Ito Gen (Orcid ID: 0000-0002-7350-6665)

Kiang Nancy, Y. (Orcid ID: 0000-0002-5730-924X)

Faluvegi Greg (Orcid ID: 0000-0001-9011-3663)

Aleinov Igor (Orcid ID: 0000-0002-7827-228X)

Ruedy Reto, Anton (Orcid ID: 0000-0002-1270-1088)

Russell Gary, L. (Orcid ID: 0000-0001-7174-5825)

Lerner Paul (Orcid ID: 0000-0003-4751-0262)

Global carbon cycle and climate feedbacks in the NASA GISS ModelE2.1

Gen Ito[1,2,3], Anastasia Romanou[1], Nancy Y. Kiang[1], Gregory Faluvegi[1,4], Igor Aleinov[1,4], Reto Ruedy[1,2], Gary Russell[1], Paul Lerner[1], Maxwell Kelley[1,2] and Ken Lo[1,2]

[1]NASA Goddard Institute for Space Studies, New York, NY, USA

[2]SciSpace, LLC, New York, NY, USA

[3]Centre de Recherches Pétrographiques et Géochimiques (CRPG), CNRS/Université de Lorraine, Vandoeuvre-lés-Nancy, France

[4]Center for Climate Systems Research, Earth Institute, Columbia University, New York, NY, USA

Corresponding author: Anastasia Romanou (anastasia.romanou@nasa.gov)

Key Points:

· Observed CO_2 and carbon stocks/fluxes were reproduced within uncertainties, with most of disagreements in transient land uptake.

· Ocean carbon-climate feedback is stronger and land feedback slightly less sensitive than CMIP5; other feedback parameters are comparable.

· Future climate shifts temperate/boreal transitions northward to ~60 °N, extend temperate deciduous forest along ~60 °N, and warm/dry Europe.

Abstract

We present results from the NASA GISS ModelE2.1-G-CC Earth System Model with coupled climate-carbon cycle simulations that were submitted to the sixth phase of the Coupled Model Intercomparison Project (CMIP6) Coupled Climate-Carbon Cycle MIP (C4MIP). Atmospheric CO_2 concentration and carbon budgets for the land and ocean in the historical simulations were generally consistent with observations. Low simulated atmospheric CO_2 concentrations during 1850-1950 were due to excess uptake from prescribed land cover change, which erroneously replaced arid shrublands with higher biomass crops, and assumed high 2004 LAI values in vegetated lands throughout the historical simulation. At the end of the historical period, slightly higher simulated CO_2 than observed resulted from the land being an insufficient net carbon sink, despite the net effect of CO_2 fertilization and warming-induced increases to leaf photosynthetic capacity. The global ocean carbon uptake agreed well with the observations with the largest discrepancies in the low latitudes. Future climate projection at 2091-2100 agreed with CMIP5 models in the northward shift, of temperate deciduous forest climate and expansion across Eurasia along 60 °N latitude, and dramatic regional biome shifts from drying and warming in continental Europe. Carbon feedback parameters were largely similar to the CMIP5 model ensemble. For our model, the variation of land feedback parameters within the uncertainty arises from the fertilization feedback being less sensitive due to lack of increased vegetation growth, and the comparably more negative ocean carbon-climate feedback is due to the large slowdown of the Atlantic overturning circulation.

Plain Language Summary

The Earth's global carbon cycle traces surface-atmosphere exchanges of CO_2 by the ocean and the land, including both natural processes (photosynthesis, respiration, ecological dynamics, gas exchange, and atmospheric transport) and anthropogenic activities (fossil fuel emissions, land use change). Scientists synthesize current understanding of carbon cycle science in Earth System Models (ESMs), to simulate the most significant processes in a highly complex system, to quantify how the land and ocean carbon sources and sinks of atmospheric CO_2 behave and how they will change in the future. The NASA GISS ModelE2.1-G-CC ESM couples the NASA Ocean Biogeochemistry Model (NOBM) in the ocean, the Ent Terrestrial Biosphere Model (Ent TBM) on land, and CO_2 transport through the atmosphere and interactions with Earth's radiation. Here we describe experiments that were contributed to the Coupled Model Intercomparison Project 6 (CMIP6). Results indicate that GISS ModelE2.1 reproduces the large scale properties of the carbon cycle generally consistently with measured data. In future warmer climates, major regional biome changes will occur in temperate to subarctic climates. The land and the ocean become net sinks, but they will have reduced ability to take up additional CO_2 and thereby ameliorate climate warming.

1   Introduction

The global carbon cycle involves the exchange and storage of carbon among numerous different reservoirs of the Earth System. This cycle traces the reduction of CO_2 from the atmosphere through photosynthesis on land and in the ocean as well as gas exchange with the ocean reservoir, and oxidation back to CO_2 through natural processes of biological growth, respiration, ecological dynamics (competition of organisms, fire disturbance), gas solubility, and atmospheric transport, and anthropogenic activities, including fossil fuel and biomass combustion and land use change. The differential rates of exchange among different natural stocks lead to temporal variations and trends in the carbon storage sizes of these stocks. On short-term time scales (from seconds to millennia), the movement of carbon is dominated by the biogeochemical exchange of carbon among the terrestrial biosphere, ocean, and atmosphere, as opposed to geological time scales (hundreds of thousands of years or greater) related to volcanic outgassing of subducted carbon (e.g., Carlson et al., 2001; Houghton, 2003; Zeebe, 2012), and geochemical time scales (100 millions of years) related to rock weathering and the carbon-silicate cycle (Berner et al. 1983). With the onset of industrialization powered by fossil fuels, humans have moved carbon from the fossil fuel reservoir, which naturally only fluctuates at the geological time scale, into the atmosphere through anthropogenic activities largely in the form of greenhouse gases such as carbon dioxide (Carlson et al., 2001; Houghton, 2003; Zeebe, 2012). This has increased the amount of atmospheric carbon dioxide on the short-term time scale from 280 ppm during the Holocene and preindustrial era to 415 ppm today, a rate of increase over the past century of ~1.0 ppm/year while for 2017 and 2018 it exceeded 2.5 ppm/year (e.g., Friedlingstein et al., 2019; Le Quéré et al., 2017).

