



Observation-based constraint suggests larger ocean carbon uptake and stronger climate feedbacks

Jens Terhaar^{1,2}, Laurent Bopp³, Pierre Friedlingstein^{3,4}, Thomas Frölicher^{1,2}, Helene Hollitzer^{1,2}, Tatiana Ilyina^{5,6,7}, Lester Kwiatkowski⁸, Paul Lerner^{9,10}, Hongmei Li^{5,7}, Anastasia Romanou^{9,10}, Jörg Schwinger¹¹, Roland Séférian¹², and Aaron Wienkers^{1,2}

¹Climate and Environmental Physics, Physics Institute, University of Bern, Bern, Switzerland

²Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland

³Laboratoire de Météorologie Dynamique, Institut Pierre-Simon Laplace, CNRS, École Normale Supérieure, Université PSL, Sorbonne Université, École Polytechnique, Paris, France

⁴Faculty of Environment, Science and Economy, University of Exeter, Exeter, EX4 4QF, UK

⁵Max Planck Institute for Marine Microbiology, Bremen, Germany

⁶University of Hamburg, Hamburg, Germany

⁷Helmholtz Centre Hereon, Geesthacht, Germany

⁸LOCEAN, Sorbonne Université-CNRS-IRD-MNH, Paris, 75005, France

⁹Applied Physics and Applied Mathematics, Columbia University, 1120 Amsterdam Ave, New York, 10027, NY, USA

¹⁰NASA Goddard Institute for Space Studies, New York, USA

¹¹NORCE Climate, Bjerknes Centre for Climate Research, Bergen, 5005, Norway

¹²CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

Correspondence: Jens Terhaar (jens.terhaar@unibe.ch)

Abstract.

By absorbing large amounts of anthropogenic carbon from the atmosphere, the ocean strongly buffers the magnitude of warming caused by CO₂ emissions. This ocean carbon uptake is often conceptually separated into two components, quantified by the feedback parameters β and γ : The uptake of anthropogenic CO₂ driven by rising atmospheric concentration (β), and the effect of climate warming (γ), which reduces this anthropogenic uptake and drives a loss of natural carbon already stored in the ocean. Together, β and γ determine how effectively the ocean absorbs CO₂ emissions and shape how much emitted carbon remains in the atmosphere. Thus, β and γ directly affect how much the climate system warms in response to emitted carbon, i.e., the transient climate response to cumulative CO₂ emissions (TCRE). However, estimates of ocean β and γ from Earth system models differ widely, leading to substantial uncertainty in the TCRE.

Here, we extend and apply a previously identified emergent constraint linking ocean anthropogenic carbon uptake to observable indicators of the formation of Southern Ocean mode and intermediate waters, North Atlantic deep water formation, and the global surface ocean chemical uptake capacity. In addition to these three predictors, we add an additional indicator for bottom water formation via open-ocean deep convection, defined as the volume of waters below 3000 m that is ventilated via extremely deep mixed layer depths. Using observations of these four indicators, the constrained β is 0.96 ± 0.05 Pg C ppm⁻¹ at doubling of atmospheric CO₂ and 0.85 ± 0.04 Pg C ppm⁻¹ at quadrupling of atmospheric CO₂. The observationally constrained β is around 4-5% larger and around 35-47% less uncertain than previously estimated. This larger ocean β is further



supported by simulated historical anthropogenic carbon uptake in most models being lower than observation-based estimates, although uncertainties in the latter remain too large to directly constrain β .

Beyond constraining ocean β , we show that ocean β and γ become increasingly correlated over time. This relationship emerges because models with higher anthropogenic carbon uptake, and hence larger β , also show a stronger suppression of that uptake in response to climate change. The stronger suppression is driven by a larger reduction in the indicators for ocean carbon uptake identified above in models with higher indicators at the beginning of the simulations, e.g., a model with stronger ventilation at present will also see a stronger reduction of that ventilation in the future. As atmospheric CO_2 increases, the climate-driven reduction in anthropogenic carbon uptake increasingly dominates over the climate-driven loss of natural carbon, causing γ to become primarily controlled by reductions in anthropogenic carbon uptake over time. As a result, the same observable indicators that constrain ocean β also constrain ocean γ at high CO_2 . At doubling of atmospheric CO_2 , this yields a constrained ocean γ of $-10.1 \pm 2.4 \text{ Pg C } ^\circ\text{C}^{-1}$, which is 23% larger in magnitude but has the same uncertainty as the unconstrained multi-model mean. Reflecting the strengthening of this relationship with increasing atmospheric CO_2 , the constraint tightens further at quadrupling of CO_2 to $-24.5 \pm 3.6 \text{ Pg C } ^\circ\text{C}^{-1}$, which is 28% larger and 42% less uncertain.

Overall, we have extended a previous constraint for the ocean carbon uptake by (1) adding an additional parameter that removed a known bias from unrealistic open-ocean deep convection, (2) transferring the constraint to γ and providing a mechanistical explanation for that constraint and evidence for that mechanism, (3) applying the constraint regionally, (4) identifying substantial biases in β and γ and reducing uncertainties, and (5) improving process understanding of drivers of changes in the ocean carbon sink. Although both the increase in atmospheric CO_2 uptake and the increase in carbon loss with climate change were biased across the model ensembles and could be constrained, the bias corrections of β and γ offset each other, leaving the best estimate of the ocean carbon sink relatively unchanged (2–3% increase compared to this ensemble mean). However, the uncertainties of β and γ are reduced and hence is the overall uncertainty of the carbon sink, which will propagate into smaller uncertainties of important climate metrics such as the TCRC.

1 Introduction

The ocean currently absorbs roughly 25–30% of anthropogenic CO_2 emissions (Friedlingstein et al., 2026b, a; DeVries et al., 2023; Terhaar et al., 2022; Terhaar, 2025), substantially moderating the rate of atmospheric CO_2 increase and global warming. The strength of the ocean carbon sink is determined by the amount of carbon taken up at the ocean surface (Broecker et al., 1979) and by the amount of carbon transported from the ocean surface to the ocean interior (Caldeira and Duffy, 2000; Orr et al., 2001; Sarmiento et al., 1992; Mikaloff Fletcher et al., 2006). Despite the importance of the ocean carbon sink, its current strength and future evolution remain uncertain (Friedlingstein et al., 2026b; Terhaar et al., 2022), contributing to uncertainty in projections of warming for a given amount of emitted carbon (Sanderson et al., 2025; Canadell et al., 2021; Matthews et al., 2021). Understanding how the ocean carbon sink will respond to rising atmospheric CO_2 and associated climate change is therefore essential for constraining the future climate response to carbon emissions.



The sensitivity of the ocean carbon sink to rising atmospheric CO₂ and to climate change is commonly quantified using the
50 feedback parameters β and γ (Friedlingstein et al., 2006; Arora et al., 2013, 2020). The parameter β describes the carbon-
concentration feedback, i.e., the increase in anthropogenic carbon uptake driven by rising atmospheric CO₂ (Sarmiento et al.,
1992; Orr et al., 2001), while γ quantifies the carbon-climate feedback, i.e., the reduction of anthropogenic carbon uptake
and loss of natural carbon associated with climate warming, driven by changes in solubility, circulation, stratification, and
biological processes (Sarmiento and Le Quéré, 1996; Plattner et al., 2001; McNeil and Matear, 2013; Joos et al., 1999; Bunsen
55 et al., 2024; Terhaar, 2024). The carbon-concentration feedback (β) dominates ocean carbon uptake, while the carbon-climate
feedback (γ) represents an important yet smaller modulation of ocean carbon fluxes. Together, these parameters determine how
effectively the ocean absorbs emitted anthropogenic CO₂ and thereby influence the transient climate response to cumulative
carbon emissions (TCRE) (Allen et al., 2009; Matthews et al., 2009; Zickfeld et al., 2009). TCRE, β , and γ are typically
quantified from idealized simulations with atmospheric CO₂ increasing by 1% per year (Arora et al., 2013, 2020). Despite
60 their central role, simulated estimates of ocean β and γ differ widely across Earth system models (Arora et al., 2013, 2020;
Schwinger et al., 2014) contributing to the multi-model spread of TCRE (Allen et al., 2009; Matthews et al., 2009; Zickfeld
et al., 2009; Sanderson et al., 2025; Canadell et al., 2021).

Large model spread in a given variable may be reduced by emergent constraints, which offer a pathway to reduce uncer-
tainties of simulated difficult-to-observe or unobservable quantities by identifying mechanistic relationships between these
65 quantities and observable present-day properties (Boé et al., 2009; Eyring et al., 2019; Hall et al., 2019). To exploit these re-
lationships with observations to constrain the simulated difficult-to-observe or unobservable quantities, the relationships must
be built on robust mechanistic understanding that is verifiable and the relationship must hold across different model ensembles
(Eyring et al., 2019; Hall et al., 2019). These criteria are important as relationships between different variables in climate
models can occur by chance (Brient, 2020; Caldwell et al., 2014), and it has been shown that some initially proposed emergent
70 constraints did not persist under extended model ensembles (Schlund et al., 2020). In the past, emergent constraints have, for
example, been used to constrain carbon-climate feedbacks for land ecosystems (Cox et al., 2013; Wenzel et al., 2014). In the
ocean, emergent constraints have so far not been applied to constrain feedback parameters but have been successfully applied
to reduce uncertainties in net primary production (Kwiatkowski et al., 2017), ocean heat uptake (Bourgeois et al., 2022; Vogt
et al., 2025), ocean acidification (Terhaar et al., 2020, 2021b, a), and past, present, and future ocean carbon uptake (Terhaar
75 et al., 2021a; Bourgeois et al., 2022; Terhaar et al., 2022, 2024).

Existing emergent constraints for ocean carbon uptake (Terhaar et al., 2021a; Bourgeois et al., 2022; Terhaar et al., 2022, 2024),
may also provide a pathway to constrain ocean β given that it dominates total ocean carbon uptake efficiency. These previ-
ous studies could trace much of the inter-model spread of ocean carbon uptake to differences in how models represent ocean
circulation, ventilation, and carbonate chemistry. Furthermore, these studies have identified mechanistic relationships between
80 observable ocean indicators of ventilation and carbonate chemistry. The indicators used for ventilation are the strength of the
Atlantic Meridional Overturning Circulation (AMOC) at 26.5°N, which is observed since 2004 (Cunningham et al., 2007;
McCarthy et al., 2015), and the average sea surface salinity between the subpolar and subtropical front in the Southern Ocean,
which dictates the strength of the formation of mode and intermediate water in this strongly ventilated ocean region (Terhaar



et al., 2021a). The North Atlantic and Southern Ocean are regions of especially strong ventilation (Dickson and Brown, 1994; Kuhlbrodt et al., 2007; Talley, 2013; Marshall and Speer, 2012) and hence large transport of carbon from the surface to the ocean interior (Sarmiento et al., 1992; Orr et al., 2001; Caldeira and Duffy, 2000; Sabine et al., 2004; Mikaloff Fletcher et al., 2006). The indicator for the chemical surface ocean uptake capacity, which is defined as the change in surface ocean DIC per change in atmospheric CO₂ (Terhaar et al., 2024) is closely related to the Revelle factor (Revelle and Suess, 1957). Terhaar et al. (2022) fitted a three-dimensional linear relationship across an ensemble of ESMs between these three indicators and past, present, and future simulated ocean carbon uptake. This relationship was then combined with observations of these indicators to constrain uncertainties in simulated ocean carbon uptake over time (Terhaar et al., 2022). The constraint reduced the uncertainty of simulated ocean carbon uptake by 50% and increased the simulated best estimate of ocean carbon uptake by 10% (Terhaar et al., 2022). In addition to the anthropogenic carbon uptake, which is strongly related to β , another study showed that the AMOC is also related to γ when quantified from historical simulations (Terhaar, 2024). However, the number of models that provided the necessary simulations to quantify γ were limited so that no formal constraint could be applied.

In this study, we examine whether previously developed emergent constraints for ocean carbon uptake can be extended to constrain ocean β and γ . We first apply the existing emergent constraint (Terhaar et al., 2022) to idealized Earth system model (ESM) simulations with atmospheric CO₂ increasing by 1% per year. We then extend this constraint by adding an indicator for bottom water formation below 3000 m via open-ocean deep convection, particularly in the Southern Ocean and North Atlantic. Building on this extended, four-dimensional constraint, we derive an observationally constrained estimate of ocean β . We then identify and explain a mechanism that increasingly links γ to β as the ocean takes up more anthropogenic carbon, and show that this link allows γ to be constrained by the same observable indicators.

2 Methods

2.1 Earth System Models and simulations

In this study, we use monthly data from several models from the Coupled Model Intercomparison Project Phase 6 (CMIP6) that provided output for air-sea CO₂ fluxes for the simulations '1pctCO₂', '1pctCO₂-bgc', and 'pi-Control', and for the North Atlantic streamfunction, potential temperature, salinity, sea surface DIC, and sea surface alkalinity in the historical simulations. In addition, streamfunction output from the Shared Socioeconomic Pathway 5-8.5 (SSP5-8.5) was used for each model. The radiatively coupled simulation '1pctCO₂-rad' is additionally required to separate γ into its natural and anthropogenic components (Sect. 2.2). It is available for all models except GFDL-ESM4, which therefore contributes to β and γ but not to γ_{nat} and γ_{ant} . All annual and globally integrated air-sea CO₂ fluxes and the AMOC strength were calculated on the original model grids. All other data were regridded on a regular 1°x1° grid using distance-weighted averaged regridding from CDO (Schulzweida, 2023) to facilitate comparison with observation-based estimates, which are also on a 1°x1° grid (section 2.4). The mixed layer depth was calculated from monthly potential temperature and salinity fields following de Boyer Montégut et al. (2004) for ESMs and observations to ensure consistency in the definition of the mixed layer depth across models and observations. Overall, the output was available for 13 ESMs from CMIP6 (models in table 1 and CMCC-ESM2 (Lovato et al.,



2022) and CNRM-ESM2-1 (Séférian et al., 2019)). All of these models have been evaluated against observation (Séférian et al., 2020) and their general behavior in future projections has been documented (Kwiatkowski et al., 2020).

We decided to exclude CNRM-ESM2-1 and CMCC-ESM2 from our analysis. CNRM-ESM2-1 was excluded because its
120 air-sea CO₂ fluxes and hence β and γ are modulated by dynamic carbon transport from the land into the ocean (Séférian et al., 2019). These changes in the land-to-ocean carbon flux are calculated in the land model component and are affected by both increasing CO₂ and climate change. The resulting changes in riverine fluxes of carbon strongly modulate air-sea CO₂ flux estimates (Terhaar et al., 2022), so that β and γ as calculated in CNRM-ESM2-1 are not directly comparable to β and γ as calculated in the remaining ESMs. In the remaining ESMs, riverine fluxes are either non-existent, constant, or only change
125 proportional to simulated changes in discharge, which result in substantially smaller changes in the fluxes of carbon, alkalinity, and nutrients with climate change than in CNRM-ESM2-1 where also the concentration in the river can change. One exception in addition to CNRM-ESM2-1 is MIROC-ES2L, where nutrient concentrations are also dynamically calculated but not carbon and alkalinity concentrations, so that the effect on the air-sea CO₂ flux remains limited. Given that most models still calculate β and γ without these changing riverine fluxes or with substantially smaller changes only driven by discharge changes, we adhere
130 to that definition. Ideally, all models would eventually implement changing riverine fluxes exceeding simple discharge scaling to enable a comprehensive evaluation of β and γ that includes the effect of varying carbon fluxes from rivers to the ocean. We also exclude CMCC-ESM2 because sea surface salinity in the Southern Ocean, the most important predictor for carbon uptake, drifts substantially in this model with the drift being larger than the climate change-driven decline in SSS at quadrupling of CO₂ (Figure S1d). In the historical simulation, sea surface salinity from 2005 to 2014 is taken after 155–164 years of drift,
135 and the historical value of CMCC-ESM2 at that time (Figure S8d) is therefore no longer representative of the conditions in the idealized simulations, which are only 140 years long and used to quantify β and γ . Including these two models leaves the constrained estimates essentially unchanged: at year 140 β shifts from 0.847 to 0.840 Pg C ppm⁻¹ and γ from -24.5 to -23.6 Pg C °C⁻¹, less than one and about three per cent respectively, well inside the constrained uncertainty. The uncertainty itself grows, for β from 0.037 to 0.042 Pg C ppm⁻¹ and for γ from 3.6 to 4.4 Pg C °C⁻¹, which is what the caveats motivating their
140 exclusion would predict (Table S1). In every case the constrained uncertainty remains well below the raw ensemble spread.

In addition to these 11 ESMs from CMIP6, we also use simulations from the GISS-E2-1-G model (Ito et al., 2020) and from one CMIP5 model (GFDL-ESM2M (Dunne et al., 2012, 2013) run with the ocean biogeochemical model COBALTv2 from CMIP6 (Stock et al., 2020)) that were made available for this study. The GISS-E2-1-G model is a slightly updated version of the original CMIP6 model (Ito et al., 2020). The main difference with the CMIP6 version is the improved spinup on both the
145 land and the ocean carbon cycle, the inclusion of oxygen and alkalinity as prognostic tracers (Lerner et al., 2021, 2024) and an updated treatment of iron scavenging (see Romanou et al. (2026) for details). Thus, we overall analyzed output from 13 fully-coupled ESMs.

All the individual model simulations follow the CMIP5 and CMIP6 standard protocols (Taylor et al., 2012; Eyring et al., 2016). The '1pctCO₂' simulation covers 140 years and in each year prescribed atmospheric CO₂ increases by 1% leading to a
150 doubling after 70 years and a quadrupling after 140 years. The '1pctCO₂-bgc' simulation is like '1pctCO₂' but only the carbon cycle *sees* the additional carbon in the atmosphere but the radiative model does not so that the climate only slightly changes



in response to changes in the land biosphere that responds to increasing atmospheric CO₂. The '1pctCO₂-rad' simulation mirrors the '1pctCO₂-bgc' simulation: climate is changing in response to increasing CO₂ but the carbon cycle does not *see* the increasing carbon in the atmosphere and assumes that the atmospheric CO₂ concentration remains at preindustrial levels.

155 The CMIP6 models do not agree on what the first archived year of the 1 % yr⁻¹ simulations represents. Ten of the fifteen models for which output was downloaded begin that year at the unperturbed concentration. The other five models begin the simulation already with an increase of atmospheric CO₂ by 1% in the first year (ACCESS-ESM1-5, CMCC-ESM2, GFDL-ESM2M, MRI-ESM2-0 and UKESM1-0-LL). We determined this for each model from the archived atmospheric CO₂ mass or mole fraction where available, from the prescribed forcing file for GFDL-ESM2M, and from the air-sea CO₂ flux response

160 over the first three years for the remainder. The five models of the second group, and the pre-industrial control aligned to them, were shifted by one year so that the first index is the unperturbed year in every model. Year 70 and year 140 then correspond to 1.99 and 3.99 times the pre-industrial concentration, which we refer to as a doubling and a quadrupling of atmospheric CO₂ throughout. In addition, GFDL-ESM2M was run with a pre-industrial concentration of 286.0 ppm rather than the CMIP6 value of 284.3 ppm, and its trajectory is referenced to that value here.

165 The historical CMIP6 simulations follow historical forcing of CO₂ and other radiative agents and stop in 2014. After 2014, we only use simulated AMOC strength to compare it to the corresponding observational record. For that time, we decided to use output from the Shared Socioeconomic Pathway 5-8.5, which describes a scenario of high emissions over the 21st century (O'Neill et al., 2016), as the simulated AMOC remains indistinguishable between scenarios until the middle of the 21st century (Weijer et al., 2020) and most models provide output for that scenario. For GFDL-ESM2M, which follows the CMIP5 protocol,

170 the historical simulation stops in 2005; we extend it using the Representative Concentration Pathway 2.6 (RCP2.6) (van Vuuren et al., 2011; Meinshausen et al., 2011), whose prescribed atmospheric CO₂ concentrations closely follow the observed evolution between 2005 and 2020 (Meinshausen et al., 2011; Lan et al., 2026), diverging only thereafter. The good agreement between observed atmospheric CO₂ and that in RCP2.6 is particularly relevant for the comparison between simulated and observed carbonate uptake capacity from 1997 to 2007 (see below), which depends strongly on atmospheric CO₂. For SST, SSS, and

175 mixed-layer depth, the choice of emission scenario has a negligible effect on the simulated trajectories at least until 2020, compared to the contribution of model differences and internal variability (Schlunegger et al., 2020). RCP2.6 is nonetheless a good choice for this period because its main forcing agent CO₂ closely aligns with observations, meaning that differences in forcing between RCP2.6 and the real world, which can only stem from non-CO₂ forcing agents, are even smaller than the already negligible differences between low- and high-emission scenarios. Lastly, the AMOC, the only simulated variable

180 analysed through 2023, is insensitive to the choice of emission scenario until approximately the middle of the 21st century (Weijer et al., 2020).

2.2 Calculating β and γ

The feedback parameters β and γ were calculated similar to previous studies (Friedlingstein et al., 2006; Arora et al., 2013, 2020) (Table 2). β was calculated as the difference in cumulative global ocean carbon uptake in the '1pctCO₂-bgc' and 'pi-Control' simulations divided by the change in atmospheric CO₂ in the given year (1% increase of CO₂ starting from pre-industrial

185



Table 1. Earth System Models used in this study. The GISS-E2-1-G and GFDL-ESM2M models are each modified version; modifications are described in Section 2.1

Model	Modeling Group	Ensemble Mem.	Reference
ACCESS-ESM1-5	Commonwealth Scientific and Industrial Research Organisation, Australia	rli1p1f1	Ziehn et al. (2020)
CanESM5	Canadian Centre for Climate Modelling and Analysis	rli1p1f1	Swart et al. (2019)
CESM2	National Center for Atmospheric Research, USA	rli1p1f1	Danabasoglu et al. (2020)
EC-Earth3-CC	EC-Earth Consortium	rli1p1f1	Döscher et al. (2022)
GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory, USA / University of Bern, Switzerland	/	Dunne et al. (2012, 2013)
GFDL-ESM4	NOAA Geophysical Fluid Dynamics Laboratory, USA	rli1p1f1	Dunne et al. (2020); Stock et al. (2020)
GISS-E2-1-G	NASA Goddard Institute for Space Studies, USA	rli1p1f1	Kelley et al. (2020)
IPSL-CM6A-LR	Institut Pierre-Simon Laplace, France	rli1p1f1	Boucher et al. (2020)
MIROC-ES2L	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute, and National Institute for Environmental Studies, Japan	rli1p1f2	Hajima et al. (2020)
MPI-ESM1-2-LR	Max Planck Institute for Meteorology, Germany	rli1p1f1	Mauritsen et al. (2019)
MRI-ESM2-0	Meteorological Research Institute, Japan	rli2p1f1	Yukimoto et al. (2019)
NorESM2-LM	Norwegian Climate Centre, Norway	rli1p1f1	Seland et al. (2020); Tjiputra et al. (2020)
UKESM1-0-LL	Met Office Hadley Centre, UK	rli1p1f2	Sellar et al. (2019)

atmospheric CO₂ (284.3 ppm)). Subtracting the 'pi-Control' simulation removes potential drifts in the simulated carbon sink that can arise from models being not yet in equilibrium.

