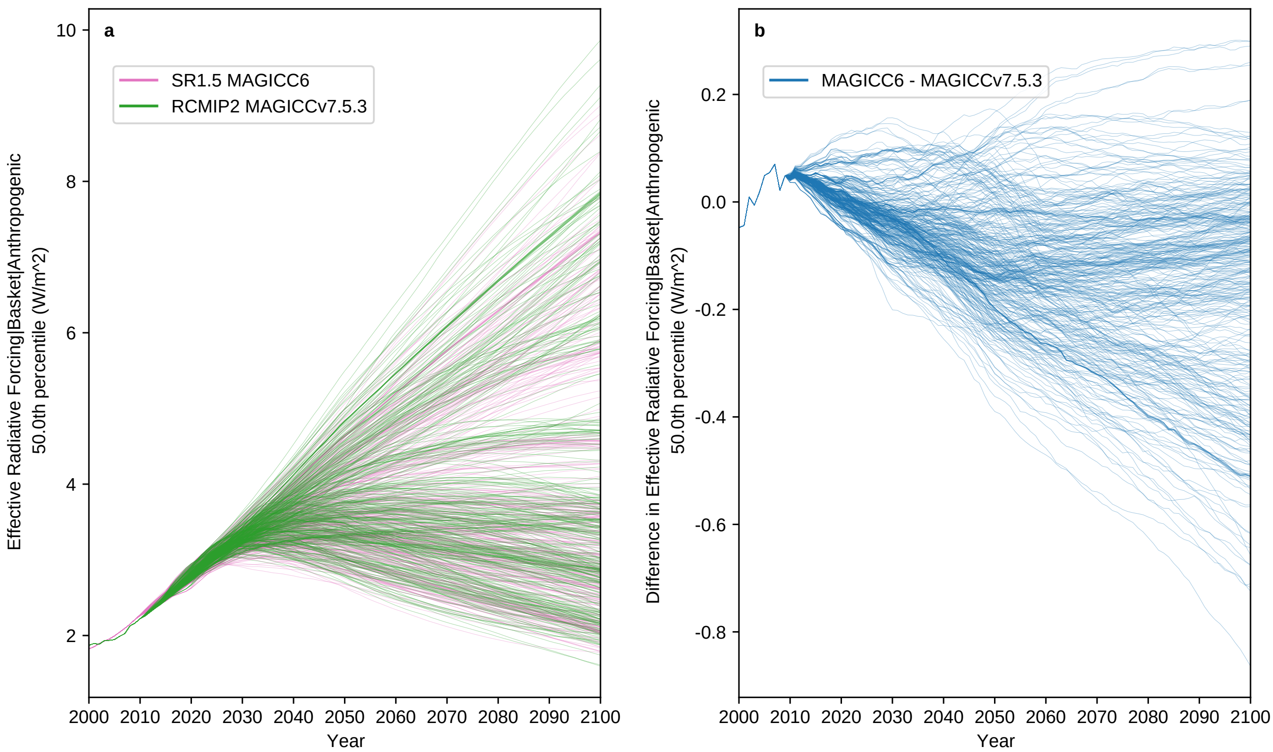

Figure S13. As in Supplementary Figure S12, except for median anthropogenic effective radiative forcing.