The carbon cycle is closely tied to climate change due to the greenhouse effect of carbon dioxide (CO_2) and methane (CH_4), the sensitivity of redox processes to warming temperatures, and the sensitivity of the ocean uptake to surface warming, changing circulation patterns and deep ocean storage. Since the industrial era, the land and ocean components of the carbon cycle have been absorbing CO_2 released by human activities, such that the rate of increase of atmospheric CO_2 is less than it would otherwise be. The ability of the future land and ocean to act as sinks of CO_2 is difficult to constrain. A major concern is that their uptake rate may decrease as the land and ocean storage capacity becomes saturated, and as net release of CO_2 may ensue if oxidation processes exceed photosynthetic uptake with climate warming, leading to positive feedback to climate warming.

Earth System Models (ESMs) offer a means to quantify the effect of anthropogenic emissions of carbon on the climate system. ESMs synthesize our understanding thus far of the interaction between the physical climate system and the carbon cycle by coupling atmosphere–ocean general circulation models (AOGCMs) to terrestrial and ocean carbon cycle models (Arora et al., 2013; Hansen et al., 2007; Meehl et al., 2005; Plattner et al., 2008; Todd-Brown et al., 2014). Furthermore, ESM projections of future climate change and the improved understanding they provide of the global carbon cycle play a vital role when forming mitigation and adaptation strategies to deal with climate change (Plattner et al., 2008). For these reasons, developing ESMs that can constrain the impact of perturbations to the carbon cycle on climate change is essential. The redistribution of elevated CO_2 into the different reservoirs of the carbon cycle is complex and was not easily modeled by the prior generation of ESMs, such that the projection of climate change that would occur in response to anthropogenically emitted carbon was uncertain and showed divergence (Friedlingstein et al. 2014; Kaushik et al., 2020). This is an imminent problem both from scientific and societal perspectives.

In this study, we use an ESM, the NASA Goddard Institute for Space Studies (GISS) ModelE 2.1, to simulate climate change and the global carbon cycle perturbed by anthropogenic emissions. We work within the scope of the sixth phase of the Coupled Model Intercomparison Project (CMIP6; Eyring et al., 2016) and follow experimental protocols established by the Coupled Climate–Carbon Cycle Model Intercomparison Project (C4MIP; Jones et al., 2016) which operates within the framework of CMIP6 and is responsible for the design, documentation, and analysis of carbon cycle feedbacks and interactions in climate simulations conducted by ESM groups around the world. Ensemble MIPs are an important community effort to assess current understanding through diverse representations. Previous model intercomparison projects have provided important guidance for modeling the coupled climate–carbon system using ESMs, and they have led to major advancements of scientific knowledge in this field (e.g., Friedlingstein et al., 2006) and contributions to the Intergovernmental Panel on Climate Change (IPCC) assessments, which shape policy on climate change (IPCC, 2013). Using the latest incarnation of the NASA GISS climate model, we have conducted key simulations identified in C4MIP that allow model estimates of the relative strengths of the different components of the carbon cycle in producing sources or sinks of CO_2. The purpose of this study is to describe the behavior of the different components of the global carbon cycle in the NASA GISS ModelE2.1 in historical, idealized 1% per year increasing CO_2, and Shared Socio-Economic Pathways (SSP) 5-8.5 scenarios and to analyze carbon cycle feedbacks.

The paper is organized as follows: In section 2, we explain the configurations of the atmosphere, land, and ocean components of our model. Section 3 outlines experimental design that followed CMIP6 C4MIP protocols. In section 4, we provide results of the simulations. Section 4.1 describes results from the fully coupled historical simulations and compares the modeled atmospheric CO_2 concentration and carbon uptake by the land and ocean to observational data. Section 4.2 shows land and ocean carbon uptake for the biogeochemically coupled historical experiments, and the results are compared to fully coupled historical experiments. Section 4.3 present land and ocean carbon uptake from the three variants of 1%/yr increasing atmospheric CO_2 simulations, and the change in carbon storage at the time of quadrupling of atmospheric CO_2 concentration. In this section, we also compute carbon cycle feedback parameters. In section 4.4, results from SSP5-8.5 scenario experiments are presented. In section 4.5, we present a Köppen-Geiger climate classification for emissions-driven historical and SSP5-8.5 simulations.