190 γ was calculated as the difference in cumulative global ocean carbon uptake in the '1pctCO₂' and '1pctCO₂-bgc' simulations divided by the difference in atmospheric warming in the '1pctCO₂'. The carbon uptake in the '1pct' simulations includes the increase in carbon uptake due to atmospheric CO₂ increase, as well as the effect of climate change on both, the natural carbon reservoir in the ocean and on the additional carbon due to increasing atmospheric CO₂. By removing the carbon uptake in the '1pct-bgc' simulation, the uptake due to atmospheric CO₂ increase is removed leaving the effect of climate change on the natural and additional carbon reservoirs in the ocean. By calculating the carbon uptake due to climate change as the difference between two simulations, any potential drift is already removed. Any drift in atmospheric warming in the '1pctCO₂' was removed by subtracting a fitted Enting spline (Enting, 1987) to the atmospheric temperature in the piControl simulation. The Enting spline was developed originally for atmospheric CO₂ observations and previously used in ESMs and climate models to remove high-frequency variability (Silvy et al., 2024; Terhaar, 2025; Terhaar et al., 2025). Here, we used a cutoff period of 100



Table 2. Simulated feedback parameters of each ESM at year 70 and year 140 of the 1 % yr⁻¹ CO₂ simulations. The unit of β is Pg C ppm⁻¹ and the unit of the three γ variables is Pg C °C⁻¹.

Model	β		γ		γ_{nat}		γ_{ant}	
	yr 70	yr 140	yr 70	yr 140	yr 70	yr 140	yr 70	yr 140
ACCESS-ESM1-5	1.031	0.879	-10.56	-22.53	-3.58	-6.17	-6.99	-16.36
CanESM5	0.911	0.772	-7.89	-14.63	-4.66	-7.66	-3.23	-6.97
CESM2	0.845	0.715	-4.37	-10.49	-1.49	-0.98	-2.88	-9.51
EC-Earth3-CC	0.859	0.742	-4.05	-11.81	-2.18	-2.37	-1.87	-9.44
GFDL-ESM2M	0.980	0.897	-9.22	-27.55	-3.58	-5.50	-5.65	-22.05
GFDL-ESM4	0.977	0.848	-7.83	-20.41	-	-	-	-
GISS-E2-1-G	0.915	0.822	-7.44	-24.33	-3.84	-6.10	-3.60	-18.23
IPSL-CM6A-LR	0.872	0.765	-6.91	-12.00	-3.70	-4.16	-3.21	-7.84
MIROC-ES2L	0.850	0.735	-9.59	-20.56	-3.30	-4.37	-6.29	-16.19
MPI-ESM1-2-LR	0.885	0.782	-9.56	-19.55	-5.10	-7.88	-4.45	-11.67
MRI-ESM2-0	1.082	0.941	-12.54	-29.88	-3.75	-6.76	-8.79	-23.12
NorESM2-LM	0.900	0.791	-10.13	-20.75	-5.11	-6.85	-5.03	-13.90
UKESM1-0-LL	0.916	0.768	-7.23	-13.22	-4.30	-6.41	-2.93	-6.81
Multi model mean	0.925	0.804	-8.26	-19.06	-3.72	-5.43	-4.58	-13.51
Standard deviation	0.073	0.069	2.39	6.23	1.07	2.10	2.03	5.71

years to remove only the overall drift and no inter-annual, decadal, or multi-decadal variability. As atmospheric temperature has large inter-annual and decadal variability that can effect the estimation of γ , we also fit an Enting spline to the undrifted modeled temperature change. This second enting spline allows us to detect warming directly in a given year without taking a 21 year average around a given year (Arora et al., 2020). We used the spline instead of an average over the years around the year in question because the 1% increase in CO₂ leads to asymmetric warming changes before and after any year. The stronger increase in the 10 years following the year in question thus sees a stronger warming rate than the years before so that a 21 year average gives a slightly too warm warming for the year in question. The warming could be calculated from the '1pctCO₂' (after removing the drift) or the difference of warming between the '1pctCO₂' and '1pctCO₂-bgc' simulations (Arora et al., 2020). Here, we have opted for the option that has been mostly used in the past (undrifted '1pctCO₂') (Arora et al., 2020; Asaadi et al., 2024). Had we chosen the other option ('1pctCO₂' and '1pctCO₂-bgc'), warming would be around 0.2°C smaller because of some small warming in '1pctCO₂-bgc' through the effect of increasing atmospheric CO₂ on the land biosphere (Arora et al., 2020) and γ would hence be around 5% larger. The small warming of around 0.2 °C in the '1pctCO₂-bgc' simulation also affects β , as it causes a small loss of carbon from the ocean that is not attributable to rising atmospheric CO₂. Removing the loss of carbon with warming would increase β by around 0.5% relative to the value calculated here (Schwinger et al., 2014;



Arora et al., 2020). We follow Friedlingstein et al. (2006) and Arora et al. (2020) in neglecting this term, which keeps our estimates directly comparable to previous assessments.

In addition to γ , we also calculated γ_{nat} and γ_{ant} , which are defined as the sensitivity of air-sea CO₂ fluxes to the effect of climate change on the natural (nat) and anthropogenic (ant) carbon reservoir. γ_{nat} was calculated here as the difference in cumulative global carbon uptake in the '1pctCO₂-rad' and 'pi-Control' simulations divided by the undrifted and smoothed atmospheric warming in the '1pctCO₂' simulation. γ_{ant} was calculated as the difference between γ and γ_{nat} . As GFDL-ESM4 did not provide the radiatively coupled simulation, γ_{nat} and γ_{ant} are available for 12 of the 13 ESMs.

2.3 Multi-variable emergent constraint

Building on Terhaar et al. (2022), we fit a multi-linear relationship between four predictor variables (three from Terhaar et al. (2022) and an additional one added here) from the historical simulations, and β derived from the '1pctCO₂-bgc' simulations across the ensemble of 13 ESMs (Table 1) analyzed here:

$$\beta(t) = a(t) \cdot \text{AMOC} + b(t) \cdot \text{SSS} + c(t) \cdot \text{CUC} + d(t) \cdot \text{MLD}_{\text{vol}}^{3000} + e(t). \quad (1)$$

Here, AMOC is the maximum of the simulated streamfunction at 26.5°N below 500 m (to avoid potential wind-driven peaks in the streamfunction that are not related to the strength of the overturning) averaged from 2005 to 2023, SSS is the mean inter-frontal sea surface salinity in the Southern Ocean averaged from 2005 to 2014 as calculated by Terhaar et al. (2021a), CUC is the global mean surface ocean chemical uptake capacity averaged from 1997 to 2007 as calculated by Terhaar et al. (2024), $\text{MLD}_{\text{vol}}^{3000}$ is the volume of waters below 3000 m to which the mixed layer extends, computed monthly, averaged to annual means and then averaged from 2005 to 2014, and a , b , c , d , and e are parameters fitted based on simulated values of the AMOC, SSS, CUC, $\text{MLD}_{\text{vol}}^{3000}$, and β across the ensemble of 13 ESMs (Table 1). The fit is calculated for each of the 140 years of the '1pctCO₂-bgc' simulation individually, making a , b , c , d , and e functions of time (t). In addition to constraining β , we test if the same predictor variables can also constrain γ , γ_{nat} , and γ_{ant} (Section 4). The CUC was calculated using mocsy2.0 (Orr and Epitalon, 2015) from simulated and observation-based estimates of temperature, salinity, dissolved inorganic carbon, total alkalinity, phosphate, and silicate using the equilibration constants recommended for best practice (Dickson et al., 2007). As not all ESMs provide output for phosphate and silicate and as both do not substantially affect differences in CUC, we decided to use values from GLODAPv2 (Lauvset et al., 2016) for all ESMs to ensure consistency. The $\text{MLD}_{\text{vol}}^{3000}$ was calculated from potential temperature and salinity using a critical density difference of 0.03 kg m⁻³ relative to the density at 10 m depth to define the depth of the mixed layer, as proposed by de Boyer Montégut et al. (2004). To prevent models with deeper bathymetry from yielding larger volumes for that reason alone, the mixed layer depth in each cell was limited to the shallowest sea floor across all models and observation-based products before computing $\text{MLD}_{\text{vol}}^{3000}$. All data was regridded on a regular 1°x1° grid. Only grid cells where all models and observation-based estimates provided data were considered when calculating the predictors in the model and observations, ensuring that differences across models and to observation-based estimates do not arise from different



data coverage or ocean masks. As the same software and constants were used for calculating the CUC, uncertainties from these sources (Orr et al., 2018) do not affect the exploitation of the constraints from ESMs with observation-based estimates.

245 For each of the four predictor variables, different time periods have been chosen based on availability of observation-based estimates and the simulation strategy of CMIP models. For the CUC, the period 1997 to 2007 was chosen to encompass the central year 2002, for which GLODAPv2 gridded data is available (Lauvset et al., 2016). For the inter-frontal SSS and mixed layer depth, data from 2005 to 2014 was chosen as the period ARGO floats substantially increased the spatial coverage of observations from 2005 onward (Roemmich et al., 2009) and because historical forcing for CMIP6 simulations stops in 2014
250 (Eyring et al., 2016) so that no choice of scenario has to be made. For the AMOC, the period with available observations over all months of a year (2005-2023) was chosen to remove as much as possible of the internal variability, which is especially strong for the AMOC (Buckley and Marshall, 2016; Jackson et al., 2022). Although this period extends beyond the end of the historical simulations, the simulated AMOC change in the initial years after the historical simulations remains indistinguishable between scenarios (Weijer et al., 2020) so that every scenario can be used to compare the simulated AMOC to observations.
255 The simulated average values of the AMOC, SSS, MLD_{vol}^{3000} , and CUC from all 13 ESMs as well as the observation-based estimates are presented in table 3 together with uncertainty estimates of the observation-based estimates which are explained in more detail in section 2.4.

The four predictors may be related to one another and thus contain overlapping information. Terhaar et al. (2022) tested for such relationships among the AMOC, SSS, and CUC in a different ensemble of ESMs. Here, we assess multicollinearity among
260 all four predictors in the present ensemble. Across the 13 ESMs, the strongest pairwise correlation is between the inter-frontal SSS and the CUC ($r = 0.51$), which is expected because both characterize the same surface water mass. All other pairwise correlations are weaker, with $|r| \leq 0.30$ (Fig. S12). Pairwise correlations, however, only consider two predictors at a time and therefore do not capture multicollinearity involving combinations of several predictors. We therefore additionally calculate the variance inflation factor (VIF) for each predictor. For each predictor i , we regress it on the other three predictors and obtain the
265 fraction of its variance explained by them, r_i^2 . The VIF is then given by

$$VIF_i = \frac{1}{1 - r_i^2}. \quad (2)$$

It quantifies the inflation of the variance of the estimated regression coefficient due to multicollinearity (Marquardt, 1970). A VIF of 1 corresponds to $r_i^2 = 0$, meaning that none of the variance of that predictor can be explained by a linear combination of the other three predictors. For example, a VIF of 5 corresponds to $r_i^2 = 0.8$, meaning that 80% of the variance across ESMs
270 in that predictor can be explained by the other three predictors. Values of 5 or 10 are commonly used as rules of thumb for identifying potentially problematic multicollinearity, although such thresholds should not be interpreted as strict criteria and higher values may still be acceptable depending on the context (O'Brien, 2007). Here, the VIFs are 1.10 for MLD_{vol}^{3000} , 1.17 for the AMOC, 1.37 for the SSS, and 1.44 for the CUC, well below any proposed threshold. Thus, at most 31% of the variance across ESMs in any one predictor can be explained by a linear combination of the other three predictors. The slightly higher
275 VIFs of the SSS and CUC are consistent with their stronger pairwise correlation. Nevertheless, all VIFs remain close to 1, indicating weak multicollinearity among the four predictors and supporting their simultaneous use in the regression.



Table 3. Mean values for the AMOC from 2005 to 2023 (Sv), the Southern Ocean inter-frontal sea surface salinity (SSS) from 2005 to 2014, the volume of waters below 3000 m ventilated by MLDs deeper than 3000 m ($\text{MLD}_{\text{vol}}^{3000}$), and the global mean sea surface chemical uptake capacity (CUC) from 1997 to 2007 ($\text{mmol m}^{-3} \text{ppm}^{-1}$) for ESMs (section 2.1) and observation-based estimates (section 2.4)

Model	AMOC (Sv) (2005–2023)	SSS (2005–2014)	$\text{MLD}_{\text{vol}}^{3000}$ (10^{12} m^3) (2005–2014)	CUC ($\text{mmol m}^{-3} \text{ppm}^{-1}$) (1997–2007)
ACCESS-ESM1-5	18.66	33.83	205.8	0.542
CanESM5	12.21	33.83	0.3	0.545
CESM2	18.25	33.63	0.0	0.537
EC-Earth3-CC	18.63	33.83	67.2	0.521
GFDL-ESM2M	18.92	33.90	116.3	0.546
GFDL-ESM4	17.96	33.98	2.0	0.555
GISS-E2-1-G	20.89	33.95	21.5	0.553
IPSL-CM6A-LR	12.86	33.85	0.4	0.529
MIROC-ES2L	11.74	33.97	0.3	0.534
MPI-ESM1-2-LR	17.71	33.90	17.0	0.519
MRI-ESM2-0	18.91	34.11	0.1	0.560
NorESM2-LM	20.87	33.82	0.0	0.536
UKESM1-0-LL	16.19	33.78	2.5	0.536
Multi model mean	17.22 ± 3.08	33.87 ± 0.12	33.3 ± 60.0	0.539 ± 0.012
Observation-based	16.91 ± 0.77	34.04 ± 0.01	0.0 ± 0.0	0.546 ± 0.003

While the VIF quantifies redundancy among the predictors, it does not address whether all four are warranted. We therefore also compare all fifteen predictor subsets using the Akaike and Bayesian information criteria (AIC and BIC) (Akaike, 1974; Schwarz, 1978) and a leave-one-out (LOO) cross validation, in which each ESM in turn is predicted from a fit that excludes it (Stone, 1974). For a Gaussian least-squares fit, $\text{AIC} = N \ln(\text{SSR}/N) + 2k$ and $\text{BIC} = N \ln(\text{SSR}/N) + k \ln N$, with $k = p + 2$ counting the p slopes, the intercept and the residual variance. Because N/k reaches about two for the four-predictor fit, we additionally report the small-sample corrected $\text{AICc} = \text{AIC} + 2k(k+1)/(N-k-1)$, which is recommended whenever N/k is below roughly 40 (Burnham and Anderson, 2002) and which penalises additional predictors more strongly than AIC at this ensemble size.

The results support the use of all four predictors. AIC, BIC and the LOO error all select the full four-predictor model, whereas AICc selects a smaller subset (Tables S7 and S8). The subset AICc selects, however, differs between the two target variables. For β , it retains CUC and $\text{MLD}_{\text{vol}}^{3000}$ at doubling and CUC, SSS, and $\text{MLD}_{\text{vol}}^{3000}$ at quadrupling. For γ , it retains SSS at doubling and SSS and the AMOC at quadrupling. Each of the four predictors is therefore retained for at least one target and we keep all four throughout to ensure consistency between β and γ .



290 Once the regression coefficients $a(t)$, $b(t)$, $c(t)$, $d(t)$, and $e(t)$ have been fitted to the ensemble of ESMs at a given time step t , we evaluate this relationship with observation-based estimates of the four predictors to obtain an observationally constrained estimate of β , yielding a single constrained value of β at each time step t .

The uncertainty of this constrained value combines two independent sources: the uncertainty in the observation-based estimates of the four predictors, and the residual spread of the ensemble around the fitted relationship. The observational contribution, $\sigma_{\text{obs}}(t)$, is obtained by propagating the uncertainty of each observation-based predictor estimate (Sect. 2.4) through the fitted regression coefficients, and combining the resulting terms in quadrature:

$$\sigma_{\text{obs}}(t) = \sqrt{(a(t)\sigma_{\text{AMOC}})^2 + (b(t)\sigma_{\text{SSS}})^2 + (c(t)\sigma_{\text{CUC}})^2 + (d(t)\sigma_{\text{MLD}})^2}, \quad (3)$$

where σ_{AMOC} , σ_{SSS} , σ_{CUC} , and σ_{MLD} are the uncertainties of the respective observation-based predictor estimates. This term reflects how much the constrained estimate would shift if the observation-based predictors were resampled within their own uncertainty, and therefore captures how precisely the predictors are known, independent of how well the fitted relationship itself performs.

The second contribution, $\sigma_{\text{model}}(t)$, quantifies the prediction interval of the fitted relationship, evaluated at the observed predictors, following standard multiple-regression theory:

$$\sigma_{\text{model}}(t) = s(t)\sqrt{1 + h(\mathbf{x}_{\text{obs}}, t)}, \quad (4)$$

305 where $s(t) = \sqrt{\text{SSR}(t)/(N - p - 1)}$ is the unbiased residual standard error of the regression, with $\text{SSR}(t)$ the residual sum of squares across the $N = 13$ ESMs and p the number of predictors (3 or 4). We use $N - p - 1$ rather than N degrees of freedom, since the residual variance would otherwise be underestimated given the small ensemble size relative to the number of fitted predictors. The leverage term $h(\mathbf{x}_{\text{obs}}, t)$ extends the univariate emergent-constraint prediction interval of Cox et al. (2013), $\sigma_f(x) = s\sqrt{1 + 1/N + (x - \bar{x})^2/(N\sigma_x^2)}$, to the multiple predictors used here: the scalar, variance-normalized squared distance $(x - \bar{x})^2/(N\sigma_x^2)$ is replaced by $h(\mathbf{x}_{\text{obs}}, t) = \mathbf{x}'_{\text{obs}}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_{\text{obs}}$, a Mahalanobis-type distance that accounts for both the spread and the correlations among the model-ensemble predictors ($\mathbf{X}$ is the design matrix of the standardized model-ensemble predictors, and $\mathbf{x}_{\text{obs}}$ the corresponding observed predictor vector, both including an intercept term). As in the univariate case, $h(\mathbf{x}_{\text{obs}}, t) = 1/N$ if the observations coincide exactly with the ensemble mean, and increases the further the observations lie from the ensemble's predictor distribution.

315 The total uncertainty of the constrained estimate, $\sigma_{\text{total}}(t)$, is obtained by combining these two independent uncertainty sources in quadrature, which is equivalent to convolving two Gaussian distributions with these standard deviations:

$$\sigma_{\text{total}}(t) = \sqrt{\sigma_{\text{obs}}(t)^2 + \sigma_{\text{model}}(t)^2}. \quad (5)$$

All uncertainties reported here correspond to one standard deviation (1σ , the likely, 68 % range). This treatment assumes Gaussian residuals rather than a Gaussian distribution of β across models itself. Indeed, a skewed β is expected if it depends on predictors that are themselves unevenly spread across the ensemble. The residuals are in fact closer to Gaussian than β itself, and refitting on log and square-root scales with the corresponding back-transformation changes the constrained β at year 140



by 0.1 and 0.06 %, with the uncertainty widening by 4 and 2 %. Rather than performing the full marginalization integral of Cox et al. (2013), $P(y) = \int P\{y|x\} P(x) dx$, we evaluate $\sigma_{\text{model}}(t)$ at the fixed observed predictor point. We verified numerically, via Monte Carlo sampling of the observational predictor uncertainty (500,000 samples, allowing the leverage $h(\mathbf{x})$ to vary with the sampled predictor value rather than being fixed), that this simplification changes $\sigma_{\text{total}}(t)$ by less than 0.5 %, since the observational uncertainty in the predictors is small relative to the spread of the model ensemble.

2.4 Observation-based estimates

To constrain the target variable using the emergent constraint derived here (Section 2.3), observation-based estimates of the four predictor variables are needed.

330 The observation-based estimate of the AMOC strength at 26.5°N is obtained from the RAPID array (Cunningham et al., 2007; McCarthy et al., 2015; Moat et al., 2025), averaged over the nineteen complete years 2005 to 2023, giving 16.91 Sv. The uncertainty on this mean combines a sampling and a measurement term. The sampling term is the standard error of the annual means, adjusted for their autocorrelation as successive years are not independent but have a lag-1 correlation of 0.28. Therefore, the effective sample size is 10.8 rather than 19 and the standard error is 0.38 Sv rather than the 0.29 Sv. The measurement term starts from the ± 0.9 Sv accuracy of an annual AMOC value (McCarthy et al., 2015), which is itself a combination of a calibration term (0.6 Sv), a gridding term (0.4 Sv), a western boundary wedge term (0.5 Sv) and a Florida Straits term (0.3 Sv). How much of this averages down over nineteen years depends on how each component correlates in time, and that differs between them. The calibration is redrawn whenever moorings are recovered and redeployed, and the Florida Straits term already falls from 1.1 Sv on ten-day means to 0.3 Sv on annual means. Therefore, these uncertainties should average down. The gridding and western boundary wedge terms are listed unchanged at both timescales, which may mean they are structural uncertainties that do not get removed over time, or alternatively that no ten-day estimate was available. We take the conservative view and assume that these two structural terms do not average down at all, while the calibration averages over deployments and the Florida Straits term over years. The resulting measurement contribution to the overall uncertainty is 0.67 Sv. If all four components were treated as independent between years, it would reduce to 0.21 Sv. Combining the sampling and measurement terms in quadrature gives an observational uncertainty of ± 0.77 Sv.

345 The observation-based estimate of the inter-frontal sea surface salinity from 2005 to 2014 of 34.04 ± 0.01 was calculated following Terhaar et al. (2021a) based on sea surface salinity and temperature data from (1) the World Ocean Atlas 2023 (Reagan et al., 2024a, b; Locarnini et al., 2024), (2) EN4 (Good et al., 2013), and (3) IAP with v2.2 for salinity (Cheng et al., 2020) and v4.2 for temperature (Cheng et al., 2024). The estimates of WOA23, EN4, and IAP are 34.038, 34.036, and 34.036, resulting in a multi-product mean estimate of 34.037. The spread across the three observational products (± 0.001 psu) provides an estimate of the uncertainty arising from differences in extrapolation and analysis technique. In addition, WOA23 and EN4 carry a per-cell error estimate. Since grid-cell errors are not independent (the objective analyses extrapolate sparse observations over broad areas, so that neighboring cells share the same underlying data), we do not divide the per-cell uncertainty by the full number of grid cells (N). Instead, we divide by the effective number of independent samples, N divided by the average number of grid cells contributing to each analysed point. This average number of contributing cells is estimated from WOA23, which



provides the number of grid squares within the smallest radius of influence around each grid square that contain a statistical mean. In the inter-frontal region that number is around 46 on average, giving roughly 78 effective independent samples in that region. This yields areal-mean sampling uncertainties of 0.002 (WOA23) and 0.008 (EN4). Combining the multi-product spread with these two sampling uncertainties in quadrature gives an overall uncertainty estimate of 0.002-0.008, which we conservatively round to 0.01. Overall, we thus use an observation-based estimate of the inter-frontal sea surface salinity in the Southern Ocean of 34.04 ± 0.01 .

The observation-based estimate of the volume of waters reached by the mixed layer depth below 3000 m depth (MLD_{vol}^{3000}) was calculated from monthly temperature and salinity data from EN4 (Good et al., 2013) and IAP with v2.2 for salinity (Cheng et al., 2020) and v4.2 for temperature (Cheng et al., 2024). The GSW-Python toolbox (McDougall and Barker, 2011) was used to compute potential density from conservative temperature and absolute salinity. Conservative temperature was derived from potential temperature for EN4 and from in situ temperature for IAP, as no potential temperature product is available for IAP. When calculating the MLD the same way as in ESMs, too deep MLDs were identified because the density varied substantially within the first 200 m and often decreased from the surface to a minimum between 50 and 100 m, which indicates an unstable water column, which is physically implausible and likely an artifact of the extrapolation of measurements. The values above the subsurface minimum are interpreted as noise from the extrapolation. As the densities were higher at the surface, interior ocean waters were never 0.03 kg m^{-3} heavier than the surface waters above and a deep MLD was identified. However, these dense water masses were more than 0.03 kg m^{-3} heavier than the density minimum in the water column. As the inversions in the upper ocean waters are likely an artifact of the extrapolation technique, we tested in the observation if the difference between the density minimum and maximum exceeds 0.03 kg m^{-3} somewhere in the water column. Such a difference would show that the water is not mixed anymore in one layer according to the criterion. A risk of only analyzing values from 100 and 2000 m is that a positive gradient in the first 100 m is not taken into account and a deep MLD is wrongly assigned in that cell if the water column below 100 m is well mixed. However, in both observational products density differences larger than 0.03 kg m^{-3} existed everywhere between 100 and 2000 m indicating that the MLD did not extend to water masses below 2000 m, in line with a previous exhaustive analysis of MLDs (Sallée et al., 2021), and that missing a potential positive gradient in the first 100 m does not wrongly assign a deep MLD to any cell. Therefore, the observation-based estimate of MLD_{vol}^{3000} is $0 \pm 0 \cdot 10^{12} \text{ m}^3$. In all ESMs, such physically implausible inversions exist in less than 0.1% of all ocean cells, likely due to monthly averaging of cells with fast changes in water masses, and never in cells with MLDs deeper than 1000 m. If we also used the difference between densities between 100 and 2000 m in ESMs as the criterium, we would remove the real simulated gradients in the first 100 m and artificially inflate the mixed layer depth, which is why we calculate the difference in density compared to the surface value.