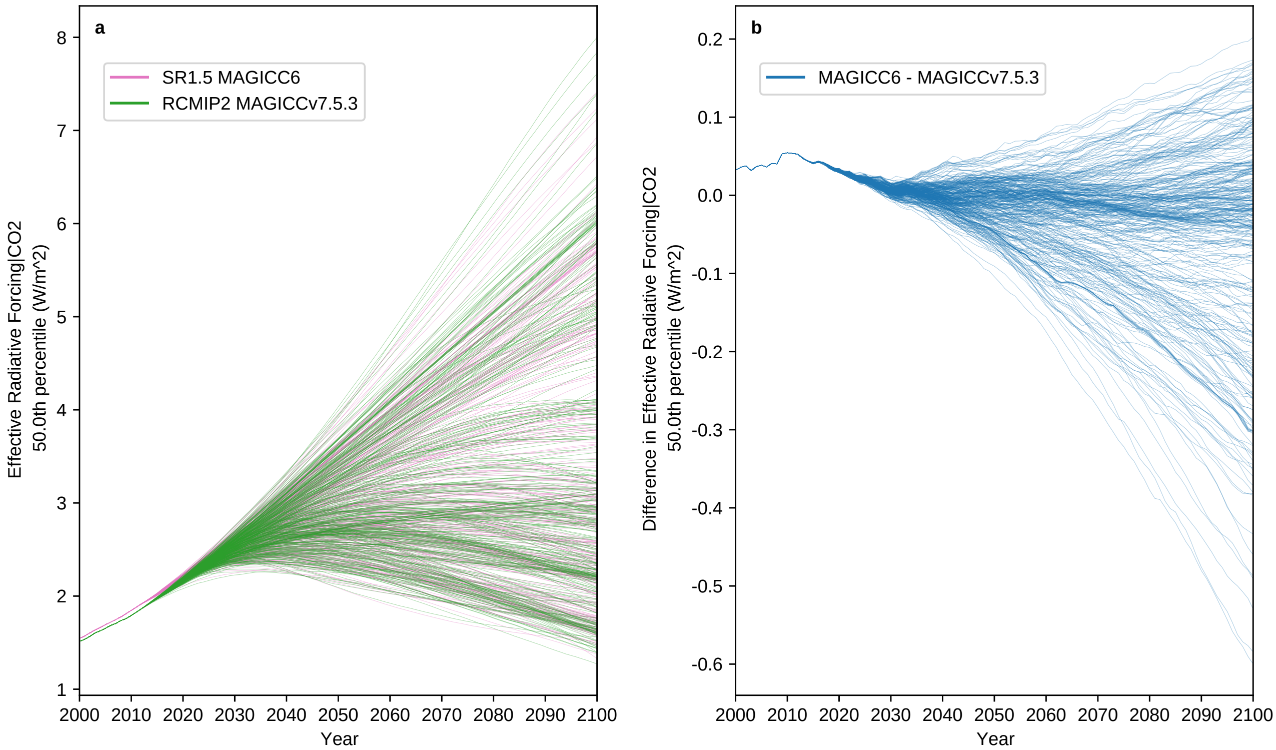

Figure S14. As in Supplementary Figure S12, except for median CO_2 effective radiative forcing.