2   Model description

2.1   The atmosphere model

The model version used here is GISS-E2.1-G, as described in Kelley et al. (2020). That paper describes the model improvements and augmentation since the CMIP5-era GISS-E2-R model described in Schmidt et al. (2014). Briefly, our configuration of the atmospheric model uses 40 vertical layers with a model top at 0.1 hPa, and a horizontal resolution of 2° latitude × 2.5° longitude. Atmospheric tracers follow the advective air mass flux of the dynamics and achieve greater than nominal resolution by transporting nine higher-order moments along with the mean tracer mass. The radiation code is called every 5 model timesteps, and we use greenhouse gas concentrations from transient historical forcing prescribed for CMIP6 (Meinshausen et al., 2017). In the ESM simulations presented here radiation experiences atmospheric emission-based CO_2 abundance changes, as described

below, and the radiation's aerosol and ozone constituents are prescribed based on output of an ensemble of GISS-E2.1 AMIP simulations which used full atmospheric chemistry and aerosol schemes (OMA; Kelley et al., 2020). The GISS ESM model version is the GISS-E2.1-G-CC in the CMIP6 repository.

2.2 The land model

The land model of the GISS GCM couples a ground hydrology model and vegetation model. The ground hydrology model is that described by Rosenzweig & Abramopoulos (1997), with updates for canopy interception of precipitation and an addition of a snow model (Schmidt et al., 2006). The ground hydrology model is responsible for computing soil water and the energy balance of the land surface, including temperatures for the vegetation canopy, three layers of snow, and six soil layers to a 3.5 m depth. Surface and subsurface runoff are routed to rivers and lakes as a function of hydraulic conductivity and local slope. Surface albedo and fluxes to the atmosphere are cover-weighted averages from subgrid fractional cover of vegetated ground, bare soil, lake, and ocean. A dynamic lakes scheme allows for the growth and shrinkage of temporary lakes due to runoff and evaporation (Schmidt et al. 2006).

The vegetation model, the Ent Terrestrial Biosphere Model (Ent TBM; Kiang et al., 2006; Kim et al., 2019; Schmidt et al., 2006; 2014), supplies the surface albedo, water vapor conductance, and CO_2 fluxes from the vegetated portion of land grid cell. The Ent TBM is a demographic dynamic global vegetation model (dDGVM), but for the CMIP6 experiments, only the biophysics of the Ent TBM is run to supply land surface and vegetation fluxes. The biophysics consists of multi-layer canopy radiative transfer model (Spitters et al., 1986) and leaf gas exchange using the Ball-Berry stomatal conductance model (Ball & Berry, 1985) coupled with the Farquhar-von Caemmer photosynthesis model (Farquhar & von Caemmerer, 1982); leaf boundary layer conductance is from Collatz et al. (1991), and autotrophic and heterotrophic respiration is parameterized as in Kim et al. (2019). Boundary conditions of natural vegetation cover and canopy structure are prescribed from a satellite-derived global vegetation structure dataset v1.0, which provides a map of subgrid fractional cover of 12 Ent TBM plant functional types (PFTs) derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) land cover and PFT products (Friedl et al. 2002, 2010), annual maximum and monthly leaf area index (LAI; Tian et al., 2002a, 2002b), and forest canopy heights derived from the Geoscience Laser Altimeter System (GLAS) aboard the ICESat (Ice, Cloud, and land Elevation Satellite; ICESat/GLAS) by Simard et al. (2011). For GISS ModelE2.1 CMIP6 experiments, monthly observed MODIS LAI of the year 2004 is prescribed for all simulated years.

Historical crop and pasture cover was prescribed annually, rescaling the subgrid fractions of natural vegetation cover. The CMIP5 historical crop and pasture cover for 1900-2005 (Miller et al. 2014) were erroneously re-used for the CMIP6 historical experiments, a comparison of which is plotted in Figure 1. For GISS ModelE2.1, these cover changes were constructed from a blend of crop cover from Pongratz et al. (2008) and crop and "pasture" cover from the CMIP5 Hurtt et al. (2006) Land Use Harmonization (LUH) data set. To maintain continuity with the Pongratz et al. (2008) crops cover for the last millenium (LUHa_t1.v1 for 1700-1850), the two data sets were linearly weighted with the Pongratz et al. (2008) weight going from 1 at 1850 to 0 at 1900. Then the LUH only was used for 1900-2014 for the historical period and 2015-2100 future scenarios. Changes in vegetation carbon stocks due to land cover change are distributed as fluxes to the atmosphere uniformly over the year.

To maintain a continuous transition in land cover change and carbon changes for the future scenarios, we continued the historical run with the LUH 2006-2014 land cover, and the

future scenarios with the LUH 2015-2100 land cover. The CMIP5 "pasture" cover of Hurtt et al. (2006) included rangeland, which alone is comparable in cover to crops and managed pasture lands combined. Because rangeland is in fact not a land cover change but a land use change (putting grazing animals on existing arid grasslands and shrublands), rescaling all natural vegetation with the crops and "pasture" cover in GISS ModelE2.1 effectively replaces these two natural types (and small amounts of forest) with higher LAI crop cover, with this change particularly significant over Australia. In the LUH2 (Hurtt et al. 2011) for CMIP6, pasture and rangeland were separated. The time course of global crop, "pasture" (including rangeland) cover area in the GISS ModelE2.1 implementation is plotted in Figure 1a, together with the CMIP6 land use change components for comparison. The global total of the CMIP5 LUH cover of crops, pasture, and rangeland is comparable to, though slightly lower than, that in the CMIP6 LUH2 sum of these types. The LUH2 also considerably revised distributions of these types within regions (not shown). We implemented only cover changes due to crops, pasture (including rangeland), but not the land use state transitions from primary forest to secondary forest, which is a standard option in the historical land cover of the Land Surface, Snow and Soil Moisture Model Intercomparison Project (LS3MIP; Lawrence et al., 2016). Therefore, in those areas where crop cover expansion replaces forest, our model will reflect deforestation, but where deforestation leads to secondary forest regrowth, our model will lack the pattern of CO_2 losses to the atmosphere followed by uptake by regrowth. The resulting time course of global cover of Ent PFTs is plotted in Figure 1b.

