Carbon cycle feedbacks in an idealized and a scenario simulation of negative emissions in CMIP6
Earth system models

Ali Asaadi¹, Jörg Schwinger², Hanna Lee¹,², Jerry Tjiputra³, Vivek Arora³, Roland Séférian⁴, Spencer Liddicoat⁵, Tomohiro Hajima⁶, Yeray Santana-Falcón⁴, Chris D. Jones⁵,⁷

¹NORCE Norwegian Research Centre & Bjerknes Centre for Climate Research, Bergen, Norway
²Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway
³Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, Victoria, BC, Canada
⁴CNRM, Université de Toulouse, Meteo-France, CNRS, Toulouse, France
⁵Met Office Hadley Centre, Exeter, United Kingdom
⁶Research Institute for Global Change, Japan Agency for Marine-Earth Science and Technology, Yokohama 236-0001, Japan
⁷School of Geographical Sciences, University of Bristol, UK

*Corresponding author, ali.asaadi@mail.mcgill.ca

Abstract

Limiting global warming to well below 2°C by the end of the century is an ambitious target that requires immediate and unprecedented emission reductions. In the absence of sufficient near term mitigation, this target will only be achieved by carbon dioxide removal (CDR) from the atmosphere later during this century, which would entail a period of temperature overshoot. Next to the socio-economic feasibility of large-scale CDR, which remains unclear, the effect on biogeochemical cycles and climate are key to assessing CDR as a mitigation option. Changes in atmospheric CO₂ concentration and climate alter the CO₂ exchange between the atmosphere and the underlying carbon reservoirs of land and the ocean. Here, we investigate carbon cycle feedbacks under idealized and more realistic overshoot scenarios in an ensemble of Earth system models. The response of oceanic and terrestrial carbon stocks to changes in atmospheric CO₂ concentration and changes in surface climate (the carbon-concentration and carbon-climate feedback, quantified by the feedback metrics β and γ, respectively) show a large hysteresis. This hysteresis leads to growing absolute values of β and γ during phases of negative emissions. We find that this growth over time occurs such that the spatial patterns of feedbacks do not change significantly for individual models. We confirm that the β and γ feedback metrics are a relatively robust tool to characterize inter-model differences in feedback strength since the relative feedback strength remains largely stable between phases of positive and negative emissions and between different simulations, although exceptions exist. When emissions become negative, we find that the model uncertainty (model disagreement) in β and γ increases stronger than expected from the assumption that the uncertainties would accumulate linearly with time. This indicates that the model response to a change from increasing to decreasing forcing introduces an additional layer of uncertainty, at least in idealized simulations with a strong signal. We also briefly discuss the existing alternative definition of feedback metrics based on instantaneous carbon fluxes instead of carbon stocks and provide recommendations for the way forward and future model intercomparison projects.
1. Introduction

Estimated remaining carbon budgets compatible with limiting anthropogenic warming to 1.5 or 2 °C above pre-industrial levels are extremely tight and will be exhausted within the next few years if the current emission rate is maintained (e.g., Rogelj et al. 2015; Goodwin et al. 2018; V. Masson-Delmotte et al. 2018; Forster et al. 2023; Smith et al. 2023). Therefore, unless CO$_2$ emissions are reduced immediately at an unprecedented rate, the 1.5 or 2°C targets can only be reached after a period of temperature overshoot (Rogelj et al. 2015; Ricke et al. 2017; Geden and Löschel 2017; Riahi et al. 2021). Although the option to remove large quantities of carbon from the atmosphere remains speculative (Gasser et al. 2015; Larkin et al. 2018; Fuss et al. 2018; Creutzig et al. 2019; Smith et al. 2023), in such overshoot pathways, too large near-term carbon emissions would be compensated by large-scale carbon dioxide removal (CDR) later in this century. Research on negative emissions exploring the reversibility of CO$_2$-induced climate change has been conducted for more than a decade (e.g., Boucher et al. 2012; Wu et al. 2015; Tokarska and Zickfeld 2015; Li et al. 2020; Jeltsch-Thömmes et al. 2020; Yang et al. 2021; Schwinger et al. 2022; Bertini and Tjiputra 2022). These studies generally report a hysteresis behavior of the Earth system under negative emission, resulting in greatly varying reversibility for different aspects of the Earth system. While the surface temperature trend follows a reduction in atmospheric CO$_2$ relatively closely (e.g., Boucher et al. 2012; Jeltsch-Thömmes et al. 2020), hysteresis can be large in the interior ocean, making for example ocean heat content and steric sea level rise as well as interior ocean oxygen content and acidification largely irreversible on policy relevant timescales (Mathesius et al. 2015; Li et al. 2020; Schwinger et al. 2022; Bertini and Tjiputra 2022). The same is true for the loss of carbon from thawing permafrost soils (MacDougall et al. 2015; Gasser et al. 2018; Park and Kug 2022; Schwinger et al. 2022).

Carbon emissions drive multiple responses of the Earth system via changes in its physical climate and the biogeochemical carbon cycle. Under increasing atmospheric CO$_2$ concentrations, increasing carbon uptake by the ocean and terrestrial biosphere slows down global climate change by removing the greenhouse gas CO$_2$ from the atmosphere, a process that is mainly driven by the dissolution of CO$_2$ into the oceans (e.g., Revelle and Suess 1957; Siegenthaler and Oeschger 1978) and the CO$_2$-fertilisation effect on the terrestrial biosphere (Schimel et al. 2015). On the other hand, Earth system model (ESM) simulations show that this carbon uptake is reduced by progressive global warming due to, among others, changes in ocean circulation and a reduction of CO$_2$ solubility in warmer waters, as well as increased respiration rates from soils (Tharammal et al. 2019; Arora et al. 2020; Canadell et al. 2021), and carbon release from degrading permafrost. These two feedback processes, the response to rising CO$_2$ concentrations and the response to climate change, are termed carbon-concentration and carbon-climate feedback, respectively (Gregory et al. 2009). In the context of overshoot pathways, carbon cycle feedbacks determine the efficiency of negative emissions as the oceans and the terrestrial biosphere will first take up carbon at reduced rates and eventually turn into sources of carbon to the atmosphere (Jones et al. 2016a; Schwinger and Tjiputra 2018).

The carbon-concentration and carbon-climate feedbacks can be characterized by feedback metrics, for example, by feedback factors $\beta$ and $\gamma$ (Friedlingstein et al. 2003) that quantify the gain/loss of carbon in terrestrial or marine reservoirs per unit change in atmospheric CO$_2$ concentration and temperature, respectively (see Section 2 for details). These feedback factors are valuable tools to compare the
feedback strength among different models (Friedlingstein et al. 2003, 2006; Yoshikawa et al. 2008; Boer and Arora 2009; Gregory et al. 2009; Roy et al. 2011; Arora et al. 2013, 2020) and can be calculated using idealized model simulations, in which the effect of CO$_2$ on the carbon cycle and the radiative effect of CO$_2$ are decoupled. In a biogeochemically coupled (BGC) simulation, the radiation code of an ESM does not respond to increasing atmospheric CO$_2$ concentrations, but the terrestrial and marine carbon cycles do. There is little climate change in such a simulation, which can therefore be used to quantify the carbon-concentration feedback. The difference between a standard (fully coupled, COU) simulation and the BGC simulation is used to quantify the carbon-climate feedback. In the last two phases of the Coupled Model Intercomparison Project (CMIP5 and CMIP6, Taylor et al. 2012; Eyring et al. 2016) carbon cycle feedbacks were addressed by conducting additional decoupled simulations of the standard 1% CO$_2$ simulation, which prescribes an increase in atmospheric CO$_2$ by 1% per year until quadrupling (Arora et al. 2013, 2020). Next to this idealized simulation, the protocol for the CMIP6 Coupled Climate-Carbon Cycle Model Intercomparison Project (C4MIP, Jones et al. 2016b) also proposes a BGC simulation for the SSP5-3.4-OS (hereafter ssp534-over) scenario (O’Neill et al. 2016). This scenario describes an overshoot pathway, in which emissions increase unmitigated until 2040, but strong mitigation (including CDR) is undertaken thereafter. In contrast to the 1% CO$_2$ simulation, where no forcing other than atmospheric CO$_2$ is varied, the quantification of feedbacks in this scenario simulation is complicated by the presence of land use change and changes in radiative forcing through emissions of aerosols and non-CO$_2$ greenhouse gasses (Melnikova et al. 2021, 2022).

One open question regarding carbon cycle feedbacks under negative emissions is relative to which state the feedbacks should be measured. A sensible definition requires that any gain or loss of carbon is calculated relative to a state where the carbon cycle is in equilibrium. Schwinger and Tjiputra (2018) have opted to keep the pre-industrial state as the reference also after the onset of negative emissions. We follow this approach here, but we note that recently Chimuka et al. (2023) proposed an alternative approach, which defines the feedbacks during the negative emission phase relative to the state at the onset of negative emissions. Since, the Earth system will be in disequilibrium at this point in time, this approach requires an additional simulation that allows to estimate and remove the lagged response of the Earth system to this disequilibrium.

Permafrost soils in the northern high latitudes have accumulated roughly 1100-1700 Pg of carbon in the form of frozen organic matter, about twice as much as currently contained in the atmosphere (Hugelius et al. 2014; Schuur et al. 2015). The release of CO$_2$ and methane (CH$_4$) from thawing permafrost will amplify global warming due to anthropogenic emissions and further accelerate permafrost degradation and microbial decomposition (Feng et al. 2020; Smith et al. 2022). This positive feedback and the fact that Arctic temperatures are increasing at a much faster rate than the global average (Liang et al. 2022; Rantaren et al. 2022) have made permafrost to be considered among the key tipping elements of the climate system, although it may not be an abrupt but irreversible process (Armstrong McKay et al. 2022; Yokohata et al. 2020; Lenton et al. 2019). A temporary temperature overshoot might entail important legacy effects as high latitude ecosystems and the state of permafrost-affected soils take several centuries to adjust to the new atmospheric condition (de Vrese and Brovkin 2021). Current generation ESMs are still in their infancy when it comes to representing the complex and often small-scale processes of permafrost carbon degradation. Here we take advantage of the fact that one of the CMIP6 ESMs considered in this study has a vertically resolved representation...
of soil carbon, which enables us to estimate the contribution of permafrost degradation to the total carbon-climate feedback for this model.

