



flat10MIP: an emissions-driven experiment to diagnose the climate response to positive, zero and negative CO₂ emissions

Benjamin M. Sanderson¹, Victor Brovkin², Rosie A. Fisher¹, David Hohn³, Tatiana Ilyina^{2,4,5}, Chris D. Jones^{6,7}, Torben Koenigk⁸, Charles Koven⁹, Hongmei Li^{2,5}, David M. Lawrence¹⁰, Peter Lawrence¹⁰, Spencer Liddicoat⁶, Andrew H. MacDougall¹¹, Nadine Mengis³, Zebedee Nicholls^{12,13,14}, Eleanor O'Rourke¹⁵, Anastasia Romanou^{16,17}, Marit Sandstad¹, Jörg Schwinger¹⁸, Roland Séférian¹⁹, Lori T. Sentman²⁰, Isla R. Simpson¹⁰, Chris Smith^{13,21}, Norman J. Steinert¹, Abigail L. S. Swann²², Jerry Tjiputra¹⁸, and Tilo Ziehn²³

¹CICERO International Center for Climate Research, Oslo, Norway

²Max Planck Institute for Meteorology, Hamburg, Germany

³GEOMAR, Helmholtz Centre for Ocean Research, Kiel, Germany

⁴Modeling the carbon cycle in the Earth system, University of Hamburg, Hamburg, Germany

⁵Helmholtz-Zentrum Hereon, Geesthacht, Germany

⁶Met Office Hadley Centre, Exeter, UK

⁷School of Geographical Sciences, University of Bristol, Bristol, UK

⁸Swedish Meteorological and Hydrological Institute (SMHI), Norrköping, Sweden

⁹Lawrence Berkeley National Laboratory, Berkeley, CA, USA

¹⁰NSF National Center for Atmospheric Research (NCAR), Boulder, CO, USA

¹¹St. Francis Xavier University, Antigonish, NS, Canada

¹²School of Geography, Earth and Atmospheric Sciences, University of Melbourne, Melbourne, Australia

¹³International Institute for Applied Systems Analysis, Laxenburg, Austria

¹⁴Climate Resource, Fitzroy, Australia

¹⁵CMIP Project Office, Harwell, Oxfordshire, United Kingdom

¹⁶NASA Goddard Institute for Space Studies, New York, NY, USA

¹⁷Applied Physics and Applied Mathematics, Columbia University, New York, USA

¹⁸NORCE Norwegian Research Centre, Bjerknes Centre for Climate Research, Bergen, Norway

¹⁹Centre National de Recherches Météorologiques (CNRM), Toulouse, France

²⁰NOAA Geophysical Fluid Dynamics Laboratory (GFDL), Princeton, NJ, USA

²¹Priestley Centre for Climate Futures, University of Leeds, Leeds, UK

²²Atmospheric and Climate Science, College of the Environment, University of Washington, Seattle, WA, USA

²³CSIRO Environment, Aspendale, Australia

Correspondence: Benjamin M. Sanderson (benjamin.sanderson@cicero.oslo.no)

Received: 30 October 2024 – Discussion started: 15 November 2024

Revised: 26 May 2025 – Accepted: 3 June 2025 – Published: 5 September 2025

Abstract. The proportionality between global mean temperature and cumulative emissions of CO₂ predicted in Earth system models (ESMs) is the foundation of carbon budgeting frameworks. Deviations from this behavior could impact estimates of required net-zero timings and negative emissions requirements to meet the Paris Agreement climate targets. However, existing ESM diagnostic experiments do not

allow for direct estimation of these deviations as a function of defined emissions pathways. Here, we perform a set of climate model diagnostic experiments for the assessment of transient climate response to cumulative CO₂ emissions (TCRE), the Zero Emissions Commitment (ZEC), and climate reversibility metrics in an emissions-driven framework. The emissions-driven experiments provide consistent inde-

pendent variables simplifying simulation, analysis and interpretation, with emissions rates more comparable to recent levels than existing protocols using model-specific compatible emissions from the CMIP DECK *1pctCO2* experiment, where emissions rates tend to increase during the experiment, such that at the time of CO₂ doubling in year 70, emissions are much greater than present-day values. A base experiment, “*esm-flat10*”, has constant emissions of CO₂ of 10 GtC per year (near-present-day values), and initial results show that the TCRE estimated in this experiment is about 0.1 K less than that obtained using *1pctCO2*. A subset of ESMs exhibit land carbon sinks that saturate during this experiment. A branch experiment, *esm-flat10-zec*, illustrates that both positive and negative ZEC effects are less pronounced under *esm-flat10* than under *1pctCO2* – the magnitude of ZEC50 in ESMs is, on average, reduced by 30 % compared with *1pctCO2* branch experiments. A final experiment, *esm-flat10-cdr*, assesses climate reversibility under negative emissions, where we find that peak warming may occur before or after net zero and that the asymmetry in temperature at a given level of cumulative emissions between the positive and negative emissions phases is well described by ZEC in most models. Further, we find that existing probabilistic simple climate model (SCM) ensembles tend to overestimate temperature reversibility compared with ESMs, highlighting the need for additional constraints. We propose a set of climate diagnostic indicators to quantify various aspects of climate reversibility. These experiments were suggested as potential candidates in CMIP7 and have since been adopted as “fast track” simulations.

1 Introduction

The concept of proportionality of global mean temperatures to cumulative carbon dioxide emissions is central to carbon budgeting frameworks and net-zero commitments (Rogelj et al., 2019b). The relationship has its origins in the recognition of a robust linear relationship in Earth system model simulations (Allen et al., 2009; Matthews et al., 2009; Zickfeld et al., 2009) and observations (Gillett et al., 2013) between the global mean temperature change and the cumulative amount of CO₂ released into the atmosphere, the slope of which we refer to as the transient climate response to cumulative emissions (TCRE) – the change in global mean temperature per trillion tonnes of carbon emitted into the atmosphere. TCRE offers a powerful, simplified lens for climate policy applications, allowing policymakers to directly equate emission budgets to projected warming levels (Lamboll et al., 2023; Rogelj et al., 2019b) and to gauge the relative impact of different emissions trajectories over time (MacDougall, 2015).

For a simulation in which temperature changes are driven by CO₂ alone,

$$\text{TCRE} = \frac{\Delta T(t)}{I_{\text{em}}(t)}, \quad (1)$$

where $\Delta T(t)$ and $I_{\text{em}}(t)$ are the temperature change and cumulative emissions at time t , respectively. For climate models, TCRE is generally calculated using results from a concentration-driven simulation *1pctCO2*, where CO₂ concentrations are prescribed and ramped up exponentially at a rate of 1 % per year. In assessments (IPCC, 2023c), the TCRE is nominally computed in year 70, when concentrations have approximately doubled:

$$\text{TCRE}_{1\text{pctCO2}} = \frac{\Delta T(70)}{I_{\text{em}}(70)} = \left(\frac{\Delta T(70)}{I_{\text{atmos}}(70)} \right) \left(\frac{I_{\text{atmos}}(70)}{I_{\text{em}}(70)} \right),$$

where $I_{\text{atmos}}(70)$ is the additional carbon in the atmosphere and $\Delta T(70)$ is the transient climate response (TCR, in practice calculated as the average over years 60–80). $\frac{I_{\text{atmos}}(70)}{I_{\text{em}}(70)}$ is the cumulative airborne fraction, the proportion of cumulative emissions which remain in the atmosphere.

In order to constrain compatible carbon emissions budgets for certain warming levels, historical human-induced warming must be calculated, along with some additional corrections (Rogelj et al., 2019b). Firstly, the temperature impact of present and future non-CO₂ emissions must be incorporated, either by assuming a ratio of future CO₂ and non-CO₂ emissions (Damon Matthews et al., 2021; Leach et al., 2018; Millar and Friedlingstein, 2018), by subtracting an estimate of non-CO₂ warming (Lamboll et al., 2023), or by defining a TCRE based on cumulative CO₂- forcing-equivalent emissions (Jenkins et al., 2021).

Secondly, any potential further carbon-induced warming after net zero has been achieved introduces additional uncertainty in remaining carbon budgets calculated using TCRE (Nicholls et al., 2020). This behavior has been characterized by the Zero Emissions Commitment (ZEC) (IPCC, 2023c; Lamboll et al., 2023). Definitions of ZEC are, to date, primarily informed by the ZECMIP CMIP6 experiment (Jones et al., 2019), which is based on an abrupt cessation of emissions branching from the *1pctCO2* experiment when 1000 PgC of CO₂ emissions have been diagnosed, where the notation ZEC_{*n*} corresponds to the warming n years after the cessation of emissions (MacDougall et al., 2020). ZEC50 and ZEC90 are thus the temperature change 50 and 90 years after the cessation of emissions, respectively. This experiment was performed by a coordinated set of Earth system models and intermediate complexity models, which led to the finding that ZEC had the potential to be either positive or negative (MacDougall et al., 2020) with a best estimate near zero.

ZECMIP experiments were designed this way to ensure consistency of ZEC and TCRE at the same point, but they do, however, have a number of limitations. Firstly, *1pctCO2* is a prescribed concentration trajectory for atmospheric CO₂,

and compatible emissions are computed as a residual term, such that each climate model has a different emissions trajectory. This poses two issues for using the run as a basis for the assessment of ZEC. Firstly, each model follows its own pathway of (implied) emissions in such experiments, obfuscating the relationship between model and ZEC response. Secondly, the compatible emissions profile in *1pctCO2* grows throughout the experiment, with the burden of cumulative emissions weighted towards the end of the experiment (Sanderson et al., 2024 and Fig. 6), whereas contemporary emissions are closer to flat (since about 2012).

Alternative frameworks have been proposed to provide a less scenario-dependent formulation for ZEC. Consideration of linear pulse-response models of the climate shows that cumulative emissions proportionality is an expected first-order response but that second-order terms allow for further temperature changes after emissions have ceased (Avakumović, 2024; Jenkins et al., 2022). This second-order behavior can be approximated by the rate of adjustment to zero emissions, or RAZE, which defines the fractional change in CO₂-induced warming after CO₂ emissions cease (Jenkins et al., 2022). In this approximation (valid for decadal timescales following net zero), RAZE can be related to ZEC for a given scenario as

$$\text{ZEC}_H = I_{\text{em}}(t = t_{\text{net-zero}}) (\text{TCRE}) (\text{RAZE}) (H),$$

where ZEC_H is the warming H years after net zero and $I_{\text{em}}(t = t_{\text{net-zero}})$ is the cumulative emissions at the time of net zero. In this framing, a linear estimate of the warming rate after net zero, if emissions are held at net zero, is given by $I_{\text{em}}(t = t_{\text{net-zero}}) (\text{TCRE}) (\text{RAZE})$.

In addition, no experiment within prior CMIP efforts has been designed to robustly understand the degree of asymmetry in the climate response to positive followed by negative CO₂ emissions. The compatible emissions from the *1pctCO2-cdr* concentration reversal experiment used in CDRMIP (Asaadi et al., 2024) are both asymmetric in time, between the positive and negative emissions periods, and have a large discontinuity of roughly 50 Pg C yr⁻¹ (Koven et al., 2023) at the point of reversal from increasing to decreasing CO₂ concentrations. Secondly, the lagged effects of the positive emission phase can complicate assessment of the response to negative emissions (Chimuka et al., 2023; Koven et al., 2023; Zickfeld et al., 2016).

These confounding effects inhibit a clear diagnosis of whether and how the general climate response to negative emissions differs from the climate response to positive emissions (MacDougall, 2019). An idealized CMIP experiment that allows for a continuous transition from positive to negative emissions, and one that is symmetric in time (so that any asymmetries that arise are driven by the coupled carbon–climate response itself), improves on this status quo (though the separation of lagged effects remains a challenge).

Here, we propose a compact set of experiments uniquely designed to cleanly assess carbon–climate dynamics relevant for mitigation. Our objectives are 3-fold:

- Re-assess the transient climate response to cumulative CO₂ emissions: assess the response of temperature change and land/ocean carbon dynamics as a function of cumulative emissions, which are the independent variable of the experiment.
- Assess the Zero Emissions Commitment across models on multiple timescales: systematically measure the unrealized warming that continues after all CO₂ emissions have been halted (again, in an experiment where emissions are the independent variable), through assessment of ZEC after 50, 90, 100 and 200 years.
- Explore climate reversibility potential: assess the impacts of global-scale carbon removals, assessing the hysteresis in the relationship between climate and cumulative CO₂ emissions.

Regional and component responses require further study beyond the scope of the globally aggregated analysis presented here. Studies in preparation will consider in detail the commitment and reversibility of ocean heat uptake, regional climatology and land carbon dynamics.

2 Flat10MIP experiment design

Sanderson et al. (2024) proposed four new experiments (Fig. 1) that would form part of a standard diagnostic suite for carbon-emissions-driven behavior in multi-model comparison activities such as CMIP. These experiments assess behavior under sustained constant carbon emissions, immediate cessation of emissions and climate reversibility under an idealized continuous climate restoration pathway where all emissions are removed by the end of the simulation (Fig. 1). Here, flat10MIP simulates three of the four experiments proposed in (Sanderson et al., 2024) using CMIP6 generation models, as a pilot study in preparation for CMIP7. Below, we briefly describe the experiments as conducted in flat10MIP and recommendations for a protocol in CMIP7 and beyond.

2.1 Experiments in flat10MIP

2.1.1 esm-flat10

The esm-flat10 experiment would serve as an emissions-driven experiment to diagnose the transient climate response to cumulative CO₂ emissions (TCRE), which is the warming from pre-industrial levels observed after the emission of 1000 PgC in a transient scenario. The *esm-flat10* experiment would branch from a stable esm-piControl simulation, with a constant annual prescribed anthropogenic flux of carbon of 10 PgC yr⁻¹ into the atmosphere, with globally homogenous emissions. In esm-flat10, the 1000 PgC threshold

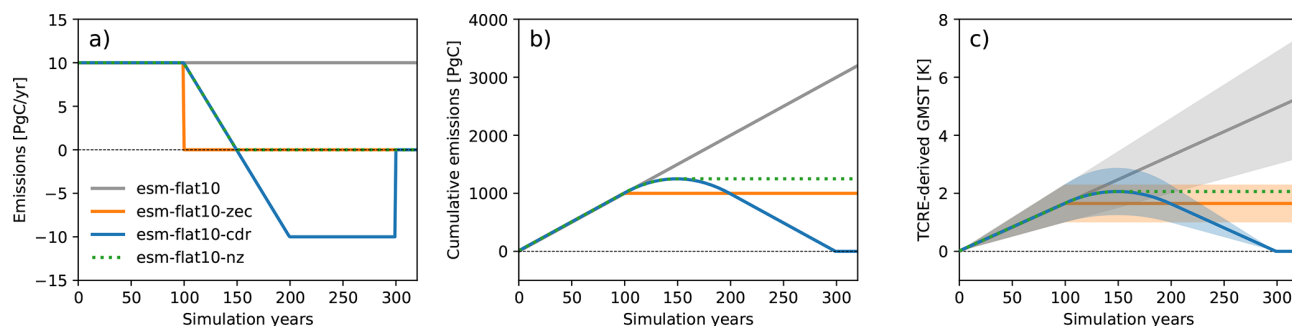


Figure 1. Experiment design. (a) and (b) show annual and cumulative carbon emissions as a function of time for the four experiments. (c) shows global mean surface temperature derived from cumulative emissions, assuming a perfectly linear TCRE relationship, with expected temperature evolution assuming cumulative emissions proportionality using the IPCC AR6 WGI best TCRE estimate (solid line, 1.65 °C per 1000 PgC) and likely range (shaded area, 1.0–2.3 °C per 1000 PgC) (IPCC, 2023b).

would occur in year 100 – such that TCRE could be estimated as the time average between global mean warming in years 90–110, sampling over internal variability in this period. As such, we refer to TCRE derived from *esm-flat10* and 1pctCO₂ as $T_{100\text{yr}}$ and $T_{1000\text{PgC}}$, respectively. The protocol for *esm-flat10* is to continue emissions at 10 PgC yr^{−1} for the duration of the experiment. In this ensemble, 150 years were considered to allow the simulation to reach 2× pre-industrial CO₂ concentrations in most cases (allowing for a wide range of plausible land and ocean carbon uptake). However, for future experiments in CMIP7 and beyond, a 300-year or longer *esm-flat10* would be useful to explore potential nonlinearities in response to higher cumulative emission levels, which have been observed in some models (Schwinger et al., 2022).

2.1.2 *esm-flat10-zec*

The *esm-flat10-zec* experiment serves as an emissions-driven experiment to diagnose ZEC, which is the additional warming seen a certain number of years after the abrupt cessation of emissions. The *esm-flat10-zec* experiment would branch from year 100 of the *esm-flat10* experiment, with an immediate cessation of emissions, and the system is then left to evolve for 220 years. ZEC50 is calculated as the average temperature change relative to that when emissions cease, averaged over a 21-year period, 50 years after the cessation of emissions (i.e., years 140–160). ZEC90 is similarly calculated using years 180–200. For CMIP7 and beyond, we recommend 300 years for *esm-flat10-zec* to allow for longer timescale comparisons with *esm-flat10-cdr*.

2.1.3 *esm-flat10-cdr*

The *esm-flat10-cdr* experiment serves as an emissions-driven experiment to diagnose the response of the climate system to reducing, and ultimately reaching, net-negative emissions and will provide a measure of climate reversibility when all cumulative anthropogenic emissions are removed (i.e., all cumulative emissions and removals sum to zero)

at the end of the experiment. The *esm-flat10-cdr* experiment would branch from year 100 of the *esm-flat10* experiment, with a linear ramp down of emissions (from a starting point of 10 PgC yr^{−1}) of −0.2 PgC yr^{−1} – such that net-zero emissions are achieved in year 150 and a negative flux of −10 PgC yr^{−1} is achieved in year 200. This negative emission flux of −10 PgC yr^{−1} would then be held constant from years 200–300, such that, by year 300, cumulative emissions from the start of the simulation would be zero. A 20-year extension follows, keeping the emissions at zero. For CMIP7 and beyond, we recommend 300 years for *esm-flat10-cdr* to allow for better evaluation of system dynamics after the termination of negative emissions.