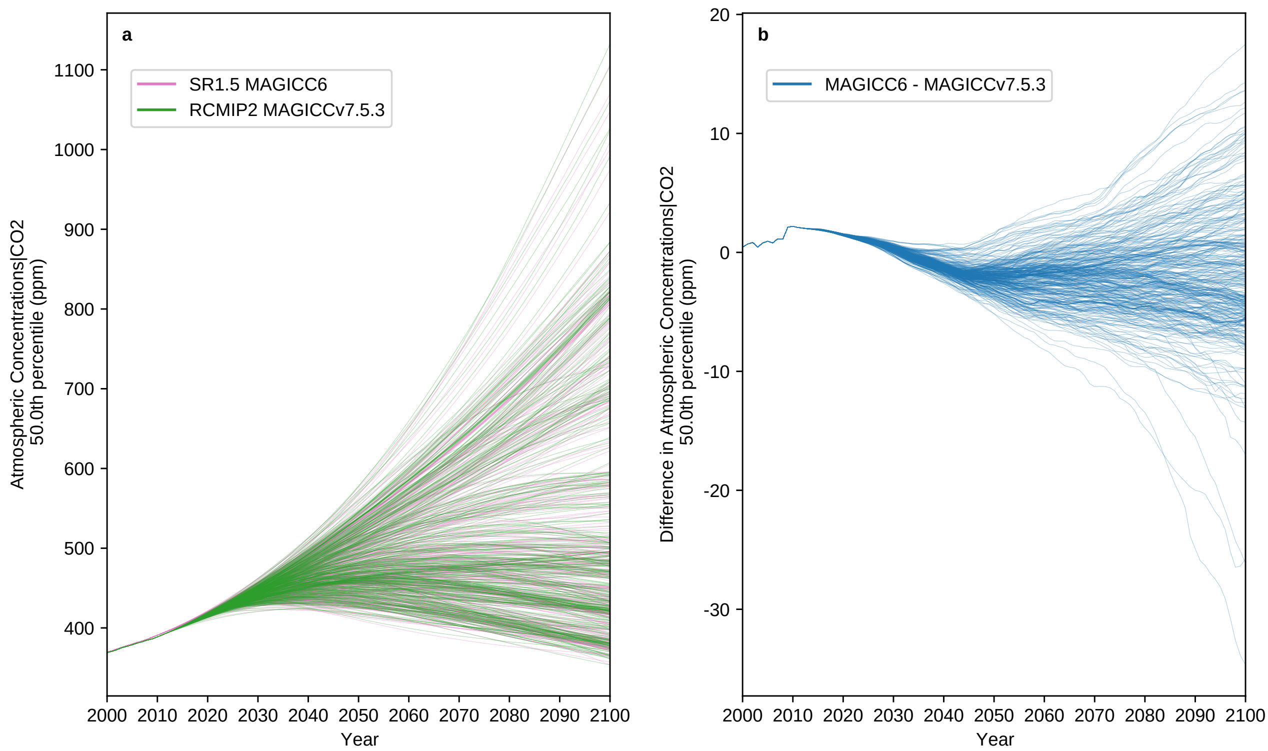

Figure S15. As in Supplementary Figure S12, except for median CO_2 atmospheric concentrations.

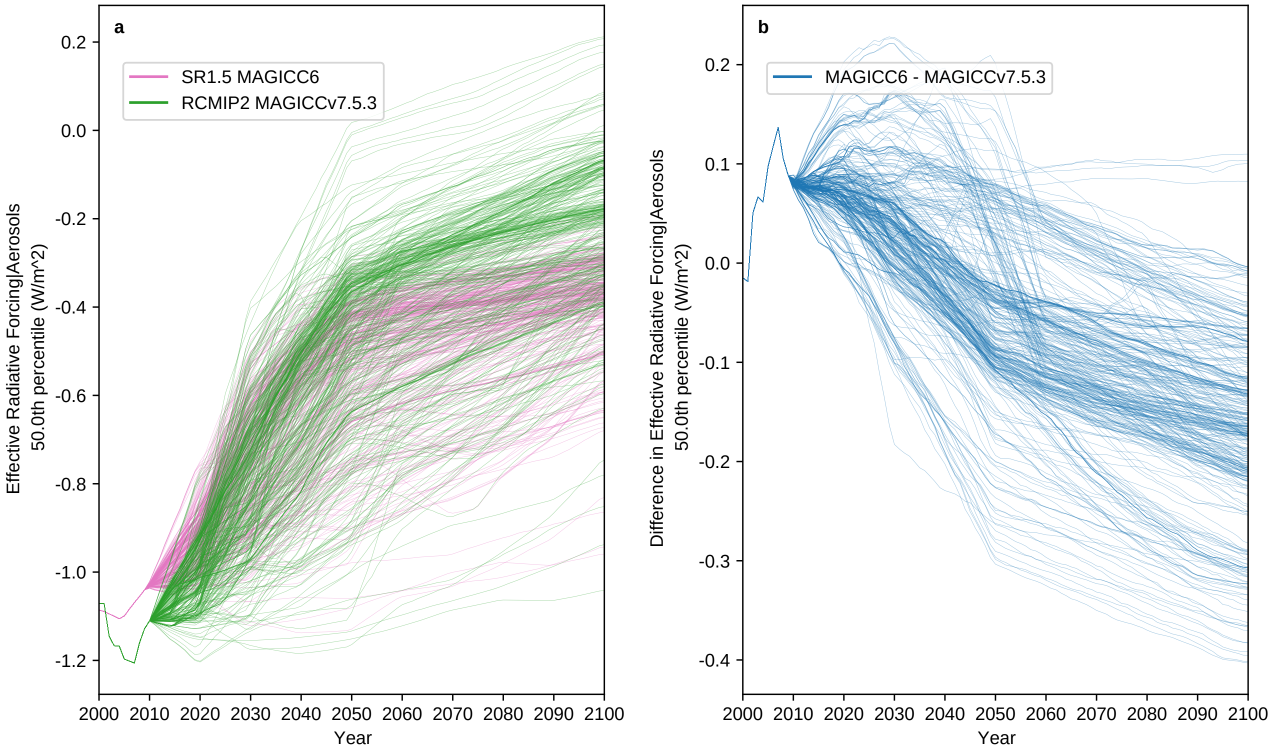

Figure S16. As in Supplementary Figure S12, except for median aerosol effective radiative forcing.