Except for the recent studies by Schwinger and Tjiputra (2018), Melnikova et al. (2021, 2022), and Chimuka et al. (2023) all previous studies that quantify carbon-concentration and carbon-climate feedbacks are based on ESM simulations with increasing atmospheric CO$_2$. Here, we take advantage of a simulation conducted for the CMIP6 Carbon Dioxide Removal Model Intercomparison Project (CDRMIP, Keller et al. 2018) that mirrors the 1% CO$_2$ simulation by prescribing a decrease of atmospheric CO$_2$ by 1% per year. For simplicity, we refer to these two simulations as 1pctCO$_2$-cdr in the following text. We complement this simulation with a BGC simulation (1pctCO$_2$-cdr-bgc) to quantify, in a manner consistent with previous feedback studies (Arora et al. 2013, 2020), carbon-concentration and carbon-climate feedbacks under negative emissions in an ensemble of CMIP6 ESMs. We complement these previous studies by a spatial analysis of feedback patterns, and compare the feedbacks from the positive and negative emission phases of the 1pctCO$_2$-cdr simulation to the feedbacks obtained from the ssp534-over scenario. For the latter, land use change has been shown to have a dominant effect over carbon-concentration or carbon-climate feedbacks by Melnikova et al. (2021, 2022), and these authors present a more detailed analysis of the role of land use change in the ssp534-over scenario. Since land use change is not a feedback process, we focus in this study on regions that are not dominated by agricultural areas when comparing feedbacks between the 1pctCO$_2$-cdr and ssp534-over simulations.

The purpose of this study is to investigate the evolution of carbon cycle feedbacks and their uncertainty under deployment of negative emissions. Since feedback metrics are known to depend on the emission (or concentration) pathway, we investigate the relative feedback strength and the spatial patterns of feedbacks between positive and negative emission phases as well as between idealized and scenario simulations. We also briefly explore the contribution of permafrost carbon losses to the carbon-climate feedback and the impact of alternative feedback metric definitions that rely on instantaneous carbon fluxes rather than carbon stocks in the context of negative emissions.

2. Description of feedback metrics, simulations, and models

2.1 Carbon cycle feedback metrics

The sensitivity of the carbon cycle to changes in atmospheric CO$_2$ concentration ([CO$_2$]) and its sensitivity to changes in physical climate can be measured using two feedback metrics, which can be based on either changes in carbon stocks (as introduced by Friedlingstein et al., 2003) or instantaneous carbon fluxes (as introduced by Boer and Arora 2009). Changes in carbon stocks are equivalent to the time-integrated carbon fluxes across the air-land and air-sea interfaces, such that for the Friedlingstein et al. approach (referred to as integrated flux-based approach), the two feedback metrics are:

1. $\beta$ (PgC/ppm), which quantifies the strength of the carbon-concentration feedback, i.e., the changes in oceanic and terrestrial carbon stocks ($\Delta C_{\text{land,sea}}$) in response to changes in atmospheric CO$_2$ concentration ($\Delta [CO_2]$), and
2. \( \gamma \) (PgC/°C), which measures the strength of the carbon-climate feedback, i.e., changes in land and ocean carbon stocks (\( \Delta C_{L,O} \)) in response to changes in global average surface temperature \((\Delta T)\), where \( \Delta T \) serves as a proxy for climate change.

Carbon feedback analysis requires, in addition to a standard fully coupled (COU) simulation, a biogeochemically (BGC) coupled simulation. In a BGC simulation, atmospheric \([CO_2]\) is kept constant at its pre-industrial values for the radiative transfer calculations, to isolate the response of land and ocean biogeochemistry to rising \([CO_2]\) in the absence of \( CO_2 \)-induced climate change. Using this pair of simulations (COU and BGC) results in the following expressions for \( \beta \) and \( \gamma \) (see Schwinger et al. 2014 for a derivation).

\[
\beta_X = \frac{1}{\Delta[CO_2]} \left( \frac{\Delta C_{BGC}^{COU} \Delta T^{COU} - \Delta C_{BGC}^{COU} \Delta T^{BGC}}{\Delta C_X^{COU} \Delta T^{COU} - \Delta C_X^{BGC} \Delta T^{BGC}} \right) 
\]

(1)

\[
\gamma_X = \frac{\Delta C_{COU}^{COU} - \Delta C_{X}^{BGC}}{\Delta C_{COU}^{COU} - \Delta C_{BGC}^{BGC}} \approx \frac{\Delta C_{X}^{COU} - \Delta C_{X}^{BGC}}{\Delta C_{X}^{COU} - \Delta C_{X}^{BGC}} 
\]

(2)

where \( X \) can be either \( L \) for land or \( O \) for ocean. Although there is no change in the radiative forcing of \( CO_2 \) in the BGC simulation (such that we could expect \( \Delta T^{BGC} = 0 \)), surface temperature can vary due to changes in other radiative forcing agents (aerosols and non-\( CO_2 \) greenhouse gases). Even in the idealized 1pct\( CO_2 \)-ced simulation, where \( CO_2 \) is the only variable forcing, there are some climatic changes over the land surface due to a reduction in latent heat fluxes associated with stomatal closure at higher \( CO_2 \) levels, as well as changes in vegetation structure, coverage, and composition (Arora et al. 2020), which result in a small temperature increase along with changes in precipitation and soil moisture. The assumption of \( \Delta T^{BGC} = 0 \) will simplify equations (1) and (2) such that the rightmost term holds. However, results presented here are calculated using the complete expression for \( \beta \) and \( \gamma \) (without the assumption \( \Delta T^{BGC} = 0 \)), unless otherwise noted. For comparison, we also provide feedback factors calculated using the simplified (rightmost) definition of \( \beta \) and \( \gamma \) in some figures. The instantaneous flux-based approach is equivalent to equations (1) to (2) except that the deviation of the carbon pools \( \Delta X \) are replaced by the instantaneous air-sea or air-land carbon fluxes \( F_X \). To distinguish these feedback metrics from the integrated flux-based ones, the capital letters \( B \) and \( \Gamma \) are used to denote them. The units of \( B \) and \( \Gamma \) are PgCyr\(^{-1}\)ppm\(^{-1}\) and PgCyr\(^{-1}\)°C\(^{-1}\), respectively.

By combining equations (1) and (2) to yield

\[
\beta_X = \frac{1}{\Delta[CO_2]} \left( \frac{\Delta C_{BGC}^{COU} - \gamma_X \Delta T^{BGC}}{\Delta C_X^{COU} - \gamma_X \Delta T^{BGC}} \right) 
\]

(3)

it can be seen that, in order to calculate \( \beta_X \), the carbon stock changes in the biogeochemically coupled simulation are corrected for global mean temperature changes using \( \gamma_X \). Hence, temperature changes in the biogeochemically coupled simulation are fully accounted for as long as the underlying assumption of linearity holds. However, this assumption might be problematic, for example, if the spatial pattern of warming in a biogeochemically coupled scenario simulation arising from non-\( CO_2 \) forcings is very...
different from the warming patterns in the fully coupled simulation, particularly if the sign of the local
temperature change is different from the global average (e.g., local cooling vs. global average warming).
Such effects could become important on regional to local scales and will be discussed in Section 3.4.

It is worth mentioning that these feedback frameworks should be seen as a technique for assessing the
relative sensitivities of models and understanding their differences (i.e., the model uncertainty of the
estimated feedbacks), rather than as absolute measures of invariant system properties (Gregory et al.
2009; Ciais et al. 2013). The values of carbon cycle feedback metrics can vary over time within a model
simulation (e.g., Arora et al. 2013) or between different scenarios (Hajima et al. 2014).

To gain insight into the reasons for differing responses among models, we apply the decomposition of
the simplified expression for $\beta_L$ (Eq. 1, assuming $\Delta T_{BGC} = 0$) following Arora et al. (2020). This allows
us to investigate the contributions from different processes to changes in vegetation and soil carbon
reservoirs ($\Delta C_V$ and $\Delta C_S$, respectively).

$$\beta_L = \frac{\Delta C_{BGC}}{[CO_2]} = \frac{\Delta C_{GPP} + \Delta C_{R}}{[CO_2]} = \left( \frac{\Delta C_{GPP}}{\Delta NPP_{BGC}} \frac{\Delta GPP_{BGC}}{\Delta GPP} \frac{\Delta R_{BGC}}{\Delta R} \frac{\Delta LF_{BGC}}{\Delta LF} \right) [CO_2]$$

$\Delta NPP$, $\Delta GPP$, $\Delta R_h$, and $\Delta LF$ represent deviations of the net primary productivity, gross primary
productivity, heterotrophic respiration, and litterfall flux, respectively, from their pre-industrial values.
The terms $\tau_{cveg}$ and $\tau_{csoil}$ are turnover times (in years) of carbon in the vegetation and litter plus
soil pools. $\frac{\Delta NPP}{\Delta GPP}$ is a measure of carbon use efficiency for the fraction of GPP (above its pre-industrial
value) that turned into NPP after subtracting autotrophic respiration losses (denoted as CUE$_A$).
$\frac{\Delta GPP}{[CO_2]}$ ($\text{PgCyr}^{-1}\text{ppm}^{-1}$) and $\frac{\Delta R_h}{\Delta LF}$ are a measure of the global CO$_2$ fertilization effect and the increase in
heterotrophic respiration per unit increase in litterfall rate, respectively. Also, $\frac{\Delta LF}{[CO_2]}$ ($\text{PgCyr}^{-1}\text{ppm}^{-1}$)
measures the global increase in litterfall rate per unit increase in CO$_2$.