2.1.4 *esm-flat10-nz*

We propose a final experiment for CMIP7 and beyond (not conducted here, but noted for its relevance), i.e., *esm-flat10-nz* (Sanderson et al., 2024), which branches from *esm-flat10-cdr* in year 150 at the point at which the simulation reaches net-zero CO₂ emissions, keeping emissions at zero thereafter. Such an experiment would provide a proxy for warming commitment after a gradual semi-idealized emission reduction to net zero and would provide additional information on ZEC. We recommend that such an experiment should run ideally for 250 years to allow for comparison of long-term dynamics with *esm-flat10-zec* and *esm-flat10-cdr*. Such an experiment could help differentiate the response of the system to negative emissions in *esm-flat10-cdr* from the delayed response to positive emissions and would provide a counterpoint to the abrupt emissions termination seen in *esm-flat10-zec* – providing an idealized scenario that might provide a more policy-relevant estimate of ZEC dynamics, reaching net zero after a period emissions reduction.

The *esm-flat10-cdr* experiment allows for a number of simple idealized diagnostics that are relevant to the net-zero transition and the response of the system to net-negative emissions (Fig. 2). Like ZEC, each of these metrics is a measure of the path dependence of the temperature to the cu-

Table 1. Experiment design for emissions-driven diagnostic runs, detailing the branch point, length and configuration of the experiments as conducted in flat10MIP (present study). The CMIP7 recommended protocol includes run lengths and experiments suggested for future multi-model comparisons, including *esm-flat10-nz*, which is not conducted in this study.

Experiment	Branches from	Years (this paper)	Years (CMIP7 recommended protocol)	CO ₂ emissions	Diagnostic metrics
<i>esm-flat10</i>	<i>esm-piControl</i>	150 years (from year 0 to year 149)	300 years	10 PgC yr ⁻¹ constant emissions, globally homogenous flux	TCRE
<i>esm-flat10-zec</i>	<i>esm-flat10</i> (branch at start of year 100 of <i>esm-flat10</i>)	220 years (from year 100 to year 319)	300 years (from year 100 to year 399)	Constant 0 PgC yr ⁻¹	ZEC50 ZEC90 ZEC100 ZEC200
<i>esm-flat10-cdr</i>	<i>esm-flat10</i> (branch at start of year 100 of <i>esm-flat10</i>)	220 years (from year 100 to year 319)	300 years (from year 100 to year 399)	– Linearly declining emissions by 2 PgC/decade from 10 PgC yr ⁻¹ (year 100) to – 10 PgC yr ⁻¹ (year 200) – Constant – 10 PgC yr ⁻¹ (years 200–299) – Zero emissions for years 300–319	TNZ TR1000 TR0 tPW
<i>esm-flat10-nz*</i>	<i>esm-flat10-cdr</i> (branch in year 150)	–	250 years (from year 150 to year 399)	0 PgC yr ⁻¹ constant	

mulative emissions relationship and thus would have a value of exactly 0 if global temperature response exactly followed TCRE proportionality. They include the following.

- Temperature difference at net zero (TNZ): this measures the error associated with assuming cumulative emissions proportionality to predict temperatures at net zero. *esm-flat10-cdr* reaches net-zero emissions in year 150, with cumulative emissions of 1250 PgC (calculated from year 1; see Fig. 1). TNZ is calculated as a 21-year average around year 150 in *esm-flat10-cdr* (i.e., 50 years after branching from *esm-flat10*) minus the expected temperature at net zero using cumulative emissions proportionality (T_{ref} , which is 1.25 times the *esm-flat10*-derived TCRE; see Fig. 2).
- Temperature asymmetry under CO₂ removal at 1000 PgC (TR1000): this measures the asymmetry in warming during positive and negative emissions at the same net cumulative emissions. It is calculated as a 21-year average around year 200 in *esm-flat10-cdr* minus a 21-year average around year 100 in *esm-flat10*. TR1000 would be a measure of hysteresis in global mean temperature when cumulative emissions return to 1000 PgC on the downward branch minus the warming at the same cumulative emissions level under *esm-flat10*. This could be calculated using a combination of the *esm-flat10* and *esm-flat10-cdr* experiments for a cumulative

carbon emissions total of 1000 PgC. *esm-flat10-cdr* reaches 1000 PgC cumulative emissions in year 200 on the downward branch (see Fig. 1). *esm-flat10* itself reaches 1000 PgC in year 100.

- Temperature asymmetry under CO₂ removal at 0 PgC (TR0): this is a measure of carbon–climate reversibility when all previously-emitted carbon has been removed from the atmosphere. It is calculated as the average of years 301–320 in *esm-flat10-cdr* minus mean global temperatures in *esm-piControl*. TR0 is a measure of hysteresis in the global mean temperature when cumulative emissions return to zero after a period of negative emissions. This is calculated using a combination of the *esm-piControl* and *esm-flat10-cdr* experiments. *esm-flat10-cdr* reaches zero cumulative emissions in year 300 on the downward branch (see Fig. 1).
- Time to peak warming (tPW): this is a measure of the difference in timing between net zero and peak warming. It is calculated as the time difference between the peak value of 20-year smoothed global mean temperatures and the point that net zero is achieved in *esm-flat10-cdr* (year 150). This metric has a clear policy-relevant translation as the expected time it will take for the climate system to achieve maximum CO₂-driven global warming after (or before) reaching net-zero

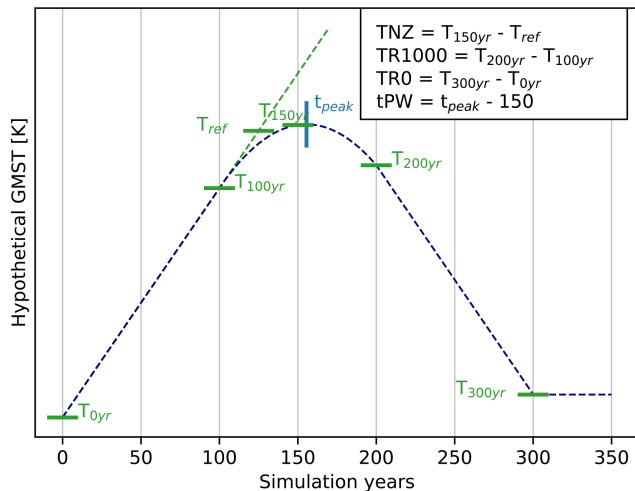


Figure 2. Schematic of metrics derived from the *esm-flat10-cdr* experiment to quantify different aspects of temperature reversibility under a continuous transition from positive to negative emissions. Dashed lines correspond to temperature trajectories for a hypothetical case where temperatures do not perfectly follow cumulative CO₂ emissions. GMST denotes global mean surface temperature.

emissions under a smooth positive-to-negative emissions transition.

2.2 Models used in flat10MIP

This ensemble provides a broad range of climate model structures and components to evaluate emissions-driven climate reversibility. We include eight CMIP6 generation Earth system models (ESMs), one CMIP3 generation model, one intermediate complexity model and the three simple climate model (SCM) ensembles used in the AR6 IPCC assessment (Forster et al., 2023). The ESMs and SCMs participating in this study are listed in Table 2 and more fully described in the Appendix. Each Earth system model has completed one ensemble member of each of the MIP experiments (*esm-flat10*, *esm-flat10-cdr* and *esm-flat10-zec*) – with supporting existing experiments from CMIP6 (C4MIP, ZECMIP and CDR-MIP). We note that metrics from Earth system models, unlike SCMs, are subject to uncertainty arising from internal variability. We would encourage centers to perform at least three members of these experiments in CMIP7 to provide better sampling and estimation of the role of initial condition uncertainty. For each SCM, approximately 1000 simulations are completed with simple climate model versions spanning a range of climate responses consistent with assessed climate uncertainty (using a combination of observational constraints, IPCC-assessed ranges and ESM data to constrain the parameter space of the simple climate models (IPCC, 2023d)).

In this study, we summarize the global mean characteristics of the simulations used to conduct the experiments, while

additional dedicated domain-specific studies will assess regional aspects of transient-emissions-driven response and reversibility.

3 Results

Figure 3 illustrates the global temperature response for the three simulations requested in flat10MIP. Throughout this section, we refer by default to $T_{100\text{yr}}$ – the warming, in units K, after 1000 PgC of cumulative emissions (which in *esm-flat10* occurs in year 100). $T_{100\text{yr}}$ is numerically equivalent to TCRE (units K/1000 PgC) but allows proper consideration of the arithmetic sum with ZECn, also in units K. Summary metrics, as defined above for each model, are detailed in Table 2. Figure 3a shows that the range of $T_{100\text{yr}}$ seen in the ESM ensembles (1.1–2.4 K) is broadly captured by the SCM ensembles considered in this study, though MAGICC shows a greater upper bound in $T_{100\text{yr}}$ (10th–90th percentile of 1.1–2.7 K) relative to FaIR or CICERO-SCM (10th–90th percentiles of 1.1–2.1 and 1.2–2.1 K, respectively). However, we see differences in the ZEC50, ZEC90 and reversibility distributions. The ESM ZEC90 distribution is best captured by FaIR (ZEC90 range of −0.1 to +0.2 K), whereas MAGICC and CICERO-SCM simulate more negative values (−0.2 to +0.1 K and −0.3 to −0.1 K, respectively). We also see that two of the three SCM ensembles (MAGICC and CICERO-SCM) tend to simulate a stronger temperature decline under negative emissions than seen in any of the ESMs, although the FaIR ensemble is broadly consistent. The intermediate complexity model, UVic-ESM, lies within the ESM distribution for both $T_{100\text{yr}}$ and ZEC90.

3.1 Earth system model responses

Figure 4 illustrates the ESM results in more detail, showing the evolution of a number of climate indicators. In *esm-flat10*, emissions are constant at 10 PgC yr^{−1} – and thus temperature change from pre-industrial to year 100 is a measure of the transient response to cumulative CO₂ emissions. Figure 4a illustrates the range of transient response in the context of the assessed TCRE range in IPCC AR6 (1.0 to 2.3 K (1000 PgC)^{−1}) (IPCC, 2023a). The models considered in the present MIP largely span this range, with values of TCRE (as calculated from 1pptCO₂) from 1.2 to 2.6 K (Table 2). The model land sink evolution varies during the extended *esm-flat10* simulations, with some models showing a saturation of the land sink (HadCM3, UVic, ACCESS) and others showing continued land uptake throughout the experiment (CESM, NorESM, GFDL, CNRM).

Figure 4b shows how temperatures evolve in *esm-flat10-zec* – showing that temperatures remain (approximately) stable following the cessation of emissions, even though atmospheric carbon dioxide concentrations decline. Different models show a diversity of evolution of land and ocean car-

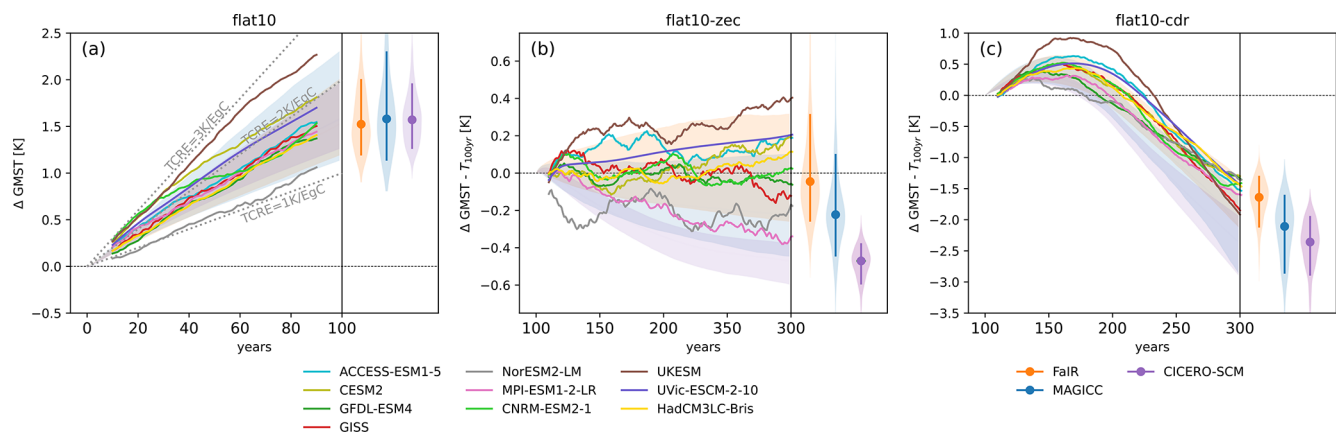


Figure 3. Summary results for global mean surface temperature (GMST) response in the trial flat10MIP. Colored lines indicate temperature change from (a) pre-industrial levels to (b, c) T100 yr (the average temperature in years 91–110 in esm-flat10) in each of the participating ESMs. Shaded regions refer to the simple climate models' probabilistic distribution ranging from the 10th to 90th percentiles. This distribution is shown as violin plots for the last time step of each scenario, where the shading shows the full range of results and the vertical line indicates the 10th–90th percentiles, with the median in the center. A 20-year moving average is applied to all time series.

bon sinks – with some models (e.g., MPI, GFDL, GISS) initially absorbing land carbon for the first 20–50 years of the zero-emissions phase before losing carbon on longer timescales, while the land sink in other models (UVic, GISS, HadCM3, UKESM) stabilize the land sink after emissions cessation. Ocean uptake is more consistent across the ensemble, with all models simulating a continued uptake of carbon in the ocean during the zero-emissions phase.

The global mean results for *esm-flat10-cdr* are summarized in Fig. 4c, showing that peak warming can occur either before or after net zero (but most models peak before), as seen in similar experiments (Koven et al., 2023) and ZECMIP experiments (Jenkins et al., 2022). By the end of the simulation, some models remain warmer than the pre-industrial period (CESM, CNRM, ACCESS, UVic), while some are cooler (GISS, MPI, NorESM, GFDL). All models are in agreement that peak CO₂ concentrations occur before net zero, and all models predict that the ocean carbon sink peaks after net zero. All models predict that the cumulative ocean carbon sink will decline but stay positive. However, models disagree on the timing of the peak land sink relative to net zero. GFDL, CESM, NorESM, GISS, MPI and CNRM show the cumulative land carbon sink peaking after net zero, whereas HadCM3, UVic and ACCESS show the cumulative land carbon sink peaking before net zero. At zero cumulative emissions in year 300, models range from the cumulative land sink being near-zero to being a slight net source of carbon over the 300-year period (model range: –50 to 0 PgC), while all models agree that the cumulative ocean sink is a net sink (model range: 120–220 PgC).

Figure 5 shows how the rate of carbon emission allocation to the atmosphere, land and ocean evolves as a function of time in the different experiments. In *esm-flat10*, we observe a transition from an initially high airborne fraction

towards increasing allocation to land and ocean pools, with the airborne fraction in year 100 ranging between 0.45 and 0.55 across models. This variation arises from inter-model differences in the representation of land and ocean carbon uptake processes. For example, some models exhibit sustained terrestrial uptake (e.g., CESM2, NorESM2), while others (e.g., ACCESS, UKESM) show land sink saturation or reversal, likely reflecting the interplay between CO₂ fertilization (Arora et al., 2020), nutrient availability (Goll et al., 2012) and warming-induced soil carbon losses (MacDougall et al., 2020; Wieder et al., 2013). Declining land uptake in some models may also reflect increasing hydrological stress or climatic constraints on productivity (Fisher et al., 2019). During the *esm-flat10-zec* experiment, atmospheric CO₂ declines following cessation of emissions, but models diverge in whether this drawdown is primarily balanced by land (e.g., GFDL, CNRM) or ocean (e.g., GISS, ACCESS) uptake. These differences reflect the distinct timescales and sensitivities of the carbon pools: the land sink responds quickly to emissions cessation but may decay as CO₂ fertilization effects diminish and heterotrophic respiration increases (Jones et al., 2013), while the ocean continues to absorb carbon due to its longer equilibration timescales and sustained pCO₂ disequilibrium (Schwinger and Tjiputra, 2018; Tjiputra et al., 2013) and the model-specific representation of deep ocean ventilation and carbon transport (Séférian et al., 2024). The resulting diversity in sink partitioning highlights key model-dependent feedbacks in the terrestrial biosphere and ocean circulation, which modulate the climate system's reversibility following net zero.