The global mean surface ocean CUC averaged from 1997 to 2007 was calculated based on observation-based estimates of surface ocean dissolved inorganic carbon, total alkalinity, temperature, and salinity from the gridded version of GLODAPv2 (Lauvset et al., 2016) and the following three $p\text{CO}_2$ products: OCEAN-SODA (Gregor and Gruber, 2021), CMEMS-LSCE-FFNN (Chau et al., 2024), and JMA-MLR (Iida et al., 2021). The best estimate ($0.546 \text{ mmol m}^{-3} \text{ ppm}^{-1}$) is calculated as the average of these four observation-based estimates ($0.548 \text{ mmol m}^{-3} \text{ ppm}^{-1}$ for GLODAPv2, $0.544 \text{ mmol m}^{-3} \text{ ppm}^{-1}$ for



OCEAN-SODA, $0.543 \text{ mmol m}^{-3} \text{ ppm}^{-1}$ for CMEMS-LSCE-FFNN, and $0.547 \text{ mmol m}^{-3} \text{ ppm}^{-1}$ for JMA-MLR). As the difference in the estimates of the CUC among these datasets is substantially larger than the uncertainty of an individual dataset based on standard errors (Terhaar et al., 2022, 2024), the uncertainty is taken as half the range across the four products ($\pm 0.003 \text{ mmol m}^{-3} \text{ ppm}^{-1}$), which is more conservative than their standard deviation of $\pm 0.002 \text{ mmol m}^{-3} \text{ ppm}^{-1}$ given that four
395 products give only three degrees of freedom.

2.5 Comparing air–sea CO_2 flux anomalies to mixed layer depths

To provide a line of evidence for the link between open-ocean deep convection in models and additional carbon uptake by the ocean, we compare MLDs across models to anomalies in the air–sea CO_2 flux at that location. The air–sea CO_2 flux anomaly was calculated as the difference to a climatology to remove any seasonal signal. In addition, the climatology was
400 only built based on months when the MLD was below 1000 m to avoid that the flux at high MLD events contributes to the climatology. Before calculating anomalies, the air–sea CO_2 flux was calculated as the average flux over the month when the MLD is analyzed and the two following months and over the cell itself and the eight surrounding cells as CO_2 gas exchange is not immediate and surface waters may move over three months. Overall, these anomalies were calculated over the last 20 years of the '1pct CO_2 -bgc' simulation, as the air–sea CO_2 flux at the beginning of the simulation, when atmospheric CO_2 is
405 rather low, is relatively more strongly influenced by the outgassing of natural carbon when DIC-rich deep waters are upwelled during the convection. At the end of the simulation, when atmospheric CO_2 is high, the anthropogenic carbon signal, which we want to compare to MLDs, should dominate. We decided to analyze 20 years to have a relatively large amount of years to capture the existing variability in MLDs (Fig. S4) but to also have an as short timeseries as possible so that anomalies are not too much driven by the ongoing increase in atmospheric CO_2 over this time.

410 2.6 Sensitivity to the mixed layer depth criterion

The critical density difference of 0.03 kg m^{-3} used to calculate $\text{MLD}_{\text{vol}}^{3000}$ (Sect. 2.3) is based on potential density referenced to the surface (σ_0). Because the compressibility of seawater depends on temperature, σ_0 becomes an increasingly poor measure of static stability far from the surface. In the weakly stratified deep Southern Ocean, σ_0 can therefore decrease with depth below roughly 2000 m although the water column is stably stratified in situ. In such columns, the σ_0 difference to the reference depth
415 may never reach 0.03 kg m^{-3} , and the mixed layer would be assigned to extend to depths that are not mixed with the surface.

To test whether this affects our results, we additionally calculated the mixed layer depth with a local criterion. Instead of comparing each level with the σ_0 at the reference depth, we compute the density difference between each pair of adjacent model levels k and $k+1$, with k increasing downward, as $\Delta\rho_k = \rho(S_{A,k+1}, \Theta_{k+1}, p_{k+1/2}) - \rho(S_{A,k}, \Theta_k, p_{k+1/2})$. Here, ρ is in situ density calculated with the GSW toolbox (McDougall and Barker, 2011) from Absolute Salinity S_A and Conservative Tem-
420 perature Θ , and both densities are evaluated at the same pressure $p_{k+1/2}$ halfway between the two levels, so that compression does not appear as stratification. These local differences, which correspond to the vertically integrated buoyancy frequency, are summed downward from the same reference depth, and the mixed layer depth is the depth at which the sum first exceeds the same threshold of 0.03 kg m^{-3} . Where a layer is unstable, the density difference is negative and reduces the sum. The sum is



not allowed to fall below zero, so that the count starts again below an inversion that removes the stratification above it entirely.
Near the surface, both criteria are nearly identical, whereas at depth the local criterion follows the actual static stability of the water column.

The constrained estimates are insensitive to the choice of the mixed layer criterion. The local criterion reduces MLD_{vol}^{3000} by 43 % in the two models with the strongest open-ocean deep convection (ACCESS-ESM1-5 and GFDL-ESM2M), and by 12 to 61 % in the three other models with smaller but still substantial deep convection (GISS-E2-1-G, MPI-ESM1-2-LR, and EC-Earth3-CC). Nevertheless, β is 0.960 ± 0.049 instead of 0.961 ± 0.048 Pg C ppm⁻¹ at year 70 and 0.846 ± 0.037 instead of 0.847 ± 0.037 Pg C ppm⁻¹ at year 140, and γ is -10.1 ± 2.3 instead of -10.1 ± 2.4 Pg C °C⁻¹ at year 70 and -24.3 ± 3.4 instead of -24.5 ± 3.6 Pg C °C⁻¹ at year 140. Here, we nevertheless keep the criterion of de Boyer Montégut et al. (2004), as it is well established and widely used.

3 Constraining ocean β with ocean ventilation and carbonate chemistry indicators

3.1 Constraining β with three established indicators for ventilation and carbonate chemistry

The ensemble of 13 ESMs used here simulates β trajectories that increase steadily over the first 70 years, before slowly decreasing over the subsequent 70 years (Figure 1). Ocean carbon uptake reflects two additive contributions: uptake driven by that year's rise in atmospheric CO₂, and uptake driven by the upwelling of old waters without any anthropogenic carbon, having resided in the ocean interior for decades to centuries, whose pCO_2 difference to the atmosphere widens as atmospheric CO₂ rises. The first contribution alone would be expected to scale roughly proportionally with the rise in atmospheric CO₂, keeping β approximately constant. The increase in β therefore reflects the growing contribution from this ventilation-driven pCO_2 gap. However, over time, the additional carbon in the ocean reduces the carbonate uptake capacity at the ocean surface (Revelle and Suess, 1957; Sabine et al., 2004; Terhaar, 2024), and upwelled waters that re-emerge after years or decades become increasingly filled with anthropogenic carbon, decreasing the pCO_2 gap between the atmosphere and newly upwelled waters and substantially lowering ocean carbon uptake through this reemergence of anthropogenic carbon (Rodgers et al., 2020). Consequently, the ocean takes up less carbon for a given increase in atmospheric CO₂ over the second 70 years of the simulation, resulting in a slowly decreasing β .

Despite this shared trajectory, β varies widely across the 13 ESMs, ranging from 0.85 to 1.08 Pg C ppm⁻¹ at year 70 (CO₂ doubling) and from 0.72 to 0.94 Pg C ppm⁻¹ at year 140 (CO₂ quadrupling). This inter-model spread reflects the similarly wide range in the strength of ocean ventilation and the surface ocean CUC, reflected by the range of their respective proxies across these models, which substantially exceed the uncertainties of the observation-based estimates (Table 3) (Terhaar et al., 2022).

To reduce the uncertainty around β , we apply the emergent-constraint approach of Terhaar et al. (2022), using first the three predictor variables from the historical simulations (AMOC, SSS, and CUC) originally identified to explain the inter-model spread in historical ocean carbon uptake. As for the historical ocean carbon uptake, these predictors explain much of the spread in β ($r^2 = 0.62$ at CO₂ doubling and $r^2 = 0.71$ at CO₂ quadrupling; Fig. 2) (see Table S3 for fitted a , b , c , and d as well as the r^2

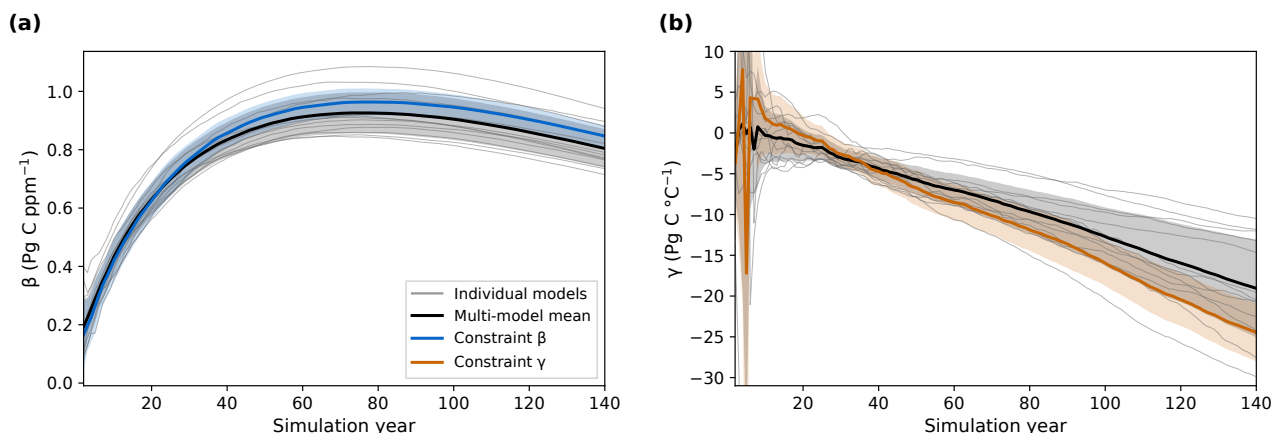


Figure 1. Evolution of β and γ over time. Simulated (a) β and (b) γ by 13 ESMs from CMIP6 over the 140 years of the simulations with atmospheric CO₂ increasing by 1% per year (grey lines), as well as the multi-model average (black line) and the multi-model standard deviation (grey shading). In addition the constraint estimates for β (blue line in (a)) and γ (orange line in (b)) using AMOC, SSS, CUC, and MLD_{vol}³⁰⁰⁰ as predictors, as well as corresponding uncertainty around these constraint estimates (blue and orange shading) are shown for (a) β and (b) γ .

in each year). The r^2 values for β here are slightly lower than the r^2 of 0.87 obtained when constraining anthropogenic carbon uptake over 1994–2007 in Terhaar et al. (2022). We see two possible explanations for this small difference of r^2 values. First, Terhaar et al. (2022) used 17 ESMs, whereas only 13 ESMs provide the diagnostics required here. The larger ensemble reduces the influence of individual outliers on the regression fit, such as ACCESS-ESM1-5 at years 70 and 140 and GFDL-ESM2M at year 140. Second, over the historical period, the predictors have already been exposed to some observed climate change whereas β is estimated by definition in the absence of climate change. However, although climate change has affected the predictors already, the inter-model differences in the predictors are nonetheless so large that the historical predictors are very good proxies for the same variables in the 1pct-bgc runs (Fig. 3 e-h; r^2 from 0.73-0.96) and even under 4×CO₂ forcing in the 1pct-cou runs (Fig. 3 a-d; r^2 from 0.65-0.97). Therefore, the observed historical predictors still constrain the same processes and therefore remain good predictors for β .

When exploiting the fitted relationship with observations, the resulting constrained β is 5–7 % larger and 20–29 % less uncertain than the multi-model mean and standard deviation (Table 4). At doubling of atmospheric CO₂, β increases from the multi-model mean of 0.92 ± 0.07 Pg C ppm⁻¹ to 0.97 ± 0.06 Pg C ppm⁻¹ after applying the constraint. Similarly, at quadrupling of atmospheric CO₂, β increases from the multi-model mean of 0.80 ± 0.07 Pg C ppm⁻¹ to 0.86 ± 0.05 Pg C ppm⁻¹ after applying the constraint. The reduction in uncertainty is dominated by how tightly the models follow the fitted relationship rather than by the precision of the observations themselves: at both target years, the fit uncertainty is roughly 5.6–6.2 times larger than the observational uncertainty, so even doubling the observational uncertainty in SSS, AMOC, and CUC would increase the constrained uncertainty in β by only about 4–5 %.

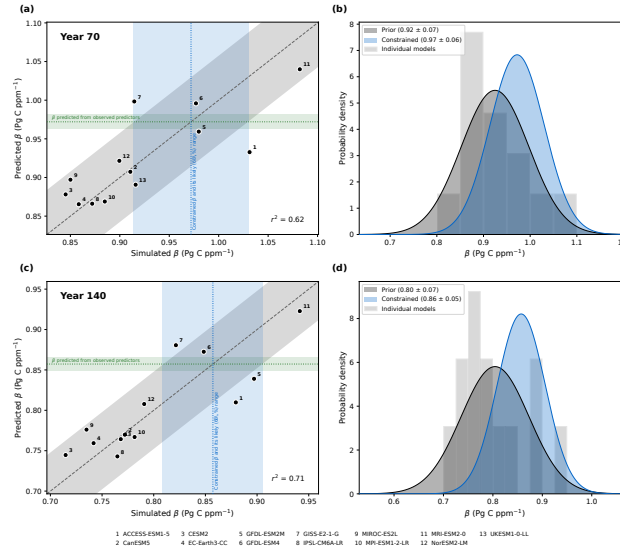


Figure 2. Constrained β using Southern Ocean frontal SSS, AMOC, and CUC as predictors, evaluated at simulation year 70 (a, b) and year 140 (c, d) following Terhaar et al. (2022). (a, c) Simulated versus predicted β for each CMIP6 model (black dots), where the predicted value is obtained from the multiple linear regression fit using the three simulated predictors in each ESM. The dashed line indicates the 1:1 relationship, and the grey shaded band shows the $\pm 1\sigma$ projection uncertainty. The green horizontal dotted line and shading show β predicted from the observed predictor values ($\pm 1\sigma$ observational uncertainty only), and the blue vertical dotted line and shading show the constrained β and its likely (68%) range, obtained by combining the observational and projection uncertainties. (b, d) Prior (grey) and constrained (blue) probability distributions of β , shown as Gaussian fits with mean $\pm 1\sigma$ given in the legend with grey bars showing the histogram of the individual (prior) model values.

475 This uncertainty reduction of 20–29% is substantial but only around half as large as the roughly twofold reduction in
uncertainty reported for the constraint on anthropogenic carbon uptake in the 2022 constraint (Terhaar et al., 2022). We see
three likely contributors to this difference. First, the predictors are less directly related to the analyzed simulations although
they are still very good proxies (Fig. 3). Second, unlike Terhaar et al. (2022), the uncertainty estimate used here explicitly
accounts for the degrees of freedom of the regression and for the distance of the observed predictors from the centre of the
model ensemble (Sect. 2.3), both of which lead to a higher $\sigma_{\text{model}}(t)$ relative to a simple residual standard deviation. Together,
480 these corrections increase $\sigma_{\text{total}}(t)$ by about 34% at both target years relative to the uncorrected estimate here. Third, this
effect is amplified compared to Terhaar et al. (2022) by the smaller ensemble used here: only 13 ESMs provide the diagnostics
required for β , compared to 17 ESMs in the 2022 constraint that provided historical carbon uptake, and the degrees-of-freedom
correction alone scales as $\sqrt{N/(N-p-1)}$, which is larger for $N = 13$ (a factor of 1.20) than for $N = 17$ (a factor of 1.14)
485 at $p = 3$ predictors. In addition, the larger number of models may also have reduced the prediction interval. Nonetheless, this
more complete treatment of the prediction-interval uncertainty gives us a more accurate estimate of how much an emergent

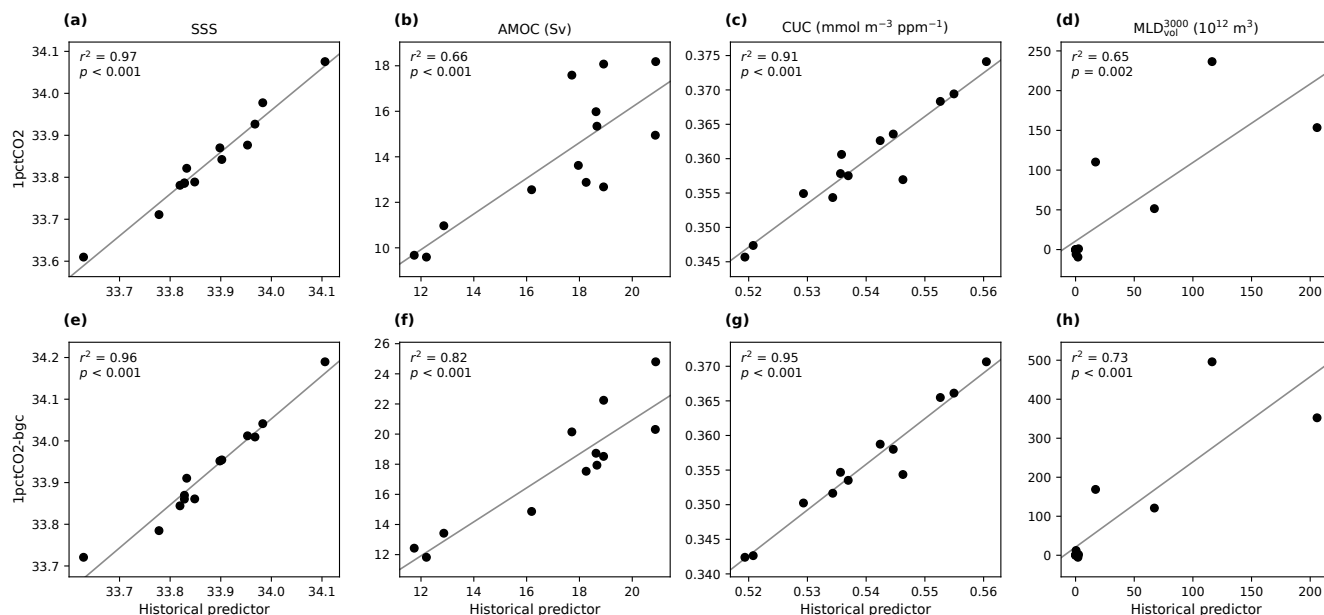


Figure 3. The four predictors (SSS, AMOC, CUC, and $\text{MLD}_{\text{vol}}^{3000}$), averaged over the $1\% \text{ yr}^{-1}$ CO_2 simulations with the fully coupled configuration (1pctCO_2 ; **a–d**) and the biogeochemically coupled one ($1\text{pctCO}_2\text{-bgc}$; **e–h**), against their historical values. The drift has been removed from the 1pctCO_2 and $1\text{pctCO}_2\text{-bgc}$ simulations with an Enting spline of 100-year cut-off period. Each dot represents one model and the grey line is the least-squares fit. The coefficient of determination and the two-sided p value of the correlation are given in each panel. The values on the y-axes are listed in Table S2 and those on the x-axes, the predictors themselves, in Table 3.

constraint of this kind can be expected to narrow the uncertainty in a quantity such as β and should be used in future constraints using multiple variables.

Although the reduction in uncertainty appears more substantial than the shift in the mean, the individual models spread widely around the constrained estimate, and they do so asymmetrically. At year 70, 7 of the 13 models (54%) fall below the lower bound of the 1σ constrained range and 2 (15%) below the lower bound of the 2σ range; at year 140 it is 8 (62%) and 3 (23%). One model, MRI-ESM2-0, lies above the 1σ range at both years. Three models lie closest to the constrained estimate at both target years: GFDL-ESM4, GFDL-ESM2M, and GISS-E2-1-G. These are also among the models whose combined SSS, AMOC and CUC values sit closest to the observations (Table 3). How closely a model matches the observations can be summarised by combining its three predictor deviations in quadrature, each expressed in units of the ensemble standard deviation of that predictor. On that measure five models stand apart from the rest: GFDL-ESM4 (0.9), GFDL-ESM2M (1.3), MRI-ESM2-0 (1.5), GISS-E2-1-G (1.6) and ACCESS-ESM1-5 (1.9), while the remaining eight lie between 2.0 and 3.6. GFDL-ESM4 is closest and is also the most evenly placed: none of its three predictors departs from the observed value by more than 0.7 ensemble standard deviations, whereas the others each have one predictor that does: the SSS in GFDL-ESM2M and ACCESS-ESM1-5, the CUC in MRI-ESM2-0, and the AMOC in GISS-E2-1-G.



Table 4. Feedback parameters at years 70 and 140 of the 1 % yr⁻¹ CO₂ simulation, before and after applying the emergent constraint. The prior is the mean and standard deviation across the 13 ESMs for β and γ , and across the 12 ESMs with a radiatively coupled simulation (all except GFDL-ESM4) for γ_{nat} and γ_{ant} . Constrained values are given for the 3-predictor constraint (SSS, AMOC, CUC) and for the 4-predictor constraint, which adds MLD_{vol}³⁰⁰⁰. Uncertainties are one standard deviation.

Parameter	Year	Prior	3 predictors		4 predictors	
			Constrained	r^2	Constrained	r^2
β (Pg C ppm ⁻¹)	70	0.925±0.073	0.972±0.058	0.62	0.961±0.048	0.78
	140	0.804±0.069	0.857±0.049	0.71	0.847±0.037	0.86
γ (Pg C °C ⁻¹)	70	-8.26±2.39	-10.35±2.31	0.43	-10.15±2.36	0.48
	140	-19.06±6.23	-25.07±3.93	0.77	-24.46±3.58	0.83
γ_{nat} (Pg C °C ⁻¹)	70	-3.72±1.07	-4.25±1.33	0.13	-4.32±1.42	0.14
	140	-5.43±2.10	-6.92±2.44	0.24	-6.89±2.63	0.24
γ_{ant} (Pg C °C ⁻¹)	70	-4.58±2.03	-6.55±1.85	0.53	-6.27±1.81	0.61
	140	-13.51±5.71	-19.15±3.75	0.76	-18.50±3.53	0.82

For a different model ensemble, the same constraint might lead to a different relative change in the best estimate. For example, the constrained estimate for the historical carbon uptake in Terhaar et al. (2022) was around 10% higher than the multi-model mean, whereas the change in β here is 4 to 5%. The relative change is largely dependent on the prior model ensemble. If the model ensemble is dominated by models with high ventilation, the constraint would even result in a smaller β than the multi-model mean of that ensemble. In contrast to the relative adjustment, the constrained β itself should not depend on the ensemble. The coefficients a , b , c , and d in Eq. (1) express the physical link between the predictors and β and should therefore not be model dependent, so a different ensemble that samples the same relationship should recover the same coefficients and, evaluated with the same observations, the same constrained β (Terhaar et al., 2022). In practice, a finite ensemble constrains these coefficients only approximately, which is reflected in the uncertainty of the constrained estimate.