2.2 Model simulations

The 1% CO$_2$ experiment is a highly idealized model experiment that prescribes a rate of 1% per year
increase in [CO$_2$] from pre-industrial values until quadrupling after 140 years. No other forcings are
varied in this experiment, i.e., land use as well as non-CO$_2$ greenhouse gasses and aerosol
concentrations are held constant at their pre-industrial levels. This experiment has already been
performed by the first coupled atmosphere-ocean general circulation models in the late 1980s, and
important climate metrics such as the transient climate response (TCR; Meehl et al. 2020) and the
transient response to cumulative emissions (TCRE; e.g., Gillett et al. 2013) are formally defined through
the 1pctCO$_2$ simulation. Likewise, the CMIP4 carbon cycle feedback analysis for the last two phases of
CMIP (Arora et al. 2013, 2020) relied on this simulation. Given the importance of the 1% CO$_2$
experiment, the CMIP6 CDRMIP protocol proposes an experiment that mirrors this simulation by
starting from its endpoint at year 140 and decreasing atmospheric CO$_2$ at a rate of 1% per year until
pre-industrial [CO$_2$] is restored. This experiment is designed to investigate the response of the Earth
system to carbon dioxide removal in an idealized fashion. As noted above, in this study we refer to the
1% CO$_2$ simulation and the mirrored -1% CO$_2$ CDRMIP simulation collectively as 1pctCO$_2$-cdr for
simplicity. We note that the implied rates of CDR in the 1pctCO$_2$-cdr simulation are huge and not
practically feasible. Also, there is a jump from very large positive to very large negative diagnosed
emissions at the end of year 140, which is clearly unrealistic. To investigate carbon cycle feedbacks
under CDR, we have complemented the 1pctCO$_2$-cdr simulation with a biogeochemical coupled
1pctCO$_2$-cdr-bgc simulation that starts from the endpoint of the 1pctCO$_2$-bgc simulation at year 140.
The family of Shared Socioeconomic Pathways (SSPs, O’Neill et al. 2014) is designed to represent
different socio-economic futures that social, demographic, political, and economic developments could
lead to. These narrative SSPs are the basis for a set of quantitative future scenarios, a selection of which
is now being used as input for scenario simulations by the latest ESMs in the frame of the CMIP6
ScenarioMIP (O’Neill et al. 2016). The ssp534-over scenario follows the high emission SSP5-8.5 pathway
until 2040 at which point strong mitigation policies are introduced to rapidly reduce emissions to zero
by about 2070 and to net-negative levels thereafter (Fig. 3 of O’Neill et al. 2016). In contrast to the
1pctCO$_2$-cdr simulation, the ssp534-over scenario includes land use change as well as time varying
forcing from aerosols and non-CO$_2$ greenhouse gases throughout the simulation period (Fig. 1 of
Liddicoat et al. 2021). For this study, we use the 1pctCO$_2$-cdr and ssp534-over simulations from the
CMIP6 archive together with the corresponding biogeochemically coupled simulations of these
experiments. We note that the biogeochemically coupled 1pctCO$_2$-cdr-bgc experiment is not part of
CMIP6, but has been performed for this study by participating modelling groups.
The CMIP simulation protocol does not allow to separate carbon release (or uptake) through land use
changes from the carbon-concentration feedback, since land use is active in the biogeochemically
coupled ssp534-over simulation. To focus on carbon cycle feedbacks, we eliminate the effect of land
use changes as much as possible by identifying regions that are mostly unaffected by human activity
(referred to as “natural land”). To accomplish this in a way that available CMIP6 output permits, we
define natural land as grid cells with a maximum cropland fraction of less than 25% at all time steps
during the period 2015-2100. The threshold of 25% used here for our heuristic approach is a
compromise between accuracy (some signal of land use change is still present) and spatial coverage
(with increasingly lower thresholds, larger areas of the globe are excluded). Our results are not very
sensitive to variations in the threshold between approximately 10 and 30%. Maps of maximum ssp534-
over cropland fraction for each model (Fig. S1) indicates that a 25% threshold reasonably identifies
hotspots of agricultural production. To make our analysis comparable between the ssp534-over and
1pctCO$_2$-cdr simulations, we use the same set of grid cells also for the 1pctCO$_2$-cdr simulation (unless
otherwise noted), even though land cover is not changed from its pre-industrial state in this simulation.
We acknowledge that our approach does not explicitly address pasture gridcells or transition from other
land use types to pasture. Nonetheless, in the ssp534-over scenario, a substantial expansion of
bioenergy crops between 2040 and 2070 is assumed to replace pasture areas, while the area of land
used as pasture remains relatively stable thereafter (see O’Neill et al. 2016). Hence, our approach, for
this specific scenario, captures the majority of gridcells with transitions from pasture to cropland, while
transitions from pasture to forest remain small.
2.3 Participating Earth System Models

Table 1 summarizes the five ESMs that contributed to this study along with the experiments used for the analyses presented here. The primary features of these models are listed in Table 2 of Arora et al. (2020). MIROC-ES2L, NorESM2-LM (which employs version 5 of the Community Land Model, CLM5), and UKESM1-0-LL have a representation of the terrestrial nitrogen cycle implemented and coupled to their carbon cycle. Only the UKESM1-0-LL model has a land component that dynamically simulates vegetation cover and competition between their plant functional types (PFTs). Fire is included in the CNRM-ESM2-1 and NorESM2-LM models. NorESM2-LM is the only ESM with vertically resolved soil carbon, which allows studying the impact of warming on the carbon stored in permafrost soils in more detail. In this study, a gridcell was considered permafrost where the pre-industrial maximum active layer thickness was shallower than three meters. A description and a comparison of the ocean biogeochemistry models used in the five ESMs can be found in the review of Séférian et al. (2020).

Table 1: List of CMIP6 ESMs used in this study, names of their biogeochemical component models, resolution and experiment variants used.

<table>
<thead>
<tr>
<th>ESM</th>
<th>Atmosphere and land resolution</th>
<th>Ocean resolution</th>
<th>Ocean biogeochemistry model name</th>
<th>Land model name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CanESM5</td>
<td>2.81°x2.81° &amp; 1.4°x1.4°</td>
<td>1° (finer in the tropics)</td>
<td>CMOC (biology); carbonate chemistry follows OMIP protocol</td>
<td>CLASS-CTEM</td>
<td>Swart et al. (2019)</td>
</tr>
<tr>
<td>CNRM-ESM2-1</td>
<td>1.4°x1.4°</td>
<td>1° (finer in the tropics)</td>
<td>PISCESv2-gas</td>
<td>ISBA–CTRIP</td>
<td>Sédérian et al. (2019)</td>
</tr>
<tr>
<td>MIROC-ES2L</td>
<td>2.81°x2.81°</td>
<td>1° (finer close to North Pole and Equator)</td>
<td>OECO2</td>
<td>MATSIRO (physics), VISIT-e (BGC)</td>
<td>Hajima et al. (2020)</td>
</tr>
<tr>
<td>NorESM2-LM</td>
<td>1.9°x2.5°</td>
<td>1° (finer near the Equator)</td>
<td>IHAMOCC</td>
<td>CLM5</td>
<td>Tjiputra et al. (2020); Seland et al. (2020)</td>
</tr>
<tr>
<td>UKESM1-0-LL</td>
<td>1.875°x1.25°</td>
<td>1°</td>
<td>MEDUSA-2.1</td>
<td>JULES-ES-1.0</td>
<td>Sellar et al. (2019)</td>
</tr>
</tbody>
</table>

*CMIP6 experiment variant used across different simulations including: piControl, historical, hist-bgc, ssp585, ssp585-bgc, ssp534-over, ssp534-over-bgc, 1pctCO2, 1pctCO2-bgc, 1pctCO2-cdr, and 1pctCO2-cdr-bgc experiments.
3. Results and Discussion

3.1 Atmospheric CO$_2$, temperature, and carbon fluxes

The atmospheric CO$_2$ concentration ([CO$_2$]) for the concentration-driven ssp534-over scenario peaks at 571 ppm (a doubling of pre-industrial CO$_2$ concentration) in the year 2062 and decreases to 497 ppm in 2100 (Fig. 1a). According to the scenario design (see O’Neill et al. 2016), strong mitigation policies (including deployment of bioenergy with carbon capture and storage (BECCS) and other carbon dioxide removal technologies) start in 2040 resulting in an immediate decrease in the CO$_2$ growth rate that peaks in 2041 (Fig. 1e). In the 1pctCO$_2$-cdr simulation, the prescribed [CO$_2$] is symmetric around its 4xCO$_2$ peak of 1133 ppm in the year 140 (Fig. 1c). The rate of change of the CO$_2$ concentration (Fig. 1e) is very different between ssp534-over and 1pctCO$_2$-cdr experiments. In particular, the CO$_2$ growth rate in the idealized 1pctCO$_2$-cdr experiment has a sudden and large jump from positive to negative values at the transition from the ramp-up to the ramp-down phase.

The five participating ESMs show large differences in global mean surface air temperature change, relative to pre-industrial values, under the ssp534-over simulation (Fig. 1b). Peak temperatures vary from 2°C in NorESM2-LM to 4.35°C in CanESM5. The timing of the global surface air temperature peak varies from 2062 for the MIROC-ES2L and UKESM1-0-LL models to 2100 for CNRM-ESM2-1. After this peak, the temperature declines again (except for CNRM-ESM2-1) reaching end-of-the-century values that range from 1.39°C above pre-industrial in NorESM2-LM to 3.47°C in CanESM5. The multi-model mean global surface air temperature is 2.66°C at the end of the 21st century. The model-mean growth rate of the global surface air temperature (Fig. 1f) plateaus at about 0.05°C/yr between approximately 2030-2050 before it starts to decline to below zero towards the end of the simulation.

Temperature changes in the BGC simulation of ssp534-over are not negligible since the non-CO$_2$ forcing agents as well as land use change do evolve in time in this scenario, in contrast to the idealized 1pctCO$_2$-cdr simulation. Positive peak temperature anomalies range from 0.37°C (CNRM-ESM2-1 in 2098) to 1.29°C (CanESM5 in 2057). UKESM1-0-LL also shows a pronounced negative temperature anomaly during the historical period of the BGC simulation of -0.80°C in the year 1990.
Figure 1: Atmospheric CO$_2$ concentration and surface air temperature changes in the fully coupled (solid lines) and biogeochemically coupled (dashed lines) configurations of the ssp534-over (a,b) and 1pctCO$_2$-cdr (c,d) experiments. The rates of change in the prescribed atmospheric CO$_2$ concentrations is shown in panel e, and the model mean rate of surface temperature change from the fully coupled simulations is shown in panel f. The dotted (solid) vertical lines in panels e and f indicate the peak of the CO$_2$ growth rate (CO$_2$ concentration) in the ssp534-over scenario. Shadings in panel f show the range across the models. An 11-year moving average has been used in panels b, d, and f.
In the 1pctCO2-cdr simulation, the peak temperature anomalies vary from 3.57°C (in year 144) in NorESM2-LM to 6.84°C (in year 151) in CanESM5 (Fig. 1d). Thereafter, temperature anomalies decline to values ranging from 0.29°C in NorESM2-LM to 2.2°C in UKESM1-0-LL at the end of the ramp-down period (year 280). The 1pctCO2-cdr BGC simulation shows, compared to the ssp534-over BGC simulation, smaller temperature anomalies ranging from -0.22°C (CNRM-ESM2-1 in year 149) to 0.79°C (UKESM1-0-LL in year 207). The relatively large magnitude of the temperature anomaly in the ssp534-over BGC simulation (peak warming of 12% - 29% of the peak warming in the fully coupled simulation) suggests that warming due to non-CO2 forcings might contribute substantially to the carbon-climate feedback in the ssp534-over scenario.