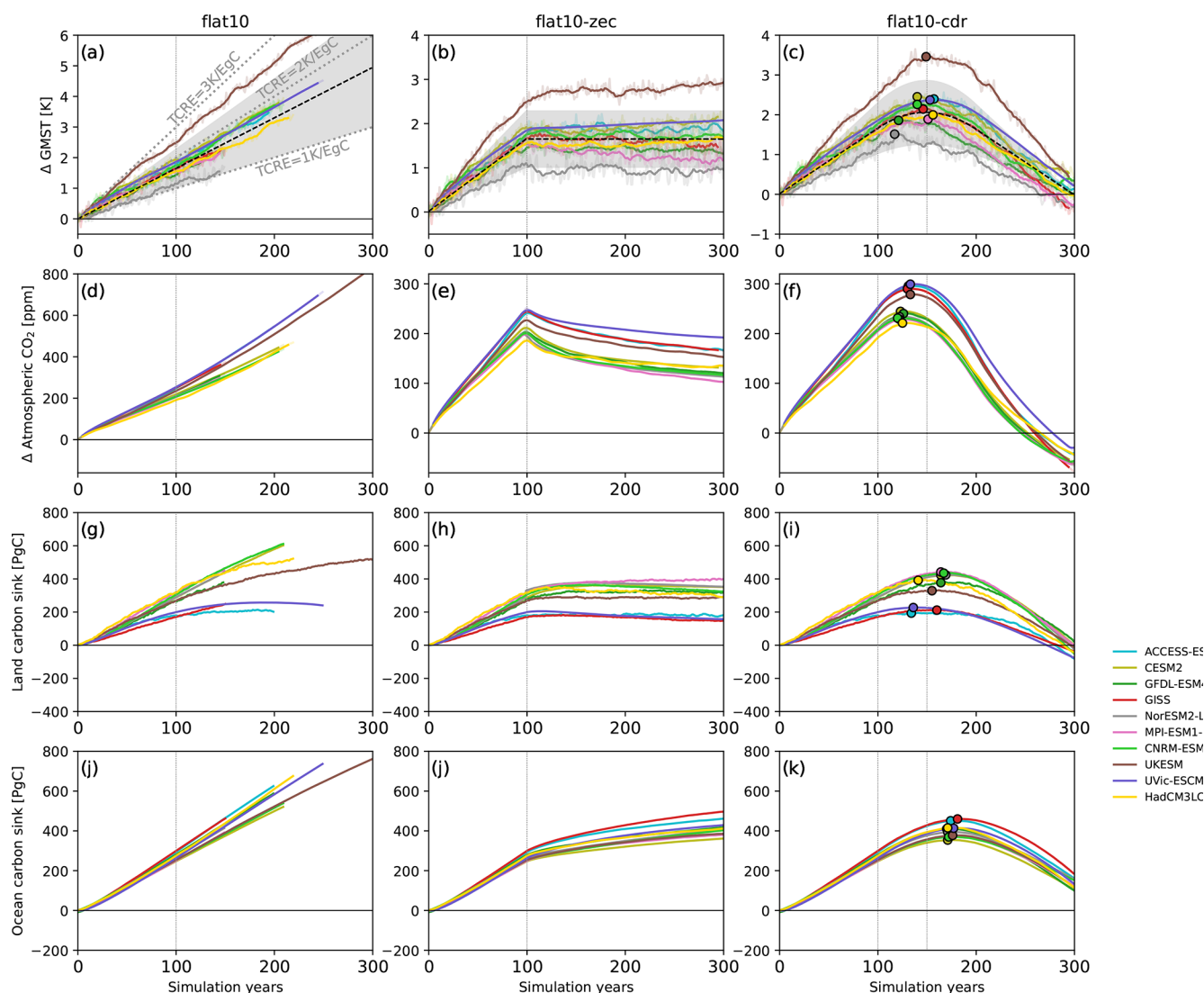


Figure 4. ESM results. Columns represent global indicators from ESM simulations running *esm-flat10* (left), *esm-flat10-zec* (center) and *esm-flat10-cdr* (right). Panels (a)–(c) show changes in GMST, with the black dashed line and gray shading denoting the central estimate and range derived from cumulative emissions, assuming a linear TCRE relationship as given in AR6 ($\text{TCRE} = 1.65 \text{ K}$, likely range of $1.0\text{--}2.3 \text{ K}$) for reference. Panels (d)–(f) illustrate changes in atmospheric CO_2 concentrations as a function of time. Panels (g)–(i) show cumulative carbon absorption by the land surface. Panels (j)–(l) show cumulative absorption of carbon by the ocean over time. The circles for the *esm-flat10-cdr* experiments indicate the maximum of each time series. A 20-year moving average is applied for the GMST time series (bold line); the faint line shows the original data.

3.2 Global response indicators in flat10 and other experiments

3.2.1 Transient climate response to positive emissions

Figure 6 and Table 2 illustrate the global trajectories and summary indicators of the ESMs that participated in the experiment set in both *esm-flat10* and *1pctCO2* (drawing on results from Arora et al., 2020). Figure 6a shows that this compatible emissions time series is time-varying and model-dependent – with typical behavior showing compatible emis-

sions growing from $\sim 10 \text{ PgC yr}^{-1}$ at the start of the experiment to between 16 and 22 PgC yr^{-1} at the time at which cumulative emissions reach 1000 PgC . As such, compatible cumulative emissions are weighted towards the end of the experiment – the mean result exceeds 500 PgC in year 39 and 1000 PgC in year 65 (Fig. 6b). Compatible emissions in *1pctCO2* are also significantly greater than current anthropogenic emissions ($11.1 \pm 0.8 \text{ PgC yr}^{-1}$ in 2023) (Friedlingstein et al., 2023).

Figure 7 compares distributions of $T_{100 \text{ yr}}$ and ZEC computed using the *1pctCO2* and *esm-flat10* approaches. We see,

Table 2. Summary diagnostics from flat10MIP and ZECMIP/1petCO2 (MacDougall et al., 2020) experiments for Earth system models and simple climate models, which reported $T_{1000}\text{PgC}$, ZEC50 and ZEC90. $T_{100\text{yr}}$ is the mean warming in years 91–110 in esm-flat10, and ZEC50/90/100 is the zero emissions commitment, measured as the warming in esm-flat10-zec between years 100 and years 150/190/200, respectively. $t - \text{PW}$ is the time difference in years of peak warming in esm-flat10-cdr relative to net zero in year 150, and (TNZ/TR1000/TR0) is the warming in years (150/200/310) in esm-flat10-cdr relative to warming in years (125/100/0) in esm-flat10 when cumulative emissions are (1250 GtC/1000 GtC/0 GtC), respectively.

Earth system models	$T_{1000}\text{PgC}$ (TCRE from 1petCO2 at 1000 PgC)	$T_{100\text{yr}}$ (TCRE from flat10)	ZEC50 (from esm-1pet-brch-1000 PgC)	ZEC90 (from esm-1pet-brch-1000 PgC)	ZEC50 (from flat10-zec)	ZEC90 (from flat10-zec)	ZEC100 (from flat10-zec)	$t - \text{PW}$ years	TNZ	TR1000	TR0
Units	K	K	K	K	K	K	K	years	K	K	K
ACCESS-ESM1-5	1.90	1.75	0.01	−0.03	0.21	0.15	0.13	7	0.08	0.23	0.17
CESM2	2.00	1.95	−0.31	−0.17	−0.27	−0.14	−0.13	−10	0.05	0.03	0.42
GFDL-ESM4	1.45	1.45	–	–	−0.21	−0.13	0.11	−29	−0.09	−0.25	−0.11
GISS	1.6 ^b	1.62	−0.11 ^c	−0.06 ^c	0.19	−0.15	−0.24	−4	0.12	0.01	−0.56
NorESM2-LM	1.32	1.18	−0.33	−0.32	−0.23	−0.21	−0.31	−33	−0.03	−0.23	−0.31
MPI-ESM1-2-LR	1.65	1.50	−0.27	−0.37	−0.14	−0.17	−0.24	1	−0.06	−0.24	−0.29
CNRM-ESM2-1	1.73 ^c	1.72	0.06 ^c	0.25 ^c	−0.01	0.067	0.11	−10	0.11	0.03	0.38
UKESM1.2	2.55 ^c	2.50	0.28 ^c	0.33 ^c	0.27	0.19	0.2	−1	0.16	0.48	–
UVic-ESCM-2-10	1.86	1.80	0.04	0.02	0.01	−0.11	0.12	3.0	0.07	0.20	0.25
HadCM3LC-Bris	1.93 ^a	1.53	–	–	−0.02	−0.04	−0.02	6.0	0.03	−0.06	0.15
Simple climate models	$T_{1000}\text{PgC}$ (TCRE from 1petCO2)	$T_{100\text{yr}}$ (TCRE from flat10)	ZEC50 (from esm-1pet-brch-1000 PgC)	ZEC100 (from esm-1pet-brch-1000 PgC)	ZEC50 (from flat10-zec)	ZEC90 (from flat10-zec)	ZEC100 (from flat10-zec)	$t - \text{PW}$	TNZ	TR1000	TR0
MAGICC6	1.71 (1.13, 2.68)	1.59 (1.05, 2.66)	−0.12 (−0.28, 0.19)	−0.18 (−0.44, 0.22)	−0.11 (−0.23, 0.12)	−0.16 (−0.35, 0.14)	−0.16 (−0.38, 0.15)	−6.00 (−13.05, 8.00)	−0.04 (−0.11, 0.11)	−0.19 (−0.43, 0.27)	−0.56 (−1.06, 0.14)
Fair	1.57 (1.16, 2.10)	1.54 (1.13, 2.18)	−0.02 (−0.19, 0.34)	−0.04 (−0.28, 0.48)	−0.02 (−0.13, 0.25)	−0.03 (−0.22, 0.37)	−0.03 (−0.22, 0.37)	0.00 (−11.0, 15.0)	0.01 (−0.07, 0.17)	−0.03 (−0.26, 0.48)	−0.15 (−0.44, 0.36)
CICERO-SCM	1.69 (1.33, 2.21)	1.58 (1.21, 2.11)	−0.05 (−0.13, 0.09)	−0.25 (−0.33, −0.11)	−0.18 (−0.26, −0.09)	−0.34 (−0.47, −0.24)	−0.34 (−0.47, −0.24)	−10.00 (−16.0, −4.0)	−0.06 (−0.11, 0.0)	−0.35 (−0.52, −0.20)	−0.79 (−1.08, −0.61)

^a Chris Jones (personal communication, 2024). ^b Anastasia Romanou (personal communication, 2024). ^c 1petCO2 responses are from a similar but non-identical previous model version.

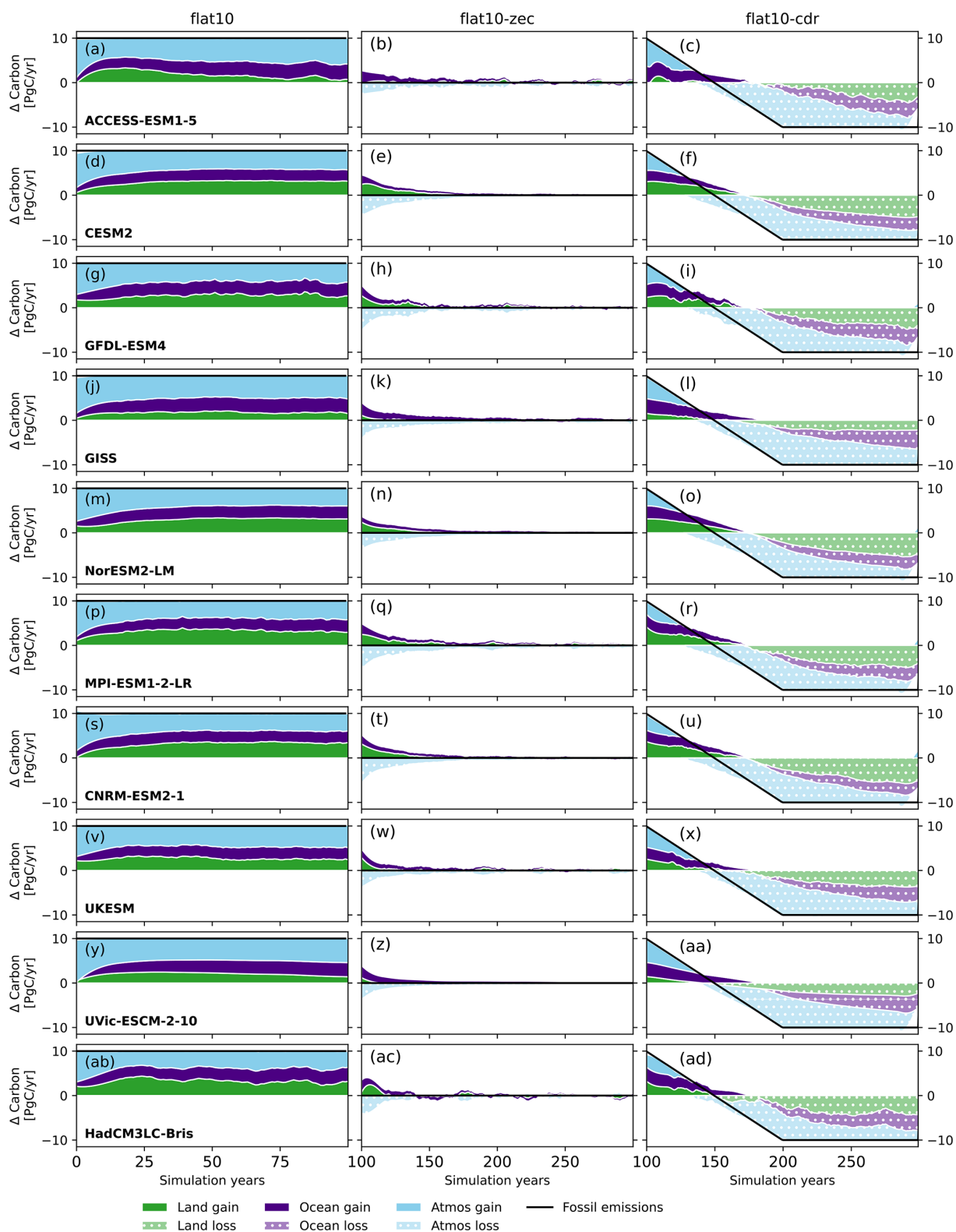


Figure 5. Evolution of carbon sinks in ESMs in flat10MIP. [Light blue/dark blue/green] shading shows the [airborne fraction/ocean fraction/land fraction] of emissions in each year as a function of time. Gains for each domain are shown in solid colors, while losses are shown in light dotted colors. The left-hand column shows the fractions for esm-flat10, where emissions total 10 PgC yr⁻¹. The central column shows results for esm-flat10-zec, where emissions are zero and atmospheric loss is compensated by gains in the land and ocean. The right-hand column shows esm-flat10-cdr, where removals are balanced by losses from each of the pools.

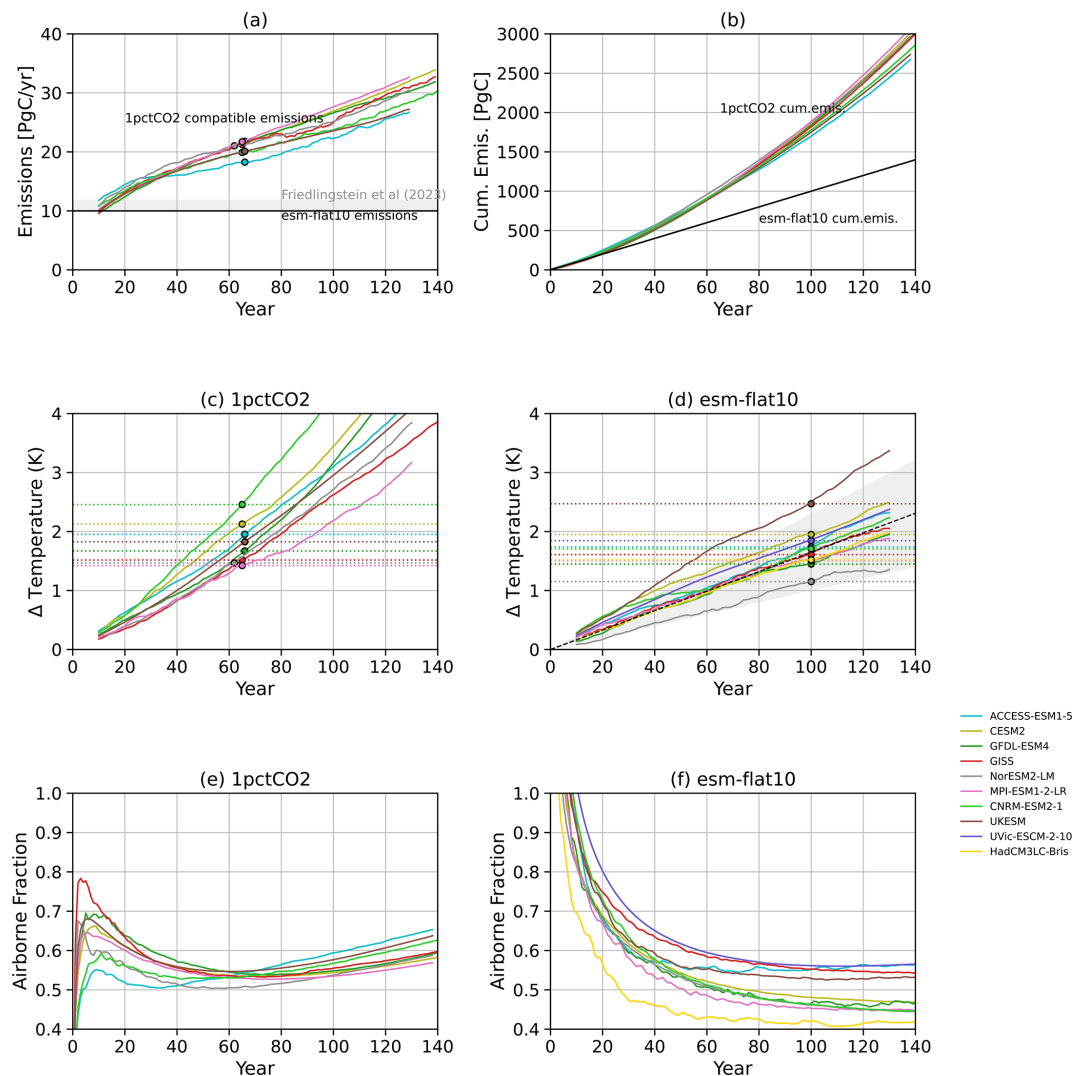


Figure 6. Comparative ESM results for TCRE calculation using 1pctCO₂ and esm-flat10, showing (a), (b) compatible [annual, cumulative] emissions in 1pctCO₂ compared with the constant 10 PgC yr⁻¹ flux in esm-flat10. Annual total anthropogenic carbon emissions in 2023 are shown for context. (c, d) show temperature evolution in [1pctCO₂, esm-flat10]. Colored lines show global model output from available ESMs with a 21-year moving average applied. (e, f) show airborne fraction in [1pctCO₂, esm-flat10]. Circles show results at the time when cumulative emissions reach 1000 PgC. Shaded region in (d) illustrates the range of warming according to the IPCC AR6-assessed likely range of TCRE.

on average, a slight offset such that TCRE estimates in the ESMs have a value that is an average of 0.12 K greater in 1pctCO₂ relative to *esm-flat10* (see Table 1, Fig. 7). This is consistent with Krasting et al. (2014), who found that TCRE estimated at high emissions rates was greater than that estimated using present-day emissions rates and attributed the difference to a greater disequilibrium between land/atmosphere and ocean response states when emissions rates are very high. Similarly, distributions in the simple climate models MAGICC and CICERO-SCM show that $T_{1000\text{PgC}}$ from 1pctCO₂ is on average about 0.1 K greater than $T_{100\text{yr}}$ from *esm-flat10*. The third simple climate model, FaIR, shows comparable values of $T_{1000\text{PgC}}$ and $T_{100\text{yr}}$ (Fig. 7). Given that

probabilistic calibration is performed independently for each SCM, it is not easy to attribute these differences to structural differences between the models or to choices of probabilistic parameter calibration strategy. Figure 8a shows correlations between $T_{1000\text{PgC}}$ and $T_{100\text{yr}}$ – re-enforcing the small average offset between the two approaches – though the gradient of the best fit line is near-unity.