Canopy albedo is a translation to Ent PFTs of the seasonal albedos prescribed by Matthews (1983), originally by biome type. Soil carbon is simulated as a 30 cm upper soil layer, with a version of the 9-pool CASA model (Potter, et al., 1993; Randerson et al., 2009), with carbon only (no prognostic nitrogen), and with soil moisture and temperature responses of soil respiration modified, as described in Kim et al. (2019).

Because vegetation cover and canopy structure are prescribed, the prognostic carbon pools on land are only the plant labile carbon and soil carbon, to enable coupling with prognostic atmospheric CO_2 and ocean carbon for a fully coupled carbon cycle. The labile carbon pool is the balance of photosynthetic uptake of carbon, respiration fluxes, growth of seasonally prescribed foliage and fine roots, and retranslocation from senescence. Because the vegetation model is run in a biophysics-only mode, carbon uptake from net photosynthesis does not drive height growth of vegetation; therefore, any carbon that would otherwise be allocated to height growth is dropped as litterfall to the soil. Carbon fluxes not included are those from fire and regrowth of secondary forest.

In the GISS ModelE2.1 C4MIP experiments, the balance of the air-to-land carbon flux is a function of three major drivers: 1) land cover change due to expanding crops, pasture, and rangeland; 2) the CO_2 fertilization effect through which vegetation gross primary productivity increases with enhanced diffusion of CO_2 into stomata with higher surface concentration of CO_2 (Ball & Berry, 1987; Sitch et al. 2008; O'Ishi et al. 2009), and an inverse relation to surface CO_2 concentration to stomatal conductance of water vapor and greater water use efficiency; and 3) the radiative warming effect of atmospheric CO_2, which enhances both leaf photosynthetic capacity and autotrophic (plant) and heterotrophic (soil) respiration due to their temperature sensitivities. Changes in climatology that affect drought, cloud cover (affecting photosynthetically active radiation), and regional temperature (affecting respiration and photosynthetic capacity) further affect the net balance of land carbon fluxes. In CMIP5 models, a small amount of climate change was seen to result from the reduction of vegetation stomatal conductance vegetation in response to elevated atmospheric CO_2, and the compensating effect of increased leaf area index and cover changes from CO_2 fertilization (Gregory et al., 2009; Wenzel et al., 2013). Because we prescribe fixed

2004 monthly LAI, our model does not capture the latter secondary negative feedback from increased leaf area index from the fertilization effect.

For CMIP6 C4MIP, soil carbon stocks were spun up according to the piControl conditions in 1850, using an acceleration scheme: 25 years of a piControl configuration was run to equilibrate the atmosphere climate and vegetation labile carbon, and then run forward another 9 years, saving these 9 years of GCM meteorology (soil moisture and temperature) and litterfall. These 9 years of forcings, averaged at a ~1 hour time step, were then used to drive the soil biogeochemistry to equilibrium, over 750 iteration × 9 years, totaling 6750 years. This achieved a state with net ecosystem productivity (NEP) fluxes (GPP − autotrophic respiration − soil respiration), averaged over the last 8 years of the spin-up, having an annual minimum −0.11, mean 0.001, and maximum of 0.10 kgC/m^2/yr of non-ice covered soil area uptake. Total global imbalance was 0.15 GtC/yr uptake. This preindustrial spin-up of soil carbon was then used as the initial condition for transient and historical runs. This spin-up therefore does not reflect state transitions from land use history prior to 1850.

2.3   The ocean model

The ocean component of the GISS GCM (GISS-E2.1-G) is a mass conserving ocean model, at 1° latitude × 1.25° longitude horizontal resolution with a free surface and natural surface boundary conditions for heat and freshwater fluxes (Russell et al., 1995; Schmidt et al., 2014). Since its CMIP5 incarnation, it has finer resolution near the surface and a total of 40 layers in the vertical, improved eddy transport parameterization also including three dimensional variation of mesoscale diffusivity, a higher-order advection scheme (Prather, 1986), and a parameterization of tidal mixing and more realistic strait-through-flows that affect property distributions in marginal seas (Kelley et al, 2020). Ocean only simulations of passive tracer uptake (CFC) showed that the model is in good agreement with observations (Romanou et al, 2017).

The ocean carbon cycle module is an update of the one used in CMIP5 simulations with the GISS model (Romanou et al, 2013; 2014) and which originated from the NASA Ocean Biogeochemistry Model (NOBM; Gregg & Casey, 2007). It is fully interactive with the atmosphere, the physical ocean, and the ice and radiation submodels, and it includes 4 phytoplankton species (diatoms, chlorophytes, cyanobacteria, and coccolithophores), four nutrient elements (nitrate, silicate, ammonium, and iron), three detrital pools (nitrate/carbon, silicate, and iron), and one heterotroph species. Carbon cycling is represented through dissolved organic (DOC) and dissolved inorganic carbon (DIC) and interacts with the atmospheric CO_2 tracer through gas exchange parameterization, following the CMIP6 ocean carbon model intercomparison protocol (Orr et al., 2017). Light profiles from the atmospheric radiation module are propagated underwater into the ocean and spectrally decomposed to 33 wavebands that are used to compute growth of the phytoplankton groups (Gregg & Conkright, 2002). Latto & Romanou (2018) showed that ocean carbon states estimated from the GISS ModelE2.1 with CMIP5 carbon cycle agreed well with observations, and that study led to the current improvements in the carbon cycle simulations.