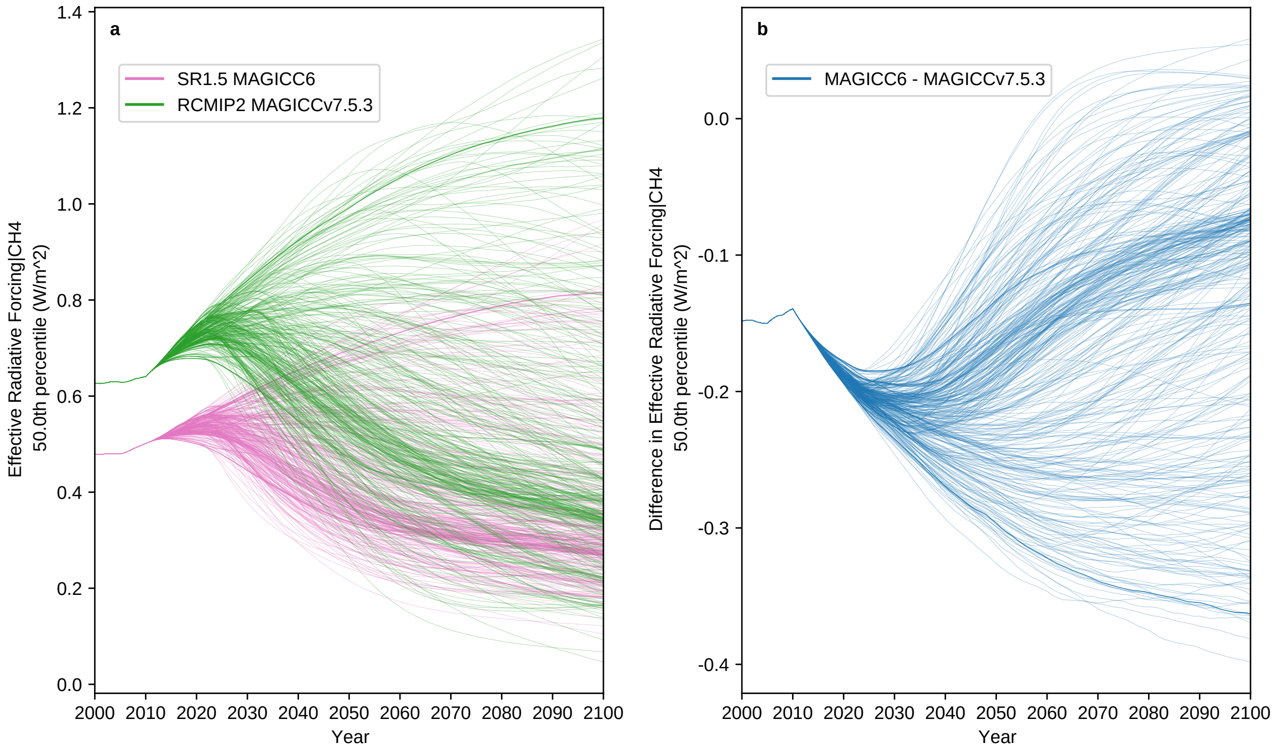

Figure S17. As in Supplementary Figure S12, except for median methane effective radiative forcing.