3.2.2 Zero Emissions Commitment

For ZEC, however, we see greater differences between the concentration-driven approach and the emissions-driven approach than for TCRE (Figs. 7, 8, 9). In the SCM ensembles,

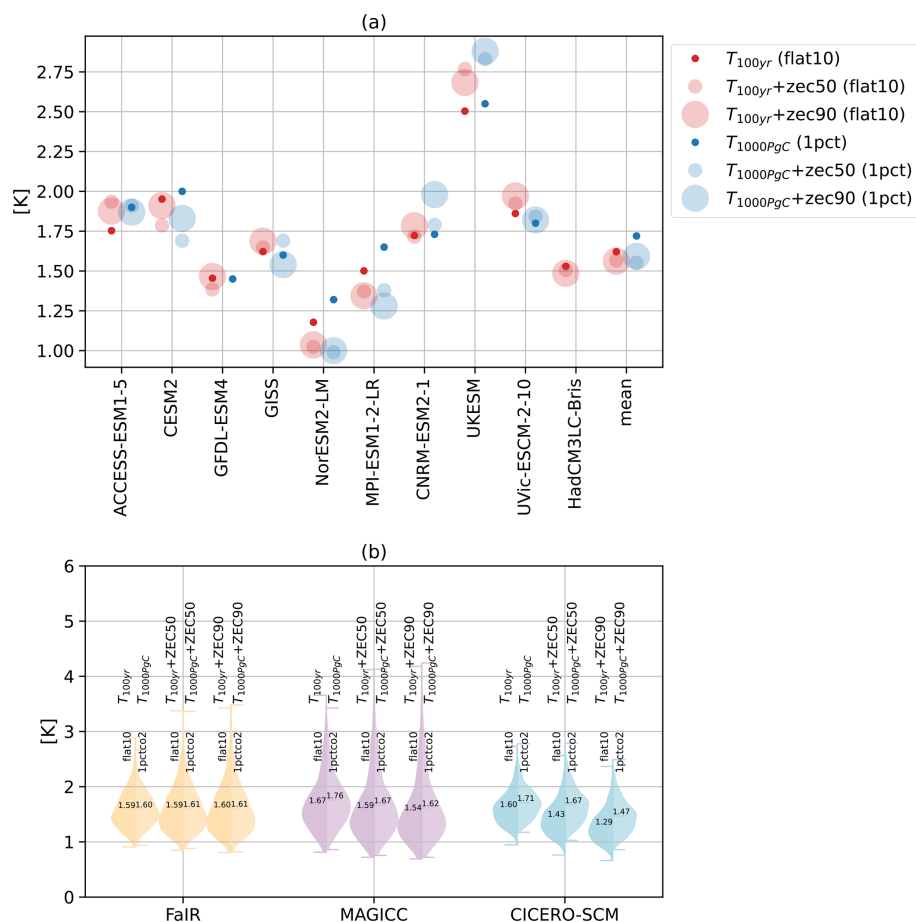


Figure 7. (a) A comparison of $T_{100\text{yr}}$ [flat10] and $T_{1000\text{PgC}}$ [1pctCO₂], $T_{100\text{yr}+\text{ZEC50}}$ [flat10-zec]/ $T_{1000\text{PgC}+\text{ZEC50}}$ [1pctCO₂] (small transparent points) and $T_{100\text{yr}+\text{ZEC90}}$ [flat10-zec]/ $T_{1000\text{PgC}+\text{ZEC90}}$ [1pctCO₂] (large transparent points) for ESMs participating in flat10MIP (red) and ZECMIP (blue, where available). The final point is the multi-model mean for cases where there exist complete runs for both ZECMIP and flat10MIP [ACCESS, CESM2, NorESM, MPI-ESM and CNRM-ESM2]. (b) For SCMs, violin plots showing distributions of $T_{100\text{yr}}$, $T_{100\text{yr}+\text{ZEC50}}$ and $T_{100\text{yr}+\text{ZEC90}}$ for *esm-flat10* (left) and *1pctCO₂* (right).

ZEC50 and ZEC90 are on the order of 25 % smaller if measured using the flat10-zec protocol relative to the ZECMIP protocol (this is true irrespective of whether ZEC is positive or negative, Fig. 8c, d). This is consistent with MacDougall et al. (2020), who found smaller ZEC in experiments with lower emission rates up to the point of net zero, proposing that both warming and carbon cycle responses are closer to equilibrium. We also note that one SCM, CICERO-SCM, shows more consistently negative values of both ZEC50 and ZEC90 when quantified via flat10MIP than by ZECMIP (Fig. 8c, d). ESMs are also consistent with the relationship of $\sim 25\%$ smaller absolute magnitudes in ZEC50 and ZEC90, albeit with larger scatter. Some models (NorESM, CESM2, MPI, CNRM) in the ZECMIP experiment suggest an apparent short-term warming pulse following the cessation of emissions, which is less pronounced in the *esm-flat10-zec* experiment (Fig. 9) – but additional ensemble members are required to properly quantify this behavior. In the MPI model,

this is consistent with findings that TCRE was higher using the ZECMIP protocol compared to flat10MIP (Fig. 1d in Winkler et al., 2024).

It is also evident that *total* warming measured from pre-industrial levels 100 years after emissions cease (i.e., $T_{100\text{yr}+\text{ZEC90}}$ from *esm-flat10* and $T_{1000\text{PgC}+\text{ZEC90}}$ from *1pctCO₂*) is more consistent between the ZECMIP and flat10MIP protocols (Fig. 8b) than either TCRE or ZEC90 independently – indicating that total warming following a period of emissions followed by cessation is path-independent in the models considered here. However, we continue to see in the mean values of the SCM distributions (Fig. 7b) for MAGICC and CICERO-SCM that $T_{1000\text{PgC}+\text{ZEC90}}$ is ~ 0.1 K greater in *esm-flat10-zec* than in *esm-1pct-brch-1000PgC*. FaIR is again consistent between the two approaches, with only a 0.01 K difference between the mean values. For the ESMs (Fig. 7a), we note that the multi-model mean $T_{1000\text{PgC}+\text{ZEC90}}$ is 0.05 K greater for *esm-1pct-brch-*

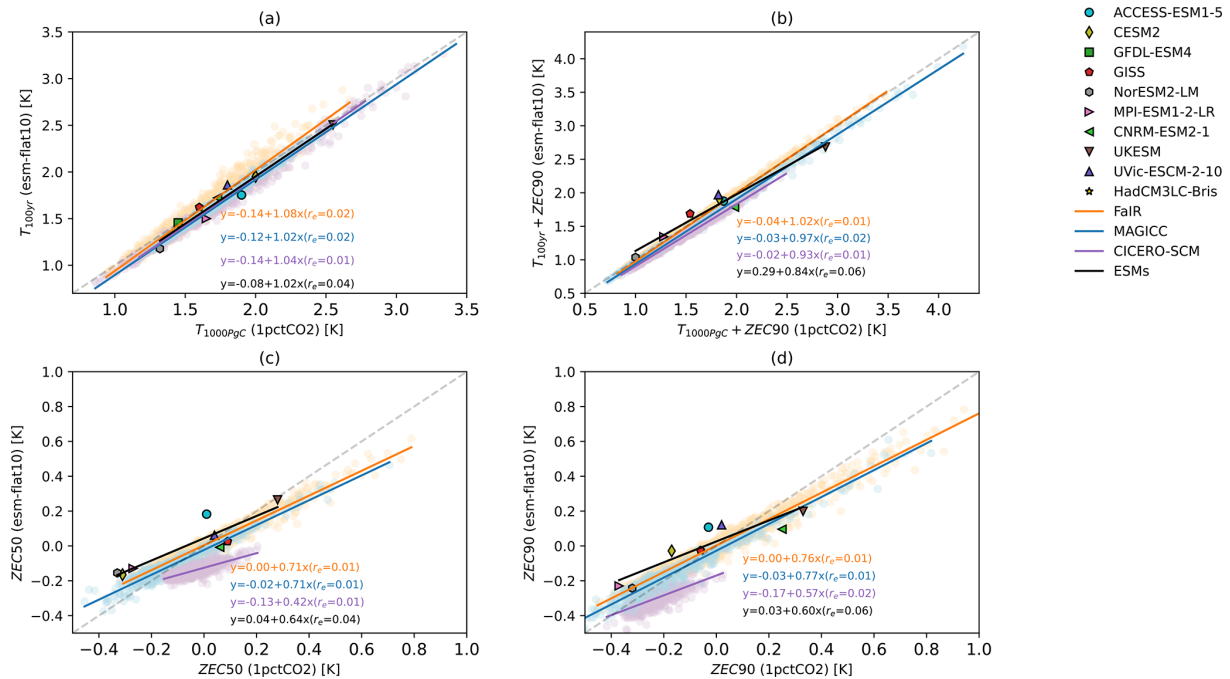


Figure 8. Comparison metrics assessed using the flat10MIP methodology and 1pctCO₂-based experiments. ESM summary metrics are $T_{100\text{yr}}$, ZEC50 and ZEC90 for *esm-flat10* and $T_{1000\text{PgC}}$ and ZEC50 and ZEC90 for 1pctCO₂. Filled shapes illustrate values assessed from ESMs, and pale dots illustrate members of the simple climate model ensembles for (FaIR, MAGICC, CICERO-SCM) in (orange, blue, purple). Straight lines show the least-squares best fits for the ESMs (black) and SCMs.

1000PgC than $T_{100\text{yr}} + \text{ZEC90}$ for *esm-flat10-zec* (whereas the mean $T_{1000\text{PgC}}$ is 0.12 K greater than $T_{100\text{yr}}$).

Our results in general suggest that the weighting of compatible emissions towards the end of the simulation in 1pctCO₂, as well as the shorter total time period over which emissions occur in 1pctCO₂ (~70 vs 100 years), has an impact on both the estimate of TCRE and the transient response following cessation of emissions. We tend to see slightly greater estimated values of TCRE in 1pctCO₂, with most models exhibiting short-term continued warming followed by cooling in the decades following cessation of emissions. In contrast, the behavior in *esm-flat10-zec* exhibits slightly less warming during the positive emissions phase and less adjustment afterwards, resulting in lower values for TCRE and smaller magnitudes (either positive or negative) of ZEC50 and ZEC90. The finding that ZEC50/90 from *esm-flat10* is lower than ZECMIP estimates is consistent with the findings of Jenkins et al. (2022), who found that ZEC is modulated by “average cumulative emissions over the period”, a metric that is different under the two experimental designs.

3.2.3 Climate reversibility experiments

The global mean results for *esm-flat10-cdr* are shown in Fig. 4. The temperature response at year 300 (when cumulative emissions return to zero) shows a range of −0.7 to +0.5 K, indicating notable deviations from cumulative emis-

sions proportionality with residual warming or cooling depending on the model. Figure 10 illustrates global-scale hysteresis in the ESM results, showing the change in global mean surface temperature as a function of cumulative emissions. Though all models broadly indicate proportionality between temperature and cumulative emissions, there are some notable deviations. Many models indicate some hysteresis, either positive (ACCESS) or negative (GFDL, NorESM, MPI-ESM), between the upward and downward branches of the simulation, and some (CESM2, GFDL, CNRM) appear to show a change in temperature/cumulative emissions response during the course of the downward branch. Overlain as dotted lines on each panel of Fig. 10 is a null hypothesis, informed only by TCRE and ZEC90 from the *esm-flat10* and *esm-flat10-zec* experiments, that temperatures in the net-negative emissions period of *esm-flat10-cdr* might be explained as a combination of the TCRE · I_{em} + ZEC terms (Koven et al., 2022, 2023). This framework explains some, but not all, of the hysteresis observed; in particular, some of the models (e.g., GFDL, UKESM) show a larger hysteresis than predicted by ZEC90, and the TCRE + ZEC framework does not predict the deviations late in the downward branch for those models that have such dynamics. Alternative frameworks such as RAZE (Jenkins et al., 2022) explain other key features – such as the expectation in a symmetrical experiment such as *esm-flat10-cdr* that half of the ZEC is manifested at the time of net zero. A unifying explanation for

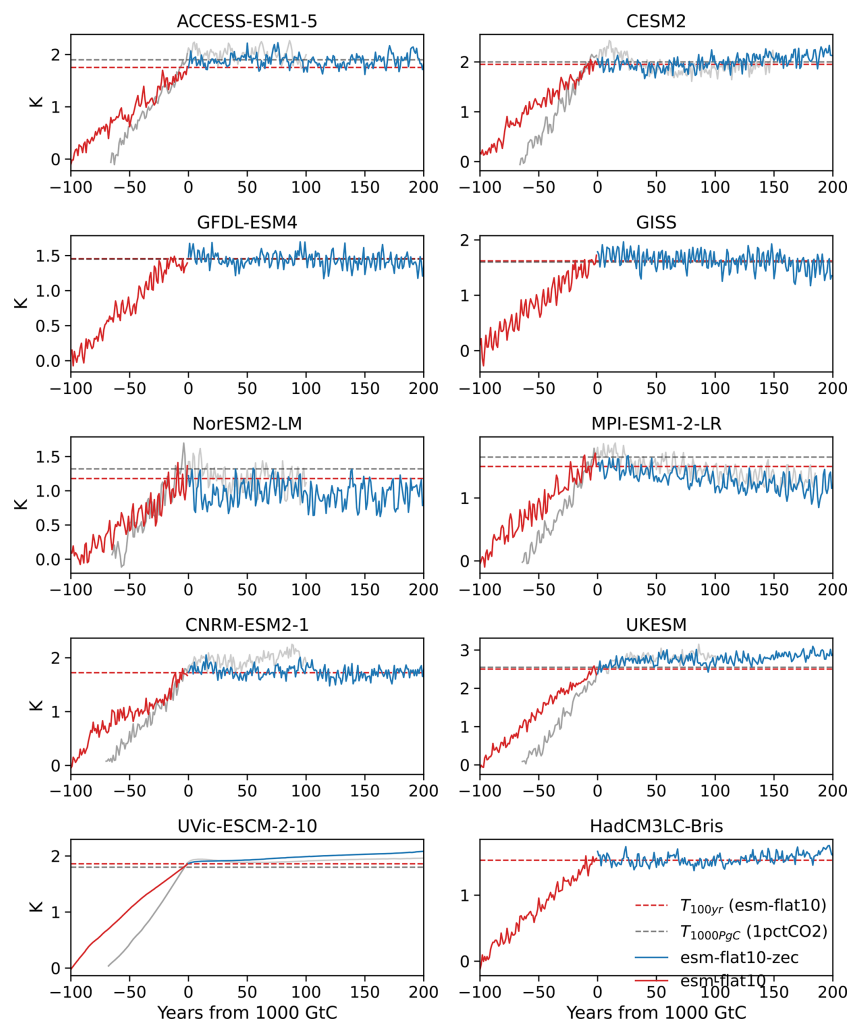


Figure 9. Global mean temperature change evolution for ESMs participating in flat10MIP (bold colors), in the context of 1pctCO₂ (gray) and ZECMIP, where comparable simulations with the same model version are available (faded colors). Red lines show the positive emissions period (10 PgC yr^{−1} for flat10 (solid red) and 1pctCO₂-compatible emissions for ZECMIP), and blue/gray lines show the zero-emissions period for *esm-flat10-zec* and *esm-1pct-brch-1000PgC*, respectively. Horizontal dashed lines show $[T_{100\text{yr}}, T_{1000\text{PgC}}]$ as estimated from [*esm-flat10* (red), 1pctCO₂ (gray)].

these frameworks that is accurate both during the net-zero transition and at timescales significantly before and after remains absent from the literature to date.

Figure 11 indicates the ESM ensemble distribution of temperature evolution in the *esm-flat10-cdr*, normalized by expected warming from TCRE. The figure shows that TCRE proportionality is consistent between models in the ensemble, with a relatively small spread during the constant positive emissions phase of the experiment. As the emissions rate reduces and becomes negative, additional spread but no systematic direction of asymmetry is seen relative to expectations from TCRE alone – and this spread remains constant throughout the negative emissions phase. We can categorize this uncertainty as approximately $\pm 10\%$ of TCRE, which remains broadly constant over time during the negative phase.

Figure 12 shows how additional climate indicators vary with cumulative emissions. Atmospheric carbon dioxide levels are consistently lower on the downward branch, but cumulative land carbon sink hysteresis varies by model – with some models showing significantly larger cumulative land carbon sinks on the downward branch (e.g. NorESM), while other models (e.g., GISS, HadCM3LC) show cumulative sinks proportional to cumulative emissions on both the upward and downward branches. Similarly, all models show a hysteresis in cumulative ocean sink strength with cumulative emissions, with between 100 and 200 PgC remaining in the ocean in year 300 of *esm-flat10-cdr*.