In the updated NOBM model, the diatom and detrital sinking speeds are modulated by an exponential profile that depends on the diatom and detrital concentrations at depth:

$$ws_d(z) = visfac(z) \times a \times exp(b \times C_d(z)) \dotfill 1$$

where $ws_d(z)$ is the sinking speed of diatoms or detritus, $visfac(z)$ is a coefficient representing that effect of viscosity on sinking speed that depends on temperature (particles sink more slowly in colder waters), a and b are constants (0.01 and 5 respectively for diatoms, 2 and 3 respectively for detrital carbon, 3 and 6 respectively for detrital silica, and 1 and 2 respectively for detrital iron), and $C_d(z)$ are the diatom or detrital concentrations. This parameterization of sinking speed differs from most models, which generally prescribe a constant sinking speed for their prognostic detrital tracers. Alkalinity is now prognostically computed following the Ocean Carbon-Cycle Model Intercomparison Project 2 (OCMIP2) protocol in order to better simulate carbonate chemistry and the oceanic carbonate pump. The ocean iron cycle is forced with the historical annual cycle climatology extracted from the online dust simulations using the GISS dust model (Miller et al, 2006) which has been expanded to include eight externally mixed minerals (illite, kaolinite, smectite, carbonates, quartz, feldspar, iron oxides, and gypsum) plus internal mixtures between minerals and iron oxides (Perlwitz et al., 2015a,b). The masses of free and structural iron and their fractions of total iron have been evaluated using measurements for location at Izaña Observatory (García-Pando et al., 2016). The GISS ocean carbon cycle is now including riverine delivery of biogeochemical constituents and is coupled to the prognostic river runoff calculated in GISS ModelE2.1 as part of the simulated global hydrological cycle. The concentrations of particulate organic carbon (POC), dissolved organic carbon (DOC), dissolved inorganic carbon (DIC), nitrate, silicate, and iron at all the major and many minor river estuaries are obtained from an annual climatology (da Cunha et al., 2007), and they modulate the biogeochemical characteristics of the freshwater outflow into the ocean at these sites. Therefore, as continental runoff increases (decreases) due to changes in precipitation over land, the delivery of these ocean biogeochemical tracers will increase (decrease). Particulate carbon burial into the sediment, although prescribed for all areas shallower than 150 m, takes place in the estuaries where detritus delivered with the riverine flow tends to accumulate. The seasonal cycle of the air-to-sea flux of CO_2 as well as the partial pressure of CO_2 in the sea water have been extensively evaluated against several observational datasets (Lerner et al, 2020).

Tracer conservation in the ocean carbon cycle model is rigorously tested with online as well as offline diagnostics. These account for the amount of carbon and other tracers that is exchanged between different pools through all the modeled processes. We ensure that each tracer is conserved to machine precision accuracy (see details in Table S2 of the Supporting Information).

The ocean carbon and the land model spun up separately. For the ocean carbon, the spin up starts from a physical climate model simulation (without carbon) that is almost at equilibrium after it has been integrated for approximately 3000 years and in which the surface of the ocean is at steady state whereas there is a small warming drift at depth of about 0.006 °C over 150 years. The ocean carbon spin up initial conditions are taken from Global Ocean Data Analysis Project version 2 (GLODAPv2) for DIC and alkalinity and from World Ocean Atlas 2013 (WOA2013) for nitrate, silicate, and iron while the atmospheric CO_2 concentration is at preindustrial levels. The model experiences an initial adjustment of about 50 years while the surface carbon distributions adjust to the physical ocean mixing and exchanges with the atmosphere and while the biological pump comes to near steady state. At that time, the air-to-sea flux of CO_2 is about −0.5 GtC/yr and slowly decreases. The model is further integrated for 700 years with prognostic CO_2 at which point surface DIC and alkalinity are shown to change very little with time. At that point, we employ a tracer acceleration technique that is based on linear extrapolation of DIC and alkalinity at each model grid point over a period of 500 years. At the end of this process, the air-to-sea flux of

CO_2 is within 0.1 GtC/yr for the duration of the experiments, meeting the C4MIP requirements (Jones et al., 2016) for submission to CMIP6.

3   Experimental Design

We have conducted a total of 12 experiments (Table 1) to contribute to the climate-carbon cycle component of CMIP6. Following the C4MIP protocol (Jones et al., 2016), three Diagnostic, Evaluation, and Characterization of Klima (DECK); two historical; two tier-1; and three tier-2 experiments were conducted. In this study, we present nine experiments (indicated in bold letters in Table 1) that enable us to assess the model's skill, describe future predictions, and examine carbon cycle feedbacks. In this section, broad overviews of the eight experiments and aspects that are unique to NASA GISS ModelE2.1 are described. The reader is referred to the C4MIP (Jones et al., 2016) for specific design information of each of the experiments.

1% per year increasing atmospheric CO_2 experiments

These are idealized scenarios where the atmospheric CO_2 concentration increases by 1% every year starting from the spun-up initial conditions (Section 2). Atmospheric CO_2 concentration is prescribed and is the only forcing present in this scenario; all other forcings (e.g., other greenhouse gases and land use) are held at their pre-industrial levels. The atmospheric CO_2 concentration doubles after 70 years and quadruples after 140 years of simulation following C4MIP design protocols (Jones et al., 2016), and the 1% per year increase of atmospheric CO_2 is applied continuously for a minimum of 150 years following CMIP6 requirements (Eyring et al., 2016).