For atmosphere-land fluxes, our analysis is complicated by the fact that land use changes are present in the ssp534-over scenario. Here, we focus on comparing fluxes and feedbacks for grid cells that are dominated by “natural land” (see Sec. 2.2 for more details). Note that, for comparability, we consider the same set of grid cells in the 1pctCO2-cdr simulation, even though land cover stays at its pre-industrial state in this simulation. In the ssp534-over simulations, the model-mean annual CO2 fluxes (Fig. 2) continue rising until the rate of change of [CO2] reaches its peak in 2041. After the peak, atmosphere-land and atmosphere-ocean fluxes start to decline rapidly in all models with little time lag. UKESM1-0-LL and MIROC-ES2L simulate negative fluxes (i.e., natural land turns into a carbon source) before the end of the 21st century in the COU simulation (Fig. 2a). Without the effect of CO2 induced warming (BGC simulation, Fig. 2b), only MIROC-ES2L shows a significant carbon source from the terrestrial biosphere before 2100, while the model-mean still shows a sink. In the fully coupled 1pctCO2-cdr experiment, sink-to-source transition of the terrestrial biosphere occurs around year 165 in the model mean, 25 years after the rate of change of [CO2] peaks (Fig. 2c). Consistent with what is seen in the biogeochemically coupled ssp534-over, the sink-to-source transition occurs 10 years later without the effect of warming in the 1pctCO2-cdr-bgc experiment. However, the terrestrial CO2 source at the end of the biogeochemically coupled 1pctCO2-cdr simulation, even though land cover stays at its pre-industrial state in this simulation. We also observe that models which take up more (less) terrestrial carbon during the CO2 ramp-up phase release more (less) carbon towards the end of the CO2 ramp-down phase (1pctCO2-cdr-bgc, Fig. 2c,d), indicating that these models have a larger (smaller) sensitivity (∂C/∂CO2) to both atmospheric CO2 increase and decrease. We therefore interpret the increased source of carbon at the end of the 1pctCO2-bgc simulation as a release of additional carbon that has been taken up in the absence of climate warming during the biogeochemically coupled simulation. The net negative emission phase of the ssp534-over scenario is too short to show this effect in 2100 (where the warming effect still reduces the model mean terrestrial carbon sink).

Likewise, the warming of the world’s oceans in both simulations tends to reduce the carbon uptake or increase the oceanic carbon source. The model spread for atmosphere-ocean carbon fluxes (Fig. 2, panels e to h) appears to be much smaller than for the atmosphere-land fluxes. In the ssp534-over simulation, the ocean remains a sink of carbon in all models until the end of the simulation in 2100. In the 1pctCO2-cdr simulation the ocean turns into a source of CO2 to the atmosphere around year 175, and in the BGC simulation without warming this transition is delayed by 7 years.
Figure 2: Time series of annual mean natural atmosphere-land (a-d) and atmosphere-ocean (e-h) carbon fluxes for the fully and biogeochemically coupled SSP5-3.4-over and 1pctCO2-cdr experiments as indicated in the panel titles. The dotted vertical lines indicate where [CO2] growth rate peaks in each experiment. An 11-year moving average has been used in all panels.

3.2 Global mean carbon cycle feedbacks

3.2.1 Ocean

In the BGC simulation, where the effect of changing atmospheric CO2 concentration on terrestrial and marine carbon uptake (the carbon-concentration feedback) is isolated, cumulative atmosphere-ocean carbon fluxes indicate an almost linear growth with [CO2] as long as atmospheric CO2 concentrations are increasing in both SSP5-3.4-over and 1pctCO2-cdr simulations (Fig. 3a-c). When [CO2] starts to decline, the atmosphere-ocean carbon flux in the 1pctCO2-cdr simulation shows pronounced hysteresis with a continued ocean carbon uptake (until the [CO2]-anomaly has been roughly reduced to 500 ppm) before starting to decrease towards the end of the ramp-down phase (Fig. 3b). In the SSP5-3.4-over-BGC simulation, where the onset of net negative emissions is more gradual, the relationship between cumulative atmosphere-ocean fluxes and [CO2] during the phase of declining atmospheric CO2 concentration also shows hysteresis; but due to the relative short period of net-negative emissions, the ocean remains a sink of carbon in all models until the end of the simulation.
Differences in the cumulative atmosphere-ocean CO$_2$ flux between the COU and the BGC simulations versus surface temperature changes (carbon-climate feedback) are shown in Fig. 3d-f. Increasing temperature results in less carbon uptake by the ocean, except for the CNRM-ESM2-1 which simulates slightly more uptake in the first half of the warming period under the ssp534-over. During the negative emission phases of the simulations when the air surface temperature is decreasing, the carbon-climate feedback still decreases the ocean carbon content, albeit at reduced rates. Even when pre-industrial CO$_2$ concentrations are restored at the end of the 1pctCO$_2$-cdr simulation all models agree that the ocean is still losing carbon due to the effect of (legacy) warming (Fig. 3e). Using the global average ocean potential temperature (averaged over the full ocean depth) instead of the surface air temperature as a proxy for oceanic climate change as proposed by Schwinger and Tjiputra (2018), gives a much more linear relationship between changes in the ocean carbon stock and changes in temperature in the majority of models (Fig. 3 g-i). At the end of the ssp534-over and 1pctCO$_2$-cdr simulations, the ocean still holds a large part of the carbon taken up from the atmosphere since pre-industrial time, between roughly 300-400 PgC in 1pctCO$_2$-cdr, and around 350 PgC in ssp534-over (Fig. S2).

Generally, the ocean carbon-concentration feedback (as indicated by the cumulative carbon uptake per unit increase of CO$_2$ concentration, Fig. 3a-c) is larger in the ssp534-over scenario, which can most likely be explained with the slower growth rate of [CO$_2$] in this scenario compared to the 1pctCO$_2$-cdr simulation (Fig. 3c). For slower growth rates, the ocean has more time to mix and partly transport the adsorbed anthropogenic carbon away from the ocean surface to the interior, increasing the capacity for more uptake. A larger carbon uptake at slower CO$_2$ growth rates has already been reported by Gregory et al. 2009 and Hajima et al. 2014, although only for combined land and ocean fluxes or land fluxes only. The ocean carbon-climate feedback, in contrast, is slightly smaller in the ssp534-over scenario, i.e., the carbon loss for a given warming is smaller.
Figure 3: Ocean carbon cycle feedbacks in the **ssp534-over** (left column) and **1pctCO2-cdr** (middle column) simulations for individual models. The model means for both simulations are shown in the right column. Global mean ocean potential temperature is used on the x-axis of panels (g-i). An 11-year moving average has been used in all panels.
Figure 4: Terrestrial carbon cycle feedbacks in the ssp534-over (left column) and 1pctCO2-cdr (middle column) simulations for grid cells that are dominated by “natural land” (less than a maximum of 25% crop fraction over the period 2015-2100 in ssp534-over). Note that we consider the same grid cells in the 1pctCO2-cdr simulation, even though land use stays at pre-industrial state. The model means for both simulations are shown in the right column. An 11-year moving average has been used in all panels.

3.2.2 Land

For grid cells representing natural land, the response of the cumulative terrestrial carbon flux to changes in [CO2] and surface temperature (Fig. 4) is qualitatively similar to the response of the atmosphere-ocean fluxes. In both ssp534-over and 1pctCO2-cdr simulations, a roughly linear relationship can be seen between the carbon flux change and both the changes in [CO2] and surface air temperature during positive emission phases. An exception is the carbon-climate feedback of the CanESM5 model, which is about zero up to 4 degrees of warming, and becomes positive for higher temperature increases. This unique behavior is caused by CanESM5’s high climate sensitivity combined with larger carbon use efficiency amongst CMIP6 models (as shown later) which causes high latitude vegetation to take up large amounts of carbon in response to warming. This more than compensates for the carbon loss elsewhere associated with climate warming. During negative emission phases both feedbacks show a considerable hysteresis behavior, as for the ocean (see also below). It is worth mentioning that, unlike for the ocean, the COU-BGC accumulated atmosphere-land flux starts to increase, albeit with a lag, in response to cooling during the negative emissions phase in most models (Figs. 3e and 4e).
The carbon-concentration feedback (as indicated by the cumulative carbon uptake per unit increase of CO₂ concentration) is slightly smaller for the ssp534-over scenario compared to the 1pctCO₂-cdr experiment (see Fig. 4c), but this difference might be attributed to the remaining influence of land-use changes. This is because, for “crop-land grid cells” (maximum crop-fraction of more than 25% in the ssp534-over scenario), the cumulative carbon fluxes are markedly smaller in the ssp534-over scenario compared to the 1pctCO₂-cdr simulation (compare panel c on Figs. S3 and 4). This indicates, consistent with Melnikova et al. (2022) who demonstrate that carbon losses from land use changes dominate over gains through CO₂ fertilization in crop dominated areas (see their Fig. 4, panels a and c), that the prescribed land use change in the SSP scenario is the driver behind the small (negative for NorESM2-LM and UKESM1-0-LL) carbon accumulation for crop land grid cells. Since grid cells that are dominated by natural land according to our separation approach, may contain up to 25% croplands, we expect a reduction of cumulative carbon fluxes due the remaining land use (changes) in the natural land grid cells. We note that land use change is externally prescribed rather than a feedback process in our simulations. It is only due to the simulation design used here (see Section 2.2 for details), that the carbon release (or uptake) due to land use changes modifies the net atmosphere-land CO₂ flux which is then seen as a carbon-concentration feedback in the ssp534-over-bgc simulation.

The model-mean carbon-climate feedback for natural land is very similar for the ssp534-over and 1pctCO₂-cdr simulations during the positive emission phases, but deviates thereafter due to hysteresis behavior (Fig. 4f). Interestingly, in contrast to the carbon-concentration feedback, the model-mean carbon-climate feedback for cropland and natural land remains very similar between the ssp534-over and 1pctCO₂-cdr simulations (Fig. S3f). This is likely due to the similar response of the soil carbon to changes in surface air temperature. (deleted: not)

### 3.2.3 Hysteresis

For the 1pctCO₂-cdr simulation, hysteresis can be defined as the difference in, for example, cumulative carbon uptake during the ramp-up and the ramp-down period at the same level of atmospheric CO₂ concentration. Here, to quantify hysteresis, we choose the years 70 and 210, which represent a state where atmospheric CO₂ has been doubled (570 ppm) or returned to this value after the overshoot. We define hysteresis as the difference between cumulative carbon uptake in year 210 minus cumulative carbon uptake in year 70 (i.e., hysteresis is positive, if cumulative carbon uptake is larger on the ramp-down side of the 1pctCO₂-cdr simulation). We refrain from quantifying hysteresis for the ssp534-over scenario, because of the relatively short period of declining [CO₂].