We identify a number of new metrics (TNZ, TR1000, TR0 and tPW; Fig. 2, Table 2), which are aimed at capturing aspects of climate reversibility and commitment from the *flat10-cdr* experiment. As noted above, each of these mea-

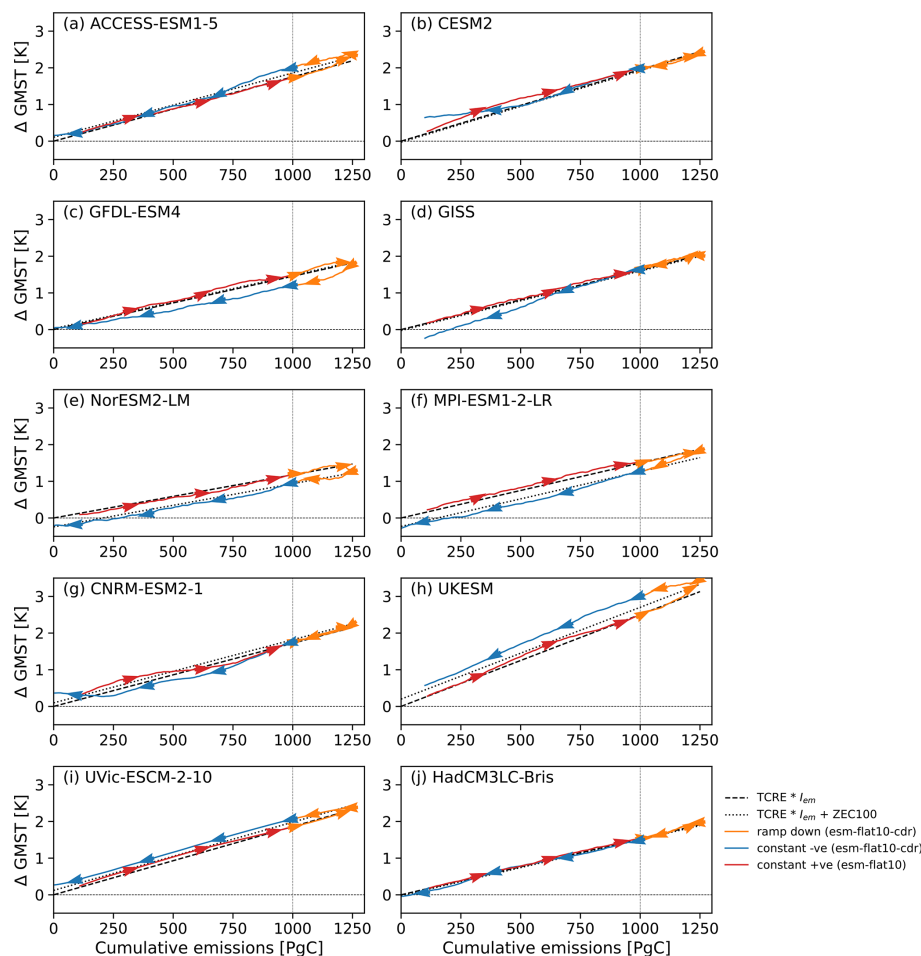


Figure 10. Global mean temperature relationship with cumulative emissions for the ESMs. A 21-year moving average is applied for the GMST time series. Arrows show the direction of time, with [red, yellow, blue] lines showing [constant positive, ramp down, constant negative] phases of the experiment. Black dashed and dotted lines show $\text{TCRE} \cdot \text{cumulative emissions}$ (I_{em}) and $\text{TCRE} \cdot \text{cumulative emissions} + \text{ZEC90}$ for each model, using TCRE and ZEC90 values as calculated from esm-flat10 and esm-flat10-zec.

sures a distinct aspect of potential deviation from perfect TCRE proportionality and thus, like ZEC, would have a value of exactly 0 if temperatures were exactly proportional to the cumulative emissions.

Two of these metrics measure the hysteresis around the net-zero transition: tPW is the time offset of peak warming relative to net zero, whereas TNZ is the difference in realized temperature at net zero relative to what one would predict through TCRE proportionality. Figure 13 shows how these metrics relate to each other and to TCRE and ZEC. With the exception of TCRE, all metrics show a positive correlation with all other metrics, particularly for SCMs. ESMs show greater scatter across a number of the pairwise relationships than the SCMs, reflecting a greater diversity of potential dynamics arising from their high complexity than are being captured in the more parsimonious relationships represented by the SCMs. For example, in each of the SCM ensembles, tPW and TNZ are highly and consistently related, but a num-

ber of ESMs (CNRM, GISS, CESM, UKESM) lie outside of the SCM distributions such that we see peak warming significantly before net zero with greater warming than one would expect from cumulative emissions proportionality (contours in Fig. 13 indicate 90 % of the ensemble distribution; for tPW vs TNZ, the ESM results lie outside of the 99th percentile (not shown)). Similar differences are seen in the relationship between tPW and ZEC50, with two ESMs showing peak warming occurring particularly early. This hints at behavior in the ESMs that might not be represented in the current generation of SCM parameter ensembles. This could potentially be due to an absence of ZEC or reversibility-related metrics used in the calibration pipelines. Alternatively, this could potentially be due to a number of different processes, e.g., ocean circulation processes such as AMOC weakening, which are not represented in current SCMs (Schwinger et al., 2022), but larger ESM initial condition ensembles are necessary to have confidence in the ESM metrics in the presence of

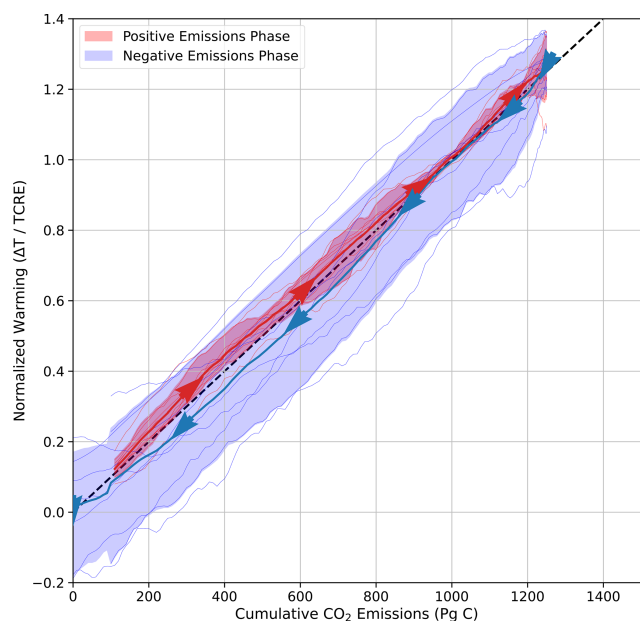


Figure 11. Global mean temperature relationship with cumulative emissions for the ESM distribution in *esm-flat10-cdr*, normalized by TCRE as estimated from *esm-flat10*. A 21-year moving average is applied for the GMST time series. Arrows show the direction of time, with [red, blue] lines showing multi-model mean [positive, negative] emissions phases of the experiment. [Red, blue] shaded regions indicate the [10th, 90th] percentiles of the ESM ensemble temperature distribution at a given cumulative emissions level. Black dashed line shows the normalized relationship between cumulative emissions and warming.

internal variability. This discrepancy could potentially be related to studies that have found inconsistencies between the temporal dynamics of the ocean heat and carbon uptake in ESM and SCM ensembles (S  f  rian et al., 2024) and would benefit from further investigation.

Another pattern that emerges in Fig. 13 is the greater correlation captured in the short-term metrics (ZEC50, TR1000) than in the longest-term metric (TR0) that shows greater scatter with the ZEC and other reversibility metrics. This high correlation (e.g., between ZEC50 and TR1000 and between ZEC100 and TR1000) has an important implication: most of the uncertainty present in the reversibility of GMST (although not necessarily regionally or in other metrics; Schleussner et al., 2024) under an idealized overshoot scenario will also be present under zero emissions at the same level of cumulative emissions that avoids the overshoot.

4 Summary and conclusions

The finding of a near-linear relationship between cumulative carbon emissions and global mean temperature (Allen et al., 2009; Matthews et al., 2009) enabled recent climate policy to link desired limits for warming to an allowable budget of

remaining carbon emissions. The years following have seen regular efforts to quantify remaining carbon budgets for the Paris Agreement goals (Lamboll et al., 2023), with scenarios built on this premise (Rogelj et al., 2019a), and refinement in the treatment of how to incorporate non-CO₂ emissions into this framework (Cain et al., 2019; Jenkins et al., 2018; Mengis and Matthews, 2020).

Further, an increased understanding has emerged that the TCRE relationship is an approximation, owing to fortuitous cancellation of terms in heat and carbon uptake in many models, but this cancellation is not perfect, and a Zero Emissions Commitment (ZEC; Palazzo Corner et al., 2023) may result in residual carbon-induced warming (or cooling) even if carbon emissions are held at net zero. This ZEC effect may cause peak temperatures to be seen before or after net zero (Koven et al., 2023). Building confidence in this timing is important; if peak temperatures occur after net zero, this may create climate adaptation challenges that might not otherwise be planned for if simple TCRE proportionality is used to predict warming outcomes.

Operational methods of quantifying TCRE and ZEC to date have utilized existing default Earth system model diagnostic experiments that have focused on the response of the Earth system to a prescribed concentration pathway – generally an exponential increase of 1 % per year – as an idealized proxy for climate change induced by carbon dioxide. It is then possible to calculate compatible CO₂ emissions, specific to a given model, to frame the output of these experiments in terms of emissions (Jones et al., 2016; Liddicoat et al., 2021) and calculate TCRE, with branched zero-emission experiments to calculate ZEC (Jones et al., 2019).

Although these experiments have been highly useful in helping to quantify TCRE and ZEC efficiently using mostly pre-existing simulations, the use of a concentration-driven diagnostic run has limitations (Gregory et al., 2015; MacDougall, 2019) – emissions are specific to a given model and are highly weighted towards the end of the experiment, when emissions rates greatly exceed present-day or projected levels. As such, given that experiments to measure ZEC seek fundamentally to measure subtle, second-order effects, there is an argument for new diagnostic experiments that cleanly measure TCRE, ZEC and climate reversibility using reproducible and cleanly interpretable benchmarks.

In this study, we have demonstrated the utility of a new set of idealized experiments that can be applied with both complex and simple Earth system models. This “flat10” framework is based upon a small number of variants around a simple core experiment, where emissions are fixed at 10 PgCyr^{−1} for 100 years, a rate that approximates current anthropogenic carbon emissions and conveniently totals 1000 PgC after 100 years of simulation, with the temperature in year 100 thus providing a direct assessment of TCRE. Branch experiments from this point can measure the Zero Emissions Commitment (with emissions set to zero in year 100) and climate reversibility (with an idealized net-

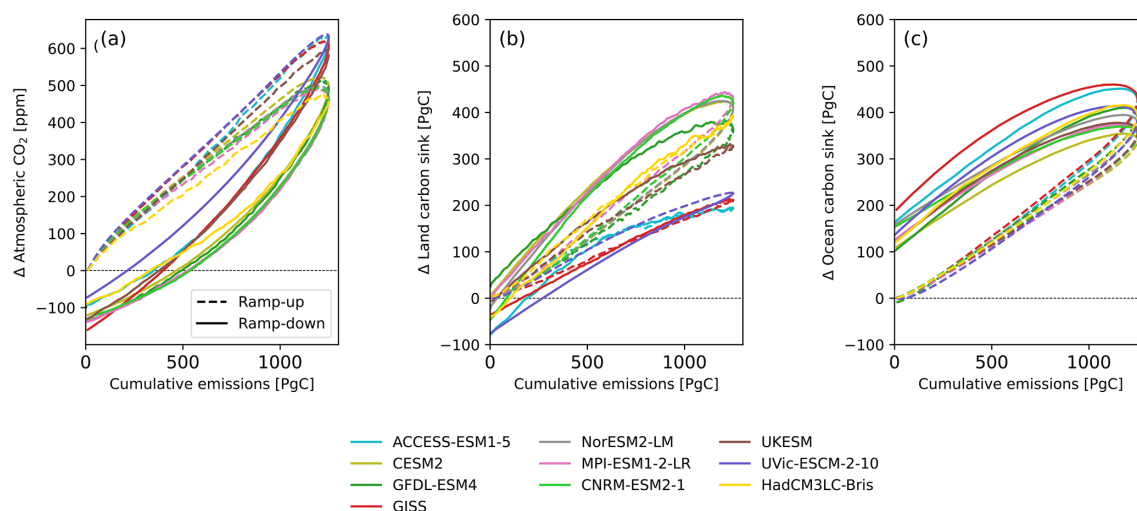


Figure 12. Climate indicators as a function of cumulative emissions for the ESMs. A 21-year moving average is applied for all time series.

zero and net-negative emissions pathway in which cumulative emissions reach zero by the end of the experiment). Along with these experiments, we propose diagnostic measures that serve to measure different aspects of non-TCRE behavior and how they relate to the likely outcomes of real-world net-zero and net-negative emissions proposals. These experiments complement a similar experimental design being developed and run by the Tipping Point Modeling Inter-comparison Project, or TIPMIP (Winkelmann et al., 2025). TIPMIP experiments also follow a prescribed constant CO₂ emission pathway, but the emissions are tailored for each model to result in a common warming rate of 2 °C per century. As such, the goal of TIPMIP is to examine the behavior of ESMs at common levels of global warming, while the goal of the flat10MIP experiments is to examine the behavior of ESMs under common external forcing. Furthermore, for future experiments using the TIPMIP protocol, the flat emissions pathway in *esm-flat10* will likely provide a more accurate TCRE estimate for calibrating the emissions rate required for constant warming rates.

These experiments form part of the “fast track” recommendation for CMIP7, through which the climate change research community will gain a greater understanding of ZEC and reversibility behavior in the next generation of climate models. Here, to illustrate the potential for these simulations to diagnose a broad suite of climate response metrics, we demonstrate the results of the flat10MIP experiments for a subset of CMIP6-generation models and the simple climate models used in the IPCC 6th Assessment Report. We find, as expected, that TCRE is first-order consistent whether calculated using the *lpctCO2* simulations or using *esm-flat10* simulations – but also that the values of ZEC estimated with *lpctCO2* tend to be greater than for *esm-flat10-zec*, indicating that the weighting of emissions towards the latter part of the *lpctCO2* experiment may increase transient warming or

cooling trends, potentially driving a larger ZEC than would be seen in a realistic emissions scenario.

We also find a large diversity of ESM behavior in the climate reversibility experiment *esm-flat10-cdr*, including that peak warming can occur before or after net-zero emissions and is not necessarily predictable from a combination of TCRE and ZEC (consistent with existing studies; Asaadi et al., 2024) with a range of carbon sink evolutions in different ESMs, in both the positive and negative emissions phases of the experiment. Models strongly disagree on the timing and amplitude of peak land carbon uptake, some showing peak uptake decades before and others decades after the net-zero transition. In addition to the difference in carbon cycle representations, the diverse transient carbon sinks behavior can also be attributed to the difference in ESMs’ pre-industrial states or initial conditions (Tjiputra et al., 2025). There is also evidence of state changes during the negative emissions phase, with some models showing a change in the rate of cooling per unit carbon removed – potentially indicating dynamical changes in ocean circulation that might impact carbon–climate dynamics.

However, in this study, our scope for understanding this diversity is limited: we present the experimental design for CMIP7 plus global-scale results from ESMs and SCMs that are now available to the community. A detailed explanation of the process will be presented in follow-up studies considering land and ocean dynamical processes from the flat10MIP ensemble, where we hope for wide community engagement.

We argue that emissions-driven diagnostic experiments are the cleanest method for diagnosing the response to climate forcings on a range of relevant timescales. In the future, we would imagine these experiments becoming elements of a wider set of idealized, yet policy-relevant emission-driven

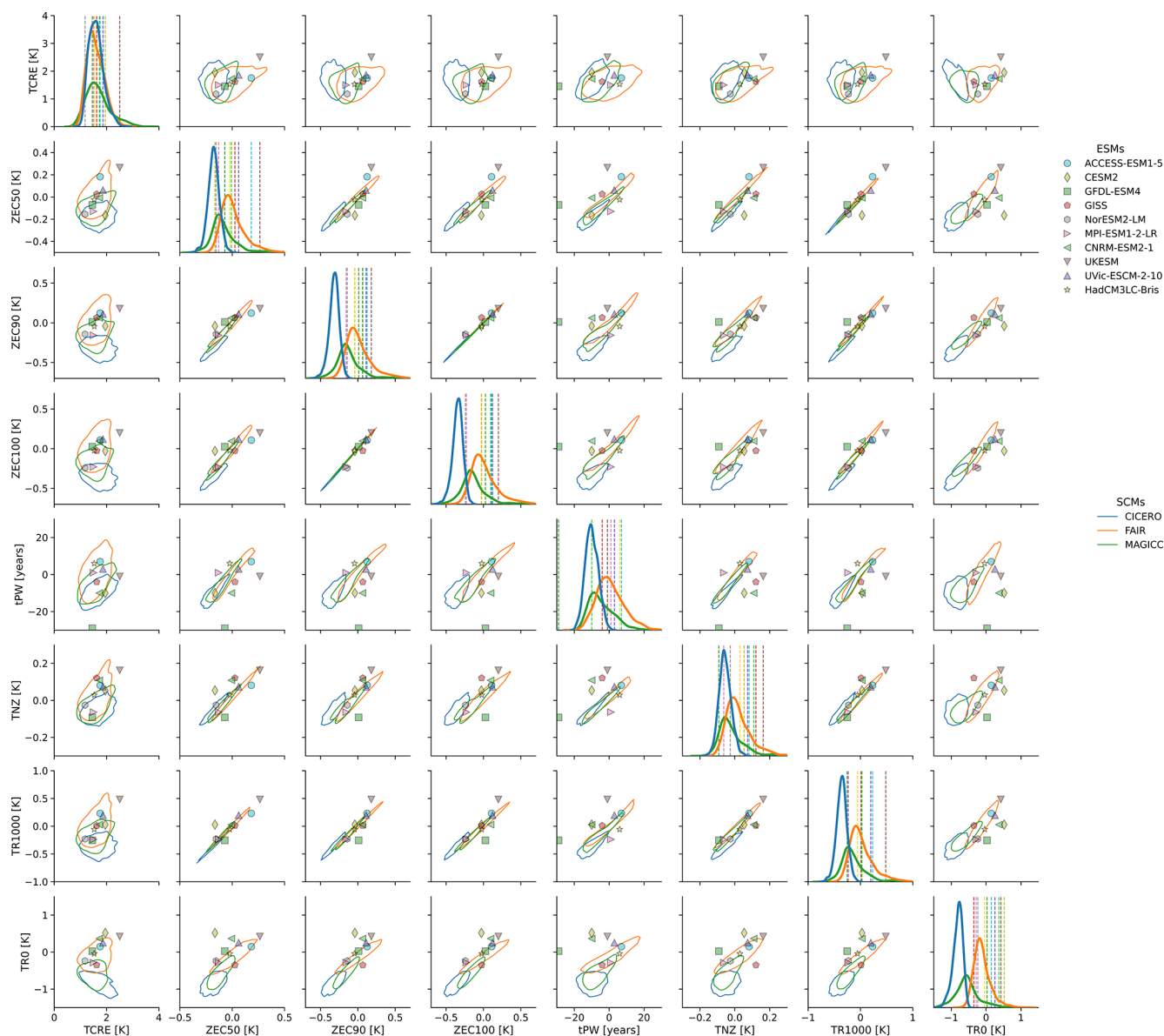


Figure 13. Matrix of relationships between metrics quantified here. Shown are pairwise plots between the following metrics: TCRE (T100yr for flat10 and T1000PgC for 1pctCO₂), ZEC50, ZEC90, ZEC100, tPW, TNZ, TR1000 and TR0. SCM ensembles are shown as contours at the 10th percentile of the joint distribution for each pairwise comparison (such that 90 % of points lie within the contours, with FaIR, MAGICC and CICERO-SCM in orange, green and blue, respectively). ESMs are shown as individual points. Diagonal panels show histograms (SCMs) and discrete values (ESMs) for each of the metrics diagnosed here.

experiments that can efficiently categorize either a simple or complex climate model's response to climate forcers.