The standard 1pctCO2 experiment is a concentration-driven coupled simulation, meaning that the terrestrial biosphere, ocean, and the radiation scheme of the model responds to the increasing atmospheric CO_2 concentration, thus promoting climate change. The 1pctCO2-bgc experiment is the biogeochemically coupled version of the standard 1pctCO2 experiment. Here, only the land and ocean carbon components of the model respond to the increasing atmospheric CO_2 level; the model's radiation component does not respond to the increasing atmospheric CO_2 level but uses instead the pre-industrial CO_2 concentration. Correspondingly, the 1pctCO2-rad is the radiatively coupled version of the standard 1pctCO2 experiment, and in this version, only radiation responds to increased atmospheric CO_2 concentrations, but the land and the ocean carbon experience pre-industrial levels of atmospheric CO_2 concentrations. The three variations of 1pctCO2 simulations are used to study carbon cycle feedbacks.

historical experiments

The standard historical experiment (hist) is a concentration-driven simulation where atmospheric CO_2 and other greenhouse gas concentrations are prescribed (CMIP6 data from Meinshausen et al., 2017). The land and ocean models are interactive as described in Sections 2.2 and 2.3. This experiment is forced by CMIP6 atmospheric CO_2 concentration data which are based on observations and reconstructs the climate from the beginning of 1850 to the end of 2014.

The esm-hist experiment (esm-hist) is the emissions-driven counterpart of the concentration-driven historical experiment. In this experiment, the atmospheric CO_2 concentration is determined prognostically from the anthropogenic CO_2 emissions from different socio-economic sectors, which are prescribed from the CMIP6 forcing data of Hoesly et al. (2018), regridded to the atmospheric model's 2º latitude × 2.5º longitude

resolution, and the carbon fluxes from land and ocean model components. These emissions sectors are energy; industrial; transportation; residential, commercial, other; international shipping; waste; and solvents sectors. It does not consider CO_2 fluxes from biomass burning, grazing, harvesting, and aircraft. Emissions data from biomass burning was not available, and grazing, harvesting, and aircraft emissions were not configured yet to be included in our model.

Both concentration-driven and emissions-driven historical experiments serve as a benchmark for assessing model performance, though the emissions-driven version is capable of more in-depth assessments of carbon cycle effects due to its prognostic determination of atmospheric CO_2 concentration from different emission sectors and the land and ocean sinks.

The hist-bgc and esm-hist-bgc experiments are the biogeochemically coupled version of the historical and esm-hist experiments. These configurations are identical to the hist and emis-hist cases except that the land and ocean components of the model respond to the increase in CO_2 concentration only while there is no radiative warming effect on climate.

SSP5-8.5 scenario experiments

The esm-ssp585 experiment is an emissions-driven future scenario simulation based on Shared Socio-Economic Pathways (SSPs) which are part of the Scenario Model Intercomparison Project (ScenarioMIP; O'Neil et al., 2016) that coordinates with C4MIP. SSP5, referred to as "Fossil-Fueled Development", is characterized by high socio-economic challenges to mitigation and low socio-economic challenges to adaptation, and SSP5-8.5 scenario achieves a forcing level of 8.5 Wm^{-2} in 2100 (Kriegler et al., 2017; O'Neill et al., 2016). The esm-ssp585 experiment starts at the end point of the esm-hist experiment and continues 85 years until 2100 using CMIP6 SSP5-8.5 anthropogenic emissions and other forcings. The ssp585-bgc experiment is the biogeochemically coupled concentration-driven version of this scenario and branches off from the historical experiment (hist).

4   Results

The global carbon budgets for the 9 described experiments and piControl experiment are listed in the Appendix in Tables B1-B4. We first summarize the simulated atmosphere, then the stocks and fluxes in the land and ocean environments.

4.1   Coupled historical simulations

In this section, we present the modeled atmospheric CO_2 concentration and CO_2 uptake by the land and ocean components of the carbon cycle in concentration-driven and emissions-driven historical experiments, and compare them to observations to assess the skill of GISS ModelE2.1 in modeling the present-day global carbon cycle.

4.1.1   Atmospheric CO_2 concentration

Figure 2a shows a comparison of the temporal evolution of simulated global mean annual atmospheric CO_2 concentration in the esm-hist experiment with the observation-based atmospheric CO_2 used to force the concentration-driven historical (hist) simulations (Meinshausen et al., 2017). In the esm-hist experiment, simulated atmospheric CO_2 concentration is lower than the observed until about 1960, and then it is in close agreement until the end of the simulation in 2014. The esm-hist low bias in the earlier part of the 20th

century is attributed to changing trends in uptake by inaccurately prescribed land cover change, detailed in Section 4.1.2, with "crop" cover replacing rangeland, creating an artificial sink from the land cover change. The convergence of the observed and simulated historical atmospheric CO_2 by 1960 is explained in Section 4.1.2 with regard to the temporal pattern of the cumulative air-to-land CO_2 flux.

After 1960, atmospheric CO_2 concentration from the esm-hist experiment increases more rapidly and eventually leads to somewhat higher values than observed towards the end of the simulation (2005) by approximately 4.1 ppm. This result is consistent with the CMIP5 multi-model-mean overestimation of the atmospheric CO_2 concentration at the end of the simulation (2005) by about 5.6 ppm higher than in observations (Hoffman et al., 2014). However, some emission sources, e.g., biomass burning and aircrafts, were not included in our simulations (see Section 3), implying that our overestimate is likely higher than reported here.