The model mean hysteresis in the carbon-concentration feedback is 443±29 PgC (model uncertainty measured as one standard deviation) for the ocean and 524±205 PgC for natural land, which for both cases is larger than the feedback at year 70 itself. Although the hysteresis of the ocean carbon-concentration feedback is smaller than the terrestrial feedback in absolute terms, it is larger in relative terms (179% of the accumulated carbon uptake at year 70 for the ocean versus 168% for land). In general, the hysteresis seems to be related to the magnitude of the carbon-concentration feedback, since models with a large (small) carbon uptake at year 140, tend to show a large (small) hysteresis at
year 210 for both ocean and land. However, towards the end of the ramp-down period, this relationship breaks down for CanESM5 and MIROC-ES2L, particularly over land.

For the carbon-climate feedback, the hysteresis in climate induced carbon loss or gain (difference between COU-BGC evaluated at years 70 and 210) is -102±22 and -158±181 PgC for ocean and natural land, respectively. As for the carbon-concentration effect, a relationship between the magnitude of carbon loss or gain at year 140 and the hysteresis is found. Models with a large (small) climate induced loss of carbon tend to have a large (small) hysteresis.

For the ocean carbon cycle, hysteresis in the carbon-concentration feedback occurs mainly due to the long time scales of ocean overturning circulation. Schwinger and Tjiputra (2018) have shown that hysteresis strongly increases with water mass age. Young waters, which reside close to the ocean surface, exchange quickly with the atmosphere and show little hysteresis, whereas old, deep ocean water masses’ responses to declining atmospheric CO$_2$ will be delayed, and thus show considerable hysteresis. Over land, both the vegetation and soil carbon pools show a lagged response to decreasing CO$_2$ due to the fact that transient changes in [CO$_2$] lead to a long term disequilibrium between the CO$_2$ fertilization effect, vegetation biomass, litterfall, and soil carbon (e.g., Krause et al. 2020). Therefore, despite declining [CO$_2$] levels at the beginning of the ramp-down phase there is still an increase in vegetation biomass due to CO$_2$ fertilization, and consequently an increase in soil carbon due to still increasing litterfall. Warming-induced hysteresis appears to be larger for soil carbon in most models. Similar to the large warming induced hysteresis (e.g., Schwinger and Tjiputra 2018; Schwinger et al. 2022; Santana-Falcón et al. 2023) in the ocean, this is caused by the fact that even though warming levels start to decline shortly after the onset of the ramp-down phase, environmental conditions remain warmer than in the pre-industrial period over the whole time of the ramp-down simulation.

3.3 Carbon cycle feedback metrics

3.3.1 Model mean global land and ocean responses

We now discuss the model-mean time evolution of the feedback metrics $\beta$ and $\gamma$ (Eqs. 1 and 2) derived from the 1pctCO$_2$-cdr and ssp534-over simulations. In the ssp534-over scenario (Fig. 5a) the model-mean feedback metric $\beta_L^*$ increases monotonically from about 0.7 to 1.9 PgC ppm$^{-1}$ during the period 2000-2100. Over the ocean, $\beta_L^*$ in the ssp534-over scenario decreases slightly until the mid-21st century, and then it rises to about 1.7 PgC ppm$^{-1}$. Due to the much larger spread in carbon fluxes over land (Fig. 2), the resulting model spread for both $\beta_L$ and $\gamma_L$ is also much larger than for $\beta_o$ and $\gamma_o$.

For the 1pctCO$_2$-cdr simulation, during the ramp-up phase over both land and ocean (Fig. 5b), $\beta$ initially increases and then decreases slightly with increasing [CO$_2$] consistent with the results of Arora et al. (2013) for the same experiment but using CMIP5 ESMs. In contrast, during the ramp-down phase of the 1pctCO$_2$-cdr experiment, $\beta$ reaches very high values over both land and ocean (Fig. 5c). This is because, during the carbon removal phase of the 1pctCO$_2$-cdr experiment, there is a much larger amount of accumulated ocean and terrestrial carbon for the same atmospheric CO$_2$ concentration due to the large hysteresis seen in Figs. 3 and 4. Eventually, while [CO$_2$] is approaching pre-industrial values (i.e., $\Delta$(CO$_2$) reaches zero), changes in cumulative fluxes (i.e., carbon stocks) relative to their pre-industrial values remain positive, making $\beta$ ill-defined towards the end of the 1pctCO$_2$-cdr ramp-down. For the same
reason, an increase of $\beta_L$ and $\beta_O$ is also seen in the ssp34-over scenario after the CO$_2$ concentration peak in 2062.

**Figure 5**: Model-mean $\beta$ (a-c) and $\gamma$ (d-f) feedback metrics in the ssp34-over and 1pctCO$_2$-cdr experiments for natural land and ocean. Panels b and e show a zoom into the ramp-up phase of the time series shown on panels c and f. Shadings show the range across the models. The dotted vertical line on panel a indicates where [CO$_2$] growth rate peaks in the fully coupled ssp34-over experiment. Dotted curves on panel d indicate the model mean with the assumption of negligible temperature change in the BGC simulation (Eq. 2). An 11-year moving average has been used in all panels.

The model mean feedback factor $\gamma$ is negative as the impact of climate change generally reduces the carbon stocks of land and ocean. In both ssp34-over and 1pctCO$_2$-cdr experiments, the carbon-climate feedback is increasing over time (more negative $\gamma_L$), similar to figure 6 of Arora et al. (2013). The carbon-climate feedback is generally much smaller for the ocean than for land, and the model uncertainty for $\gamma_O$ is only a small fraction of $\gamma_L$. Note that the same globally averaged surface air temperature anomaly is being used for the calculation of both $\gamma_O$ and $\gamma_L$ (Eq. 2). As noted above, the CanESM5 model simulates a globally increasing land uptake due to climate change towards the end of the 1pctCO$_2$-cdr simulation (Fig. 4e), resulting in a positive $\gamma_L$ for this model. During the ramp-down phase of the 1pctCO$_2$-cdr experiment (Fig. 5f), $\gamma$ reaches very large negative values. Similar to $\beta$, this is caused by the large hysteresis of the climate change impact on cumulative carbon stock while the surface temperature change becomes small (see Eq. 2). The assumption of $\Delta T^{BGC} = 0$ generally works...
well except for $\gamma_L$ in the ssp534-over scenario where non-CO$_2$ forcings have a significant contribution to $\Delta T^{BGCC}$ (dashed curves in Fig. 5d).

The global feedback factors $B$ and $\Gamma$ for the ssp534-over and 1pctCO$_2$-cdr simulations are shown in Fig. S4. This feedback metric directly reflects the instantaneous fluxes, not cumulative fluxes, and is therefore less influenced by the history of carbon fluxes, unlike $\beta$ and $\gamma$. Consistent with Fig. 2, the model-mean $B$ remains positive during the ssp534-over simulation and during the positive emission phase of the 1pctCO$_2$-cdr both over natural land and ocean. Only one model indicates a negative carbon-concentration feedback over natural land towards the very end of the ssp534-over simulation during its relatively short negative emission phase. $B$ reflects the saturation of carbon sinks in the 1pctCO$_2$-cdr simulation with time and decreases monotonically during the positive emission phase. Similar to what we have seen earlier for $\beta$, $B$ shows large but negative values towards the end of the 1pctCO$_2$-cdr ramp-down phase (Fig. S4c).

An interesting difference between the $\gamma$ and $\Gamma$ feedback metrics is seen towards the end of the 1pctCO$_2$-cdr negative emissions phase (Fig. S4f), where $\Gamma$ turns positive around year 180. This indicates that the land biosphere starts gaining carbon that was previously lost due to the impacts of climate change. In contrast, $\Gamma_O$ remains negative indicating that the ocean continues to lose carbon due to warmer than pre-industrial conditions until the end of the 1pctCO$_2$-cdr ramp-down phase. Because they are based on cumulative emissions, both $\gamma_O$ and $\gamma_L$ remain negative throughout the 1pctCO$_2$-cdr ramp-down. This illustrates that the use of a feedback metric based on time-integrated carbon fluxes might obscure changes in important processes during net-negative emission phases. Eventually, both approaches for calculating feedback metrics become ill-defined when the deviation of [CO$_2$] or temperature from their pre-industrial values becomes small. This implies that both feedback metrics are not suited to describe feedbacks towards the end (and beginning) of a concentration driven simulation set-up where pre-industrial concentrations are restored. We note that this problem is connected to the choice of the reference relative to which the feedbacks are calculated. In the approach of Chimuka et al. (2023), where the reference is chosen to be at the transition from positive to negative emissions, singularities towards the end of the 1pctCO$_2$-cdr simulation are avoided.

### 3.3.2 Model uncertainties and relative feedback strength in global feedback metrics

Figure 6 shows the model spread of feedback metrics at different points in time for the 1pctCO$_2$-cdr simulation and the ssp534-over scenario (see also Table 2). The larger model-mean values during the negative emission phases have been discussed in the previous section, but Fig. 6 also shows a strong increase in model uncertainty (measured as the standard deviation around the model mean, Table 2) between the ramp-up and ramp-down phase of the 1pctCO$_2$-cdr simulation. For both $\beta_L$ and $\beta_O$, there is either no ($\beta_O$) or only a small ($\beta_L$) increase in model uncertainty between the years 70 and 140 of the 1pctCO$_2$-cdr simulation, whereas at year 210 uncertainty has increased by about a factor of four. This “jump” in uncertainty in $\beta$ is solely caused by differences in how the models react to the sharp change in forcing from increasing to decreasing CO$_2$ at year 140 (see Eq. 1, note that atmospheric CO$_2$ is prescribed and $\Delta T^{BGCC}$ is small). A similar behavior is seen for $\gamma_O$, while for $\gamma_L$ the increase in model
uncertainty is more gradual, i.e., the increase between years 70 and 140 is about the same as between years 140 and 210. There is also a consistent increase in model uncertainty in all feedback metrics from the positive to the negative emissions phase in the ssp534-over scenario.

The relative strength of the feedback among the models remains relatively stable over time, between positive and negative emission phases, and between the different experiments. Model A having a stronger (weaker) feedback than model B at one of the instances depicted in Fig. 6, indicates that model A will have a stronger (weaker) feedback than model B for the other instances with only few exceptions. Most of these exceptions arise because modeled feedbacks are very similar such that small changes in feedback strength can lead to a different ranking. In a few cases relative feedback strength evolves differently in the models. For example, NorESM2-LM evolves from having a weaker than average $\gamma_L$ in the positive emission phase of the 1pctCO$_2$-cdr simulation to having a stronger than average $\gamma_L$ in the negative emission phase.