In the present study, this has been limited to a specific trajectory of carbon emissions that has been chosen pragmatically to minimize computational burden. Future understanding would be increased by adding to this archive, in terms of both larger initial condition ensembles to improve confidence in ZEC and reversibility metrics and perturbed parameter ensembles in ESMs to understand conditionalities

on model calibration choices, including longer simulations to understand longer timescales of commitment.

Despite these caveats, the present effort has indicated that some models exhibit nonlinear and threshold behavior. Further experiments would be required to fully document the conditions under which such transitions occur. As such, future CMIP activities might consider a range of flat-n-type experiments spanning warming levels and decarbonization rates to categorize the response of the carbon–climate dynamics to different types of overshoot pathways. Also,

as ESMs increasingly seek to represent the response to a range of activities (land use change, methane and nitrous oxide emissions, among others), it will become necessary to cleanly categorize the response to each of these in a reproducible fashion – creating a necessity for well-crafted experiments to cleanly represent model responses to non-fossil- CO_2 forcings. A shift towards emissions-driven modeling is essential to produce relevant climate simulations for increasingly specific emissions pathways referred to in climate policy, and this requires a new generation of emissions-driven diagnostic experiments.

Appendix A: Participating models

A1 Earth system models

The flat10MIP experiments are included in the recommended CMIP7 “fast track” – a subset of experiments highlighted for particular relevance as input for climate change assessments. In preparation for this recommendation, a trial model inter-comparison conducted the *esm-flat10* experiment set for a collection of eight Earth system models from the CMIP6 ensemble (Eyring et al., 2016) and one intermediate complexity model.

- *ACCESS-ESM1-5* (Ziehn et al., 2020). Atmosphere: UM7.3 (Walters et al., 2019) at $1.875^\circ \times 1.25^\circ$ resolution
Ocean: MOM5 (Griffies, 2012) at $1^\circ \times 1^\circ$ resolution
Land: CABLE2.4 (Kowalczyk et al., 2013)
ACCESS-ESM1-5 features a coupled carbon–nitrogen–phosphorus cycle in the land component (CABLE2.4), with an ocean provided by the GFDL MOM5 model.
- *CESM2* (Danabasoglu et al., 2020). Atmosphere: CAM6 (Bogenschütz et al., 2018) at 1° resolution
Ocean: POP2 (Smith et al., 2010) at $1^\circ \times 1^\circ$ resolution
Land: CLM5 (Lawrence et al., 2019)
CESM2 includes updated aerosol–cloud interactions in CAM6, while CLM5 provides new parameterizations for carbon and nitrogen interactions in terrestrial ecosystems, and POP2 emphasizes ocean–ice dynamics.
- *GFDL-ESM4* (Dunne et al., 2020). Atmosphere: AM4.1 (Horowitz et al., 2020) at $1^\circ \times 1^\circ$ resolution
Ocean: MOM6 (Adcroft et al., 2019)
Land: LM4.1 (Shevliakova et al., 2024)
GFDL-ESM4 uses MOM6 for advanced representations of ocean circulation and biogeochemical processes, with AM4.1 providing a fully coupled aerosol

and cloud interaction system. LM4.1 emphasizes nutrient constraints on land carbon cycles.

- *GISS-E2-1-G* (Kelley et al., 2020). Atmosphere: ModelE (Schmidt et al., 2014) at $2^\circ \times 2.5^\circ$ resolution (Kelley et al., 2020)
Ocean: GISS Ocean v1 at $1^\circ \times 1^\circ$ resolution
Land: The vegetation model is the Ent Terrestrial Biosphere Model (Kim et al., 2015) with a prescribed leaf area index and a prescribed interannual variation of land use and land cover (LULC) change; interactive with the carbon cycle (Ito et al., 2020)
Ocean carbon: NASA Ocean Biogeochemical Model (GISS version NOBMg, Romanou et al., 2013; Ito et al., 2020; Lerner et al., 2021)
- *HadCM3LC-Bris*. Atmosphere: HadAM3 (Pope et al., 2000), $3.75^\circ \times 2.5^\circ$ resolution, 19 vertical levels
Ocean: HadCM3L (Cox et al., 2000), $3.75^\circ \times 2.5^\circ$ resolution, 20 vertical levels
Land: MOSES-2 (Essery et al., 2003), with dynamic vegetation and nine plant functional types (Cox, 2001)
Ocean BGC: HadOCC (Palmer and Totterdell, 2001) marine biogeochemistry with NPZD biology model
HadCM3LC-Bris is based on the HadCM3 climate model (Gordon et al., 2000), adapted for use with an interactive carbon cycle by adopting lower ocean resolution (Cox et al., 2000) and subsequently modified slightly for use on Bristol HPC (Valdes et al., 2017).
- *NorESM2-LM* (Seland et al., 2020). Atmosphere: CAM6 (Bogenschütz et al., 2018) at $2^\circ \times 2^\circ$ resolution (with modifications)
Ocean: BLOM-iHAMOCC (Tjiputra et al., 2020)
Land: CLM5 (Lawrence et al., 2019)
NorESM2-LM shares land and some atmosphere elements with CESM2 but modifies CAM6 to include updated aerosol and cloud microphysical schemes and uses the isopycnal-coordinate BLOM for ocean processes, which improves deep ocean mixing simulations.
- *MPI-ESM1-2-LR* (Mauritsen et al., 2019; MPI, 2024). Atmosphere: ECHAM6.3 at $1.875^\circ \times 1.875^\circ$ resolution
Ocean: MPIOM (Jungclaus et al., 2013) at $1.5^\circ \times 1.5^\circ$ resolution
Land: JSBACH3 (Reick et al., 2021)
MPI-ESM1-2-LR utilizes ECHAM6.3, featuring updates in atmospheric chemistry processes, while MPIOM improves ocean heat transport. JSBACH3 integrates biogeophysical and biogeochemical interactions.

– *CNRM-ESM2-1* (Séférian et al., 2019).

Atmosphere: ARPEGE-Climat version 6 (Roehrig et al., 2020) at $1.4^\circ \times 1.4^\circ$ resolution

Ocean: NEMO (Madec et al., 2017) version 3.6 at $1^\circ \times 1^\circ$ resolution

Land: ISBA (Decharme et al., 2019)

CNRM-ESM2-1 features NEMO 3.6, which includes advanced parameterizations of ocean mixing, and ARPEGE-Climat for atmospheric dynamics, with updates in stratospheric processes and land–atmosphere coupling through ISBA.

– *UKESM1* (Sellar et al., 2019).

Atmosphere: HadGEM3-GA7.1 (Walters et al., 2019) at $1.875^\circ \times 1.25^\circ$ resolution

Ocean: NEMO3.6 (Madec et al., 2017) at $1^\circ \times 1^\circ$ resolution

Land: JULES (Best et al., 2011)

UKESM1 includes JULES, which features dynamic vegetation and coupled nitrogen cycles, along with HadGEM3-GA7.1, which provides improved stratosphere–troposphere interactions and cloud–aerosol physics relative to previous versions.

A2 Intermediate-complexity models

– *UVic ESCM 2.10* (Mengis et al., 2020).

Atmosphere: 2D energy moisture balance model with $3.6^\circ \times 1.8^\circ$ resolution (Fanning and Weaver, 1996)

Ocean: MOM2 $3.6^\circ \times 1.8^\circ$ (Pacanowski et al., 1998) with thermodynamic–dynamic sea ice model (Bitz et al., 2001)

Land: Dynamic vegetation with five plant functional types (Meissner et al., 2003), 14 layers of soil, permafrost (MacDougall and Knutti, 2016) and no N, P cycle

Ocean: NPZD model with two nutrients (N, P) and the Fe limitation scheme (Keller et al., 2012).

A3 Simple climate models

We also include simulations from three simple climate models that provided climate assessments in the IPCC AR6 WG3 assessment (IPCC, 2023c).

– *MAGICC6* (Meinshausen et al., 2011).

MAGICC6 is a reduced-complexity model that uses simplified representations of global carbon cycles and radiative forcing, allowing for rapid simulation of emissions-driven climate pathways.

– *FaIR* (Smith et al., 2018).

FaIR uses simplified equations to model temperature responses and radiative forcing – using pulse-response assumptions to model carbon and thermal responses to climate forcings, with flexible configurations that allow it to mimic the behavior of more complex models in emissions-driven scenarios.

– *CICERO-SCM* (Sandstad et al., 2024).

CICERO-SCM is a reduced-complexity model that focuses on simplified representations of carbon cycle and climate feedbacks but with extensively developed short-lived climate forcer parameterizations. It emphasizes flexibility in handling uncertainties in emissions scenarios and climate sensitivity. Calibration and run-scripts for Flat10MIP are archived here (Sanderson and Sandstad, 2024).

Code and data availability. All code to reproduce plots in this study is permanently available at <https://doi.org/10.5281/zenodo.15267556> (Sanderson et al., 2025). All data to reproduce this study are included at <https://doi.org/10.5281/zenodo.15267556> (Sanderson et al., 2025).

Author contributions. Analysis/plots were performed by BMS, NS and CK. Model simulations were conducted by TI, CDJ, TK, HL, PL, SL, NM, ZM, AR, MS, JS, RS, LS, CS, JT, BMS and TZ. Framing and scoping were performed by BMS, VB, TI, CDK, DML, AM, EOR, IRS and ALSS.

Competing interests. At least one of the (co-)authors is a member of the editorial board of *Geoscientific Model Development*. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. Benjamin M. Sanderson, Chris D. Jones, Roland Séférian, Spencer K. Liddicoat and Zebedee Nicholls acknowledge support from the European Union's Horizon 2020 research and innovation program under grant agreement no. 101003536 (ESM2025). Benjamin M. Sanderson and Marit Sandstad acknowledge the Research Council of Norway under grant agreement 334811 (TRIFECTA). Benjamin M. Sanderson and Norman J. Steinert acknowledge support from the European Union's Horizon 2020 research and innovation program under grant agree-

ment 101003687 (PROVIDE). Chris D. Jones and Spencer K. Liddicoat were supported by the Joint UK BEIS/Defra Met Office Hadley Centre Climate Programme (GA01101). CDK acknowledges support from the Director, Office of Science, Office of Biological and Environmental Research of the US Department of Energy under contract DE-AC02-05CH11231 through the Regional and Global Model Analysis Program (RUBISCO SFA). Abigail L. S. Swann acknowledges support from the National Science Foundation under grant no. AGS-2330096 and the US Department of Energy Regional and Global Model Analysis Program under grant number DE-SC0021209. The work of David M. Lawrence, Peter Lawrence, and Isla R. Simpson is supported by the NSF National Center for Atmospheric Research, which is a major facility sponsored by the NSF under Cooperative Agreement no. 1852977. David Hohn and Nadine Mengis are grateful to be funded under the Emmy Noether scheme by the German Research Foundation (DFG) in the project “FOOTPRINTS – From carbOn remOval To achieving the PaRI's agreemeNt's goal: Temperature Stabilisation” (project number 459765257). Anastasia Romanou acknowledges support from the NASA Modeling Analysis and Prediction (NASA-MAP) program under grant NNX16AC93. Andrew H. MacDougall is supported by the Natural Science and Engineering Research Council of Canada Discovery grant program. Tatiana Ilyina and Hongmei Li acknowledge support from the European Union's Horizon 2020 research and innovation program (4C, grant no. 821003; ESM2025, grant no. 101003536) and the Deutsche Forschungsgemeinschaft (Germany's Excellence Strategy – EXC 2037 “CLICCS – Climate, Climatic Change, and Society” – project no. 390683824). The MPI-ESM1-2-LR simulations used resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project ID bm1124. Roland Séférian acknowledges support from the European Union's Horizon Europe research and innovation program under grant agreement no. 101081193 (OptimESM). Tilo Ziehn receives funding from the Australian Government under the National Environmental Science Program (NESP).

Financial support. This research has been supported by the Norges Forskningsråd (grant no. 334811), the HORIZON EUROPE European Research Council (grant nos. 101003687, 821003, 101003536 and 101081193), the Biological and Environmental Research (grant nos. DE-AC02-05CH11231 and DESC0021209), the National Science Foundation (grant nos. AGS2330096 and 1852977), the Deutsche Forschungsgemeinschaft (grant nos. 459765257 and 390683824), the Earth Sciences Division (grant no. NNX16AC93 G), the Australian National Environmental Science Program 2 (NESP2) Climate Systems Hub Project 5.1 “ACCESS development and delivery to CMIP7”, the Deutsches Klimarechenzentrum (DKRZ) (project ID bm1124), the Natural Sciences and Engineering Research Council of Canada, the U.S. Department of Energy Regional and Global Model Analysis Program (grant number DE-SC0021209), the US National Science Foundation (grant no. AGS-2330096), and the joint UK Department for Business, Energy and Industrial Strategy, UK Government/Defra Met Office Hadley Centre Climate Programme (grant no. GA01101).

Review statement. This paper was edited by Bo Zheng and reviewed by Kirsten Zickfeld and one anonymous referee.

References

- Adcroft, A., Anderson, W., Balaji, V., Blanton, C., Bushuk, M., Dufour, C. O., Dunne, J. P., Griffies, S. M., Hallberg, R., Harrison, M. J., Held, I. M., Jansen, M. F., John, J. G., Krasting, J. P., Langenhorst, A. R., Legg, S., Liang, Z., McHugh, C., Radhakrishnan, A., Reichl, B. G., Rosati, T., Samuels, B. L., Shao, A., Stouffer, R., Winton, M., Wittenberg, A. T., Xiang, B., Zadeh, N., and Zhang, R.: The GFDL global ocean and sea ice model OM4.0: Model description and simulation features, *J. Adv. Model. Earth Sy.*, 11, 3167–3211, <https://doi.org/10.1029/2019ms001726>, 2019.
- Allen, M. R., Frame, D. J., Huntingford, C., Jones, C. D., Lowe, J. A., Meinshausen, M., and Meinshausen, N.: Warming caused by cumulative carbon emissions towards the trillionth tonne, *Nature*, 458, 1163–1166, <https://doi.org/10.1038/nature08019>, 2009.
- Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., Schwinger, J., Bopp, L., Boucher, O., Cadule, P., Chamberlain, M. A., Christian, J. R., Delire, C., Fisher, R. A., Hajima, T., Ilyina, T., Joetzjer, E., Kawamiya, M., Koven, C. D., Krasting, J. P., Law, R. M., Lawrence, D. M., Lenton, A., Lindsay, K., Pongratz, J., Raddatz, T., Séférian, R., Tachiiri, K., Tjiputra, J. F., Wiltshire, A., Wu, T., and Ziehn, T.: Carbon-concentration and carbon-climate feedbacks in CMIP6 models and their comparison to CMIP5 models, *Biogeosciences*, 17, 4173–4222, <https://doi.org/10.5194/bg-17-4173-2020>, 2020.
- Asaadi, A., Schwinger, J., Lee, H., Tjiputra, J., Arora, V., Séférian, R., Liddicoat, S., Hajima, T., Santana-Falcón, Y., and Jones, C. D.: Carbon cycle feedbacks in an idealized simulation and a scenario simulation of negative emissions in CMIP6 Earth system models, *Biogeosciences*, 21, 411–435, <https://doi.org/10.5194/bg-21-411-2024>, 2024.
- Avakumović, V.: Carbon budget concept and its deviation through the pulse response lens, *Earth Syst. Dynam.*, 15, 387–404, <https://doi.org/10.5194/esd-15-387-2024>, 2024.
- Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, *Geosci. Model Dev.*, 4, 677–699, <https://doi.org/10.5194/gmd-4-677-2011>, 2011.
- Bitz, C. M., Holland, M. M., Weaver, A. J., and Eby, M.: Simulating the ice-thickness distribution in a coupled climate model, *J. Geophys. Res.*, 106, 2441–2463, <https://doi.org/10.1029/1999JC000113>, 2001.
- Bogenschütz, P. A., Gettelman, A., Hannay, C., Larson, V. E., Neale, R. B., Craig, C., and Chen, C.-C.: The path to CAM6: coupled simulations with CAM5.4 and CAM5.5, *Geosci. Model Dev.*, 11, 235–255, <https://doi.org/10.5194/gmd-11-235-2018>, 2018.
- Cain, M., Lynch, J., Allen, M. R., Fuglestedt, J. S., Frame, D. J., and Macey, A. H.: Improved calculation of warming-equivalent emissions for short-lived climate pollutants, *NPJ Clim. Atmos. Sci.*, 2, 29, <https://doi.org/10.1038/s41612-019-0086-4>, 2019.
- Chimuka, V. R., Nzotungicimpaye, C.-M., and Zickfeld, K.: Quantifying land carbon cycle feedbacks under negative CO₂ emissions, *Biogeosciences*, 20, 2283–2299, <https://doi.org/10.5194/bg-20-2283-2023>, 2023.