Figure 2b shows the simulated surface air temperatures in the hist and esm-hist experiments as an anomaly with respect to the 1951-1980 mean temperatures obtained from observation-based results from GISS Surface Temperature Analysis (GISTEMP) version 4 (GISTEMP Team, 2020; Lenssen et al., 2019). Both hist and esm-hist experiments capture the overall slow rate of surface air temperature increase in the 1850-1970 period and the faster rate of increase after 1970. The esm-hist experiment agrees particularly well with GISTEMP v4 results in the 1880-1940 period where the surface air temperature is low. These patterns are consistent with the atmospheric CO_2 concentration patterns in which CO_2 increases gradually in the 1850-1960 period and increases sharply after about 1960. As described before, atmospheric CO_2 concentration in the 1850-1960 period is lower in the esm-hist experiment which likely leads to lower surface air temperatures than the historical experiment.

The 2010-2014 mean column-averaged atmospheric CO_2 concentration at each grid point from the esm-hist experiment is compared with column-averaged NOAA CarbonTracker CT2017 atmospheric CO_2 mole fraction data (Peters et al., 2007) and GOSAT (Greenhouse gases Observing SATellite; JAXA/NIES/MOE) data for the same 5 year period to assess the spatial distribution of atmospheric CO_2 (Figure 3a-c). The GISS ModeE2.1 esm-hist global mean atmospheric CO_2, about 400 ppm, is much higher than observed, which is about 390 ppm, so these maps show the difference relative to those respective means, to reveal differences in the distribution patterns. Seasonal averages over December-January-February (DJF), March-April-May (MAM), June-July-August (JJA), and September-October-November (SON) are also shown in Figure 3d-o. The model captures the hemispheric contrast in atmospheric CO_2 concentration, the Northern Hemisphere (NH) higher winter concentrations, and summer growing season draw-down of atmospheric CO_2. However, the NH seasonal amplitude is relatively damped compared to observations, while the Southern Hemisphere (SH) concentration is consistently lower relative to the mean (albeit still higher than observed). The model does not reproduce MAM having the highest relative NH atmospheric column CO_2, which is due to the model not capturing a respiration pulse prior to spring leafout, because the GISS ModelE2.1 high latitude temperatures are biased low by ~8 °C (see Kelley et al., 2020), causing a later soil thawing and suppressed respiration. This low temperature bias also causes a damped summer NH uptake of atmospheric CO_2 in JJA, due to cold-suppressed summer GPP at the same time with a tardy soil respiration pulse. Because of the low land uptake in JJA, soil respiration losses with fall senescence manifest as higher NH atmospheric CO_2 in SON relative to the model mean compared to observations. The consistently low SH atmospheric column CO_2 is a result of the damped NH cycle and reduced concentration gradient for equatorial transport as well as

due to high CO_2 uptake by the oceans associated with overestimated NPP (Lerner et al, 2020).

At the continental scale, due to localized emissions sites, the model exhibits high concentrations of atmospheric CO_2 in winter due to fossil fuel emissions over the Eastern United States, Eastern China, and Europe, but because we do not simulate fire or biomass burning, the locally elevated atmospheric CO_2 concentrations over central Africa and the Amazon are absent in our model.

4.1.2   Terrestrial biosphere CO_2 uptake

Table B1 summarizes the components of land carbon fluxes and stocks in the final year of each experiment. Gross primary productivity (GPP) and net primary productivity (NPP) values of the historical (116.7 GtC/yr and 43.8 GtC/yr, respectively) and esm-hist (117.6 GtC/yr and 44.7 GtC/yr) simulations are all well within the range of observed GPP and NPP uncertainty and prior MIP GPP ranges of 90-150 GtC/yr (Houghton et al, 2007; Welp et al., 2011; Piao et al, 2013), with NPP being just over about a third of GPP. The GISS ModelE2.1 historical and esm-hist experiments show that the combined CO_2 fertilization effect and warming-induced increases to leaf photosynthetic capacity result in enhanced GPP; we analyze their separate contributions later in Section 4.2. The net effect with warming-induced increases to respiration (both autotrophic and heterotrophic) leads to ~30% higher NPP in the historical and esm-hist experiments compared to the preindustrial, only slightly higher in the emissions-driven experiments. Compared to inventory estimates of ~380-536 (mean 450) GtC in global vegetation biomass (Erb et al., 2018), the model biomass estimate without the labile pool is low at 340 GtC, but with the labile pool the total is high at 563 GtC. The total soil carbon stocks for the historical and esm-hist experiments are close to the median of ~1500 GtC of inventory map estimates and well within the 504-3000 GtC literature range (Scharlemann et al., 2014).

We examine next the transient behavior of these experiments to understand the final year outcome. The annual mean CO_2 exchange between the terrestrial biosphere and the atmosphere for the historical and esm-hist experiments during the 1850-2014 period is shown in Figure 4a,c. The concentration-driven and emissions-driven experiments yield very similar evolution of fluxes. Cumulative uptake for the two experiments is also similar at the end of 2014; however, the historical experiment has higher uptake until approximately 1980 because atmospheric CO_2 concentrations are higher.

The changes in trends of air-to-land CO_2 flux in the historical and esm-hist experiments are strongly tied to rates of land conversion to human land uses (Figure 1), which accelerated with population growth starting in the colonial era ~1700 A.D., with the industrialized world plateauing in the late 1990's (Pongratz et al. 2008). With regard to vegetation behavior, after decades of regrowth, CO_2 fertilization, and fire suppression (Pongratz et al., 2008), satellite observations have shown that the Earth experienced a greening trend from the 1980's to late 1990's or 2000's due to CO_2 fertilization (Zhu et al., 2016), but the trend has reversed since then, with declining LAI due to water stress from increasing vapor pressure deficits (Yuan et al., 2019).