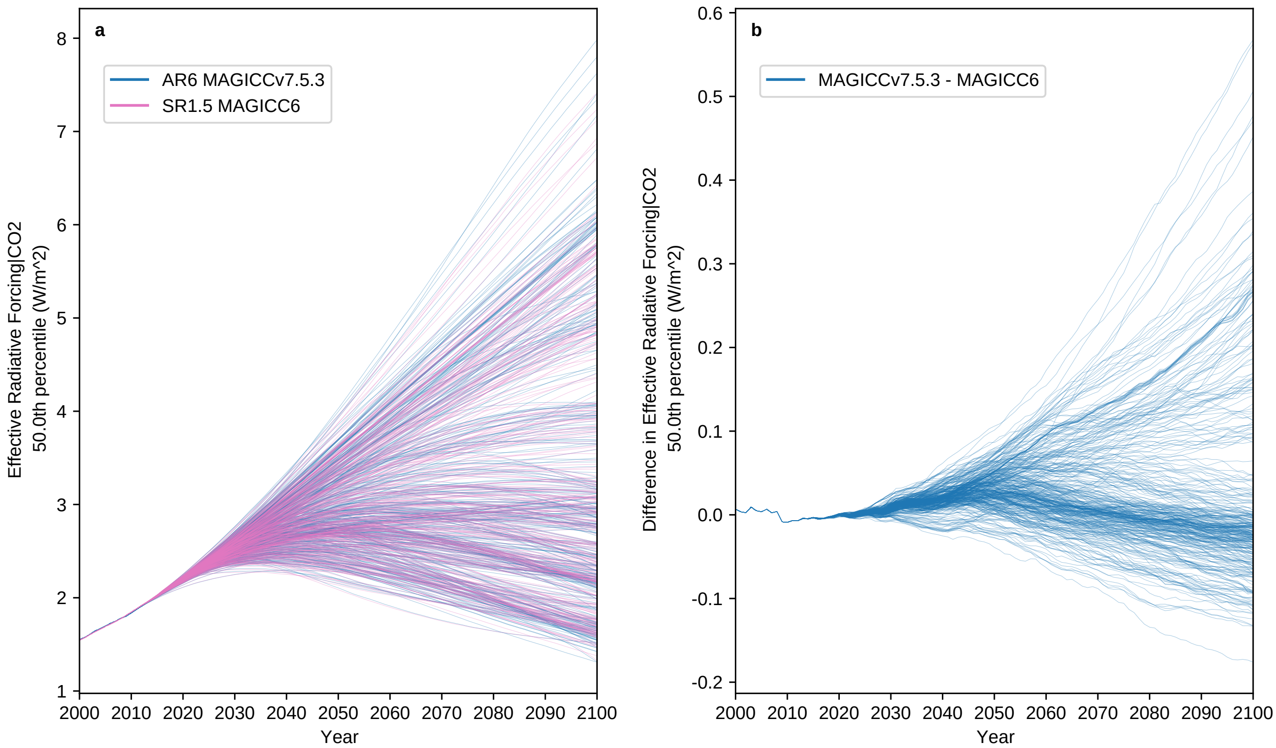

Figure S18. As in Supplementary Figure S8, except for median CO_2 effective radiative forcing.