- Cox, P. M.: Description of the TRIFFID dynamic global vegetation model, Hadley Cent, Hadley Cent. Tech. Note, 24, 16 pp., <https://library.metoffice.gov.uk/Portal/Default/en-GB/RecordView/Index/252319> (last access: 26 August 2025), 2001.
- Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A., and Totterdell, I. J.: Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model, *Nature*, 408, 184–187, <https://doi.org/10.1038/35041539>, 2000.
- Damon Matthews, H., Tokarska, K. B., Rogelj, J., Smith, C. J., MacDougall, A. H., Haustein, K., Mengis, N., Sippel, S., Forster, P. M., and Knutti, R.: An integrated approach to quantifying uncertainties in the remaining carbon budget, *Communications Earth & Environment*, 2, 1–11, <https://doi.org/10.1038/s43247-020-00064-9>, 2021.
- Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The community earth system model version 2 (CESM2), *J. Adv. Model. Earth Sy.*, 12, e2019MS001916, <https://doi.org/10.1029/2019ms001916>, 2020.
- Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J.-P., Alias, A., Saint-Martin, D., Séférian, R., Sénéci, S., and Voldoire, A.: Recent changes in the ISBA-CTRIP land surface system for use in the CNRM-CM6 climate model and in global off-line hydrological applications, *J. Adv. Model. Earth Sy.*, 11, 1207–1252, <https://doi.org/10.1029/2018ms001545>, 2019.
- Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P. G., Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng, Y., and Zhao, M.: The GFDL earth system model version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics, *J. Adv. Model. Earth Sy.*, 12, e2019MS002015, <https://doi.org/10.1029/2019ms002015>, 2020.
- Essery, R. L. H., Best, M. J., Betts, R. A., Cox, P. M., and Taylor, C. M.: Explicit representation of subgrid heterogeneity in a GCM land surface scheme, *J. Hydrometeorol.*, 4, 530–543, [https://doi.org/10.1175/1525-7541\(2003\)004<0530:eroshi>2.0.co;2](https://doi.org/10.1175/1525-7541(2003)004<0530:eroshi>2.0.co;2), 2003.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev.*, 9, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>, 2016.
- Fanning, A. F., and Weaver, A. J.: An atmospheric energy-moisture balance model: Climatology, interpentadal climate change, and coupling to an ocean general circulation model, *J. Geophys. Res.*, 101, 15111–15128, <https://doi.org/10.1029/96JD01017>, 1996.
- Fisher, R. A., Wieder, W. R., Sanderson, B. M., Koven, C. D., Oleson, K. W., Xu, C., Fisher, J. B., Shi, M., Walker, A. P., and Lawrence, D. M.: Parametric controls on vegetation responses to biogeochemical forcing in the CLM5, *J. Adv. Model. Earth Sy.*, 11, 2879–2895, <https://doi.org/10.1029/2019ms001609>, 2019.
- Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J., Frame, D., Lunt, D., Mauritsen, T., Palmer, M., Watanabe, M., Wild, M., and Zhang, H.: The earth's energy budget, climate feedbacks and climate sensitivity, in: *Climate Change 2021 – The Physical Science Basis*, Cambridge University Press, 923–1054, <https://doi.org/10.1017/9781009157896.009>, 2023.
- Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Barbero, L., Bates, N. R., Becker, M., Bellouin, N., Decharme, B., Bopp, L., Brasika, I. B. M., Cadule, P., Chamberlain, M. A., Chandra, N., Chau, T.-T.-T., Chevallier, F., Chini, L. P., Cronin, M., Dou, X., Enyo, K., Evans, W., Falk, S., Feely, R. A., Feng, L., Ford, D. J., Gasser, T., Ghattas, J., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Heinke, J., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jacobson, A. R., Jain, A., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Joos, F., Kato, E., Keeling, R. F., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Körtzinger, A., Lan, X., Lefèvre, N., Li, H., Liu, J., Liu, Z., Ma, L., Marland, G., Mayot, N., McGuire, P. C., McKinley, G. A., Meyer, G., Morgan, E. J., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O'Brien, K. M., Olsen, A., Omar, A. M., Ono, T., Paulsen, M., Pierrot, D., Pockock, K., Poulter, B., Powis, C. M., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T. M., Schwinger, J., Séférian, R., Smallman, T. L., Smith, S. M., Sospedra-Alfonso, R., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tans, P. P., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., van Ooijen, E., Wanninkhof, R., Watanabe, M., Wimart-Rousseau, C., Yang, D., Yang, X., Yuan, W., Yue, X., Zaehle, S., Zeng, J., and Zheng, B.: Global Carbon Budget 2023, *Earth Syst. Sci. Data*, 15, 5301–5369, <https://doi.org/10.5194/essd-15-5301-2023>, 2023.
- Gillett, N. P., Arora, V. K., Matthews, D., and Allen, M. R.: Constraining the ratio of global warming to cumulative CO₂ emissions using CMIP5 simulations, *J. Climate*, 26, 6844–6858, <https://doi.org/10.1175/jcli-d-12-00476.1>, 2013.
- Goll, D. S., Brovkin, V., Parida, B. R., Reick, C. H., Kattge, J., Reich, P. B., van Bodegom, P. M., and Niinemets, Ü.: Nutrient limitation reduces land carbon uptake in simulations with a model of combined carbon, nitrogen and phosphorus cycling, *Biogeosciences*, 9, 3547–3569, <https://doi.org/10.5194/bg-9-3547-2012>, 2012.
- Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B., and Wood, R. A.: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, *Clim. Dynam.*, 16, 147–168, <https://doi.org/10.1007/s003820050010>, 2000.
- Gregory, J. M., Andrews, T., and Good, P.: The inconsistency of the transient climate response parameter under

- increasing CO₂, *Philos. T. R. Soc. A*, 373, 20140417, <https://doi.org/10.1098/rsta.2014.0417>, 2015.
- Griffies, S. M.: Elements of the Modular Ocean Model (MOM), https://mom-ocean.github.io/assets/pdfs/MOM5_manual.pdf (last access: 26 August 2025), 2012.
- Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J. H., von Storch, J.-S., Brüggemann, N., Haak, H., and Stössel, A.: Max Planck Institute Earth System Model (MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP), *Geosci. Model Dev.*, 12, 3241–3281, <https://doi.org/10.5194/gmd-12-3241-2019>, 2019.
- Horowitz, L. W., Naik, V., Paulot, F., Ginoux, P. A., Dunne, J. P., Mao, J., Schnell, J., Chen, X., He, J., John, J. G., Lin, M., Lin, P., Malyshev, S., Paynter, D., Shevliakova, E., and Zhao, M.: The GFDL Global Atmospheric Chemistry–Climate Model AM4.1: Model Description and Simulation Characteristics, *J. Adv. Model Earth Sy.*, 12, e2019MS002032, <https://doi.org/10.1029/2019MS002032>, 2020.
- Intergovernmental Panel on Climate Change (IPCC): Climate Change 2021 – The Physical Science Basis – July 2023, 147–286, <https://doi.org/10.1017/9781009157896.003>, 2023a.
- Intergovernmental Panel on Climate Change (IPCC): Global Carbon and Other Biogeochemical Cycles and Feedbacks, in: Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 673–816, <https://doi.org/10.1017/9781009157896.007>, 2023b.
- Intergovernmental Panel on Climate Change (IPCC): Mitigation Pathways Compatible with Long-term Goals, in: Climate Change 2022 – Mitigation of Climate Change: Working Group III Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 295–408, <https://doi.org/10.1017/9781009157926.005>, 2023c.
- Intergovernmental Panel on Climate Change (IPCC): Technical Summary, in: Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 35–144, <https://doi.org/10.1017/9781009157896.002>, 2023d.
- Ito, G., Romanou, A., Kiang, N. Y., Faluvegi, G., Aleinov, I., Ruedy, R., Russell, R., Lerner, P., Kelley, M., and Lo, K.: Global carbon cycle and climate feedbacks in the NASA GISS ModelE2.1, *J. Adv. Model. Earth Sy.*, 12, e2019MS002030, <https://doi.org/10.1029/2019MS002030>, 2020.
- Jenkins, S., Millar, R. J., Leach, N., and Allen, M. R.: Framing climate goals in terms of cumulative CO₂-forcing-equivalent emissions, *Geophys. Res. Lett.*, 45, 2795–2804, <https://doi.org/10.1002/2017gl076173>, 2018.
- Jenkins, S., Cain, M., Friedlingstein, P., Gillett, N., Walsh, T., and Allen, M. R.: Quantifying non-CO₂ contributions to remaining carbon budgets, *npj Climate and Atmospheric Science*, 4, 1–10, <https://doi.org/10.1038/s41612-021-00203-9>, 2021.
- Jenkins, S., Sanderson, B., Peters, G., Frölicher, T. L., Friedlingstein, P., and Allen, M.: The multi-decadal response to net zero CO₂ emissions and implications for emissions policy, *Geophys. Res. Lett.*, 49, e2022GL101047, <https://doi.org/10.1029/2022gl101047>, 2022.
- Jones, C., Robertson, E., Arora, V., Friedlingstein, P., Shevliakova, E., Bopp, L., Brovkin, V., Hajima, T., Kato, E., Kawamiya, M., Liddicoat, S., Lindsay, K., Reick, C. H., Roelandt, C., Segschneider, J., and Tjiputra, J.: Twenty-First-Century Compatible CO₂ Emissions and Airborne Fraction Simulated by CMIP5 Earth System Models under Four Representative Concentration Pathways, *J. Climate*, 26, 4398–4413, <https://doi.org/10.1175/JCLI-D-12-00554.1>, 2013.
- Jones, C. D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., Graven, H., Hoffman, F., Ilyina, T., John, J. G., Jung, M., Kawamiya, M., Koven, C., Pongratz, J., Raddatz, T., Randerson, J. T., and Zaehle, S.: C4MIP – The Coupled Climate–Carbon Cycle Model Intercomparison Project: experimental protocol for CMIP6, *Geosci. Model Dev.*, 9, 2853–2880, <https://doi.org/10.5194/gmd-9-2853-2016>, 2016.
- Jones, C. D., Frölicher, T. L., Koven, C., MacDougall, A. H., Matthews, H. D., Zickfeld, K., Rogelj, J., Tokarska, K. B., Gillett, N. P., Ilyina, T., Meinshausen, M., Mengis, N., Séférian, R., Eby, M., and Burger, F. A.: The Zero Emissions Commitment Model Intercomparison Project (ZECMIP) contribution to C4MIP: quantifying committed climate changes following zero carbon emissions, *Geosci. Model Dev.*, 12, 4375–4385, <https://doi.org/10.5194/gmd-12-4375-2019>, 2019.
- Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and Storch, J. S.: Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model: Mpiom CMIP5 Ocean Simulations, *J. Adv. Model. Earth Sy.*, 5, 422–446, <https://doi.org/10.1002/jame.20023>, 2013.
- Keller, D. P., Oschlies, A., and Eby, M.: A new marine ecosystem model for the University of Victoria Earth System Climate Model, *Geosci. Model Dev.*, 5, 1195–1220, <https://doi.org/10.5194/gmd-5-1195-2012>, 2012.
- Kelley, M., Schmidt, G. A., Nazarenko, L. S., Bauer, S. E., Ruedy, R., Russell, G. L., Ackerman, A. S., Aleinov, I., Bauer, M., Bleck, R., Canuto, V., Cesana, G., Cheng, Y., Clune, T. L., Cook, B. I., Cruz, C. A., Del Genio, A. D., Elsaesser, G. S., Faluvegi, G., Kiang, N. Y., Kim, D., Lacis, A. A., Leboissetier, A., LeGrande, A. N., Lo, K. K., Marshall, J. L., Matthews, E. E., McDermid, S., Mezuman, K., Miller, R. L., Murray, L. T., Oinas, V., Orbe, C., Garcia-Pando, C. P., Perlwitz, J. P., Puma, M. J., Rind, D., Romanou, A., Shindell, D. T., Sun, S., Tausnev, N., Tsigaridis, K., Tselioudis, G., Weng, E., Wu, J., and Yao, M.-S.: GISS-E2.1: Configurations and Climatology, *J. Adv. Model Earth Syst.*, 12, e2019MS002025, <https://doi.org/10.1029/2019MS002025>, 2020.
- Kim, Y., Moorcroft, P. R., Aleinov, I., Puma, M. J., and Kiang, N. Y.: Variability of phenology and fluxes of water and carbon with observed and simulated soil moisture in the Ent Terrestrial Biosphere Model (Ent TBM version 1.0.1.0.0), *Geosci. Model Dev.*, 8, 3837–3865, <https://doi.org/10.5194/gmd-8-3837-2015>, 2015.
- Koven, C. D., Arora, V. K., Cadule, P., Fisher, R. A., Jones, C. D., Lawrence, D. M., Lewis, J., Lindsay, K., Mathesius, S., Meinshausen, M., Mills, M., Nicholls, Z., Sanderson, B. M., Séférian, R., Swart, N. C., Wieder, W. R., and Zickfeld, K.: Multi-century dynamics of the climate and carbon cycle under both high and net negative emissions scenarios, *Earth Syst. Dynam.*, 13, 885–909, <https://doi.org/10.5194/esd-13-885-2022>, 2022.
- Koven, C. D., Sanderson, B. M., and Swann, A. L. S.: Much of zero emissions commitment occurs before reach-

- ing net zero emissions, *Environ. Res. Lett.*, 18, 014017, <https://doi.org/10.1088/1748-9326/acab1a>, 2023.
- Kowalczyk, E. A., Stevens, L., Law, R. M., Dix, M. R., Wang, Y., Harman, I. N., Haynes, K. D., Sribnovsky, J., Pak, B., and Ziehn, T.: The land surface model component of ACCESS: description and impact on the simulated surface climatology, *Aust. Meteorol. Ocean.*, 63, 65–82, 2013.
- Krasting, J. P., Dunne, J. P., Shevliakova, E., and Stouffer, R. J.: Trajectory sensitivity of the transient climate response to cumulative carbon emissions, *Geophys. Res. Lett.*, 41, 2520–2527, <https://doi.org/10.1002/2013gl059141>, 2014.
- Lamboll, R. D., Nicholls, Z. R. J., Smith, C. J., Kikstra, J. S., Byers, E., and Rogelj, J.: Assessing the size and uncertainty of remaining carbon budgets, *Nat. Clim. Change*, 13, 1360–1367, <https://doi.org/10.1038/s41558-023-01848-5>, 2023.
- Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., van den Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., and Zeng, X.: The community land model version 5: Description of new features, benchmarking, and impact of forcing uncertainty, *J. Adv. Model. Earth Sy.*, 11, 4245–4287, <https://doi.org/10.1029/2018ms001583>, 2019.
- Leach, N. J., Millar, R. J., Haustein, K., Jenkins, S., Graham, E., and Allen, M. R.: Current level and rate of warming determine emissions budgets under ambitious mitigation, *Nat. Geosci.*, 11, 574–579, <https://doi.org/10.1038/s41561-018-0156-y>, 2018.
- Lerner, P., Romanou, A., Kelley, M., Romanski, J., Ruedy, R., and Russell, G.: Drivers of air-sea CO₂ flux seasonality and its long-term changes in the NASA-GISS model CMIP6 submission, *J. Adv. Model. Earth Sy.*, 13, e2019MS002028, <https://doi.org/10.1029/2019MS002028>, 2021.
- Liddicoat, S. K., Wiltshire, A. J., Jones, C. D., Arora, V. K., Brovkin, V., Cadule, P., Hajima, T., Lawrence, D. M., Pongratz, J., Schwinger, J., Séférian, R., Tjiputra, J. F., and Ziehn, T.: Compatible Fossil Fuel CO₂ Emissions in the CMIP6 Earth System Models' Historical and Shared Socioeconomic Pathway Experiments of the Twenty-First Century, *J. Climate*, 34, 2853–2875, <https://doi.org/10.1175/JCLI-D-19-0991.1>, 2021.
- MacDougall, A. H.: The Transient Response to Cumulative CO₂ Emissions: a Review, *Current Climate Change Reports*, 2, 39–47, <https://doi.org/10.1007/s40641-015-0030-6>, 2015.
- MacDougall, A. H. and Knutti, R.: Projecting the release of carbon from permafrost soils using a perturbed parameter ensemble modelling approach, *Biogeosciences*, 13, 2123–2136, <https://doi.org/10.5194/bg-13-2123-2016>, 2016.
- MacDougall, A. H.: Limitations of the 1% experiment as the benchmark idealized experiment for carbon cycle inter-comparison in C⁴MIP, *Geosci. Model Dev.*, 12, 597–611, <https://doi.org/10.5194/gmd-12-597-2019>, 2019.
- MacDougall, A. H., Frölicher, T. L., Jones, C. D., Rogelj, J., Matthews, H. D., Zickfeld, K., Arora, V. K., Barrett, N. J., Brovkin, V., Burger, F. A., Eby, M., Eliseev, A. V., Hajima, T., Holden, P. B., Jeltsch-Thömmes, A., Koven, C., Mengis, N., Menviel, L., Michou, M., Mokhov, I. I., Oka, A., Schwinger, J., Séférian, R., Shaffer, G., Sokolov, A., Tachiiri, K., Tjiputra, J., Wiltshire, A., and Ziehn, T.: Is there warming in the pipeline? A multi-model analysis of the Zero Emissions Commitment from CO₂, *Biogeosciences*, 17, 2987–3016, <https://doi.org/10.5194/bg-17-2987-2020>, 2020.
- Madec, G., Bourdallé-Badie, R., Bouttier, P.-A., et al.: NEMO ocean engine, Zenodo, <https://doi.org/10.5281/zenodo.1464816>, 2017.
- Matthews, H. D., Gillett, N. P., Stott, P. A., and Zickfeld, K.: The proportionality of global warming to cumulative carbon emissions, *Nature*, 459, 829–832, <https://doi.org/10.1038/nature08047>, 2009.
- Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenez-de-la-Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjaeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO₂, *J. Adv. Model. Earth Sy.*, 11, 998–1038, <https://doi.org/10.1029/2018ms001400>, 2019.
- Meinshausen, M., Raper, S. C. B., and Wigley, T. M. L.: Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6 – Part 1: Model description and calibration, *Atmos. Chem. Phys.*, 11, 1417–1456, <https://doi.org/10.5194/acp-11-1417-2011>, 2011.
- Meissner, K. J., Weaver, A. J., Matthews, H. D., and Cox, P. M.: The role of land surface dynamics in glacial inception: a study with the UVic Earth System Model, *Clim. Dynam.*, 21, 515–537, <https://doi.org/10.1007/s00382-003-0352-2>, 2003.
- Mengis, N. and Matthews, H. D.: Non-CO₂ forcing changes will likely decrease the remaining carbon budget for 1.5 °C, *npj Climate and Atmospheric Science*, 3, 1–7, <https://doi.org/10.1038/s41612-020-0123-3>, 2020.
- Mengis, N., Keller, D. P., MacDougall, A. H., Eby, M., Wright, N., Meissner, K. J., Oschlies, A., Schmittner, A., MacIsaac, A. J., Matthews, H. D., and Zickfeld, K.: Evaluation of the University of Victoria Earth System Climate Model version 2.10 (UVic ESCM 2.10), *Geosci. Model Dev.*, 13, 4183–4204, <https://doi.org/10.5194/gmd-13-4183-2020>, 2020.
- Millar, R. J. and Friedlingstein, P.: The utility of the historical record for assessing the transient climate response to cumulative emissions, *Philos. T. R. Soc. A*, 376, 20160449, <https://doi.org/10.1098/rsta.2016.0449>, 2018.
- MPI: MPI-ESM 1.2.01p7, <https://doi.org/10.17617/3.H44EN5>, 2024.