3.2 A fourth indicator for open-ocean deep convection removes a bias and further reduces the uncertainty of β

After constraining β with a previously identified constraint (Terhaar et al., 2022) and corroborating that existing constraint, we go a step further by adding another predictor accounting for remaining biases. Comparing simulated and predicted β (based on SSS, AMOC, and CUC) shows that ACCESS-ESM1-5 and GFDL-ESM2M consistently have simulated β being above the predicted β (Fig. 2). These are also two of the models with the most vigorous open-ocean deep convection to depth deeper than 3000 m (Figure S4). Thus, this extremely deep convection, which is not yet captured by the predictors that account for carbonate chemistry, the AMOC strength that is mostly related to deep-water formation above 3000 m, and the formation of



mode and intermediate water in the Southern Ocean that is also mostly above 3000 m, may be the reason why GFDL-ESM2M and ACCESS-ESM1-5 simulate a higher β than predicted from SSS, AMOC, and CUC.

Here, we present two lines of evidence that support that the deep mixing below 3000 m in models, quantified by MLD_{vol}^{3000} , indeed enhances the uptake of anthropogenic carbon and hence β . First, the remaining inter-model uncertainty after accounting for biases in SSS, AMOC, and CUC is related to the MLD_{vol}^{3000} (Fig. 4 a). However, this relationship is largely driven by two models with strong deep convection, GFDL-ESM2M and ACCESS-ESM1-5, so we test how far it depends on any individual one. Removing GFDL-ESM2M or ACCESS-ESM1-5 leaves the fitted MLD_{vol}^{3000} coefficient at year 140 within -13 to $+8\%$ of its value across all 13 ESMs, and neither is exceptional. Across all 13 leave-one-out refits the coefficient stays within -13 to $+10\%$. Adding MLD_{vol}^{3000} also improves the prediction of a held-out model, which a predictor fitting only noise could not do. Removing both models together does collapse the relationship, but this is a loss of range rather than a change of sign: the ensemble then spans 0 to 67 instead of 0 to $206 \times 10^{12} \text{ m}^3$ and the standard error quadruples. The predictor therefore describes the models that convect below 3000 m collectively rather than either of these two in particular, though it remains a statistical indication. Therefore, our second line of evidence is more mechanistically based.

Therefore, our second line of evidence is more mechanistically based. Across models, we compared the relationship between MLD in a given month and cell against the air-sea CO_2 flux anomaly (Fig. 4 b). This comparison between MLD and air-sea CO_2 flux (Fig. 4 b) shows that it is indeed during times of deep-water convection that the highest carbon uptake occurs. With increasing MLD, carbon uptake already strongly increases when mixing exceeds 2000 m and further increases when it exceeds 3000 m (Fig. 4 b). As the AMOC and SSS represent already processes linked to water masses in the upper 3000 m, we decided to use the volume mixed below 3000 m as a fourth indicator to avoid double counting. Based on these lines of evidence, we add MLD_{vol}^{3000} as a fourth predictor. As such large-scale open-ocean deep convection is not observed in the real world (Sallée et al., 2021) and even rare polynyas in the Southern Ocean do not convect below 2000 m (Campbell et al., 2019; Gordon, 1978), the observed value is $0 \pm 0 \cdot 10^{12} \text{ m}^3$. Therefore, this additional predictor term in the constraint can be interpreted as a bias correction term for too much ocean carbon uptake driven by overly strong open-ocean deep convection in ESMs that is often simulated by ESMs (Heuzé, 2021) and by many ESMs in this ensemble (Figure S4). MLD_{vol}^{3000} is the only predictor whose observation-based estimate carries no uncertainty, so that it does not contribute to $\sigma_{obs}(t)$ in Eq. (3). This reflects that the observational constraint here is not a measured value with an error range but the absence of open-ocean convection below 3000 m in both observational products analysed (Sect. 2.4) and in previous exhaustive analyses of mixed layer depths (Sallée et al., 2021).

Previously, Roy et al. (2011) had already shown across an ensemble of older and coarser models with a substantially larger spread in simulated β and MLDs that large parts of β can be explained by the spread in MLDs. In the model ensemble used here, the spread in MLDs and β is substantially removed so that the overall MLD bias has been removed. However, the open-water deep convection bias is still present and our constraint now provides one way to account for it.

When fitting the relationship between all four predictors and β , the strength of the relationship increases, uncertainties of β are consequentially reduced, and constrained β gets smaller as the bias of unrealistic open-ocean deep convection is removed (Tab. 4 and S4, Fig. 5). The constrained β by all four predictors is $0.96 \pm 0.05 \text{ Pg C ppm}^{-1}$ in year 70 and $0.85 \pm 0.04 \text{ Pg C}$

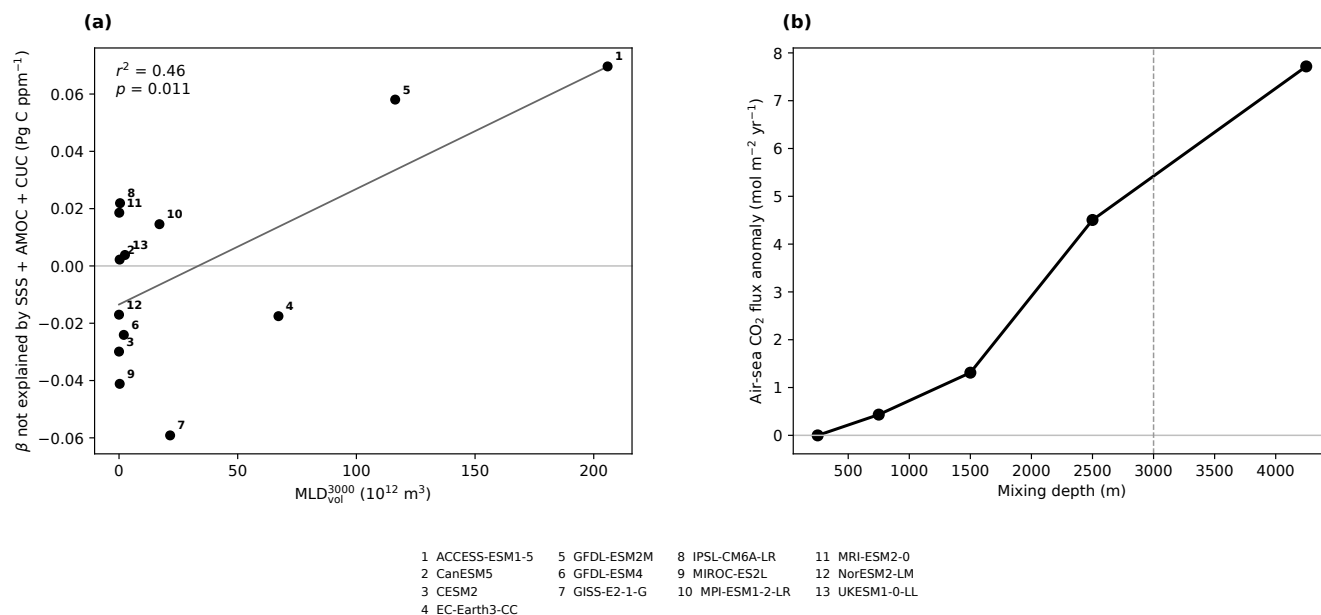


Figure 4. The contribution of the deep-mixing predictor to β , from two independent directions. **(a)** The part of β at year 140 that the other three predictors do not explain, against MLD_{vol}^{3000} from the historical simulation, that is the same value used as the predictor (Table 3). The dark grey line is the least-squares fit and the light grey line indicates zero. **(b)** The mean air-sea CO₂ flux anomaly as a function of the mixing depth of the same cell and month, over the last twenty years of the 1pctCO₂-bgc simulations and averaged across all models and cells. Anomalies are relative to a calendar-month climatology built from the months in which that cell mixes shallower than 1000 m, so that a cell which convects every winter is not compared with itself. The flux is averaged over the month and the two following months and over the cell and its eight neighbours, weighted by area, accounting for the fact that the water in the recently mixed cell may move and that air-sea CO₂ exchange is not immediate. The dashed line marks the 3000 m threshold used for the predictor.

ppm⁻¹ in year 140. The mean value is reduced by 0.01 Pg C ppm⁻¹ in each year, which quantifies the bias of the overly strong open-ocean deep convection. More importantly, the uncertainty around β is reduced by 19–25% when adding MLD_{vol}^{3000} compared to the constraint value without MLD_{vol}^{3000} . This reduction is a direct consequence of the increased r^2 for β from 0.62 to 0.78 in year 70 and from 0.71 to 0.86 in year 140 with MLD_{vol}^{3000} included. Adding this additional predictor has shown how an existing emergent constraint can miss a process and hence have a biased result, as earlier warned for by Sanderson et al. (2021) and as happened in previous cases, for example, in the case of ocean acidification where alkalinity changes with freshening had not been identified across CMIP5 models and changed the constrained results (Terhaar et al., 2020, 2021b). However, in this case, the new constrained estimate lays still within the previous one showing that the bias was small in this case. Consequently, the constrained estimate of historic and future ocean carbon uptake (Terhaar et al., 2022) may also be slightly smaller. However, GFDL-ESM2M was not part of that study and was one of the models with strong open-ocean deep

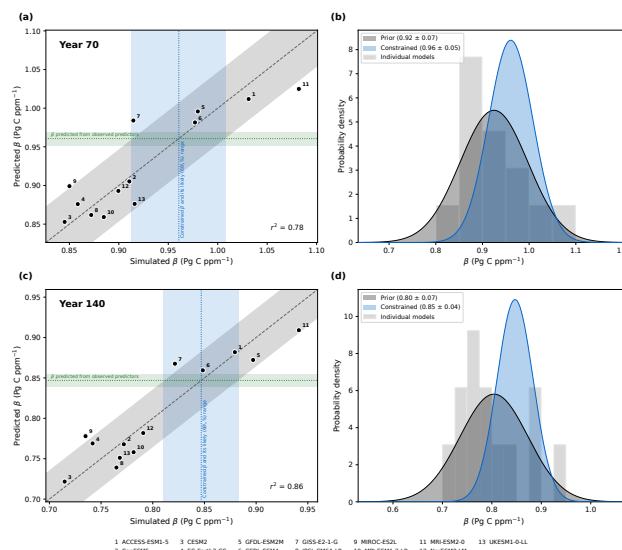


Figure 5. Constrained β using Southern Ocean frontal SSS, AMOC, CUC, and MLD_{vol}^{3000} as predictors, evaluated at simulation year 70 (a, b) and year 140 (c, d). (a, c) Simulated versus predicted β for each CMIP6 model (black dots), where the predicted value is obtained from the multiple linear regression fit using the four simulated predictors in each ESM. The dashed line indicates the 1:1 relationship, and the grey shaded band shows the $\pm 1\sigma$ projection uncertainty. The green horizontal dotted line and shading show β predicted from the observed predictor values ($\pm 1\sigma$ observational uncertainty only), and the blue vertical dotted line and shading show the constrained β and its likely (68%) range, obtained by combining the observational and projection uncertainties. (b, d) Prior (grey) and constrained (blue) probability distributions of β , shown as Gaussian fits with mean $\pm 1\sigma$ given in the legend with grey bars showing the histogram of the individual (prior) model values.

convection. Therefore, the bias in historic and future fluxes is likely even smaller as less bias from open-ocean deep convection had to be removed from the multi-model mean.

As the new predictor has removed a bias, we recommend using this constrained estimate, which is also why we show it in Fig. 1a. Compared to the prior multi-model mean and standard deviation, the constraint reduced uncertainties by 35-47% and increases β by 4–5% (Fig. 1). The reduced uncertainty is again only slightly smaller as the one for historical and future uptake (Terhaar et al., 2022). Compared to Terhaar et al. (2022), uncertainties of the constrained estimate had been increased because of the increase in predictor variables and the reduction in models with the required output and the resulting change in degrees of freedom, as well as the less direct link between historical predictors and the air–sea CO_2 flux in idealized 1pct simulations. The additional predictor has now again reduced some of this added uncertainty.

Adding the deep-mixing predictor also reorders the models by their distance from the observations, measured by combining the four predictor deviations in quadrature, each expressed in units of the ensemble standard deviation of that predictor. GFDL-ESM4 remains the closest (0.9) and is the only model whose four predictors are all within 0.7 standard deviations of the observed values. However, proximity in the predictors does not by itself place a model near the constrained estimate. MRI-



ESM2-0 (1.5) is the next closest, yet it sits above the constrained range at both years, because the deviations of its simulated predictors from the observed ones all point the same way: every predictor is biased high, so that nothing cancels and the biases accumulate. The same holds for MIROC-ES2L (2.0), whose SSS, AMOC and CUC are all biased low while its MLD_{vol}^{3000} is not biased at all. Conversely, GISS-E2-1-G (1.6) and GFDL-ESM2M (2.3) fall close to the constrained estimate, their negative salinity deviations being offset by positive deviations in the AMOC and in MLD_{vol}^{3000} . ACCESS-ESM1-5, among the closest models under the three-predictor constraint, becomes the most distant of all (3.8) once deep mixing is included, since its MLD_{vol}^{3000} exceeds the observation-based estimate of zero by 3.3 ensemble standard deviations.

4 Constraining ocean γ and relating it to β

The loss of carbon from the ocean or the reduction of carbon uptake per °C warming (γ) and the associated uncertainties increase over time with stronger warming and increasing atmospheric CO₂ (Figures 1b & 6a). After the beginning of the simulation when warming becomes large enough to quantify γ , the simulated multi-model mean γ decreases almost linearly to -19.1 Pg C °C⁻¹ at the end of the simulation period. Together with the multi-model mean, the absolute uncertainty quantified as the standard deviation across models also increases over time and reaches ± 6.2 Pg C °C⁻¹ at the end of the simulation at quadrupling of atmospheric CO₂, representing a relative uncertainty of 33%. This relative uncertainty is substantially larger than the relative uncertainty of β (9%). The larger relative uncertainty associated with γ is likely due to different atmospheric warming across the models, whereas the increase in atmospheric CO₂ that drives β is prescribed identically for all models. Although γ is normalized by temperature, the carbon loss per degree of warming increases with warming (Friedlingstein et al., 2006), so that the difference in warming across models contributes to the difference in γ . Furthermore γ quantifies climate change-driven changes on natural carbon and anthropogenic carbon (Figure 6a), whereas β only quantifies the additional anthropogenic carbon from increasing atmospheric CO₂. Thus, the inter model differences with respect to the loss of natural carbon from the ocean increase the uncertainties of γ but not the uncertainties of β .

The decrease in γ is mainly driven by reductions in anthropogenic carbon uptake per °C over time, whereas the loss of natural carbon per °C (γ_{nat}) decreases less (Figure 6a). The small decrease of γ_{nat} from -2.3 ± 1.0 Pg C °C⁻¹ in year 20 to -5.4 ± 2.1 Pg C °C⁻¹ at a quadrupling of atmospheric CO₂ in year 140 indicates that the natural carbon loss is enhanced with stronger warming over time, as circulation (Fig. S1,S2, and S4) and biology changes become larger. The small magnitude of γ_{nat} is a result of compensating effects of lower CO₂ solubility due to warmer surface ocean waters and a higher efficiency of the biological pump due to more sluggish ocean circulation (Schwinger et al., 2014; Arora et al., 2020). In contrast, the large suppression of anthropogenic carbon uptake per °C (γ_{ant}) from $+0.6 \pm 2.2$ Pg C °C⁻¹ in year 20 to -13.5 ± 5.7 Pg C °C⁻¹ at a quadrupling of atmospheric CO₂ in year 140, is driven by decreased transport of anthropogenic carbon to depth due to decreased deep ocean ventilation (Fig. S1,S3, and S4).

Not only does the strength of γ increase with the amount of carbon taken up by the ocean over time, but models with higher carbon uptake with increasing atmospheric CO₂ (β) also show the strongest reduction in anthropogenic carbon uptake due to climate change (Figure 6b). Over time, the relationship between γ and β becomes stronger and reaches an r^2 of 0.73 by the

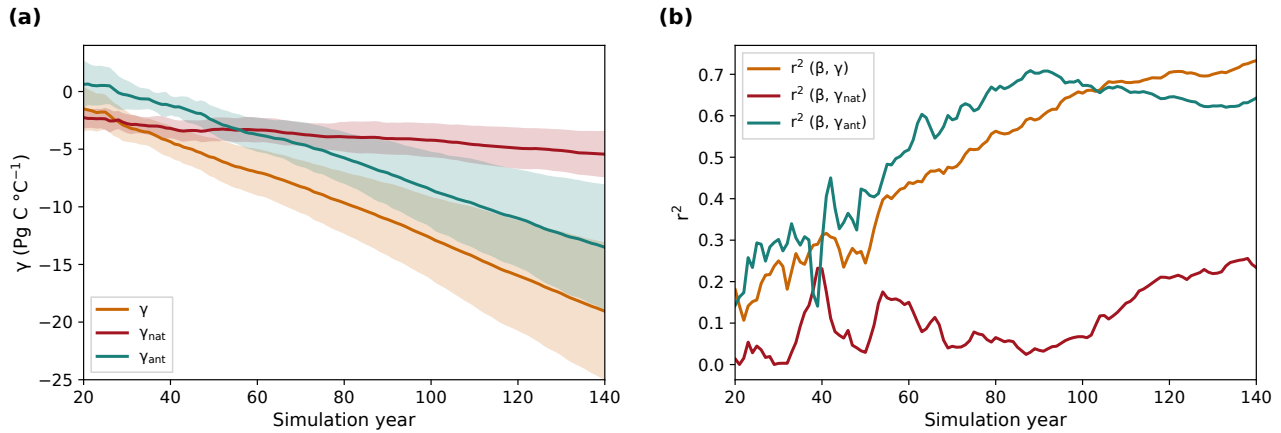


Figure 6. Evolution of γ , γ_{nat} , and γ_{ant} and of their correlation with β over time. **(a)** Multi-model mean (solid lines) and standard deviation (shading) of γ (orange), γ_{nat} (red) and γ_{ant} (teal) from the CMIP6 ESMs over 140 years of simulations with atmospheric CO₂ increasing by 1 % per year. **(b)** Temporal evolution of the coefficient of determination (r^2) between β and each of the three, showing how the relationship between the carbon–concentration and carbon–climate feedbacks changes over the course of the simulations. The first 20 years are not shown as the spread in γ is large (Figure 1) as warming is close to zero so that any correlation in these years remains purely random.

end of the simulation at a quadrupling of atmospheric CO₂ (Figure 6b). The relationship between γ and β is driven by the increasing suppression of anthropogenic carbon uptake due to climate change (γ_{ant}), which drives the increase of γ with time (Figure 6a) and not the loss of natural carbon due to climate change (γ_{nat}), as only a weak relationship ($r^2 = 0.23$) emerges between γ_{nat} and β over time (Figure 6b).

The link between β and γ emerges not only in the global averages but also regionally. The patterns of inter-model spread in β , γ , γ_{nat} , and γ_{ant} resemble each other closely (Figs. 7). At year 140 the spatial correlation between the inter-model spread in β and in γ is 0.75. It is carried by γ_{ant} , which correlates with β at 0.87, while γ_{nat} does so at 0.66. The main departures between the spread in β and in γ_{nat} are in upwelling regions along the west coast of South America and west of South Africa, where the models disagree substantially more about γ_{nat} than about β and γ_{ant} . Although the spatial correlation between the strength of inter-model differences between β and γ_{nat} is relatively high, this does not translate into a relationship of the globally integrated β and γ_{nat} , indicating that the size of inter-model difference is at similar places but that it is not the models with a high β that also have a high γ_{nat} and vice versa.

That γ_{ant} is the more strongly related of the two (Fig. 6b) suggests that the processes governing β also govern the suppression of anthropogenic carbon uptake under warming. As γ_{ant} is the change in β with global warming and β can be approximated by the four predictors identified here using equation (1), γ_{ant} should be determined by the changes of the predictors of β with warming. This means that models with high ventilation at present will also see a larger reduction of this ventilation with warming. Across the model ensemble used here, the predictors and their change with warming are indeed related (Fig. 8) apart from CUC, which does not represent ventilation and only has small changes with warming and experiences most of its changes

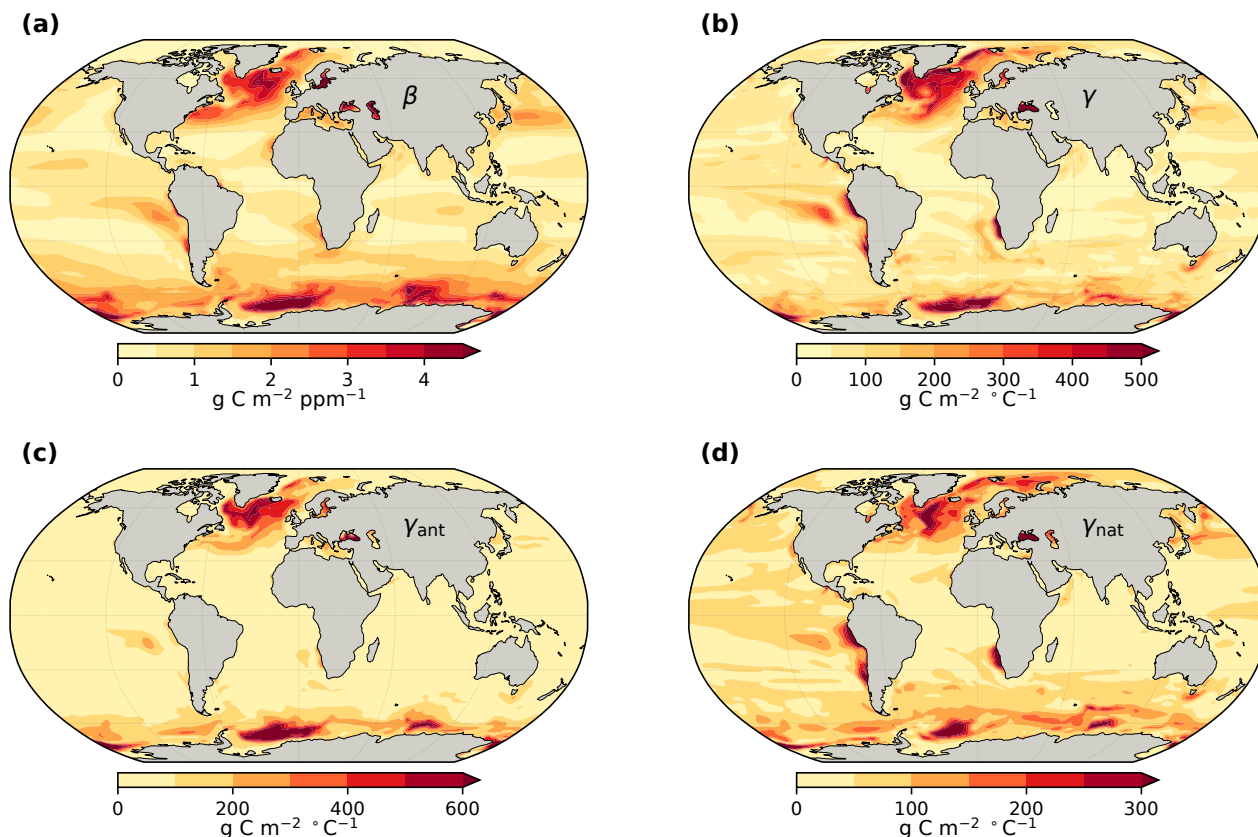


Figure 7. Inter-model spread of the simulated climate–carbon cycle feedbacks. Standard deviation across the 12 ESMs with a radiatively coupled simulation (all except GFDL-ESM4; Table 2) of (a) β , (b) γ , (c) γ_{ant} , and (d) γ_{nat} , evaluated cell by cell at year 140 of the 1 % yr⁻¹ CO₂ simulations. The global rather than the local warming is used, so that the maps show where the carbon response to a given global warming occurs rather than a local sensitivity, consistent with how the global parameters are defined.

with increasing atmospheric CO₂ (Fig. S2). While the relationships exist, some appear relatively weak because some historical predictors, such as MLD_{vol}³⁰⁰⁰ in GFDL-ESM2M, are within a particular phase of internal variability (Fig. S4) and hence do not perfectly represent the general state of that variable. When the mean value of each predictor over the biogeochemically coupled (1pct-bgc) simulation without climate change is calculated instead and the internal variability is removed, the relationships emerge even clearer (Fig. S13). Although few outliers in these relationships still exists even after removing internal variability, the predictors that constrain β also become informative for γ towards the end of the simulation (Table 4).