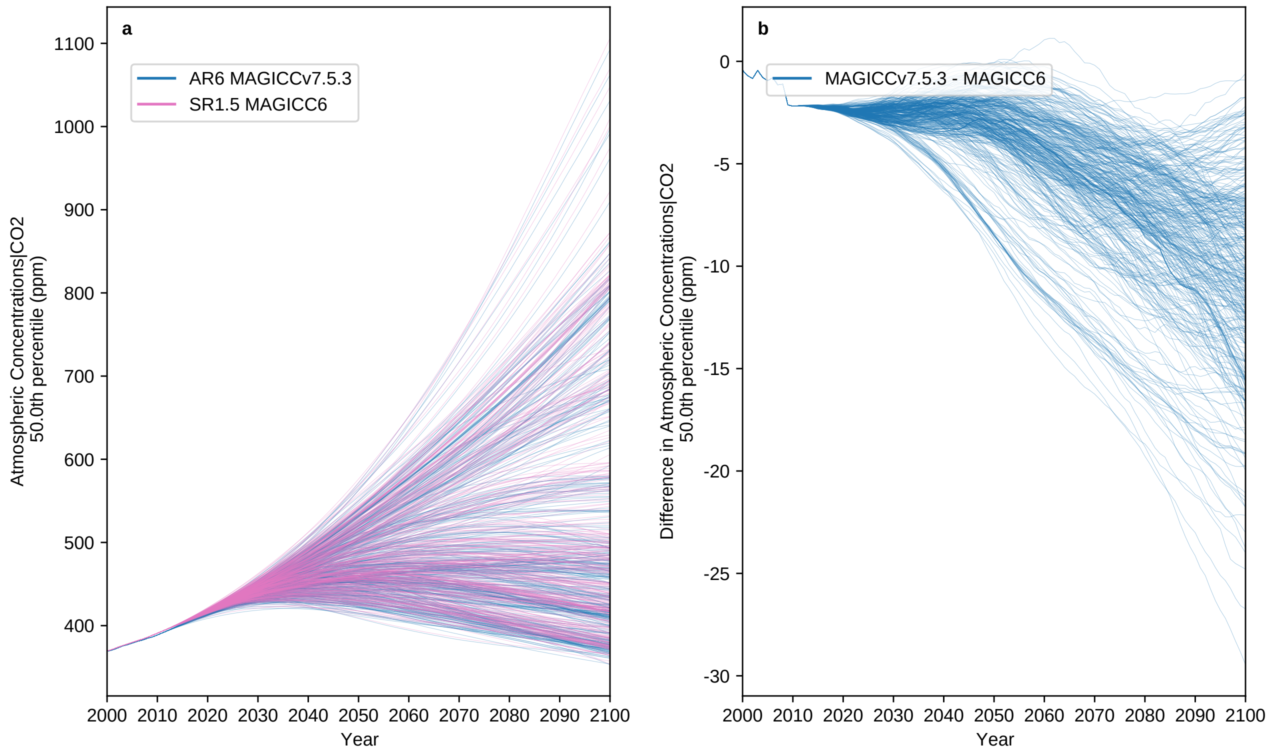

Figure S19. As in Supplementary Figure S8, except for median CO_2 atmospheric concentrations.