- Nicholls, Z. R. J., Gieseke, R., Lewis, J., Nauels, A., and Meinshausen, M.: Implications of non-linearities between cumulative CO₂ emissions and CO₂-induced warming for assessing the remaining carbon budget, *Environ. Res. Lett.*, 15, 074017, <https://doi.org/10.1088/1748-9326/ab83af>, 2020.
- Pacanowski, Ronald C., and Gnanadesikan, A.: Transient Response in a Z-Level Ocean Model That Resolves Topography with Partial Cells, *Mon. Weather Rev.*, 126, 3248–3270, [https://doi.org/10.1175/1520-0493\(1998\)126<3248:TRIAZL>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<3248:TRIAZL>2.0.CO;2), 1998.
- Palazzo Corner, S., Siegert, M., Ceppi, P., Fox-Kemper, B., Frölicher, T. L., Gallego-Sala, A., Haigh, J., Hegerl, G. C., Jones, C. D., Knutti, R., Koven, C. D., MacDougall, A. H., Meinshausen, M., Nicholls, Z., Sallée, J. B., Sanderson, B. M., Séférian, R., Turetsky, M., Williams, R. G., Zaehle, S., and Rogelj, J.: The Zero Emissions Commitment and climate stabilization, *Front. Sci. Ser.*, 1, 1170744, <https://doi.org/10.3389/fsci.2023.1170744>, 2023.
- Palmer, J. R. and Totterdell, I. J.: Production and export in a global ocean ecosystem model, *Deep-Sea Res. Pt. I*, 48, 1169–1198, [https://doi.org/10.1016/S0967-0637\(00\)00080-7](https://doi.org/10.1016/S0967-0637(00)00080-7), 2001.
- Pope, V. D., Gallani, M. L., Rowntree, P. R., and Stratton, R. A.: The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3, *Clim. Dynam.*, 16, 123–146, <https://doi.org/10.1007/s003820050009>, 2000.
- Reick, C. H., Gayler, V., Goll, D., Hagemann, S., and Heidkamp, M.: JSBACH 3-The land component of the MPI Earth System Model: documentation of version 3.2, *Berichte zur Erdsystemforschung*, 240, <https://doi.org/10.17617/2.3279802>, 2021.
- Roehrig, R., Beau, I., Saint-Martin, D., Alias, A., Decharme, B., Guérémy, J.-F., Voldoire, A., Abdel-Lathif, A. Y., Bazile, E., Belamari, S., Blein, S., Bouniol, D., Bouteloup, Y., Cattiaux, J., Chauvin, F., Chevallier, M., Colin, J., Douville, H., Marquet, P., Michou, M., Nabat, P., Oudar, T., Peyrillé, P., Piriou, J.-M., Salas y Mélia, D., Séférian, R., and Sénési, S.: The CNRM global atmosphere model ARPEGE-climat 6.3: Description and evaluation, *J. Adv. Model. Earth Sy.*, 12, e2020MS002075, <https://doi.org/10.1029/2020ms002075>, 2020.
- Rogelj, J., Huppmann, D., Krey, V., Riahi, K., Clarke, L., Gidden, M., Nicholls, Z., and Meinshausen, M.: A new scenario logic for the Paris Agreement long-term temperature goal, *Nature*, 573, 357–363, <https://doi.org/10.1038/s41586-019-1541-4>, 2019a.
- Rogelj, J., Forster, P. M., Kriegler, E., Smith, C. J., and Séférian, R.: Estimating and tracking the remaining carbon budget for stringent climate targets, *Nature*, 571, 335–342, <https://doi.org/10.1038/s41586-019-1368-z>, 2019b.
- Romanou, A., Gregg, W. W., Romanski, J., Kelley, M., Bleck, R., Healy, R., Nazarenko, L., Russell, G., Schmidt, G. A., Sun, S., and Tausnev, N.: Natural air–sea flux of CO₂ in simulations of the NASA-GISS climate model: Sensitivity to the physical ocean model formulation, *Ocean Model.*, 66, 26–44, 2013.
- Sanderson, B. and Sandstad, M.: *ciceroOslo/cicerooscm: v1.1.3-flat10*, Zenodo [code], <https://doi.org/10.5281/zenodo.13939554>, 2024.
- Sanderson, B., Koven, C., Sandstad, M., and Nicholls, Z.: *benmsanderson/esm-deck: GMD revised version*, Zenodo [code and data set], <https://doi.org/10.5281/zenodo.15267556>, 2025.
- Sanderson, B. M., Booth, B. B. B., Dunne, J., Eyring, V., Fisher, R. A., Friedlingstein, P., Gidden, M. J., Hajima, T., Jones, C. D., Jones, C. G., King, A., Koven, C. D., Lawrence, D. M., Lowe, J., Mengis, N., Peters, G. P., Rogelj, J., Smith, C., Snyder, A. C., Simpson, I. R., Swann, A. L. S., Tebaldi, C., Ilyina, T., Schleussner, C.-F., Séférian, R., Samset, B. H., van Vuuren, D., and Zaehle, S.: The need for carbon-emissions-driven climate projections in CMIP7, *Geosci. Model Dev.*, 17, 8141–8172, <https://doi.org/10.5194/gmd-17-8141-2024>, 2024.
- Sandstad, M., Aamaas, B., Johansen, A. N., Lund, M. T., Peters, G. P., Samset, B. H., Sanderson, B. M., and Skeie, R. B.: CICERO Simple Climate Model (CICERO-SCM v1.1.1) – an improved simple climate model with a parameter calibration tool, *Geosci. Model Dev.*, 17, 6589–6625, <https://doi.org/10.5194/gmd-17-6589-2024>, 2024.
- Schleussner, C.-F., Ganti, G., Lejeune, Q., Zhu, B., Pfeleiderer, P., Prütz, R., Ciais, P., Frölicher, T. L., Fuss, S., Gasser, T., Gidden, M. J., Kropf, C. M., Lacroix, F., Lamboll, R., Martyr, R., Maussion, F., McCaughey, J. W., Meinshausen, M., Mengel, M., Nicholls, Z., Quilcaille, Y., Sanderson, B., Seneviratne, S. I., Sillmann, J., Smith, C. J., Steinert, N. J., Theokritoff, E., Warren, R., Price, J., and Rogelj, J.: Overconfidence in climate overshoot, *Nature*, 634, 366–373, <https://doi.org/10.1038/s41586-024-08020-9>, 2024.
- Schmidt, G. A., Kelley, M., Nazarenko, L., Ruedy, R., Russell, G. L., Aleinov, I., Bauer, M., Bauer, S. E., Bhat, M. K., Bleck, R., Canuto, V., Chen, Y.-H., Cheng, Y., Clune, T. L., Del Genio, A., de Fainchtein, R., Faluvegi, G., Hansen, J. E., Healy, R. J., Kiang, N. Y., Koch, D., Lacis, A. A., LeGrande, A. N., Lerner, J., Lo, K. K., Matthews, E. E., Menon, S., Miller, R. L., Oinas, V., Olosio, A. O., Perlwitz, J. P., Puma, M. J., Putman, W. M., Rind, D., Romanou, A., Sato, M., Shindell, D. T., Sun, S., Syed, R. A., Tausnev, N., Tsigaridis, K., Unger, N., Yao, M.-S., and Zhang, J.: Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive, *J. Adv. Model. Earth Sy.*, 6, 141–184, <https://doi.org/10.1002/2013MS000265>, 2014.
- Schwinger, J. and Tjiputra, J.: Ocean carbon cycle feedbacks under negative emissions, *Geophys. Res. Lett.*, 45, 5062–5070, <https://doi.org/10.1029/2018gl077790>, 2018.
- Schwinger, J., Asaadi, A., Goris, N., and Lee, H.: Possibility for strong northern hemisphere high-latitude cooling under negative emissions, *Nat. Commun.*, 13, 1095, <https://doi.org/10.1038/s41467-022-28573-5>, 2022.
- Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., Decharme, B., Delire, C., Berthet, S., Chevallier, M., Sénési, S., Franchisteguy, L., Vial, J., Mallet, M., Joetzjer, E., Geoffroy, O., Guérémy, J.-F., Moine, M.-P., Msadek, R., Ribes, A., Rocher, M., Roehrig, R., Salas-y-Mélia, D., Sanchez, E., Terray, L., Valcke, S., Waldman, R., Aumont, O., Bopp, L., Deshayes, J., Éthé, C., and Madec, G.: Evaluation of CNRM earth system model, CNRM-ESM2-1: Role of earth system processes in present-day and future climate, *J. Adv. Model. Earth Sy.*, 11, 4182–4227, <https://doi.org/10.1029/2019ms001791>, 2019.
- Séférian, R., Bossy, T., Gasser, T., Nichols, Z., Dorheim, K., Su, X., Tsutsui, J., and Santana-Falcón, Y.: Physical inconsistencies in the representation of the ocean heat-carbon nexus in simple climate models, *Commun. Earth Environ.*, 5, 1–10, <https://doi.org/10.1038/s43247-024-01464-x>, 2024.
- Seland, Ø., Bentsen, M., Olivieri, D., Toniazio, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y.,

- Griesfeller, J., Grini, A., Guo, C., Ilıcak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, *Geosci. Model Dev.*, 13, 6165–6200, <https://doi.org/10.5194/gmd-13-6165-2020>, 2020.
- Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahayan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and evaluation of the U.K. earth system model, *J. Adv. Model. Earth Sy.*, 11, 4513–4558, <https://doi.org/10.1029/2019ms001739>, 2019.
- Shevliakova, E., Malyshev, S., Martinez-Cano, I., Milly, P. C. D., Pacala, S. W., Ginoux, P., Dunne, K. A., Dunne, J. P., Dupuis, C., Findell, K. L., Ghannam, K., Horowitz, L. W., Knutson, T. R., Krasting, J. P., Naik, V., Philipps, P., Zadeh, N., Yu, Y., Zeng, F., and Zeng, Y.: The land component LM4.1 of the GFDL Earth System Model ESM4.1: Model description and characteristics of land surface climate and carbon cycling in the historical simulation, *J. Adv. Model. Earth Sy.*, 16, e2023MS003922, <https://doi.org/10.1029/2023ms003922>, 2024.
- Smith, C. J., Forster, P. M., Allen, M., Leach, N., Mil- lar, R. J., Passerello, G. A., and Regayre, L. A.: FAIR v1.3: a simple emissions-based impulse response and carbon cycle model, *Geosci. Model Dev.*, 11, 2273–2297, <https://doi.org/10.5194/gmd-11-2273-2018>, 2018.
- Smith, R., Jones, P., Briegleb, B., Bryan, F., Danabasoglu, G., Dennis, J., Dukowicz, J., Eden, C., Fox-Kemper, B., Gent, P., Hecht, M., Jayne, S., Jochum, M., Large, W., Lindsay, K., Maltrud, M., Norton, N., Peacock, S., Vertenstein, M., and Yeager, S.: The Parallel Ocean Program (POP) reference manual, Ocean component of the Community Climate System Model (CCSM), LANL Tech, <https://n2t.org/ark:/85065/d70g3j4h> (last access: 26 August 2025), 2010.
- Tjiputra, J. F., Roelandt, C., Bentsen, M., Lawrence, D. M., Lorentzen, T., Schwinger, J., Seland, Ø., and Heinze, C.: Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM), *Geosci. Model Dev.*, 6, 301–325, <https://doi.org/10.5194/gmd-6-301-2013>, 2013.
- Tjiputra, J. F., Schwinger, J., Bentsen, M., Morée, A. L., Gao, S., Bethke, I., Heinze, C., Goris, N., Gupta, A., He, Y.-C., Olivié, D., Seland, Ø., and Schulz, M.: Ocean biogeochemistry in the Norwegian Earth System Model version 2 (NorESM2), *Geosci. Model Dev.*, 13, 2393–2431, <https://doi.org/10.5194/gmd-13-2393-2020>, 2020.
- Tjiputra, J. F., Couespel, D., and Sanders, R.: Marine ecosystem role in setting up preindustrial and future climate, *Nat. Commun.*, 16, 2206, <https://doi.org/10.1038/s41467-025-57371-y>, 2025.
- Valdes, P. J., Armstrong, E., Badger, M. P. S., Bradshaw, C. D., Bragg, F., Crucifix, M., Davies-Barnard, T., Day, J. J., Farnsworth, A., Gordon, C., Hopcroft, P. O., Kennedy, A. T., Lord, N. S., Lunt, D. J., Marzocchi, A., Parry, L. M., Pope, V., Roberts, W. H. G., Stone, E. J., Tourte, G. J. L., and Williams, J. H. T.: The BRIDGE HadCM3 family of climate models: HadCM3@Bristol v1.0, *Geosci. Model Dev.*, 10, 3715–3743, <https://doi.org/10.5194/gmd-10-3715-2017>, 2017.
- Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Ged- ney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whittall, M., Williams, K., and Zerroukat, M.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, *Geosci. Model Dev.*, 12, 1909–1963, <https://doi.org/10.5194/gmd-12-1909-2019>, 2019.
- Wieder, W. R., Bonan, G. B., and Allison, S. D.: Global soil carbon projections are improved by modelling microbial processes, *Nat. Clim. Change*, 3, 909–912, <https://doi.org/10.1038/nclimate1951>, 2013.
- Winkelmann, R., Dennis, D. P., Donges, J. F., Loriani, S., Klose, A. K., Abrams, J. F., Alvarez-Solas, J., Albrecht, T., Armstrong McKay, D., Bathiany, S., Blasco Navarro, J., Brovkin, V., Burke, E., Danabasoglu, G., Donner, R. V., Driike, M., Georgievski, G., Goelzer, H., Harper, A. B., Hegerl, G., Hirota, M., Hu, A., Jackson, L. C., Jones, C., Kim, H., Koenigk, T., Lawrence, P., Lenton, T. M., Liddy, H., Licón-Saláiz, J., Menthon, M., Montoya, M., Nitzbon, J., Nowicki, S., Otto-Bliesner, B., Pausata, F., Rahmstorf, S., Ramin, K., Robinson, A., Rockström, J., Romanou, A., Sakschewski, B., Schädel, C., Sherwood, S., Smith, R. S., Steinert, N. J., Swingedouw, D., Willeit, M., Weijer, W., Wood, R., Wyser, K., and Yang, S.: The Tipping Points Modelling Intercomparison Project (TIPMIP): Assessing tipping point risks in the Earth system, *EGU sphere* [preprint], <https://doi.org/10.5194/egusphere-2025-1899>, 2025.
- Winkler, A. J., Myneni, R., Reimers, C., Reichstein, M., and Brovkin, V.: Carbon system state determines warming potential of emissions, *PLoS One*, 19, e0306128, <https://doi.org/10.1371/journal.pone.0306128>, 2024.
- Zickfeld, K., Eby, M., Matthews, H. D., and Weaver, A. J.: Setting cumulative emissions targets to reduce the risk of dangerous climate change, *P. Natl. Acad. Sci. USA*, 106, 16129–16134, <https://doi.org/10.1073/pnas.0805800106>, 2009.
- Zickfeld, K., MacDougall, A. H., and Damon Matthews, H.: On the proportionality between global temperature change and cumulative CO₂ emissions during periods of net negative CO₂ emissions, *Environ. Res. Lett.*, 11, 055006, <https://doi.org/10.1088/1748-9326/11/5/055006>, 2016.
- Ziehn, T., Chamberlain, M. A., Law, R. M., Lenton, A., Bodman, R. W., Dix, M., Stevens, L., Wang, Y.-P., and Sribnovsky, J.: The Australian Earth System Model: ACCESS-ESM1.5, *JSHES*, 70, 193–214, <https://doi.org/10.1071/ES19035>, 2020.