The strong relationship between γ and the AMOC, SSS, CUC, and MLD_{vol}³⁰⁰⁰, driven by models with strong ventilation at present experiencing a stronger reduction with warming, allows us to constrain γ (Figure 9 and 1b). When constraining γ with the same multi-linear relationship that was used for β (Equation 1, Table S6), the multi-model mean estimate of γ decreases from -8.3 ± 2.4 Pg C °C⁻¹ to -10.1 ± 2.4 Pg C °C⁻¹ at a doubling of atmospheric CO₂ ($r^2=0.48$) and from -19.1 ± 6.2

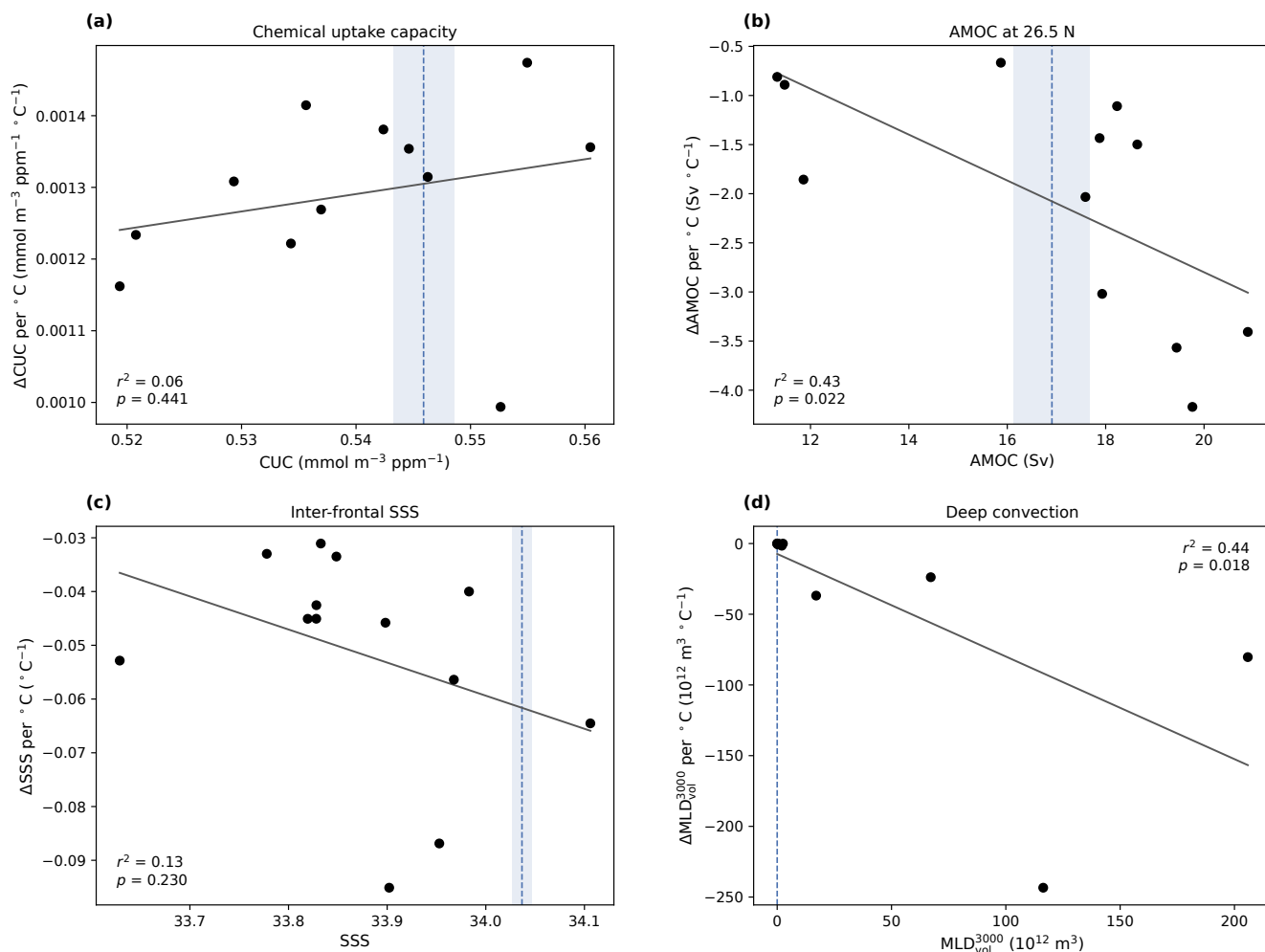


Figure 8. Present-day value of each predictor against its change under warming, across the ESMs. **(a)** the chemical uptake capacity (CUC), **(b)** the AMOC at 26.5° N, **(c)** the inter-frontal sea surface salinity (SSS), and **(d)** the volume of water mixed deeper than 3000 m (MLD_{vol}^{3000}). The change under warming is calculated as the difference between the fully coupled (1pctCO₂) and biogeochemically coupled (1pctCO₂-bgc) simulations, averaged over the last 30 years of the 140-year ramp and divided by the difference in their global mean warming. The dashed vertical line and the shading are the observation-based estimate and its one-standard-deviation uncertainty. Grey lines are least-squares fits, and the coefficient of determination and the two-sided p value of the correlation are given in each panel. Simulated changes in the predictors per warming for each model are listed in table S5.

Pg C °C⁻¹ to -24.5 ± 3.6 Pg C °C⁻¹ at a quadrupling of atmospheric CO₂ ($r^2=0.83$) (Table 4, Figure 9 and 1b). As such, the magnitude of γ increases by 23–28%, while the uncertainty of γ decreases by 1% at doubling of atmospheric CO₂ and by 42% at quadrupling. The increasing reduction of the uncertainty with increasing atmospheric CO₂ is a consequence of γ_{ant} dominating γ with increasing atmospheric CO₂ (Fig. 6a). When constraining γ_{ant} and γ_{nat} separately (Tab. 4), it becomes

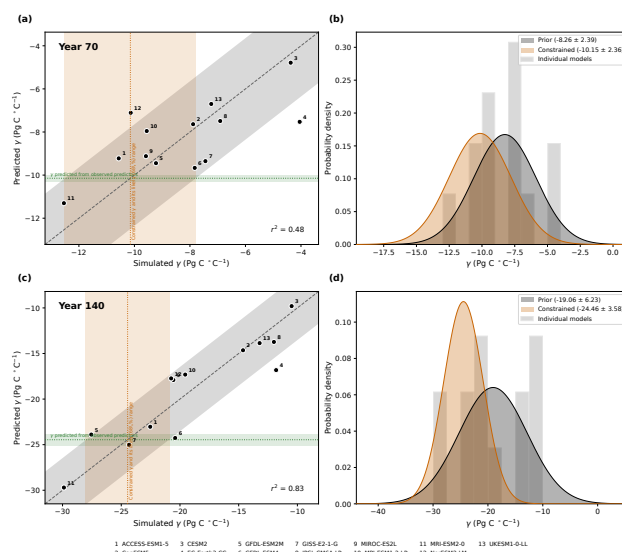


Figure 9. Constrained γ using Southern Ocean frontal SSS, AMOC, CUC, and MLD_{vol}^{3000} as predictors, evaluated at simulation year 70 (a, b) and year 140 (c, d). (a, c) Simulated versus predicted γ for each CMIP6 model (black dots), where the predicted value is obtained from the multiple linear regression fit using the four simulated predictors in each ESM. The dashed line indicates the 1:1 relationship, and the grey shaded band shows the $\pm 1\sigma$ projection uncertainty. The green horizontal dotted line and shading show γ predicted from the observed predictor values ($\pm 1\sigma$ observational uncertainty only), and the orange vertical dotted line and shading show the constrained γ and its likely (68 %) range, obtained by combining the observational and projection uncertainties. (b, d) Prior (grey) and constrained (orange) probability distributions of γ , shown as Gaussian fits with mean $\pm 1\sigma$ given in the legend with grey bars showing the histogram of the individual (prior) model values.

apparent that most of the changes in γ by the constraint are realized by better constraining γ_{ant} and not γ_{nat} . For γ_{nat} the predictors explain little of the inter-model spread ($r^2 = 0.14$ at doubling and 0.24 at quadrupling), so that the constraint shifts the best estimate but widens rather than narrows its uncertainty, as the uncertainty of the observation-based predictor estimates is added without a corresponding reduction in the model spread. This is consistent with γ_{nat} being governed by processes other than the ventilation and carbonate chemistry captured by the predictors.

5 Constraining β and γ regionally

As most of the deep-ocean ventilation and carbon uptake occurs in the North Atlantic Ocean (between 26.5° N and the Arctic Ocean limited by the Davis Strait, Fram Strait, and Barents Sea Opening) and Southern Ocean (limited by 30° S), the proxies for ventilation in each basin can also be used to constrain regional estimates of β and partly γ . Here, regional γ and β are the respective cumulative air–sea CO_2 fluxes divided by global changes in atmospheric CO_2 and atmospheric warming, respectively.



5.1 Southern Ocean and Southern hemisphere

Following and extending Terhaar et al. (2021a), we constrain β in the Southern Ocean south of 30°S using the SSS, as well as the CUC and MLD_{vol}^{3000} south of 30°S as predictors. The relationship between the predictors and β increases over time from an r^2 of 0.62 at a doubling of atmospheric CO_2 to an r^2 of 0.78 at a quadrupling of atmospheric CO_2 . When exploiting the relationship between SSS, MLD_{vol}^{3000} , and CUC and β in the Southern Ocean, the multi-model mean β of 0.39 ± 0.04 Pg C ppm^{-1} at a doubling of atmospheric CO_2 and 0.35 ± 0.04 Pg C ppm^{-1} at a quadrupling of atmospheric CO_2 becomes 0.42 ± 0.03 Pg C ppm^{-1} and 0.38 ± 0.02 Pg C ppm^{-1} , respectively.

The new predictor (MLD_{vol}^{3000}) substantially improves this regional constraint for β in the Southern Ocean. Without MLD_{vol}^{3000} as a predictor, the r^2 for β reduces to 0.38 at doubling of CO_2 and 0.49 at quadrupling. The improvement is even clearer in the uncertainty. Without MLD_{vol}^{3000} the constraint reduces the uncertainty of β by only 3% at doubling and 12% at quadrupling, against 20% and 38% when it is included. MLD_{vol}^{3000} has a large influence in the Southern Ocean as it is the ocean region where 93% of all MLD_{vol}^{3000} is simulated in the ESMs. Taking into account MLD_{vol}^{3000} thus removes a bias from unrealistic open-ocean deep convection in models (Heuzé, 2021), which largely exceed observed mixed layer depths (Sallée et al., 2021).

In contrast to β , the three predictors are not strongly related to γ in the Southern Ocean. The r^2 of the relationship between SSS, MLD_{vol}^{3000} , CUC, and γ is 0.10 at a doubling of atmospheric CO_2 and 0.41 at a quadrupling of atmospheric CO_2 . The constrained estimate does not reduce uncertainty but leads to a slight increase in magnitude of γ in the Southern Ocean of about 10% from -2.3 ± 1.4 Pg C $^{\circ}C^{-1}$ to -2.6 ± 1.7 Pg C $^{\circ}C^{-1}$ at doubling of CO_2 and from -6.1 ± 3.3 Pg C $^{\circ}C^{-1}$ to -6.8 ± 3.4 Pg C $^{\circ}C^{-1}$ at quadrupling of CO_2 . However, the three predictors can constrain γ_{ant} because the strength of the predictors is linked to their change over time (Fig. 8). When constraining, the Southern Ocean γ_{ant} changes from -1.3 ± 1.2 Pg C $^{\circ}C^{-1}$ to -2.2 ± 1.1 Pg C $^{\circ}C^{-1}$ at doubling of CO_2 ($r^2=0.56$) and from -4.5 ± 3.0 Pg C $^{\circ}C^{-1}$ to -5.9 ± 2.0 Pg C $^{\circ}C^{-1}$ at quadrupling of CO_2 ($r^2=0.77$), an increase in magnitude of 32-68% and a reduction in uncertainty of 9-31%. The reason that γ_{ant} can be constrained but γ cannot is the response of natural carbon to warming (γ_{nat}), which is not related to the predictors used here. Its inter-model spread is particularly large in the Southern Ocean (-1.7 ± 1.9 Pg C $^{\circ}C^{-1}$ at quadrupling of CO_2), so it contributes relatively more to the uncertainty in γ here than elsewhere.

5.2 North Atlantic

In the North Atlantic, we constrain β and γ with the AMOC and regional MLD_{vol}^{3000} and CUC building on Terhaar et al. (2022) and expanding it. The constraint based on AMOC and regional MLD_{vol}^{3000} and CUC changes the multi-model means of β of 0.108 ± 0.020 Pg C ppm^{-1} at a doubling of atmospheric CO_2 and 0.092 ± 0.021 Pg C ppm^{-1} at a quadrupling of atmospheric CO_2 to 0.111 ± 0.014 Pg C ppm^{-1} and 0.093 ± 0.017 Pg C ppm^{-1} , respectively. As only 7% of global MLD_{vol}^{3000} occurs in the North Atlantic, North Atlantic MLD_{vol}^{3000} contributes only very little to the constraints in that region.

Similarly, γ in the North Atlantic is constrained from -2.6 ± 1.0 Pg C $^{\circ}C^{-1}$ at a doubling of atmospheric CO_2 and -6.4 ± 3.1 Pg C $^{\circ}C^{-1}$ at a quadrupling of atmospheric CO_2 to -3.0 ± 1.1 Pg C $^{\circ}C^{-1}$ and -7.3 ± 3.2 Pg C $^{\circ}C^{-1}$, respectively. As the relationships are weak ($r^2=0.25-0.28$), the uncertainty is not reduced but increases even slightly due to the additional uncertainty from the



observation-based estimates of the predictors and the weak relationships. However, the constraint identifies a slight bias in γ by about 14–15%. As in the Southern Ocean, the link between the predictors and γ is driven by γ_{ant} , which is better linked to the predictors ($r^2=0.43\text{--}0.44$) as models with higher ventilation also show a stronger change in that ventilation with warming (Fig. 8). The constraint changes γ_{ant} in the North Atlantic from $-1.9\pm0.8 \text{ Pg C } ^\circ\text{C}^{-1}$ at a doubling of atmospheric CO_2 and $-4.8\pm2.8 \text{ Pg C } ^\circ\text{C}^{-1}$ at a quadrupling of atmospheric CO_2 to $-2.5\pm0.8 \text{ Pg C } ^\circ\text{C}^{-1}$ and $-5.9\pm2.7 \text{ Pg C } ^\circ\text{C}^{-1}$, respectively. This is no reduction in uncertainty but highlights that γ_{ant} is likely 24–29% larger than the multi-model mean.

6 Discussion

6.1 Robustness of the applied emergent constraints

A statistical relationship that is exploited with observations to constrain a simulated difficult-to-observe or unobservable quantities must build on robust mechanistic understanding that is verifiable and the relationship must hold across different model ensembles (Eyring et al., 2019; Hall et al., 2019; Sanderson et al., 2021). The mechanism behind the constraint for β by the AMOC, SSS, CUC, and $\text{MLD}_{\text{vol}}^{3000}$ (Equation 1) is based on previous knowledge about ocean ventilation (Broecker et al., 1979; Sarmiento et al., 1992; Orr et al., 2001; Caldeira and Duffy, 2000) and carbonate chemistry (Revelle and Suess, 1957; Broecker et al., 1979) and has been verified in detail and tested across different model ensembles (Terhaar et al., 2022). The added predictor, i.e., the ventilated volume below 3000 m by open-ocean deep convection ($\text{MLD}_{\text{vol}}^{3000}$), has been linked to the remaining inter-model differences in β after adjusting models for the biases in the remaining three predictors (Fig. 4a). Furthermore, we demonstrated that such deep mixing in a given cell leads to large positive anomalies in the ocean carbon uptake (Fig. 4b) and demonstrated that the mechanism of deep mixing that allows for more carbon uptake is indeed occurring in the models. In addition, the relationships were additionally tested by leave-one-out cross validation across the ensemble (Sect. 2.3, Tables S7 and S8) Thus, the emergent constraint for β can be assessed as robust and exploited with observations. The mechanism behind the constraint for γ by the AMOC, SSS, CUC, and $\text{MLD}_{\text{vol}}^{3000}$ is that models with high ventilation at present also simulate a strong reduction of that ventilation with warming in the future, which is a plausible physical mechanism and is indeed simulated across the model ensemble (Fig. 8 and S13). As γ is increasingly dominated by γ_{ant} , which is the change of β with warming, it is the change of the predictors of β with warming that also constrain γ_{ant} . As the predictors of β predict their own change with warming, these predictors can also constrain γ_{ant} and hence γ . We thus consider also this constraint as robust.

However, unknown or not-implemented processes might still affect the relationships for β and γ , as the addition of $\text{MLD}_{\text{vol}}^{3000}$ has slightly decreased the magnitude of constraint β and γ . Other examples of missing processes may be mesoscale or sub-mesoscale processes, which have been shown to affect the ventilation and ocean carbon sink globally and regionally (Lachkar et al., 2007, 2009; Dufour et al., 2015; Griffies et al., 2015; Couespel et al., 2024; Terhaar et al., 2019b; Ford et al., 2023; Mortenson et al., 2025; Salinas-Matus et al., 2025; Guo and Timmermans, 2024; Carnier Casaroli et al., 2026). So far, it is not clear yet if these processes have a substantial effect on the globally integrated ocean carbon sink and on longer time-scales as some studies find no difference in the net carbon uptake (Guo and Timmermans, 2024; Mortenson et al., 2025) whereas others find substantial changes (Couespel et al., 2024). However, when explicitly simulating mesoscale processes at higher



resolution, both the ocean carbon uptake change and the predictors used to constrain the global or regional uptake can change. In the Southern Ocean, Lachkar et al. (2009) show that increasing resolution results in a reduction in sea surface salinity and volume of ventilated waters and also a weaker CFC-11 uptake and hence weaker carbon uptake. In the Arctic, the effect goes in the opposite direction so that higher resolution leads to higher anthropogenic carbon storage and also higher surface densities, which are used to constrain that storage (Terhaar et al., 2020). In both cases the predictor moves in line with the predicted variable, even though resolution changes them in opposite directions, which suggests that resolving mesoscale processes shifts the indicators and the carbon uptake together rather than altering the relationship between them. However, only an ensemble of models that resolve mesoscale processes will be able to definitely show that the constraint does not change with resolution.

Another process that might affect the relationship and the constrained results are evolving river fluxes of carbon and nutrients over time (Dai et al., 2022; Lacroix et al., 2021; Gao et al., 2023; Richardson et al., 2025; Regnier et al., 2013, 2022) that may affect the carbon uptake regionally and globally (Lacroix et al., 2021; Gao et al., 2023; Terhaar et al., 2019a; Séférian et al., 2019; Regnier et al., 2022). Riverine carbon and nutrient fluxes can be altered by climate change, e.g., precipitation and temperature changes, and with increasing CO₂, which affects the primary production on land and hence the export of carbon and nutrients to the ocean. Only one of the ESMs from CMIP6, CNRM-ESM2-1 (Séférian et al., 2019), has implemented riverine carbon fluxes that change dynamically over time with changes on land. In the historical and SSP simulations, riverine carbon fluxes weaken over time reducing the ocean carbon content from 1850 to 1995 by around 20 Pg C. After 1995, the riverine carbon fluxes strongly strengthen and increase the interior ocean carbon content by 70-80 Pg C between 1995 and 2100 (Terhaar et al., 2022). As this additional carbon is partly outgassed and affects air-sea CO₂ fluxes (Lacroix et al., 2021; Gao et al., 2023; Terhaar et al., 2019a; Séférian et al., 2019; Regnier et al., 2022), it also affects β and γ . Due to this effect, we had not included CNRM-ESM2-1 in the constraint here. While adding CNRM-ESM2-1 to this study would change constrained estimates of β and γ only slightly but add additional uncertainty (β at quadrupling of atmospheric CO₂ of 0.84 ± 0.04 Pg C ppm⁻¹ with CNRM-ESM2-1 instead of 0.85 ± 0.04 Pg C ppm⁻¹ without CNRM-ESM2-1; γ at quadrupling of atmospheric CO₂ of -23.9 ± 4.4 Pg C °C⁻¹ with CNRM-ESM2-1 instead of -24.5 ± 3.6 Pg C °C⁻¹ without CNRM-ESM2-1), adding dynamic carbon and nutrient fluxes in each model may well change the simulated unconstrained β and γ , the fit of the relationship, and hence the constrained estimates of β and γ . To get a first idea of the effect of changing riverine carbon fluxes on β and γ in at least one model, we can compare the simulated β and γ by CNRM-ESM2-1 to the expected β and γ based on the simulated AMOC, SSS, CUC, and MLD_{vol}³⁰⁰⁰ in that model and the fitted relationship across the ESMs used here (Table 1). The simulated β is with 0.70 Pg C ppm⁻¹ substantially smaller than the reconstructed β of 0.78 Pg C ppm⁻¹. Similarly the magnitude of the simulated γ is with -9.3 Pg C °C⁻¹ smaller than the magnitude of the reconstructed γ of -17.5 Pg C °C⁻¹ based on the fitted relationship and simulated AMOC, SSS, CUC, and MLD_{vol}³⁰⁰⁰. Although these differences between expected and simulated β and γ may also be due to other processes, the size of the effect of rivers in the historical simulation in CNRM-ESM2-1 suggests that the changes in riverine fluxes are indeed the main driver of the here quantified differences between simulated and predicted β and γ . Thus, riverine changes in carbon fluxes thus decrease the magnitudes of β and γ in CNRM-ESM2-1. However, it is unclear if sign and magnitude is robust as we only have one model estimate yet and observation-based estimates of changes in riverine fluxes and burial remain still to uncertain to determine the effect of changing rivers on the carbon sink (Regnier et al., 2022).



Yet, even the potential effect of dynamic riverine fluxes of carbon and nutrients on the constrained and unconstrained estimates of β and γ highlights again that an emergent constraint can adjust biases that affect the predictor and predicted variable and substantially improve simulated estimates but cannot account for additional processes that occur independently of the predictor variables. Examples of this effect have been identified earlier, for example in the Arctic Ocean, where a constraint identified biases in the main pathway of Arctic Ocean ventilation but could not account for the very small scale processes along the slopes along the shelf break that transport additional carbon into the Arctic Ocean (Terhaar et al., 2020, 2021b).

Another effect that is not present in the here constrained β is the skin-layer effect (Watson et al., 2020; Bellenger et al., 2023). The skin-layer of the ocean is an extremely thin layer (~ 1 mm), which is 0.1 to 0.2°C cooler than the underlying water (Donlon et al., 2002; Woolf et al., 2016; Watson et al., 2020). This temperature contributes to the estimate of the partial pressure of CO₂ and hence the estimate of the air-sea CO₂ fluxes. However, models do not simulate such a thin skin layer but represent the surface ocean as grid cells of depths of usually 1 to 10 m. The average temperature in these cells is also the sea surface temperature that is used to calculate the partial pressure of CO₂ and hence the estimate of the air-sea CO₂ fluxes. As this temperature is warmer than the overlying skin layer temperature, the air-sea CO₂ uptake is underestimated (Watson et al., 2020; Bellenger et al., 2023). As this effect is common across all models, it would increase β across the ensemble and an emergent constraint cannot account for it. Unfortunately, only one model has tested the effect of the skin temperature and shown that accounting for this skin layer, increases carbon uptake over the historical period by around 5%. Unfortunately, this estimate so far was only done for one model and the conditions over the historical period are substantially different to the idealized 1% simulations analyzed here, with higher atmospheric CO₂ than in the historical simulations and much stronger climate change. It remains therefore uncertain if this estimate can be transferred to β and γ quantitatively. However, if we assumed that the adjustment was 5% throughout the simulations, which is likely not the case, both γ and β would increase by 5% in magnitude.