Finally, it is worth noting that while the model uncertainty in $\gamma_o$ is much smaller than in $\gamma_L$ during the ramp-up phase of the 1pctCO$_2$-cdr simulation (uncertainty in $\gamma_o$ is only 15% of those in $\gamma_L$ at year 140), this situation has changed for the ramp-down phase. At year 210, the uncertainty in the ocean carbon-climate feedbacks has grown much stronger than the uncertainties of the terrestrial carbon-climate feedback, such that model uncertainties in $\gamma_o$ are 42% of those in $\gamma_L$. 
Figure 6: Globally averaged values of $\beta$ (a and c) and $\gamma$ (b and d) in the 1pctCO$_2$-cdr (years 70, 140, and 210) and ssp534-over (years 2045 and 2090) experiments for natural land and ocean. The bars show the mean ± 1 standard deviation range, and the individual colored dots represent individual models.

Table 2: Globally averaged values of $\beta$ (Pg C ppm$^{-1}$) and $\gamma$ (Pg C °C$^{-1}$) at years 70, 140, and 210 of the 1pctCO$_2$-cdr simulation and years 2045 and 2090 of the ssp534-over experiment for natural land and ocean.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\beta_{L(70)}$</th>
<th>$\beta_{L(140)}$</th>
<th>$\gamma_{L(70)}$</th>
<th>$\gamma_{L(140)}$</th>
<th>Mean $\beta$</th>
<th>Mean $\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIROC-ES2L</td>
<td>1.15</td>
<td>0.94</td>
<td>1.02</td>
<td>0.76</td>
<td>1.11 (SD=0.16)</td>
<td>0.99 (SD=0.23)</td>
</tr>
</tbody>
</table>
### 3.3.3 Model differences in the terrestrial carbon-concentration feedback

Figure 7 shows the individual components of the decomposition of $\beta$ (Eq. 4), separately for tropical and subtropical (30°S-30°N) and higher latitudes (between 30°N/S and the poles), both on the ramp-up and ramp-down phases (years 70 and 210, respectively) of the 1pctCO$_2$-cdr-bgc experiment. The time periods are selected such that the atmospheric CO$_2$ concentration is the same (569 ppm, a doubling of pre-industrial CO$_2$ concentration). All models consistently show increases in both $r_{veg}^{a}$ and $r_{soild}^{a}$ during the ramp-down compared to the ramp-up phase, since these metrics are based on cumulative

<table>
<thead>
<tr>
<th>$L(210)$</th>
<th>$\beta_{L(210)}$</th>
<th>$\gamma_{L(210)}$</th>
<th>$\beta_{O(210)}$</th>
<th>$\gamma_{O(210)}$</th>
<th>$\beta_{O(2090)}$</th>
<th>$\gamma_{O(2090)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.48</td>
<td>2.43</td>
<td>4.51</td>
<td>3.05</td>
<td>2.88</td>
<td>3.07 (SD=0.85)</td>
</tr>
<tr>
<td>$L(2045)$</td>
<td>1.04</td>
<td>0.56</td>
<td>1.59</td>
<td>0.88</td>
<td>0.57</td>
<td>0.93 (SD=0.42)</td>
</tr>
<tr>
<td>$L(2090)$</td>
<td>1.55</td>
<td>1.22</td>
<td>2.63</td>
<td>1.59</td>
<td>1.14</td>
<td>1.63 (SD=0.59)</td>
</tr>
<tr>
<td>$\gamma_{L(70)}$</td>
<td>-42.14</td>
<td>-19.54</td>
<td>-41.58</td>
<td>0.82</td>
<td>-14.12</td>
<td>-23.31 (SD=18.5)</td>
</tr>
<tr>
<td>$\gamma_{L(140)}$</td>
<td>-44.17</td>
<td>-31.19</td>
<td>-74.97</td>
<td>17.78</td>
<td>-16.31</td>
<td>-29.77 (SD=34.3)</td>
</tr>
<tr>
<td>$\gamma_{L(210)}$</td>
<td>-77.26</td>
<td>-60.45</td>
<td>-144.01</td>
<td>12.77</td>
<td>-71.64</td>
<td>-68.12 (SD=55.8)</td>
</tr>
<tr>
<td>$\gamma_{L(2045)}$</td>
<td>-41.08</td>
<td>-6.80</td>
<td>-48.46</td>
<td>-10.93</td>
<td>-15.78</td>
<td>-24.61 (SD=18.9)</td>
</tr>
<tr>
<td>$\gamma_{L(2090)}$</td>
<td>-56.43</td>
<td>-34.76</td>
<td>-69.66</td>
<td>-10.41</td>
<td>-25.95</td>
<td>-39.44 (SD=23.7)</td>
</tr>
<tr>
<td>$\beta_{O(70)}$</td>
<td>0.85</td>
<td>0.93</td>
<td>0.82</td>
<td>0.92</td>
<td>0.90</td>
<td>0.88 (SD=0.05)</td>
</tr>
<tr>
<td>$\beta_{O(140)}$</td>
<td>0.76</td>
<td>0.81</td>
<td>0.73</td>
<td>0.81</td>
<td>0.82</td>
<td>0.78 (SD=0.04)</td>
</tr>
<tr>
<td>$\beta_{O(210)}$</td>
<td>2.49</td>
<td>2.60</td>
<td>2.35</td>
<td>2.50</td>
<td>2.70</td>
<td>2.53 (SD=0.13)</td>
</tr>
<tr>
<td>$\beta_{O(2045)}$</td>
<td>1.08</td>
<td>1.09</td>
<td>0.97</td>
<td>1.09</td>
<td>1.08</td>
<td>1.06 (SD=0.05)</td>
</tr>
<tr>
<td>$\beta_{O(2090)}$</td>
<td>1.59</td>
<td>1.57</td>
<td>1.43</td>
<td>1.55</td>
<td>1.60</td>
<td>1.55 (SD=0.07)</td>
</tr>
<tr>
<td>$\gamma_{O(70)}$</td>
<td>-10.09</td>
<td>-7.95</td>
<td>-3.60</td>
<td>-10.13</td>
<td>-10.84</td>
<td>-8.52 (SD=2.96)</td>
</tr>
<tr>
<td>$\gamma_{O(140)}$</td>
<td>-22.40</td>
<td>-14.56</td>
<td>-9.44</td>
<td>-16.77</td>
<td>-20.48</td>
<td>-16.61 (SD=5.10)</td>
</tr>
<tr>
<td>$\gamma_{O(210)}$</td>
<td>-58.94</td>
<td>-29.78</td>
<td>-20.16</td>
<td>-32.75</td>
<td>-76.28</td>
<td>-43.59 (SD=23.2)</td>
</tr>
<tr>
<td>$\gamma_{O(2045)}$</td>
<td>-7.88</td>
<td>-5.43</td>
<td>0.78</td>
<td>-6.75</td>
<td>-9.56</td>
<td>-5.77 (SD=3.96)</td>
</tr>
<tr>
<td>$\gamma_{O(2090)}$</td>
<td>-14.50</td>
<td>-11.98</td>
<td>-1.10</td>
<td>-11.05</td>
<td>-21.50</td>
<td>-12.03 (SD=7.35)</td>
</tr>
</tbody>
</table>
vegetation and soil carbon (Eq. 4), which are slower than NPP and GPP in reacting to decreasing [CO$_2$].

Lower (higher) latitudes are associated with higher $\tau_{\text{vegA}}$ ($\tau_{\text{CO2lat}}$). Likewise, the litterfall term $\frac{\Delta LFF}{[\text{CO}_2]}$ is larger during the ramp-down phase in all models due to lagged reaction of vegetation carbon to the decrease in [CO$_2$], with this effect being generally most pronounced at low latitudes. There is also a consistent but small increase in the term $\frac{\Delta GPP}{[\text{CO}_2]}$, which represents the CO$_2$ fertilization effect. This increase implicitly includes the effect of changes (typically an increase) in standing vegetation biomass and leaf area index for all models but also changes in vegetation cover as [CO$_2$] varies for UKESM1-0-LL that simulates dynamic vegetation cover. For the dimensionless fractions $\frac{\Delta R}{\Delta LFF}$ and CUE$_{\Delta}$, changes between ramp-up and ramp-down phases are less consistent between the models. For CUE$_{\Delta}$, three models show an increase and two models a decrease, although the changes between ramp-up and ramp-down phases are generally small. For $\frac{\Delta R}{\Delta LFF}$, changes range from a 115% increase (CNRM-ESM2-1) at low latitudes to a small decrease (UKESM1-0-LL). It is worth noting that for four out of six terms of Eq. 4 ($\tau_{\text{vegA}}, \tau_{\text{CO2lat}}, \frac{\Delta R}{\Delta LFF}$ and $\frac{\Delta LFF}{[\text{CO}_2]}$) the model disagreement is significantly larger during the ramp-down phase of the 1pctCO$_2$-cdr simulation, indicating that changes in these processes are responsible for the strong increase in model uncertainty in $\beta_L$ between positive and negative emission phases pointed out in the previous section.

The decomposition applied here helps to understand some of the model differences visible in Fig. 4. As already pointed out in Arora et al. (2020), the high accumulation of terrestrial carbon by the CNRM-ESM2-1 model in the BGC simulation (Fig. 4b) is not caused by a particularly strong CO$_2$ fertilization effect or CUE$_{\Delta}$ but rather by relatively high values of $\tau_{\text{vegA}}$ and $\tau_{\text{CO2lat}}$, indicating long residence timescales in vegetation and soil. Likewise, CanESM5’s higher than average atmosphere-land C flux (Fig. 4b), despite its near-average strength of the CO$_2$ fertilization effect and soil and vegetation turnover times is due to its high CO$_2$ fertilization effect at lower latitudes and also its high CUE$_{\Delta}$ through which the model converts a much larger fraction of GPP to NPP. Compared to the other models, CanESM5 shows the largest relative increase (85% and 134% for lower and higher latitudes, respectively) in $\tau_{\text{CO2lat}}$ between years 70 and 210.
Figure 7: Individual terms of Eq. (4) contributing to $\beta_C$. Values for tropical and subtropical (between 30°S and 30°N) regions are in green, and for northern and southern latitudes (above 30°S and 30°N) are in blue. Lighter (darker) color on each panel corresponds to the middle of the ramp-up (ramp-down) phase of the 1pctCO$_2$-cdr-bgc experiment (years 70 and 210, respectively).

3.3.4 Northern hemisphere high-latitude permafrost and non-permafrost regions

Of the ESMs considered here, only NorESM2-LM has a terrestrial model that vertically resolves soil carbon (CLM5, Lawrence et al. 2019). Since this is a prerequisite to skillfully simulate carbon release during gradual permafrost degradation, we restrict our analysis of high latitude and permafrost feedbacks to the NorESM2-LM model. If only natural land is considered, the area associated with permafrost and non-permafrost regions north of 45°N is about 14.7 and 17.5 x10$^6$ km$^2$, respectively (total area is 14.7 and 24.1 x10$^6$ km$^2$).