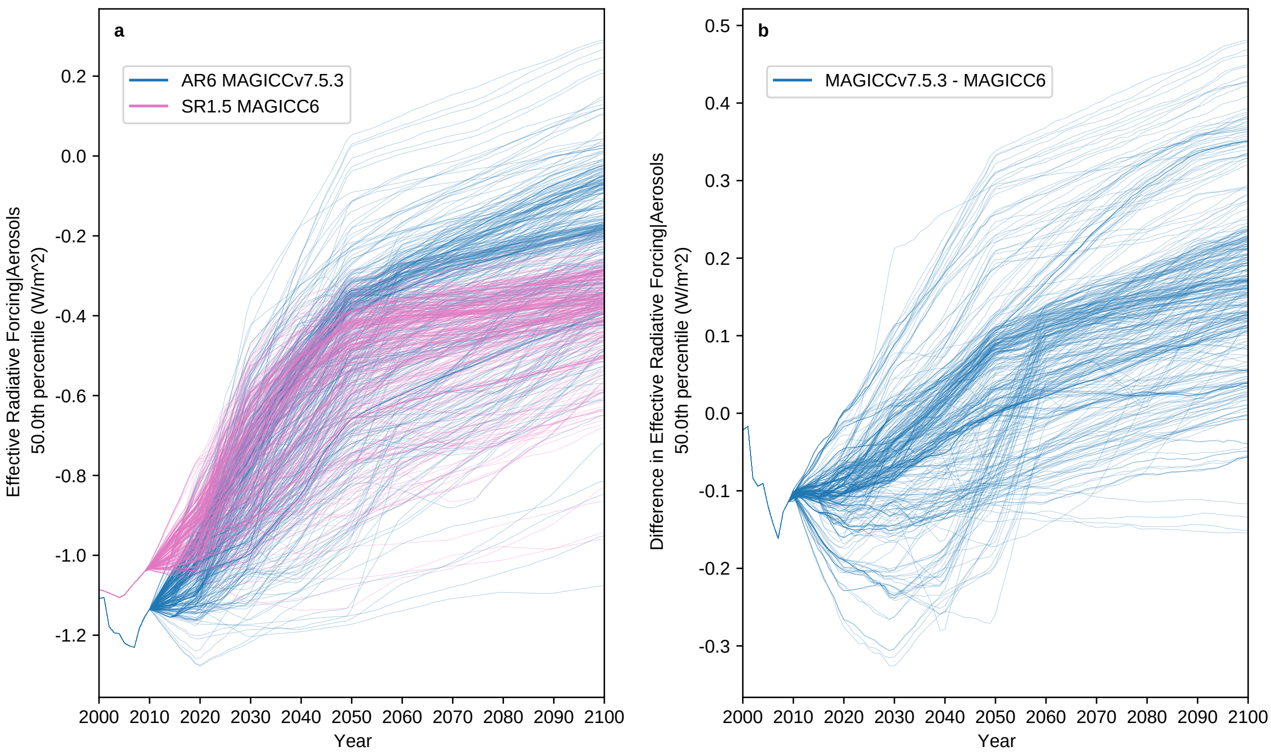

Figure S20. As in Supplementary Figure S8, except for median aerosol effective radiative forcing.

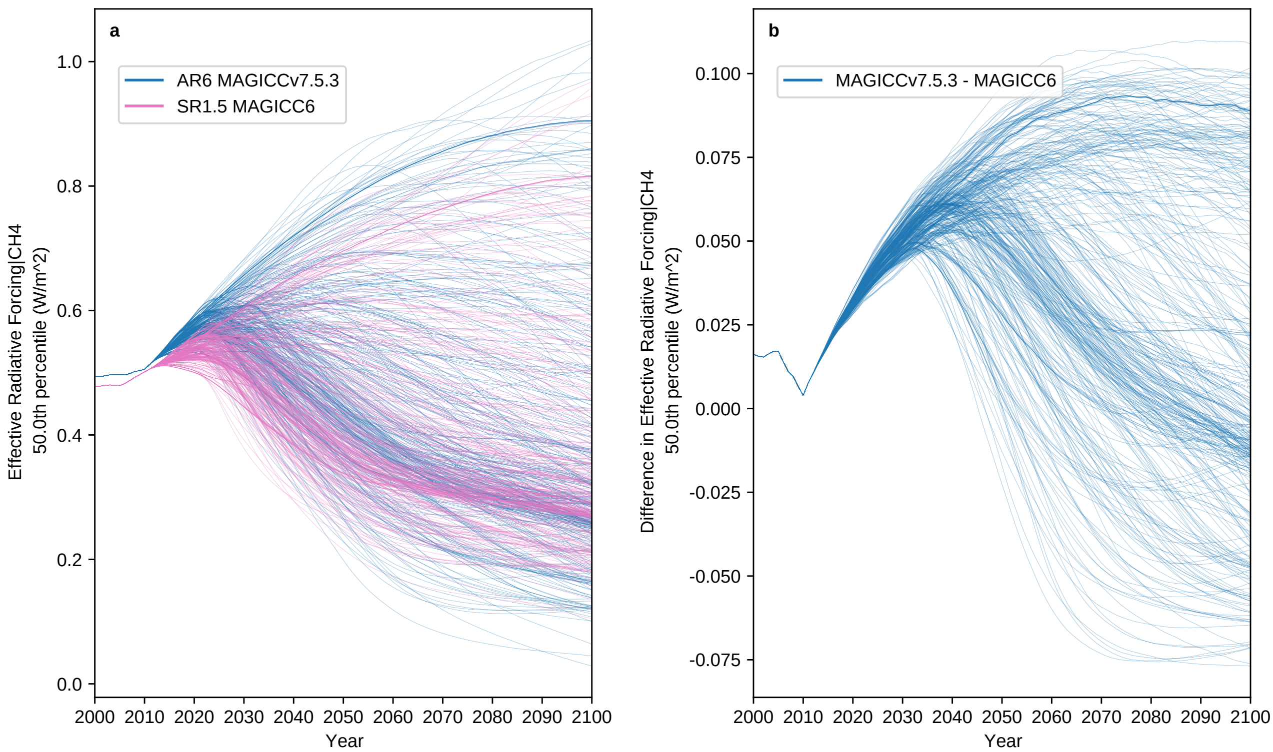

Figure S21. As in Supplementary Figure S8, except for median methane effective radiative forcing.