Overall, our constraint identified and corrected an independent important model bias, i.e., models underestimate the ocean ventilation and CUC and hence the magnitude of β and γ . However, it cannot correct for biases not represented in the models yet, which are, for example, dynamic riverine fluxes that may increase or decrease β and γ and the skin layer effect, which will likely increase both β and γ in magnitude. However, so far it remains highly uncertain how these riverine fluxes will change in the future (Gao et al., 2023; Richardson et al., 2025) and how these changes will affect the ocean carbon sink (Lacroix et al., 2021; Gao et al., 2023; Terhaar et al., 2019a; Séférian et al., 2019) so that it remains impossible to adequately quantify the effect of changing riverine fluxes on β and γ yet. A better quantification of the effect of these fluxes could be achieved by following the example of CNRM-ESM2-1 and to implement dynamic changes in riverine carbon and nutrient fluxes in more earth system models. Similarly, we cannot yet quantify the bias in β and γ from not accounting for the skin layer in models without having models implementing these processes. Thus, our constraint estimates of β and γ are the best estimates given the processes included in state-of-the-art models and directly comparable to previous assessments of climate-carbon cycle feedbacks (Arora et al., 2013, 2020; Friedlingstein et al., 2006), which also relied on models without dynamically changing riverine fluxes of carbon and nutrients and without the skin layer effect. However, the "real-world" β and γ in the ocean likely differ from these and it remains critical to quantify these biases to better understand the future evolution of the ocean carbon sink and the TCRE.



6.2 Potential of constraining ocean β with observation-based estimates of past ocean carbon uptake

In addition to the here-applied mechanistic constraint for β and γ based on indicators for ocean ventilation and surface ocean carbonate chemistry (Terhaar et al., 2021a, 2022), β and γ may also be constrainable by observation-based estimates of past and present carbon uptake as ESMs with a high past carbon uptake also simulate a high future carbon uptake (Terhaar et al., 2021a, 2022; Kessler and Tjiputra, 2016). The link between simulated past and future carbon uptake remains strong because biases in the AMOC, SSS, and CUC persist even under increasing atmospheric CO_2 and climate change (Figure 3). Compared to observation-based estimates of air-sea CO_2 fluxes (Fay et al., 2021; Friedlingstein et al., 2026b) or interior anthropogenic ocean carbon changes (Sabine et al., 2004; Gruber et al., 2019; Müller et al., 2023; Lauvset et al., 2016), simulated carbon uptake in ESMs and global ocean biogeochemical models forced with historic atmospheric conditions from reanalysis data have been shown to be consistently lower by 8-25% (Terhaar et al., 2024, 2022). How well the here used model ensemble simulates past carbon uptake will determine the magnitude of the resulting estimated upward or downward correction of both climate-carbon cycle feedback parameters.

However, observation-based estimates can not yet robustly constrain the future ocean carbon sink and β and γ because their uncertainties remain relatively large and because they are not directly comparable. Observation-based estimates of the air-sea CO_2 fluxes, for example, quantify the total air-sea CO_2 flux and not the change in the air-sea CO_2 flux with climate change and increasing atmospheric CO_2 , which is quantified by ESMs. To use the change in the air-sea CO_2 flux with climate change and increasing atmospheric CO_2 over the historical period as a predictor for the future ocean carbon sink or β , the pre-industrial air-sea CO_2 flux has to be subtracted from the observation-based estimates of the total air-sea CO_2 flux. This flux is mainly driven by the imbalance of carbon and alkalinity influx from rivers and burial in ocean sediments (Sarmiento and Sundquist, 1992; Regnier et al., 2022). Moreover, this flux may have changed since pre-industrial times (Lacroix et al., 2021; Regnier et al., 2022), and this change also affects the carbon sink. Unfortunately, estimates of these pre-industrial outgassing due to the difference between riverine input and burial of carbon and alkalinity, as well as their changes over the last decades and centuries, remain uncertain and differ across estimates (Aumont et al., 2001; Regnier et al., 2013, 2022; Sarmiento et al., 1992; Lacroix et al., 2020; Jacobson et al., 2007; Planchat et al., 2025; Resplandy et al., 2018). Additionally, it remains uncertain how the effect of the skin layer temperature affects the estimates of the carbon sink based on measurements from the Surface Ocean CO_2 Atlas (SOCAT), possibly contributing to an underestimation of the ocean carbon sink by observation-based estimates (Friedlingstein et al., 2026a). In addition, there are diurnal warm layers, for which, under conditions of high insolation and low wind speed, waters close to the surface are heated relative to underlying waters and which are not captured by measurements below (Fairall et al., 1996; Ye et al., 2023; Ford et al., 2024). The uncertainty of the pre-industrial adjustment combined with non-negligible uncertainty and inter-product differences of the observation-based estimates of the air-sea CO_2 flux (Fay et al., 2021; Friedlingstein et al., 2026b) result in uncertainties of the observation-based estimates of the air-sea CO_2 flux that are similar or larger than inter-model differences (Terhaar et al., 2022) and hence do not constrain the future ocean carbon sink or β yet.



820 Similar to observation-based estimates of air-sea CO₂ fluxes, observation-based estimates of historical interior ocean anthropogenic carbon accumulation (Sabine et al., 2004; Gruber et al., 2019; Müller et al., 2023) are not directly comparable to simulated increases over the historical simulations and their uncertainties also remain relatively high. The interior ocean estimates quantify the change in anthropogenic carbon in the absence of warming and do not account for changes in anthropogenic and natural carbon with global warming (Joos et al., 1999; Bunsen et al., 2024; McNeil and Matear, 2013). However, when
825 constraining β and γ with simulated historical changes in ocean interior carbon, these simulated changes include differences due to atmospheric CO₂ increases and climate change. Thus, the loss of natural and anthropogenic carbon with climate change would have to be subtracted from the anthropogenic carbon interior ocean changes. However, estimates of these changes remain highly uncertain (Bunsen et al., 2024; Joos et al., 1999; McNeil and Matear, 2013). One reason is that global biogeochemical models forced with atmospheric reanalysis data are often too warm as they are spun up with climate conditions from the
830 second half of the 20th century (Huguenin et al., 2022). Thus, they underestimate the observed transient warming (Huguenin et al., 2022; Hollitzer et al., 2024; Takano et al., 2023) and associated oxygen loss (Takano et al., 2023; Hollitzer et al., 2024) and potentially also the loss of carbon with climate change. Ideally, global biogeochemical models forced with reanalysis data would be spun up as suggested by Huguenin et al. (2022) to better quantify the loss of carbon with climate change. Alternatively, one could use ESMs to quantify the loss of natural carbon as they simulate the observed long-term loss of carbon
835 with climate and the associated deoxygenation more accurately as they are spun-up under pre-industrial conditions (Hollitzer et al., 2024). However, ESMs simulate their own internal climate variability and might not capture the observed one. As the variability can be high on decadal timescales (Bunsen et al., 2024; Terhaar et al., 2024; DeVries et al., 2023), ESM-derived estimates of natural carbon loss due to climate change on decadal timescales might lead to over- or underestimation of that loss and increase the uncertainties. Another option to make simulated and observation-based estimates comparable would be
840 to simulate the historical atmospheric CO₂ increase without associated climate change in future CMIP versions, similar to the '1pct-BGC' simulations. These simulations have been done by a limited number of ESMs in CMIP6. Thus, the simulated interior ocean change in anthropogenic carbon could be related to β and directly be exploited with observation-based estimates of the interior ocean anthropogenic carbon changes (Sabine et al., 2004; Gruber et al., 2019; Müller et al., 2023). This second approach is likely the most promising although uncertainties of observation-based estimates of anthropogenic carbon over past
845 decades are still larger than inter-model differences (Terhaar et al., 2022; Gruber et al., 2019; Müller et al., 2023) and strong decadal variability of anthropogenic carbon (Müller et al., 2023) might limit uncertainty reductions of the simulated future carbon sink or β with observation-based interior ocean anthropogenic carbon estimates.

Overall, observation-based estimates of air-sea CO₂ fluxes and interior ocean anthropogenic carbon estimates thus suggest an upward adjustment of the magnitude of β and γ . However, uncertainties of these estimates themselves and necessary but
850 uncertain adjustments to the observation-based estimates for natural outgassing of carbon during pre-industrial times or loss of carbon with climate change make a constraint based on these observation-based estimates that reduces the uncertainty of β and γ so far not possible. On regional scales, the adjustment of the observation-based estimates becomes even more uncertain (Aumont et al., 2001; Regnier et al., 2013, 2022; Sarmiento et al., 1992; Lacroix et al., 2020; Jacobson et al., 2007; Planchat



et al., 2025; Resplandy et al., 2018; Joos et al., 1999; Bunsen et al., 2024; McNeil and Matear, 2013) and further complicates
855 constraining regional β or γ .

6.3 The enhanced coupling of β and γ and its implications

The increasing relation between β and γ over time (Figure 6b) acts like a negative feedback and keeps the impact of the biases in the AMOC, SSS, CUC, and MLD_{vol}^{3000} on the net carbon sink smaller. The stronger suppression of anthropogenic carbon uptake with climate change (larger magnitude of γ) in models with higher carbon uptake (larger β) reduces the inter-model
860 differences in the ocean carbon sink that arise due to inter-model differences in rising atmospheric CO_2 -driven carbon uptake. Treating β and γ as independent variables or only constraining one thus leads to inaccurate adjustments of the overall carbon sink at higher atmospheric CO_2 .

This coupling also influences the carbon sink over time. Observations of the AMOC, SSS, CUC, and MLD_{vol}^{3000} combined with the here identified constraints indicate that models on average initially underestimate the ocean carbon sink and once
865 sufficient carbon has been taken up and climate change has become stronger, subsequently underestimate the suppression of anthropogenic carbon uptake. When constraining the ocean carbon sink in each year individually as done here and in Terhaar et al. (2022), this change over time is explicitly accounted for.

While β and the suppression of anthropogenic carbon uptake with climate change (γ_{ant}) are connected, the loss of natural carbon with climate change (γ_{nat}) is only weakly ($r^2=0.23$) related to both variables and seems to be driven by other processes.
870 Although γ_{nat} has limited influence on γ at high atmospheric CO_2 , it still has relatively large impact at atmospheric CO_2 levels such as today (Figure 6a). To further reduce uncertainties of γ , especially in the past and at present, it would thus be important to be able to further constrain the uncertainties of γ_{nat} .

6.4 Importance of the Southern Ocean for the global ocean carbon sink

Within the framework of this multi-variate constraint, the Southern Ocean has the most prominent role. Two of the four
875 predictors are essentially located in the Southern Ocean, the SSS and 93% of MLD_{vol}^{3000} . In addition, the Southern Ocean also contributes to the globally averaged CUC. This concentration of predictors in the Southern Ocean is expected as it is the region with the highest carbon uptake and largest inter-model differences in the carbon sink across all major ocean basins (Frölicher et al., 2015). As a consequence, the Southern Ocean stands out in its contribution to changes in global β . The identified bias of the regional β in the Southern Ocean at a quadrupling of CO_2 is with $0.027 \text{ Pg C ppm}^{-1}$ around 60% of the global bias of
880 $0.043 \text{ Pg C ppm}^{-1}$. Similarly, the uncertainty reduction of β in the Southern Ocean ($0.015 \text{ Pg C ppm}^{-1}$) is almost half as large as the global reduction in uncertainty ($0.032 \text{ Pg C ppm}^{-1}$). While the constraint presented here helps to adjust for these biases and reduce uncertainties, it would eventually be ideal if ESMs would better simulate the circulation in the Southern Ocean and also its biogeochemistry.



6.5 Implications for TCRE and remaining carbon budgets

885 The changes in β and γ due to the constraint impact the overall ocean carbon sink and thus also the TCRE. The more carbon taken out of the atmosphere per emitted CO_2 , the less remains in the atmosphere and the less it contributes to global warming. Thus, a higher β reduces the TCRE and a lower γ increases the TCRE. Quantitatively, the increase of β (Figure 1a) corresponds to additional ocean carbon uptake of carbon of 10 Pg C at a doubling of atmospheric CO_2 and 36 Pg C at quadrupling. At the same time, the increase in the magnitude of γ (Figure 1b) corresponds to an additional carbon loss of 4 Pg C and 25 Pg C, respectively. Thus overall, the carbon sink increases by 6 Pg C at a doubling of atmospheric CO_2 and by 11 Pg C at a quadrupling of atmospheric CO_2 after applying the constraint, resulting in only a tiny decrease of the TCRE leaving it essentially unchanged. Although the net change is small, the uncertainty stemming from ocean carbon uptake is substantially reduced and changes of β and γ individually as identified by our constraint are large as well.

7 Conclusions

895 In this study, we constrained the climate-carbon cycle feedback parameters β and γ using observations of the Atlantic Meridional Overturning Circulation at 26.5°N , the average sea surface salinity in the Southern Ocean between the polar and subtropical fronts, the globally-averaged chemical uptake capacity of the surface ocean, and the volume of waters below 3000 m that is mixed with surface waters in models through open-ocean deep convection. The first three indicators for ocean ventilation and capacity of the ocean to absorb carbon from the atmosphere have previously been identified as variables that can constrain the historical, present, and future ocean carbon sink under different scenarios over the 21st century (Terhaar et al., 2022). By adding the volume of mixed waters below 3000 m, we have corrected β and γ for biases in unrealistic open-ocean deep convection (Heuzé, 2021) and further reduced uncertainties. Thus, accurately simulating the Southern Ocean, North Atlantic, and the carbonate chemistry should be a continued priority of the modeling community.

We further showed that the capacity of the ocean to take up carbon with increasing atmospheric CO_2 is also related to how much carbon uptake is suppressed with climate change. Over time, ventilation is more strongly reduced in models that have a strong present-day ventilation. Hence, the more carbon is taken up with increasing atmospheric CO_2 (larger β), the more carbon is lost due to climate change (larger magnitude of γ). Due to the link between β and γ in the ocean, it is important that β and γ are not considered alone and that a constraint or bias correction of one variable also affects the other variable. This link was especially important as both β and γ were biased in the CMIP6 ensemble considered here. Moreover, the relative bias in γ was an order of magnitude larger than that in β , so that both biases affect the overall carbon sink in almost similar magnitude, despite β contributing much more to the magnitude of the ocean carbon sink than γ . While the biases being of similar magnitude, the increase in the ocean carbon sink from a larger β is still slightly larger than the carbon loss with climate change from a larger magnitude of γ , so that overall carbon sink with increasing atmospheric CO_2 and climate change (β and γ combined) increases still slightly after applying the constraint.



915 *Code availability.* The core functions used for the analyses in this study are available at Zenodo, <https://doi.org/10.5281/zenodo.22884358> (Terhaar, 2026). All NASA-GISS ModelE components are open source and available at <https://www.giss.nasa.gov/tools/modelE/>. GFDL-ESM2M is described by Dunne et al. (2012) and Dunne et al. (2013); its ocean component (MOM5) is publicly available at <https://github.com/mom-ocean/MOM5>

Data availability. The Earth system model output used in this study is available via the Earth System Grid Federation (<https://esgf-node.ipsl.upmc.fr/project> ipsl/, last access: 30 November 2025). For further information, please see Table 1. The output from GISS-E2-1-G and GFDL-ESM2M that is necessary for analyses, figures, and tables will be made available before publication. The observation-based products used to constrain the predictors are publicly available: World Ocean Atlas 2023 (Reagan et al., 2024a, b; Locarnini et al., 2024), EN4 (Good et al., 2013), IAP (Cheng et al., 2020, 2024), GLODAPv2 (Lauvset et al., 2016), OceanSODA-ETHZ (Gregor and Gruber, 2021), CMEMS-LSCE-FFNN (Chau et al., 2024), JMA-MLR (Iida et al., 2021), and the RAPID array (Cunningham et al., 2007; McCarthy et al., 2015; Moat et al., 2025).

925 *Author contributions.* J.T. conceptualized and conceived the study, developed the methodology, performed the formal analysis, and prepared the original manuscript draft and visualisations. A.W. set up the GFDL-ESM2M model configuration on the Swiss National Supercomputing Centre (CSCS) and H.H. performed the model simulations of GFDL-ESM2M. P.L. and A.R. performed the model simulations of GISS-E2-1-G. All co-authors contributed to reviewing and editing the manuscript and are listed alphabetically.

Competing interests. No competing interests are present

930 *Acknowledgements.* J.T. and H.H. have been supported by the Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung (grant no. PZ00P2_209044 (ArcticECO)). This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s1323 on Alps. P.F. acknowledge support from the CALIPSO project supported by Schmidt Sciences. J.S. was supported by the Research Council of Norway through the project NorESM4CMIP7 (grant no. 352204). R.S. acknowledges support from Agence Nationale de la Recherche - France 2030 as part of the PEPR TRACCS programme under grant number ANR-22-EXTR-0008 and ANR-22-EXTR-0009. A.W. acknowledges support by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number MB22.00069. All authors thank the WCRP Safe Landing Climate Light House activity and ESMO that provide support for the WCRP TCRE assessment, which led to this assessment here. Claude (Anthropic) was used to restructure and document the analysis code, making it clearer and easier to use, and to assist with language editing. The core analyses, the methodological choices, and the interpretation are the authors' own.



940 References

- Akaike, H.: A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, 19, 716–723, <https://doi.org/10.1109/TAC.1974.1100705>, 1974.
- 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.
- 945 Arora, V. K., Boer, G. J., Friedlingstein, P., Eby, M., Jones, C. D., Christian, J. R., Bonan, G., Bopp, L., Brovkin, V., Cadule, P., Hajima, T., Ilyina, T., Lindsay, K., Tjiputra, J. F., and Wu, T.: Carbon-Concentration and Carbon-Climate Feedbacks in CMIP5 Earth System Models, *Journal of Climate*, 26, 5289–5314, <https://doi.org/10.1175/JCLI-D-12-00494.1>, 2013.
- 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.
- 950 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.
- 955 Aumont, O., Orr, J. C., Monfray, P., Ludwig, W., Amiotte-Suchet, P., and Probst, J.-L.: Riverine-driven interhemispheric transport of carbon, *Global Biogeochemical Cycles*, 15, 393–405, <https://doi.org/10.1029/1999GB001255>, 2001.
- Bellenger, H., Bopp, L., Ethé, C., Ho, D., Duvel, J. P., Flavoni, S., Guez, L., Kataoka, T., Perrot, X., Parc, L., and Watanabe, M.: Sensitivity of the Global Ocean Carbon Sink to the Ocean Skin in a Climate Model, *Journal of Geophysical Research: Oceans*, 128, e2022JC019479, <https://doi.org/https://doi.org/10.1029/2022JC019479>, e2022JC019479 2022JC019479, 2023.
- 960 Boé, J., Hall, A., and Qu, X.: September sea-ice cover in the Arctic Ocean projected to vanish by 2100, *Nature Geoscience*, 2, 341–343, <https://doi.org/10.1038/ngeo467>, 2009.
- Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D’Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, L. E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and evaluation of the IPSL-CM6A-LR climate model, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS002010, <https://doi.org/10.1029/2019MS002010>, 2020.
- 970 Bourgeois, T., Goris, N., Schwinger, J., and Tjiputra, J. F.: Stratification constrains future heat and carbon uptake in the Southern Ocean between 30°S and 55°S, *Nature Communications*, 13, 340, <https://doi.org/10.1038/s41467-022-27979-5>, 2022.
- Brient, F.: Reducing uncertainties in climate projections with emergent constraints: concepts, examples and prospects, *Advances in Atmospheric Sciences*, 37, 1–15, <https://doi.org/10.1007/s00376-019-9140-8>, 2020.
- 975



- Broecker, W. S., Takahashi, T., Simpson, H. J., and Peng, T.-H.: Fate of fossil fuel carbon dioxide and the global carbon budget, *Science*, 206, 409–418, <https://doi.org/10.1126/science.206.4417.409>, 1979.
- Buckley, M. W. and Marshall, J.: Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review, *Reviews of Geophysics*, 54, 5–63, <https://doi.org/10.1002/2015RG000493>, 2016.
- 980 Bunsen, F., Nissen, C., and Hauck, J.: The impact of recent climate change on the global ocean carbon sink, *Geophysical Research Letters*, 51, e2023GL107 030, <https://doi.org/10.1029/2023GL107030>, 2024.
- Burnham, K. P. and Anderson, D. R.: *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, Springer, New York, 2nd edn., <https://doi.org/10.1007/b97636>, 2002.
- Caldeira, K. and Duffy, P. B.: The role of the Southern Ocean in uptake and storage of anthropogenic carbon dioxide, *Science*, 287, 620–622, <https://doi.org/10.1126/science.287.5453.620>, 2000.
- 985 Caldwell, P. M., Bretherton, C. S., Zelinka, M. D., Klein, S. A., Santer, B. D., and Sanderson, B. M.: Statistical significance of climate sensitivity predictors obtained by data mining, *Geophysical Research Letters*, 41, 1803–1808, <https://doi.org/10.1002/2014GL059205>, 2014.
- Campbell, E. C., Wilson, E. A., Moore, G. W. K., Riser, S. C., Brayton, C. E., Mazloff, M. R., and Talley, L. D.: Antarctic offshore polynyas linked to Southern Hemisphere climate anomalies, *Nature*, 570, 319–325, <https://doi.org/10.1038/s41586-019-1294-0>, 2019.
- 990 Canadell, J. G., Monteiro, P. M. S., Costa, M. H., Cotrim da Cunha, L., Cox, P. M., Eliseev, A. V., Henson, S., Ishii, M., Jaccard, S., Koven, C., Lohila, A., Patra, P. K., Piao, S., Rogelj, J., Syampungani, S., Zaehle, S., and Zickfeld, K.: Global Carbon and other Biogeochemical Cycles and Feedbacks, in: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., pp. 673–816, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, <https://doi.org/10.1017/9781009157896.007>, 2021.
- Carnier Casaroli, L., Serra, N., Chegini, F., and Ilyina, T.: Tropical Instability Vortices reduce Pacific Ocean ENSO-Driven CO₂ outgassing, *EGUsphere*, pp. 1–35, <https://doi.org/10.5194/egusphere-2026-504>, preprint, 2026.
- 1000 Chau, T.-T.-T., Gehlen, M., Metzl, N., and Chevallier, F.: CMEMS-LSCE: a global, 0.25°, monthly reconstruction of the surface ocean carbonate system, *Earth System Science Data*, 16, 121–160, <https://doi.org/10.5194/essd-16-121-2024>, 2024.
- Cheng, L., Trenberth, K. E., Gruber, N., Abraham, J. P., Fasullo, J. T., Li, G., Mann, M. E., Zhao, X., and Zhu, J.: Improved Estimates of Changes in Upper Ocean Salinity and the Hydrological Cycle, *Journal of Climate*, 33, 10 357–10 381, <https://doi.org/10.1175/JCLI-D-20-0366.1>, 2020.
- 1005 Cheng, L., Pan, Y., Tan, Z., Zheng, H., et al.: IAPv4 ocean temperature and ocean heat content gridded dataset, *Earth System Science Data*, 16, 3517–3546, <https://doi.org/10.5194/essd-16-3517-2024>, 2024.
- Couespel, D., Lévy, M., and Bopp, L.: Stronger Oceanic CO₂ Sink in Eddy-Resolving Simulations of Global Warming, *Geophysical Research Letters*, 51, e2023GL106 172, <https://doi.org/10.1029/2023GL106172>, 2024.
- Cox, P. M., Pearson, D., Booth, B. B., Friedlingstein, P., Huntingford, C., Jones, C. D., and Luke, C. M.: Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability, *Nature*, 494, 341–344, <https://doi.org/10.1038/nature11882>, 2013.
- 1010 Cunningham, S. A., Kanzow, T., Rayner, D., Baringer, M. O., Johns, W. E., Marotzke, J., Longworth, H. R., Grant, E. M., Hirschi, J. J.-M., Beal, L. M., Meinen, C. S., and Bryden, H. L.: Temporal variability of the Atlantic meridional overturning circulation at 26.5°N, *Science*, 317, 935–938, <https://doi.org/10.1126/science.1141304>, 2007.