The effect of warming on carbon uptake in the high-latitude non-permafrost region is positive ($\gamma > 0$, increased uptake) in NorESM2-LM in both the ssp534-over and the 1pctCO$_2$-cdr simulation (Fig. 8a-c, blue lines). Within the permafrost region, $\gamma$ is close to zero for the ssp534-over simulation up to 2100 and the ramp-up phase of the 1pctCO$_2$-cdr simulation (Fig. 8a-b, red line), albeit with a decreasing (more negative) trend. This is due to a compensation of vegetation carbon gain and soil carbon losses (Fig. 8). During the ramp-down phase of the 1pctCO$_2$-cdr simulation, permafrost soil carbon losses increase approximately until year 210 of the simulation (Fig. S5). Thereafter, permafrost soil carbon stays roughly constant with a cumulative loss of about 55 PgC over the simulation. Vegetation carbon over the permafrost region still increases for the first 30 years of the ramp-down phase of the 1pctCO$_2$-cdr...
simulation, after which it decreases mainly due to decreasing temperature (Fig. S5g). The $\gamma$ value calculated for the permafrost region, therefore, shows a sharp decrease during the ramp-down period of the 1pctCO$_2$-cdr simulation (Fig. 8c). Eventually, when $\Delta \mathcal{T}$ approaches small values $\gamma$ loses its significance as seen before for the global feedback factors.

**Figure 8:** $\gamma$ (a-c) and $\beta$ (d-f) for northern hemisphere high latitude (above 45°N) natural land permafrost and non-permafrost regions in the ssp534-over and 1pctCO$_2$-cdr simulations using the NorESM2-LM model. An 11-year moving average has been used in all panels.

In both the ssp534-over scenario and the 1pctCO$_2$-cdr simulations, $\beta$ is positive (except initially in the ssp534-over simulation) although the absolute values remain very small. The carbon-concentration feedback is stronger over the non-permafrost area, where both soil and vegetation carbon increase, than over the permafrost area, where soil and vegetation carbon stay almost constant in the BGC simulation (Fig. S5).

NorESM2-LM has the smallest transient climate response (TCR) of the models considered here, and it can be expected that the permafrost carbon-climate feedback estimated here would be larger in a model with higher TCR. Nevertheless, the permafrost carbon loss of 26.9 Pg C °C$^{-1}$ in the year 210 of the simulation contributes 38% of the total carbon-climate feedback at this point in time in NorESM2-LM.
3.4 Geographical pattern of carbon cycle feedback metrics

We have calculated $\beta$ and $\gamma$ feedback factors at grid-scale to assess the spatial patterns of feedbacks over the land and ocean (Figs. 9 and 10). In order to compare positive and negative emission phases, we selected 21-year time intervals centered at years 70 and 210 of the ramp-up and ramp-down phases of the 1pctCO$_2$-cdr simulation, at an atmospheric CO$_2$ concentration of 570 ppm (corresponding to a doubling of pre-industrial CO$_2$ concentration). We also selected a 21-year time-interval centered at year 2045 (corresponding to CO$_2$ concentration of 523 ppm), shortly before the CO$_2$ peak of the SSP534-over scenario. We have also analyzed a 21-year time interval during the net-negative emission phase of the SSP scenario (centered at year 2090), but since the time-period of net-negative emissions in the SSP-scenario is relatively short, we focus on comparing the feedbacks during the positive and negative emission phases of the 1pctCO$_2$-cdr simulation alongside with the feedbacks during the positive emission phase of SSP534-over. For completeness, Fig. S6 shows the spatially resolved feedback during the net-negative emission phase of SSP534-over.

In the 1pctCO$_2$-cdr simulation, rising [CO$_2$] increases the modeled carbon sinks almost everywhere (i.e., positive $\beta$) over the land and ocean (Fig. 9a-e). CanESM5 shows a weak negative $\beta$ over northern high-latitude land areas, and there are some spurious negative values of $\beta$ over desert areas in some models. For the ocean, all models agree that the regions with the strongest increase of the oceanic CO$_2$ sinks in response to higher [CO$_2$] are the North Atlantic and the Southern Ocean. As seen for the global average (Fig. 5), $\beta$ remains positive and increases in magnitude during the ramp-down phase (Fig. 9 f-j, note the different color scale). As an overarching observation, the large scale patterns of the carbon-concentration feedback are remarkably similar during the ramp-up and ramp-down phases of the 1pctCO$_2$-cdr simulation (with spatial correlations, averaged across all the models, of 0.93 and 0.80 over land and the ocean, respectively) but the magnitude of the feedback is about two times larger in the ramp-down phase, consistent with the lagged response of cumulative carbon uptake to the decrease in atmospheric CO$_2$ (Figs. 3 and 4). The most prominent change in the spatial pattern of $\beta$ occurs in the equatorial Pacific. All models consistently show that this area has turned from a cumulative carbon sink at year 70 to a cumulative carbon source at year 210.

We find the largest values of $\beta$ over tropical land and to a lesser extent over northern hemisphere temperate and boreal ecosystems coincident with areas of large biomass (forests). For three of the models (NorESM2-LM, CanESM5, and UKESM1-0-LL), the feedback is clearly dominated by tropical and subtropical regions, while for MIROC-ES2L the feedback is approximately of the same strength in northern temperate and high-latitude regions. For CNRM-ESM2-1, the carbon-concentration feedback is on average stronger north of 30° latitude than in tropical/subtropical regions. For NorESM2-LM and UKESM1-0-LL, the tropical dominance of the carbon-concentration feedback stems from vegetation carbon, while for CanESM5 both vegetation and soil carbon contribute about equally (Figs. S7 and S8).

The results presented in Section 3.3.3 provide to some extent a mechanistic understanding of these model differences. CNRM-ESM2-1 has the highest CO$_2$ fertilization effect $\frac{\Delta GPP}{[CO_2]}$ in high latitudes and...
the lowest CUE_{Δ} at low latitudes. This, combined with a large high-latitude τ_{soilΔ} leads to a larger carbon accumulation in vegetation and soil in higher latitudes than in the tropics/subtropics in this model. The three models with tropical dominance of β (NorESM2-LM, CanESM5, and UKESM1-0-LL) have a relatively high τ_{regΔ} and relatively low τ_{soilΔ}. CanESM5, shows the strongest tropical/subtropical CO₂ fertilization effect, but also a large response of the litterfall term leading to large responses in both vegetation and soil carbon.

In the ssp534-over simulation, the ocean β magnitude is similar to that of the 1pctCO₂-cdr simulation and the spatial distribution of the ocean response to the [CO₂] rise is roughly consistent between the models (Fig. 9k-o). In contrast, the feedback pattern over natural land is different in some regions and models between the SSP scenario simulation and the idealized 1pctCO₂-cdr experiment. UKESM1-0-LL, CanESM5, and to a lesser extent NorESM2-LM project negative β values in some northern high latitude regions (e.g., Siberia). These negative β values are either not seen at all (UKESM1-0-LL, NorESM2-LM) or are weaker (CanESM5) in the 1pctCO₂-cdr simulation, and they originate from a combination of vegetation and soil carbon pools (Figs. S7 and S8). Unlike in the 1pctCO₂-cdr experiment, temperature changes are not negligible in the BGC simulation of the ssp534-over experiment (Fig. 1). Furthermore, the spatial pattern of temperature changes is very different for some models, particularly for UKESM1-0-LL, NorESM2-LM, and CNRM-ESM2-1, which show local cooling that is not present (or much weaker) in the fully coupled simulations (Fig. S9). This cooling (and other changes in surface climate related to non-CO₂ forcings) lead to local carbon losses and negative β-values in UKESM1-0-LL and NorESM2-LM in northern high latitudes. In addition, according to Eq. 3, these negative values are reinforced by positive γ-values in this region and a positive global mean temperature change in ssp534-over in these models (see Eq. 3). In contrast, CNRM-ESM2-1 does not show negative values of β in northern high latitudes (despite local cooling), which can be explained by much larger β-values to begin with, and a smaller (and negative) temperature sensitivity γ in high latitudes.
Figure 9: The spatial distribution of $\beta$ (kg C m$^{-2}$ ppm$^{-1}$) at year 70 of the ramp-up phase of the 1pctCO$_2$-cdr simulation (a-e), at year 210 of the ramp-down phase of the 1pctCO$_2$-cdr simulation (f-j), and at year 2045 (natural land only, gray areas are crop-dominated grid cells) during the positive emission phase of the ssp534-over scenario (k-o).

Figure 10 indicates that the ESMs considered here simulate predominantly negative values of $\gamma_e$ over the ocean. Positive values of $\gamma_e$ are found in the Arctic, and in the Southern Ocean most models simulate a banded pattern of positive (adjacent to Antarctica), negative (centered between 60 and 50°S), and positive (between approximately 50 and 40°S) values. In the region adjacent to Antarctica, climate change increases the ocean CO$_2$ sink mainly due to a reduction in sea ice coverage (Roy et al.,...
The North Atlantic Ocean and the Southern Ocean have the largest negative values due to changes in ocean circulation and deep water formation. In tropical and subtropical ocean regions, the reduced oceanic carbon uptakes can be attributed to warming-induced decreased CO₂ solubility and increased stratification (Roy et al. 2011).

Over land, climate change generally reduces carbon sinks in the tropics and mid-latitudes. In the high latitudes models disagree on the strength and the sign of the carbon-climate feedback. CNRM-ESM2-1 shows relatively strong soil carbon losses in northern high latitudes, which overcome vegetation carbon gains (Figs. S10 and S11) leading to mostly negative values of γ_L in this region. As mentioned above, CanESM5’s carbon-climate feedback switches from weakly negative at 2xCO₂ to positive at 4xCO₂.

Figure 10c clearly shows that the positive global γ values originate from the northern hemisphere high latitudes. Also, the positive γ_L in CanESM5 over the northern high latitudes is seen in both vegetation and soil carbon reservoirs, but with a time lag for soil carbon. Consistent with our analysis in Sect. 3.3.4, NorESM2-LM shows permafrost carbon loss in north-eastern Siberia and northern Alaska, but these losses become significant only during the ramp-down phase of the 1pctCO₂-cdr simulation (Fig. 10).

The spatial pattern of the carbon-climate feedback is similar during the ramp-up and ramp-down phases of the 1pctCO₂-cdr simulation, but the magnitude has roughly doubled during the ramp-down phase, consistent with the cumulative nature of the γ feedback metric used here (note the different color-scales in Fig. 10). The correlations of the spatial patterns (at years 70 and 210) are lower than for β and range from 0.41 (MIROC-ES2L) to 0.66 (UKESM1-0-LL) for γ₀ and from 0.49 (NorESM2-LM) to 0.88 (UKESM1-0-LL) for γ_L.