- Dai, M., Su, J., Zhao, Y., Hofmann, E. E., et al.: Carbon Fluxes in the Coastal Ocean: Synthesis, Boundary Processes, and Future Trends, *Annual Review of Earth and Planetary Sciences*, 50, 593–626, <https://doi.org/10.1146/annurev-earth-032320-090746>, 2022.
- 1015 Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C.,
- 1020 Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), *Journal of Advances in Modeling Earth Systems*, 12, e2019MS001916, <https://doi.org/10.1029/2019MS001916>, 2020.
- de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology, *Journal of Geophysical Research*, 109, C12003, <https://doi.org/10.1029/2004JC002378>, 2004.
- 1025 DeVries, T., Yamamoto, K., Wanninkhof, R., Gruber, N., Hauck, J., Müller, J. D., Bopp, L., Carroll, D., Carter, B., Chau, T.-T.-T., Doney, S. C., Gehlen, M., Gloege, L., Gregor, L., Henson, S., Kim, J. H., Iida, Y., Ilyina, T., Landschützer, P., Le Quéré, C., Munro, D., Nissen, C., Patara, L., Pérez, F. F., Resplandy, L., Rodgers, K. B., Schwinger, J., Séférian, R., Sicardi, V., Terhaar, J., Triñanes, J., Tsujino, H., Watson, A., Yasunaka, S., and Zeng, J.: Magnitude, Trends, and Variability of the Global Ocean Carbon Sink From 1985 to 2018, *Global Biogeochemical Cycles*, 37, e2023GB007780, <https://doi.org/https://doi.org/10.1029/2023GB007780>, e2023GB007780 2023GB007780,
- 1030 2023.
- Dickson, A. G., Sabine, C. L., and Christian, J. R., eds.: Guide to best practices for ocean CO₂ measurements, North Pacific Marine Science Organization, Sidney, British Columbia, <https://doi.org/10.25607/OBP-1342>, pICES Special Publication 3; IOCCP Report No. 8, 2007.
- Dickson, R. R. and Brown, J.: The production of North Atlantic Deep Water: Sources, rates, and pathways, *Journal of Geophysical Research: Oceans*, 99, 12319–12341, <https://doi.org/10.1029/94JC00530>, 1994.
- 1035 Donlon, C. J., Minnett, P. J., Gentemann, C., Nightingale, T. J., Barton, I. J., Ward, B., and Murray, M. J.: Toward Improved Validation of Satellite Sea Surface Skin Temperature Measurements for Climate Research, *Journal of Climate*, 15, 353 – 369, [https://doi.org/10.1175/1520-0442\(2002\)015<0353:TIVOSS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<0353:TIVOSS>2.0.CO;2), 2002.
- Dufour, C. O., Griffies, S. M., de Souza, G. F., Frenger, I., Morrison, A. K., Palter, J. B., Sarmiento, J. L., Galbraith, E. D., Dunne, J. P., Anderson, W. G., and Slater, R. D.: Role of mesoscale eddies in cross-frontal transport of heat and biogeochemical tracers in the Southern
- 1040 Ocean, *Journal of Physical Oceanography*, 45, 3057–3081, <https://doi.org/10.1175/JPO-D-14-0240.1>, 2015.
- Dunne, J. P., John, J. G., Adcroft, A. J., Griffies, S. M., Hallberg, R. W., Shevliakova, E., Stouffer, R. J., Cooke, W. F., Dunne, K. A., Harrison, M. J., Krasting, J. P., Malyshev, S. L., Milly, P. C. D., Philipps, P. J., Sentman, L. T., Samuels, B. L., Spelman, M. J., Winton, M., Wittenberg, A. T., and Zadeh, N.: GFDL's ESM2 Global Coupled Climate-Carbon Earth System Models. Part I: Physical Formulation and Baseline Simulation Characteristics, *Journal of Climate*, 25, 6646–6665, <https://doi.org/10.1175/JCLI-D-11-00560.1>, 2012.
- 1045 Dunne, J. P., John, J. G., Shevliakova, E., Stouffer, R. J., Krasting, J. P., Malyshev, S. L., Milly, P. C. D., Sentman, L. T., Adcroft, A. J., Cooke, W. F., Dunne, K. A., Griffies, S. M., Hallberg, R. W., Harrison, M. J., Levy, H., Wittenberg, A. T., Philipps, P. J., and Zadeh, N.: GFDL's ESM2 Global Coupled Climate-Carbon Earth System Models. Part II: Carbon System Formulation and Baseline Simulation Characteristics, *Journal of Climate*, 26, 2247–2267, <https://doi.org/10.1175/JCLI-D-12-00150.1>, 2013.
- 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,
- 1050 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, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS002015, <https://doi.org/10.1029/2019MS002015>, 2020.
- Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L.-P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., Fuentes-Franco, R., Gröger, M., v. Hardenberg, J., Hieronymus, J., Karami, M. P., Keskinen, J.-P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégos, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., van Noije, T., Nolan, P., O'Donnell, D., Ollinaho, P., van den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F., Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., and Zhang, Q.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6, *Geoscientific Model Development*, 15, 2973–3020, <https://doi.org/10.5194/gmd-15-2973-2022>, eC-Earth3-CC configuration, 2022.
- Enting, I. G.: On the use of smoothing splines to filter CO₂ data, *Journal of Geophysical Research: Atmospheres*, 92, 10977–10984, <https://doi.org/10.1029/JD092iD09p10977>, 1987.
- 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, *Geoscientific Model Development*, 9, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>, 2016.
- Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., Collins, W. D., Gier, B. K., Hall, A. D., Hoffman, F. M., Hurtt, G. C., Jahn, A., Jones, C. D., Klein, S. A., Krasting, J. P., Kwiatkowski, L., Lorenz, R., Maloney, E., Meehl, G. A., Pendergrass, A. G., Pincus, R., Ruane, A. C., Russell, J. L., Sanderson, B. M., Santer, B. D., Sherwood, S. C., Simpson, I. R., Stouffer, R. J., and Williamson, M. S.: Taking climate model evaluation to the next level, *Nature Climate Change*, 9, 102–110, <https://doi.org/10.1038/s41558-018-0355-y>, 2019.
- Fairall, C. W., Bradley, E. F., Godfrey, J. S., Wick, G. A., Edson, J. B., and Young, G. S.: Cool-skin and warm-layer effects on sea surface temperature, *Journal of Geophysical Research: Oceans*, 101, 1295–1308, <https://doi.org/10.1029/95JC03190>, 1996.
- Fay, A. R., Gregor, L., Landschützer, P., McKinley, G. A., Gruber, N., Gehlen, M., Iida, Y., Laruelle, G. G., Rödenbeck, C., Roobaert, A., and Zeng, J.: SeaFlux: harmonization of air–sea CO₂ fluxes from surface pCO₂ data products using a standardized approach, *Earth System Science Data*, 13, 4693–4710, <https://doi.org/10.5194/essd-13-4693-2021>, 2021.
- Ford, D. J., Tilstone, G. H., Shutler, J. D., Kitidis, V., Sheen, K. L., Dall’Olmo, G., and Orselli, I. B. M.: Mesoscale Eddies Enhance the Air–Sea CO₂ Sink in the South Atlantic Ocean, *Geophysical Research Letters*, 50, e2022GL102137, <https://doi.org/10.1029/2022GL102137>, 2023.
- Ford, D. J., Shutler, J. D., Blanco-Sacristán, J., Corrigan, S., Bell, T. G., Yang, M., Kitidis, V., Nightingale, P. D., Brown, I., Wimmer, W., Woolf, D. K., Casal, T., Donlon, C., Tilstone, G. H., and Ashton, I.: Enhanced ocean CO₂ uptake due to near-surface temperature gradients, *Nature Geoscience*, 17, 1135–1140, <https://doi.org/10.1038/s41561-024-01570-7>, 2024.
- Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate-Carbon Cycle Feedback Analysis: Results from the C4MIP Model Intercomparison, *Journal of Climate*, 19, 3337–3353, <https://doi.org/10.1175/JCLI3800.1>, 2006.



- Friedlingstein, P., Le Quéré, C., O'Sullivan, M., Hauck, J., Landschützer, P., Luijkx, I. T., Li, H., van der Woude, A., Schwingshackl, C.,
1090 Pongratz, J., Regnier, P., Andrew, R. M., Bakker, D. C. E., Canadell, J. G., Ciais, P., Gasser, T., Jones, M. W., Lan, X., Morgan, E., Olsen, A., Peters, G. P., Peters, W., Sitch, S., and Tian, H.: Emerging climate impact on carbon sinks in a consolidated carbon budget, *Nature*, 649, 98–103, <https://doi.org/10.1038/s41586-025-09802-5>, 2026a.
- Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Li, H., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Aas, K., Alin, S. R., Anthoni, P.,
1095 Barbero, L., Bates, N. R., Bellouin, N., Benoit-Cattin, A., Berghoff, C. F., Bernardello, R., Bopp, L., Brasika, I. B. M., Chamberlain, M. A., Chandra, N., Chevallier, F., Chini, L. P., Collier, N. O., Colligan, T. H., Cronin, M., Djeutchouang, L. M., Dou, X., Enright, M. P., Enyo, K., Erb, M., Evans, W., Feely, R. A., Feng, L., Ford, D. J., Foster, A., Fransner, F., Gasser, T., Gehlen, M., Gkritzalis, T., Goncalves De Souza, J., Grassi, G., Gregor, L., Gruber, N., Guenet, B., Gürses, Ö., Harrington, K., Harris, I., Heinke, J., Hurtt, G. C., Iida, Y., Ilyina, T., Ito, A., Jacobson, A. R., Jain, A. K., Jarníková, T., Jersild, A., Jiang, F., Jones, S. D., Kato, E., Keeling, R. F., Klein Goldewijk, K.,
1100 Knauer, J., Kong, Y., Korsbakken, J. I., Koven, C., Kunimitsu, T., Lan, X., Liu, J., Liu, Z., Liu, Z., Lo Monaco, C., Ma, L., Marland, G., McGuire, P. C., McKinley, G. A., Melton, J. R., Monacci, N., Monier, E., Morgan, E. J., Munro, D. R., Müller, J. D., Nakaoka, S.-I., Nayagam, L. R., Niwa, Y., Nutz, T., Olsen, A., Omar, A. M., Pan, N., Pandey, S., Pierrot, D., Qin, Z., Regnier, P., Rehder, G., Resplandy, L., Roobaert, A., Rosan, T. M., Rödenbeck, C., Schwinger, J., Skjelvan, I., Smallman, T. L., Spada, V., Sreeush, M. G., Sun, Q., Sutton, A. J., Sweeney, C., Swingedouw, D., Séférian, R., Takao, S., Tatebe, H., Tian, H., Tian, X., Tilbrook, B., Tsujino, H., Tubiello, F., van
1105 Ooijen, E., van der Werf, G. R., van de Velde, S. J., Walker, A. P., Wanninkhof, R., Yang, X., Yuan, W., Yue, X., and Zeng, J.: Global Carbon Budget 2025, *Earth System Science Data*, 18, 3211–3288, <https://doi.org/10.5194/essd-18-3211-2026>, 2026b.
- Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P., and Winton, M.: Dominance of the Southern Ocean in Anthropogenic Carbon and Heat Uptake in CMIP5 Models, *Journal of Climate*, 28, 862–886, <https://doi.org/10.1175/JCLI-D-14-00117.1>, 2015.
- 1110 Gao, S., Schwinger, J., Tjiputra, J., Bethke, I., Hartmann, J., Mayorga, E., and Heinze, C.: Riverine impact on future projections of marine primary production and carbon uptake, *Biogeosciences*, 20, 93–119, <https://doi.org/10.5194/bg-20-93-2023>, 2023.
- Good, S. A., Martin, M. J., and Rayner, N. A.: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, *Journal of Geophysical Research: Oceans*, 118, 6704–6716, <https://doi.org/10.1002/2013JC009067>, 2013.
- Gordon, A. L.: Deep Antarctic Convection West of Maud Rise, *Journal of Physical Oceanography*, 8, 600–612, [https://doi.org/10.1175/1520-0485\(1978\)008<0600:DACWOM>2.0.CO;2](https://doi.org/10.1175/1520-0485(1978)008<0600:DACWOM>2.0.CO;2), 1978.
1115
- Gregor, L. and Gruber, N.: OceanSODA-ETHZ: a global gridded data set of the surface ocean carbonate system for seasonal to decadal studies of ocean acidification, *Earth System Science Data*, 13, 777–808, <https://doi.org/10.5194/essd-13-777-2021>, 2021.
- Griffies, S. M., Winton, M., Anderson, W. G., Benson, R., Delworth, T. L., Dufour, C. O., Dunne, J. P., Goddard, P., Morrison, A. K., Rosati, A., Wittenberg, A. T., Yin, J., and Zhang, R.: Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models, *Journal of Climate*, 28, 952–977, <https://doi.org/10.1175/JCLI-D-14-00353.1>, 2015.
1120
- Gruber, N., Clement, D., Carter, B. R., Feely, R. A., van Heuven, S., Hoppema, M., Ishii, M., Key, R. M., Kozyr, A., Lauvset, S. K., Lo Monaco, C., Mathis, J. T., Murata, A., Olsen, A., Perez, F. F., Sabine, C. L., Tanhua, T., and Wanninkhof, R.: The oceanic sink for anthropogenic CO₂ from 1994 to 2007, *Science*, 363, 1193–1199, <https://doi.org/10.1126/science.aau5153>, 2019.
- Guo, Y. and Timmermans, M.-L.: The Role of Ocean Mesoscale Variability in Air-Sea CO₂ Exchange: A Global Perspective, *Geophysical Research Letters*, 51, e2024GL108373, <https://doi.org/10.1029/2024GL108373>, 2024.
1125



- Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., Ohgaito, R., Ito, A., Yamazaki, D., Okajima, H., Ito, A., Takata, K., Ogochi, K., Watanabe, S., and Kawamiya, M.: Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks, *Geoscientific Model Development*, 13, 2197–2244, <https://doi.org/10.5194/gmd-13-2197-2020>, 2020.
- 1130 Hall, A., Cox, P., Huntingford, C., and Klein, S.: Progressing emergent constraints on future climate change, *Nature Climate Change*, 9, 269–278, <https://doi.org/10.1038/s41558-019-0436-6>, 2019.
- Heuzé, C.: Antarctic Bottom Water and North Atlantic Deep Water in CMIP6 models, *Ocean Science*, 17, 59–90, <https://doi.org/10.5194/os-17-59-2021>, 2021.
- Hollitzer, H. A. L., Patara, L., Terhaar, J., and Oschlies, A.: Competing effects of wind and buoyancy forcing on ocean oxygen trends in recent decades, *Nature Communications*, 15, 9212, <https://doi.org/10.1038/s41467-024-53557-y>, 2024.
- 1135 Huguenin, M. F., Holmes, R. M., and England, M. H.: Drivers and distribution of global ocean heat uptake over the last half century, *Nature Communications*, 13, 4921, <https://doi.org/10.1038/s41467-022-32591-0>, 2022.
- Iida, Y., Takatani, Y., Kojima, A., and Ishii, M.: Global trends of ocean CO₂ sink and ocean acidification: an observation-based reconstruction of surface ocean inorganic carbon variables, *Journal of Oceanography*, 77, 323–358, <https://doi.org/10.1007/s10872-020-00571-5>, 2021.
- 1140 Ito, G., Romanou, A., Kiang, N. Y., Faluvegi, G., Aleinov, I., Ruedy, R., Russell, G., Lerner, P., Kelley, M., and Lo, K.: Global Carbon Cycle and Climate Feedbacks in the NASA GISS ModelE2.1, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS002030, <https://doi.org/10.1029/2019MS002030>, 2020.
- Jackson, L. C., Biastoch, A., Buckley, M. W., Desbruyères, D. G., Frajka-Williams, E., Moat, B., and Robson, J.: The evolution of the North Atlantic Meridional Overturning Circulation since 1980, *Nature Reviews Earth & Environment*, 3, 241–254, <https://doi.org/10.1038/s43017-022-00263-2>, 2022.
- 1145 Jacobson, A. R., Mikaloff Fletcher, S. E., Gruber, N., Sarmiento, J. L., and Gloor, M.: A joint atmosphere–ocean inversion for surface fluxes of carbon dioxide: 1. Methods and global-scale fluxes, *Global Biogeochemical Cycles*, 21, GB1019, <https://doi.org/10.1029/2005GB002556>, 2007.
- Joos, F., Plattner, G.-K., Stocker, T. F., Marchal, O., and Schmittner, A.: Global warming and marine carbon cycle feedbacks on future atmospheric CO₂, *Science*, 284, 464–467, <https://doi.org/10.1126/science.284.5413.464>, 1999.
- 1150 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., Matthews, E. E., McDermid, S., Mezuman, K., Miller, R. L., Murray, L. T., Oinas, V., Orbe, C., García-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, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS002025, <https://doi.org/10.1029/2019MS002025>, 2020.
- 1155 Kessler, A. and Tjiputra, J.: The Southern Ocean as a constraint to reduce uncertainty in future ocean carbon sinks, *Earth System Dynamics*, 7, 295–312, <https://doi.org/10.5194/esd-7-295-2016>, 2016.
- Kuhlbrodt, T., Griesel, A., Montoya, M., Levermann, A., Hofmann, M., and Rahmstorf, S.: On the driving processes of the Atlantic meridional overturning circulation, *Reviews of Geophysics*, 45, RG2001, <https://doi.org/10.1029/2004RG000166>, 2007.
- 1160 Kwiatkowski, L., Bopp, L., Aumont, O., Ciais, P., Cox, P. M., Laufkötter, C., Liao, Y., and Roëdenbeck, C.: Emergent constraints on projections of declining primary production in the tropical oceans, *Nature Climate Change*, 7, 355–358, <https://doi.org/10.1038/nclimate3265>, 2017.



- Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J. R., Dunne, J. P., Gehlen, M., Ilyina, T., John, J. G., Lenton, A., Li, H., Lovenduski, N. S., Orr, J. C., Palmieri, J., Santana-Falcón, Y., Schwinger, J., Séférian, R., Stock, C. A., Tagliabue, A., Takano, Y., Tjiputra, J., Toyama, K., Tsujino, H., Watanabe, M., Yamamoto, A., Yool, A., and Ziehn, T.: Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections, *Biogeosciences*, 17, 3439–3470, <https://doi.org/10.5194/bg-17-3439-2020>, 2020.
- Lachkar, Z., Orr, J. C., Dutay, J.-C., and Delecluse, P.: Effects of mesoscale eddies on global ocean distributions of CFC-11, CO₂, and $\Delta^{14}\text{C}$, *Ocean Science*, 3, 461–482, <https://doi.org/10.5194/os-3-461-2007>, 2007.
- Lachkar, Z., Orr, J. C., and Dutay, J.-C.: Seasonal and mesoscale variability of oceanic transport of anthropogenic CO₂, *Biogeosciences*, 6, 2509–2523, <https://doi.org/10.5194/bg-6-2509-2009>, 2009.
- Lacroix, F., Ilyina, T., and Hartmann, J.: Oceanic CO₂ outgassing and biological production hotspots induced by pre-industrial river loads of nutrients and carbon in a global modeling approach, *Biogeosciences*, 17, 55–88, <https://doi.org/10.5194/bg-17-55-2020>, 2020.
- Lacroix, F., Ilyina, T., Mathis, M., Laruelle, G. G., and Regnier, P.: Historical increases in land-derived nutrient inputs may alleviate effects of a changing physical climate on the oceanic carbon cycle, *Global Change Biology*, 27, 5491–5513, <https://doi.org/10.1111/gcb.15822>, 2021.
- Lan, X., Tans, P., and Thoning, K. W.: Trends in Globally-Averaged CO₂ Determined from NOAA Global Monitoring Laboratory Measurements, <https://doi.org/10.15138/9N0H-ZH07>, version 2026-03, 2026.
- Lauvset, S. K., Key, R. M., Olsen, A., van Heuven, S., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F. F., Suzuki, T., and Watelet, S.: A new global interior ocean mapped climatology: the 1° × 1° GLODAP version 2, *Earth System Science Data*, 8, 325–340, <https://doi.org/10.5194/essd-8-325-2016>, 2016.
- 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, *Journal of Advances in Modeling Earth Systems*, 13, e2019MS002028, <https://doi.org/10.1029/2019MS002028>, 2021.
- Lerner, P., Romanou, A., Nicholson, D., Kelley, M., Ruedy, R., and Russell, G.: The Sensitivity of the Equatorial Pacific ODZ to Particulate Organic Matter Remineralization in a Climate Model Under Pre-Industrial Conditions, *Ocean Modelling*, 188, 102303, <https://doi.org/10.1016/j.ocemod.2023.102303>, 2024.
- Locarnini, R. A., Mishonov, A. V., Baranova, O. K., Reagan, J. R., Boyer, T. P., Seidov, D., Wang, Z., García, H. E., Bouchard, C., Cross, S. L., Paver, C. R., and Dukhovskoy, D.: World Ocean Atlas 2023, Volume 1: Temperature, NOAA Atlas NESDIS 89, NOAA, <https://doi.org/10.25923/54bh-1613>, technical Editor, 2024.
- Lovato, T., Peano, D., Butenschön, M., Materia, S., Iovino, D., Scoccimarro, E., Fogli, P. G., Cherchi, A., Bellucci, A., Gualdi, S., Masina, S., and Navarra, A.: CMIP6 simulations with the CMCC Earth System Model (CMCC-ESM2), *Journal of Advances in Modeling Earth Systems*, 14, e2021MS002814, <https://doi.org/10.1029/2021MS002814>, 2022.
- Marquardt, D. W.: Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation, *Technometrics*, 12, 591–612, <https://doi.org/10.1080/00401706.1970.10488699>, 1970.
- Marshall, J. and Speer, K.: Closure of the meridional overturning circulation through Southern Ocean upwelling, *Nature Geoscience*, 5, 171–180, <https://doi.org/10.1038/ngeo1391>, 2012.
- 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.



- Matthews, H. D., 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, 7, <https://doi.org/10.1038/s43247-020-00064-9>, 2021.
- Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., 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, *Journal of Advances in Modeling Earth Systems*, 11, 998–1038, <https://doi.org/10.1029/2018MS001400>, 2019.
- McCarthy, G. D., Smeed, D. A., Johns, W. E., Frajka-Williams, E., Moat, B. I., Rayner, D., Baringer, M. O., Meinen, C. S., Collins, J., and Bryden, H. L.: Measuring the Atlantic Meridional Overturning Circulation at 26°N, *Progress in Oceanography*, 130, 91–111, <https://doi.org/10.1016/j.pocean.2014.10.006>, 2015.
- McDougall, T. J. and Barker, P. M.: *Getting Started with TEOS-10 and the Gibbs Seawater (GSW) Oceanographic Toolbox*, Tech. rep., SCOR/IAPSO WG127, ISBN 978-0-646-55621-5, 2011.
- McNeil, B. I. and Matear, R. J.: The non-steady state oceanic CO₂ signal: its importance, magnitude and a method of detection, *Biogeosciences*, 10, 2219–2228, <https://doi.org/10.5194/bg-10-2219-2013>, 2013.
- Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and van Vuuren, D. P. P.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, *Climatic Change*, 109, 213–241, <https://doi.org/10.1007/s10584-011-0156-z>, 2011.
- Mikaloff Fletcher, S. E., Gruber, N., Jacobson, A. R., Doney, S. C., Dutkiewicz, S., Gerber, M., Follows, M., Joos, F., Lindsay, K., Mene-menlis, D., Mouchet, A., Müller, S. A., and Sarmiento, J. L.: Inverse estimates of anthropogenic CO₂ uptake, transport, and storage by the ocean, *Global Biogeochemical Cycles*, 20, GB2002, <https://doi.org/10.1029/2005GB002530>, 2006.
- Moat, B. I., Smeed, D. A., Rayner, D., Johns, W. E., Smith, R., Volkov, D., Elipot, S., Petit, T., Kajtar, J., Baringer, M. O., and Collins, J.: Atlantic meridional overturning circulation observed by the RAPID-MOCHA-WBTS (RAPID-Meridional Overturning Circulation and Heatflux Array-Western Boundary Time Series) array at 26N from 2004 to 2024 (v2024.1), <https://doi.org/10.5285/3f24651e-2d44-dee3-e063-7086abc0395e>, 2025.
- Mortenson, E., Lenton, A., Shadwick, E., Matear, R., Chamberlain, M., and Baird, M.: Meridional asymmetry in the evolution of the Southern Ocean carbon sink over the 21st century in a high-resolution model, *Frontiers in Marine Science*, 12, 1589735, <https://doi.org/10.3389/fmars.2025.1589735>, 2025.
- Müller, J. D., Gruber, N., Carter, B., Feely, R., Ishii, M., Lange, N., Lauvset, S. K., Murata, A., Olsen, A., Pérez, F. F., Sabine, C. L., Tanhua, T., Wanninkhof, R., and Zhu, Y.: Decadal trends in the oceanic storage of anthropogenic carbon from 1994 to 2014, *AGU Advances*, 4, e2023AV000875, <https://doi.org/10.1029/2023AV000875>, 2023.
- O'Brien, R. M.: A Caution Regarding Rules of Thumb for Variance Inflation Factors, *Quality & Quantity*, 41, 673–690, <https://doi.org/10.1007/s11135-006-9018-6>, 2007.



- O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, Geoscientific Model Development, 9, 3461–3482, <https://doi.org/10.5194/gmd-9-3461-2016>, describes the SSP scenario experiments (SSP1-2.6, SSP2-4.5, SSP3-7.0, SSP5-8.5, etc.), 2016.
- Orr, J. C. and Epitalon, J.-M.: Improved routines to model the ocean carbonate system: mocsy 2.0, Geoscientific Model Development, 8, 485–499, <https://doi.org/10.5194/gmd-8-485-2015>, 2015.
- Orr, J. C., Maier-Reimer, E., Mikolajewicz, U., Monfray, P., Sarmiento, J. L., Toggweiler, J. R., Taylor, N. K., Palmer, J., Gruber, N., Sabine, C. L., Le Quéré, C., Key, R. M., and Boutin, J.: Estimates of anthropogenic carbon uptake from four three-dimensional global ocean models, Global Biogeochemical Cycles, 15, 43–60, <https://doi.org/10.1029/2000GB001273>, 2001.
- Orr, J. C., Epitalon, J.-M., Dickson, A. G., and Gattuso, J.-P.: Routine uncertainty propagation for the marine carbon dioxide system, Marine Chemistry, 207, 84–107, <https://doi.org/10.1016/j.marchem.2018.10.006>, 2018.
- Planchat, A., Bopp, L., and Kwiatkowski, L.: A fresh look at the pre-industrial air–sea carbon flux using the alkalinity budget, Biogeosciences, 22, 6017–6055, <https://doi.org/10.5194/bg-22-6017-2025>, 2025.
- Plattner, G.-K., Joos, F., Stocker, T. F., and Marchal, O.: Feedback mechanisms and sensitivities of ocean carbon uptake under global warming, Tellus B, 53, 564–592, <https://doi.org/10.3402/tellusb.v53i5.16637>, 2001.
- Reagan, J. R., Boyer, T. P., García, H. E., Locarnini, R. A., Baranova, O. K., Bouchard, C., Cross, S. L., Mishonov, A. V., Paver, C. R., Seidov, D., Wang, Z., and Dukhovskoy, D.: World Ocean Atlas 2023, <https://doi.org/10.25921/va26-hv25>, nCEI Accession 0270533, 2024a.
- Reagan, J. R., Seidov, D., Wang, Z., Dukhovskoy, D., Boyer, T. P., Locarnini, R. A., Baranova, O. K., Mishonov, A. V., García, H. E., Bouchard, C., Cross, S. L., and Paver, C. R.: World Ocean Atlas 2023, Volume 2: Salinity, NOAA Atlas NESDIS 90, NOAA, <https://doi.org/10.25923/70qt-9574>, technical Editor, 2024b.
- Regnier, P., Friedlingstein, P., Ciais, P., Mackenzie, F. T., Gruber, N., Janssens, I. A., Laruelle, G. G., Lauerwald, R., Luyssaert, S., Andersson, A. J., Arndt, S., Arnosti, C., Borges, A. V., Dale, A. W., Gallego-Sala, A., Godderis, Y., Goossens, N., Hartmann, J., Heinze, C., Ilyina, T., Joos, F., LaRowe, D. E., Leifeld, J., Meysman, F. J. R., Munhoven, G., Raymond, P. A., Spahni, R., Suntharalingam, P., and Thullner, M.: Anthropogenic perturbation of the carbon fluxes from land to ocean, Nature Geoscience, 6, 597–607, <https://doi.org/10.1038/ngeo1830>, 2013.
- Regnier, P., Resplandy, L., Najjar, R. G., and Ciais, P.: The land-to-ocean loops of the global carbon cycle, Nature, 603, 401–410, <https://doi.org/10.1038/s41586-021-04339-9>, 2022.
- Resplandy, L., Keeling, R. F., Rödenbeck, C., Stephens, B. B., Khatiwala, S., Rodgers, K. B., Long, M. C., Bopp, L., and Tans, P. P.: Revision of global carbon fluxes based on a reassessment of oceanic and riverine carbon transport, Nature Geoscience, 11, 504–509, <https://doi.org/10.1038/s41561-018-0151-3>, 2018.
- Revelle, R. and Suess, H. E.: Carbon dioxide exchange between atmosphere and ocean and the question of an increase of atmospheric CO₂ during the past decades, Tellus, 9, 18–27, <https://doi.org/10.1111/j.2153-3490.1957.tb01849.x>, 1957.
- Richardson, C. M., Peucker-Ehrenbrink, B., Wyatt, S., Bourbonnais, A., Hatje, V., Frey, C., Sanders, T., Varela, D. E., and Paytan, A.: Effects of climate change on river and groundwater nutrient inputs to the coastal ocean, Communications Earth & Environment, 6, 761, <https://doi.org/10.1038/s43247-025-02594-6>, published 24 September 2025, 2025.
- Rodgers, K. B., Schlunegger, S., Slater, R. D., Ishii, M., Frölicher, T. L., Toyama, K., Plancherel, Y., Aumont, O., and Fassbender, A. J.: Reemergence of Anthropogenic Carbon Into the Ocean's Mixed Layer Strongly Amplifies Transient Climate Sensitivity, Geophysical Research Letters, 47, e2020GL089275, <https://doi.org/10.1029/2020GL089275>, 2020.



- Roemmich, D., Johnson, G. C., Riser, S., Davis, R., Gilson, J., Owens, W. B., Garzoli, S. L., Schmid, C., and Ignaszewski, M.: The Argo Program: Observing the global ocean with profiling floats, *Oceanography*, 22, 34–43, <https://doi.org/10.5670/oceanog.2009.36>, 2009.
- Romanou, A., Lerner, P., Kiang, N., Aleinov, I., Kelley, M., Miller, R., Russell, G., Ruedy, R., Schmidt, G., Hakuba, M., and Wang, O.: The NASA-GISS ModelE2.1-CC2 ESM: Development and Simulations, *EGUsphere*, pp. 1–113, <https://doi.org/10.5194/egusphere-2025-4839>, preprint under review for Geoscientific Model Development, 2026.
- 1280 Roy, T., Bopp, L., Gehlen, M., Schneider, B., Cadule, P., Frölicher, T. L., Segschneider, J., Tjiputra, J., Heinze, C., and Joos, F.: Regional Impacts of Climate Change and Atmospheric CO₂ on Future Ocean Carbon Uptake: A Multimodel Linear Feedback Analysis, *Journal of Climate*, 24, 2300–2318, <https://doi.org/10.1175/2010JCLI3787.1>, 2011.
- Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R., Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F. J., Peng, T.-H., Kozyr, A., Ono, T., and Rios, A. F.: The oceanic sink for anthropogenic CO₂, *Science*, 305, 367–371, <https://doi.org/10.1126/science.1097403>, 2004.
- 1285 Salinas-Matus, M., Serra, N., Chugini, F., and Ilyina, T.: Mesoscale eddies heterogeneously modulate CO₂ fluxes in eddy-rich regions of the Southern Ocean, *Biogeosciences*, 22, 7519–7534, <https://doi.org/10.5194/bg-22-7519-2025>, 2025.
- Sallée, J.-B., Pellichero, V., Akhondas, C., Pauthenet, E., Vignes, L., Schmidtko, S., Naveira Garabato, A., Sutherland, P., and Kuusela, M.: Summertime increases in upper-ocean stratification and mixed-layer depth, *Nature*, 591, 592–598, <https://doi.org/10.1038/s41586-021-03303-x>, 2021.
- 1290 Sanderson, B. M., Pendergrass, A. G., Koven, C. D., Brient, F., Booth, B. B. B., Fisher, R. A., and Knutti, R.: The potential for structural errors in emergent constraints, *Earth System Dynamics*, 12, 899–918, <https://doi.org/10.5194/esd-12-899-2021>, 2021.
- Sanderson, B. M., Brovkin, V., Fisher, R. A., Hohn, D., Ilyina, T., Jones, C. D., Koenigk, T., Koven, C., Li, H., Lawrence, D. M., Lawrence, P., Liddicoat, S., MacDougall, A. H., Mengis, N., Nicholls, Z., O'Rourke, E., Romanou, A., Sandstad, M., Schwinger, J., Séférian, R., Sentman, L. T., Simpson, I. R., Smith, C., Steinert, N. J., Swann, A. L. S., Tjiputra, J., and Ziehn, T.: flat10MIP: an emissions-driven experiment to diagnose the climate response to positive, zero and negative CO₂ emissions, *Geoscientific Model Development*, 18, 5699–5724, <https://doi.org/10.5194/gmd-18-5699-2025>, 2025.
- 1295 Sarmiento, J. L. and Le Quéré, C.: Oceanic carbon dioxide uptake in a model of century-scale global warming, *Science*, 274, 1346–1350, <https://doi.org/10.1126/science.274.5291.1346>, 1996.
- 1300 Sarmiento, J. L. and Sundquist, E. T.: Revised budget for the oceanic uptake of anthropogenic carbon dioxide, *Nature*, 356, 589–593, <https://doi.org/10.1038/356589a0>, 1992.
- Sarmiento, J. L., Orr, J. C., and Siegenthaler, U.: A perturbation simulation of CO₂ uptake in an ocean general circulation model, *Journal of Geophysical Research: Oceans*, 97, 3621–3645, <https://doi.org/10.1029/91JC02849>, 1992.
- 1305 Schlund, M., Lauer, A., Gentine, P., Sherwood, S. C., and Eyring, V.: Emergent constraints on equilibrium climate sensitivity in CMIP5: do they hold for CMIP6?, *Earth System Dynamics*, 11, 1233–1258, <https://doi.org/10.5194/esd-11-1233-2020>, 2020.
- Schlunegger, S., Rodgers, K. B., Sarmiento, J. L., Ilyina, T., Dunne, J. P., Takano, Y., Christian, J. R., Long, M. C., Frölicher, T. L., Slater, R., and Lehner, F.: Time of Emergence and Large Ensemble Intercomparison for Ocean Biogeochemical Trends, *Global Biogeochemical Cycles*, 34, e2019GB006453, <https://doi.org/10.1029/2019GB006453>, 2020.
- 1310 Schulzweida, U.: CDO User Guide, <https://doi.org/10.5281/zenodo.10020800>, 2023.
- Schwarz, G.: Estimating the dimension of a model, *The Annals of Statistics*, 6, 461–464, <https://doi.org/10.1214/aos/1176344136>, 1978.



- Schwinger, J., Tjiputra, J. F., Heinze, C., Bopp, L., Christian, J. R., Gehlen, M., Ilyina, T., Jones, C. D., Salas-Mélia, D., Segschneider, J., and Séférian, R.: Nonlinearity of Ocean Carbon Cycle Feedbacks in CMIP5 Earth System Models, *Journal of Climate*, 27, 3869–3888, <https://doi.org/10.1175/JCLI-D-13-00452.1>, 2014.
- 1315 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., Guéremy, 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, *Journal of Advances in Modeling Earth Systems*, 11, 4182–4227, <https://doi.org/10.1029/2019MS001791>, 2019.
- 1320 Séférian, R., Berthet, S., Yool, A., Palmiéri, J., Bopp, L., Tagliabue, A., Kwiatkowski, L., Aumont, O., Christian, J., Dunne, J., Gehlen, M., Ilyina, T., John, J. G., Li, H., Long, M. C., Luo, J. Y., Nakano, H., Romanou, A., Schwinger, J., Stock, C., Santana-Falcón, Y., Takano, Y., Tjiputra, J., Tsujino, H., Watanabe, M., Wu, T., Wu, F., and Yamamoto, A.: Tracking Improvement in Simulated Marine Biogeochemistry Between CMIP5 and CMIP6, *Current Climate Change Reports*, 6, 95–119, <https://doi.org/10.1007/s40641-020-00160-0>, 2020.
- Seland, Ø., Bentsen, M., Olivie, D., Toniazzo, 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., Ilicak, 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, *Geoscientific Model Development*, 13, 6165–6200, <https://doi.org/10.5194/gmd-13-6165-2020>, 2020.
- 1325 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., de 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., Siahhaan, 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, *Journal of Advances in Modeling Earth Systems*, 11, 4513–4558, <https://doi.org/10.1029/2019MS001739>, 2019.
- 1330 Silvy, Y., Frölicher, T. L., Terhaar, J., Joos, F., Burger, F. A., Lacroix, F., Allen, M., Bernardello, R., Bopp, L., Brovkin, V., Buzan, J. R., Cadule, P., Dix, M., Dunne, J., Friedlingstein, P., Georgievski, G., Hajima, T., Jenkins, S., Kawamiya, M., Kiang, N. Y., Lapin, V., Lee, D., Lerner, P., Mengis, N., Monteiro, E. A., Paynter, D., Peters, G. P., Romanou, A., Schwinger, J., Sparrow, S., Stofferahn, E., Tjiputra, J., Tourigny, E., and Ziehn, T.: AERA-MIP: emission pathways, remaining budgets, and carbon cycle dynamics compatible with 1.5 and 2 °C global warming stabilization, *Earth System Dynamics*, 15, 1591–1628, <https://doi.org/10.5194/esd-15-1591-2024>, 2024.
- 1340 Stock, C. A., Dunne, J. P., Fan, S., Ginoux, P., John, J., Krasting, J. P., Laufkötter, C., Paulot, F., and Zadeh, N.: Ocean Biogeochemistry in GFDL's Earth System Model 4.1 and Its Response to Increasing Atmospheric CO₂, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS002 043, <https://doi.org/10.1029/2019MS002043>, 2020.
- Stone, M.: Cross-validatory choice and assessment of statistical predictions, *Journal of the Royal Statistical Society: Series B (Methodological)*, 36, 111–147, <https://doi.org/10.1111/j.2517-6161.1974.tb00994.x>, 1974.
- 1345 Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), *Geoscientific Model Development*, 12, 4823–4873, <https://doi.org/10.5194/gmd-12-4823-2019>, 2019.



- 1350 Takano, Y., Ilyina, T., Tjiputra, J., Eddebbar, Y. A., Berthet, S., Bopp, L., Buitenhuis, E., Butenschön, M., Christian, J. R., Dunne, J. P., Gröger, M., Hayashida, H., Hieronymus, J., Koenigk, T., Krasting, J. P., Long, M. C., Lovato, T., Nakano, H., Palmieri, J., Schwinger, J., Séférian, R., Suntharalingam, P., Tatebe, H., Tsujino, H., Urakawa, S., Watanabe, M., and Yool, A.: Simulations of ocean deoxygenation in the historical era: insights from forced and coupled models, *Frontiers in Marine Science*, 10, 1139 917, <https://doi.org/10.3389/fmars.2023.1139917>, 2023.
- 1355 Talley, L. D.: Closure of the global overturning circulation through the Indian, Pacific, and Southern Oceans: Schematics and transports, *Oceanography*, 26, 80–97, <https://doi.org/10.5670/oceanog.2013.07>, 2013.
Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, *Bulletin of the American Meteorological Society*, 93, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>, 2012.
Terhaar, J.: Drivers of decadal trends in the ocean carbon sink in the past, present, and future in Earth system models, *Biogeosciences*, 21, 3903–3926, <https://doi.org/10.5194/bg-21-3903-2024>, 2024.
- 1360 Terhaar, J.: Composite model-based estimate of the ocean carbon sink from 1959 to 2022, *Biogeosciences*, 22, 1631–1649, <https://doi.org/10.5194/bg-22-1631-2025>, 2025.
Terhaar, J.: Code for: Observation-based constraint suggests larger ocean carbon uptake and stronger climate feedbacks, <https://doi.org/10.5281/zenodo.22884358>, 2026.
- 1365 Terhaar, J., Orr, J. C., Ethé, C., Regnier, P., and Bopp, L.: Simulated Arctic Ocean response to doubling of riverine carbon and nutrient delivery, *Global Biogeochemical Cycles*, 33, 1048–1070, <https://doi.org/10.1029/2019GB006200>, 2019a.
Terhaar, J., Orr, J. C., Gehlen, M., Ethé, C., and Bopp, L.: Model constraints on the anthropogenic carbon budget of the Arctic Ocean, *Biogeosciences*, 16, 2343–2367, <https://doi.org/10.5194/bg-16-2343-2019>, 2019b.
Terhaar, J., Kwiatkowski, L., and Bopp, L.: Emergent constraint on Arctic Ocean acidification in the twenty-first century, *Nature*, 582, 379–383, <https://doi.org/10.1038/s41586-020-2360-3>, 2020.
- 1370 Terhaar, J., Frölicher, T. L., and Joos, F.: Southern Ocean anthropogenic carbon sink constrained by sea surface salinity, *Science Advances*, 7, eabd5964, <https://doi.org/10.1126/sciadv.abd5964>, 2021a.
Terhaar, J., Torres, O., Bourgeois, T., and Kwiatkowski, L.: Arctic Ocean acidification over the 21st century co-driven by anthropogenic carbon increases and freshening in the CMIP6 model ensemble, *Biogeosciences*, 18, 2221–2240, <https://doi.org/10.5194/bg-18-2221-2021>, 2021b.
- 1375 Terhaar, J., Frölicher, T. L., and Joos, F.: Observation-constrained estimates of the global ocean carbon sink from Earth system models, *Biogeosciences*, 19, 4431–4457, <https://doi.org/10.5194/bg-19-4431-2022>, 2022.
Terhaar, J., Goris, N., Müller, J. D., DeVries, T., Gruber, N., Hauck, J., Perez, F. F., and Séférian, R.: Assessment of Global Ocean Biogeochemistry Models for Ocean Carbon Sink Estimates in RECCAP2 and Recommendations for Future Studies, *Journal of Advances in Modeling Earth Systems*, 16, e2023MS003 840, <https://doi.org/https://doi.org/10.1029/2023MS003840>, 2024.
- 1380 Terhaar, J., Burger, F. A., Vogt, L., Frölicher, T. L., and Stocker, T. F.: Record sea surface temperature jump in 2023–2024 unlikely but not unexpected, *Nature*, 639, 942–946, <https://doi.org/10.1038/s41586-025-08674-z>, 2025.
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, , and Schulz, M.: Ocean biogeochemistry in the Norwegian Earth System Model version 2 (NorESM2), *Geoscientific Model Development*, 13, 2393–2431, <https://doi.org/10.5194/gmd-13-2393-2020>, 2020.
- 1385



- van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., and Rose, S. K.: The Representative Concentration Pathways: An Overview, *Climatic Change*, 109, 5–31, <https://doi.org/10.1007/s10584-011-0148-z>, 2011.
- Vogt, L., de Lavergne, C., Sallée, J.-B., Kwiatkowski, L., Frölicher, T. L., and Terhaar, J.: Increased future ocean heat uptake constrained by Antarctic sea ice extent, *Earth System Dynamics*, 16, 1453–1482, <https://doi.org/10.5194/esd-16-1453-2025>, 2025.
- Watson, A. J., Schuster, U., Shutler, J. D., Holding, T., Ashton, I. G. C., Landschützer, P., Woolf, D. K., and Goddijn-Murphy, L.: Revised estimates of ocean-atmosphere CO₂ flux are consistent with ocean carbon inventory, *Nature Communications*, 11, 4422, <https://doi.org/10.1038/s41467-020-18203-3>, 2020.
- Weijer, W., Cheng, W., Garuba, O. A., Hu, A., and Nadiga, B. T.: CMIP6 models predict significant 21st century decline of the Atlantic Meridional Overturning Circulation, *Geophysical Research Letters*, 47, e2019GL086075, <https://doi.org/10.1029/2019GL086075>, 2020.
- Wenzel, S., Cox, P. M., Eyring, V., and Friedlingstein, P.: Emergent constraints on climate-carbon cycle feedbacks in the CMIP5 Earth system models, *Journal of Geophysical Research: Biogeosciences*, 119, 794–807, <https://doi.org/10.1002/2013JG002591>, 2014.
- Woolf, D. K., Land, P. E., Shutler, J. D., Goddijn-Murphy, L. M., and Donlon, C. J.: On the calculation of air-sea fluxes of CO₂ in the presence of temperature and salinity gradients, *Journal of Geophysical Research: Oceans*, 121, 1229–1248, <https://doi.org/https://doi.org/10.1002/2015JC011427>, 2016.
- Ye, W., Li, Y., Wen, J., Zhang, J., Shakhova, N., Liu, J., Wu, M., Semiletov, I., and Zhan, L.: Enhanced Transport of Dissolved Methane From the Chukchi Sea to the Central Arctic, *Global Biogeochemical Cycles*, 37, e2022GB007368, <https://doi.org/10.1029/2022GB007368>, 2023.
- Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deuschi, M., Tanaka, T., Hosaka, M., Yabu, S., Yoshimura, H., Shindo, E., Mizuta, R., Obata, A., Adachi, Y., and Ishii, M.: The Meteorological Research Institute Earth System Model version 2.0, MRI-ESM2.0: Description and basic evaluation of the physical component, *Journal of the Meteorological Society of Japan. Ser. II*, 97, 931–965, <https://doi.org/10.2151/jmsj.2019-051>, 2019.
- Zickfeld, K., Eby, M., Matthews, H. D., and Weaver, A. J.: Setting cumulative emissions targets to reduce the risk of dangerous climate change, *Proceedings of the National Academy of Sciences*, 106, 16 129–16 134, <https://doi.org/10.1073/pnas.0805800106>, 2009.
- Ziehn, T., Chamberlain, M. A., Law, R. M., Lenton, A., Bodman, R. W., Dix, M., Stevens, L., Wang, Y.-P., and Srbinovsky, J.: The Australian Earth System Model: ACCESS-ESM1.5, *Journal of Southern Hemisphere Earth Systems Science*, 70, 193–214, <https://doi.org/10.1071/ES19035>, 2020.