The value of the γ feedback metric in the SSP5-3.4-OS scenario simulation is less affected by land-use change, since the same land-use changes are imposed in both the COU and the BGC simulation. In contrast to β, which is directly altered by carbon stock changes due to land-use changes, γ is only influenced indirectly, possibly by different sensitivities of the new vegetation cover after a land-use transition, or by changes in local to regional climatic conditions. In the global mean, the carbon-climate feedback during the positive emission phase is very similar for the SSP scenario and the 1pctCO₂-cdr simulation (Fig. 5d and e). Also, the spatial patterns of γ_L are largely similar between the SSP5-3.4-OS and the ramp-up phase of the 1pctCO₂-cdr simulation with correlations ranging from 0.71 (NorESM2-LM) to 0.84 (CNRM-ESM2-1). The largest difference between the two simulations is an enhanced positive feedback over northern high-latitude land in the UKESM1-0-LL model in the SSP scenario compared to the 1pctCO₂-cdr simulation, which is seen in both vegetation and soil carbon pools (Figs. S10 and S11). These differences are related to the negative β-values (discussed above) for these models, which make the carbon gain due to warming (the difference ΔC_COU − ΔC_BGC) considerably larger than in the 1pctCO₂ simulation. Again, this is reinforced by the fact that the global average temperature change in the SSP5-3.4-OS simulation is positive and thus (ΔT_COU − ΔT_BGC) is smaller than the actual (local) temperature differences. This indicates that, if the global mean temperature change due to non-CO₂ forcings does not broadly reflect local changes correctly (e.g., local cooling vs. global warming), regional scale feedback factors might show unexpected results.
Over the ocean the global mean carbon-climate feedback is slightly smaller in ssp534-over compared to the 1pctCO\textsubscript{2}-cdr simulation (Fig. 3f), but again, the spatial pattern is largely similar with correlations ranging from 0.47 (CNRM-ESM2-1) to 0.78 (MIROC-ES2L).

**Figure 10:** same as Fig. 9 but for \( \gamma \) (kg C m\textsuperscript{-2} °C\textsuperscript{-1}). Note that cropland areas are not excluded from panels (k-o) as in Fig. 9.

**4. Summary and conclusions**

We have investigated carbon cycle feedbacks in a highly idealized model experiment with exponentially increasing and decreasing atmospheric CO\textsubscript{2} concentration (1pctCO\textsubscript{2}-cdr) and in a more realistic
overshoot scenario simulation (ssp534-over). We employ an ensemble of five CMIP6 ESMs that have
run additional (biogeochemically coupled) simulations that allow us to separate the effects of changing
atmospheric CO₂ and of changing surface climate on the simulated carbon cycle. We find that both the carbon-concentration (\( \beta \)) and the carbon-climate (\( \gamma \)) feedbacks show a
calorish hysteresis behavior during negative emission phases. The well-known reduction of ocean
and land carbon uptake with increasing temperatures continues long into the negative emissions
phases of the simulations (when temperature is decreasing), albeit at a reduced rate. For the ocean,
there is still a reduction in carbon stocks due to legacy warming when pre-industrial atmospheric CO₂
is restored in the 1pctCO₂-cdr simulation, consistent with the single-model studies of Schwinger and
Tjiputra (2018) and Bertini and Tjiputra (2022). In contrast, all models agree that the effect of legacy
warming is less important for the terrestrial carbon-climate feedback as the reduction of global mean
surface temperature leads to a reduction in temperature-induced losses of terrestrial carbon towards
the end of the 1pctCO₂-cdr simulation.

Carbon cycle feedback metrics vary over time, and between different scenarios. When the deviations
in surface temperature and atmospheric CO₂ become small towards the end of a modeled negative emission scenario, the magnitude of these feedback metrics “explodes” since they are defined as the
ratio between the deviations in carbon stocks and the change in temperature and atmospheric CO₂,
respectively. Arguably, the latter is mainly a problem due to the strongly idealized simulation design of
the 1pctCO₂-cdr experiment, not for more realistic scenarios as the ssp534-over. Also, a different
definition of the reference state for the feedback metrics, as proposed by Chimuka et al. (2023), avoids
this problem.

We find that the relative strength of the feedback remains relatively robust between positive and
negative emission phases and between the different simulations considered here. For example, a model
with a stronger than average terrestrial carbon-concentration feedback (\( \beta_2 \)) during the positive
emission phase of the 1pctCO₂-cdr simulation will also show a stronger than average \( \beta_1 \) during the
negative emission phase or for the ssp534-over scenario. Regarding the model uncertainty of feedback
metrics we find that there is an increase in uncertainty in all feedback metrics between the positive and
negative emission phases of the 1pctCO₂-cdr simulation. Except for \( \gamma_2 \), this increase is much larger than
expected from an accumulation of uncertainty over time. This indicates that there is an additional
component of model uncertainty resulting from differences in the lagged model responses to the
change from increasing to decreasing radiative forcing.

The geographical patterns of terrestrial \( \beta \) and \( \gamma \) feedback metrics highlight differences in the responses
of tropical/subtropical versus temperate/boreal ecosystems as a major source of model disagreement.
For individual models, however, the spatial feedback patterns are remarkably similar during phases of
increasing CO₂ compared to phases of decreasing CO₂ concentrations, indicating that the increase of
global mean values of \( \beta \) and \( \gamma \) due to lagged responses of the carbon cycle during negative emissions
phases does not stem from a particular region but is generally seen over the whole globe. We estimate
the contribution of permafrost carbon release to the carbon-climate feedback only for one of the five
ESMs (NorESM2-LM, which vertically resolves soil carbon). Permafrost carbon release is clearly seen as
a strong positive feedback (i.e., negative \( \gamma \)) over the permafrost area, but it emerges only relatively late.
in the simulations. Permafrost carbon release accounts for 38% of NorESM2-LM’s carbon-climate feedback at the midpoint of the negative emission phase of the 1pctCO$_2$-cdr simulation.

In the ssp534-over simulation, the presence of land-use change complicates the analysis of feedbacks. Land-use change is not a feedback process, yet owing to the C4MIP simulation design, carbon losses (or gains) due to land use change are confounded with the carbon-concentration feedback derived from a biogeochemically coupled scenario simulation. If we disregard agricultural areas, terrestrial carbon cycle feedback patterns in the ssp534-over scenario are largely similar to those in the 1pctCO$_2$-cdr simulation, although some differences particularly in high northern latitudes due to the influences of non-CO$_2$ forcings exist.

We conclude with some recommendations for future research and the design of future model intercomparison projects (MIPs) like C4MIP and CDRMIP. Identifying and better understanding the causes of differences in the lagged model response to decreasing emissions, which we have shown to increase the model disagreement under negative emissions should be pursued further with high priority. Both the integrated-flux ($\beta$ and $\gamma$) and instantaneous-flux (B and $\Gamma$) based feedback metrics and their uncertainties become difficult to interpret in scenarios where atmospheric CO$_2$ concentration decreases, particularly in the extreme case when atmospheric CO$_2$ concentration and surface temperature approach their pre-industrial level. In the light of the discussion around CDR perhaps it is timely to rethink other but related forms of these metrics (e.g., see Chimuka et al. 2023) that describe the response of land and ocean carbon systems in scenarios of decreasing atmospheric CO$_2$ in a more robust manner.

The 1pctCO$_2$ simulation combined with the 1pctCO$_2$-cdr simulation is an extremely idealized model experiment with huge (and infeasible) amounts of implied net-negative emissions and a discontinuity at year 140, where implied emissions jump from large positive to large negative values. As we know that carbon cycle feedbacks are scenario dependent, it would be preferable to assess these feedbacks using model simulations that have a more realistic emission pathway and that include more realistic amounts of net-negative emissions. Alternative idealized simulation designs that include negative emissions have been proposed in the literature (MacDougall 2019; Schwinger et al. 2022) and we have also considered the ssp534-over scenario in this study. However, the presence of land-use change and variable non-CO$_2$ forcings in SSP scenarios complicates the quantification of carbon cycle feedbacks. Whether this problem can be solved for future phases of C4MIP by providing more detailed model output or by requesting additional idealized experiments (e.g., scenario simulations with fixed land use) should be discussed in the C4MIP community.

Finally, most proposed negative emission options would be realized by manipulating the terrestrial or oceanic carbon sinks (e.g., bioenergy with carbon capture and storage, afforestation or ocean alkalinization), thereby not only changing the atmospheric CO$_2$ concentration and possibly the surface climate but also the carbon cycle feedbacks themselves. Such interactions go beyond what can be addressed with the traditional C4MIP design of fully- and biogeochemically coupled ESM simulations. Consequently, a new framework for determining feedbacks caused by large scale CDR in realistic scenarios of CDR deployment is needed and should be developed in close collaboration with the integrated assessment modeling community that will create such scenarios.
Data availability

All CMIP6 model output data is freely available through the Earth System Grid Federation (for example, under https://esgf-data.dkrz.de/search/cmip6-dkrz/). The model output data of the 1pctCO₂-cdr-bgc simulation will be made publicly available upon final acceptance of this manuscript.

Competing interests

None of the authors has any competing interests.

Acknowledgements

A.A., J.S., and H.L. were supported by the Research council of Norway through the project IMPOSE (grant no. 294930). J.S. and H.L. also received funding from the European Union’s Horizon Europe research and innovation programme (project RESCUE, grant agreement no. 101056939). Supercomputing and storage resources for additional NorESM2 simulations were provided by UNINETT Sigma2 (projects nn9708k/ns9708k). T.H. was supported by the Integrated Research Program for Advancing Climate Models (TOUGOU, grant number JPMXD0717935715) and the Program for the Advanced Studies of Climate Change Projection (SENTAN, grant number JPMXD0722681344) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. C.D.J. and S.L. were supported by the Joint UK BEIS/Defra Met Office Hadley Centre Climate Programme (GA01101), and the European Union’s Horizon 2020 research and innovation programme under Grant Agreement No 101003536 (ESM2025 – Earth System Models for the Future). R.S. and Y.S.-F. are grateful for the support of the team in charge of the CNRM-CM climate model. Supercomputing time was provided by the Meteo-France/DSI supercomputing center. R.S. acknowledges the European Union’s Horizon 2020 research and innovation program under grant agreement No. 101003536 (ESM2025 – Earth System Models for the Future). Y.S.-F. acknowledges the TRIATLAS project under the grant agreement No. 817578 and the COMFORT project under the grant agreement No. 820989. J.T. acknowledges the OceanICU project under the grant agreement no. 101083922.

We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF.
The work reflects only the authors’ view; the European Commission and their executive agency are not responsible for any use that may be made of the information the work contains.

References


Gasser, T., and Coauthors, 2018: Path-dependent reductions in CO2 emission budgets caused by...