The effect of land cover changes is reflected in the land use change term that comprises a component of air-to-land flux. In our model, the terrestrial biosphere flux is determined as

$$F_L = npp - rh - cProduct \dotfill 2$$

where F_L is the air-to-land flux (positive into surface), npp is net primary productivity, rh is heterotrophic respiration, and $cProduct$ is carbon flux in products of land use change, which corresponds to the change in vegetation biomass in our model. The magnitudes of npp and rh terms are much larger than $cProduct$; however, the difference of npp and rh is small and on the same order as the $cProduct$ term (Figure 5a,b). For this reason, the $cProduct$ term has a relatively large influence on the evolution of carbon uptake by the terrestrial biosphere. The $cProduct$ magnitudes can be attributed to erroneous specification of land cover change due to the LUH inclusion of rangeland in the "pasture" cover type.

The annual fluxes due to net ecosystem productivity $npp - rh$ in Figure 5a are low compared to the magnitude of land CO_2 fluxes (excluding land use change) estimated in a multi-model DGVM compilation in Friedlingstein et al. (2019), which ranges ~ −6 to −3 GtC/yr by 2010-2018. The CO_2 fluxes from land use change in Friedlingstein et al. (2019) are based on the LUH2 data set and do not have the 1850-1900 carbon sink as in our $cProduct$ term, while the trends through the 20th and early 21st centuries reflect still the stronger loss at 1950-1960 and steady losses thereafter, but restricted within 1-1.5 GtC/yr. The net effect in our model is that the cumulative air-to-land flux remains negative after about 1920 onwards, with the land as a steady source.

The departures of the esm-hist simulated atmospheric CO_2 concentration from observations can be explained by how our land model does or does not deal with the above known forcings and responses. The annual and cumulative air-to-land flux (Figure 4a,c) trends to closely follow our prescribed crop, pasture, and rangeland cover changes as plotted earlier in Figure 1, with the consequent changes in biomass carbon stocks quantified as $cProduct$ in Figure 5; the land carbon uptake is positive over 1850-1900, negative over 1900-1960, with a spike in loss to the atmosphere over 1950-1960, then becomes a net sink after 1960. In comparison to the historical experiment, the lower cumulative uptake in the esm-hist experiment over 1850-1900, where both have otherwise the same cover changes, is due to several factors: 1) The largest driver is from using the CMIP5 LUH land cover data set, in which the rangeland included with "pasture" replaces a large fraction of less productive natural shrubland cover with our crops PFT, particularly in the Western U.S., South America, and Australia. This adds biomass and LAI to these areas and creates a significant sink. In the LUH2, crops and pasture without rangeland in fact mostly replaces forest in the Eastern U.S. during 1850-1900, the losses from which would be more consistent with observed CO_2. 2) Our prescribed LAI from the year 2004 occurs during a period of peak LAI, such that monthly LAI prescribed from 2004 values is higher than actual LAI earlier in the century (Yuan et al. 2019), thus driving overestimated CO_2 uptake. Finally, 3) a positive feedback results from the lower simulated atmospheric CO_2 concentration causing less of a fertilization effect. The low simulated atmospheric CO_2 concentration over the pre-industrial to mid-20th century is therefore due to an overestimated land sink. Together, these lead to the land gaining biomass carbon and serving as a net sink.

Over 1900-1950, the switch of the land to being a net source corresponds to a marked increase in losses of biomass carbon due to land cover change replacing natural grasslands and small areas of forest with crops. Shrubland cover also is replaced at an accelerated rate, but the net effect is of a net loss, as shown in the negative $cProduct$ values over this period in Figure 5a.

Over 1950-1960, there is a jump in crop/pasture/rangeland cover expansion and conversion of natural vegetation, not only shrublands but also grasslands and forest to the crop PFT (Figure 1), affecting all continents. This cover change is expressed as the strong dip

in the air-to-land flux as seen in Figures 4a and 4c and quantified as *cProduct* in Figure 5. This cover replacement continues at a slower rate through the end of the historical period.

The recovery to historical atmospheric CO_2 concentration levels by 1990 and slightly exceeding the historical levels of atmospheric CO_2 concentration by the 2000's appears to be due to lower than actual land CO_2 uptake rates, when compared to the CT2017 estimate. After 2000, while CO_2 fertilization leads to land being a net sink, simulated atmospheric CO_2 concentration is still higher than observed due to insufficient land CO_2 uptake. The lower simulated air-to-land flux compared to the CarbonTracker estimate since 2000 could reflect underestimated GPP or higher respiration, but the model estimate is within the lower uncertainty bound of CarbonTracker. This latter period is observed to have experienced a reduction in global vegetation growth due to drying (Yuan et al. 2019), but also high latitude increases in LAI since 2004 (Zeng et al. 2018; Zhu et al. 2016). The insufficient simulated land CO_2 sink may be because our experiments lack the high latitude enhanced LAI growth, or because they do not distinguish primary from secondary forest and miss the enhanced CO_2 sink from regrowth of the latter. Given the decline otherwise of global vegetation growth, however, if the modeled land CO_2 uptake is low, it is most likely because soil respiration is high, or our biophysics has low GPP due to either drought stress or an insufficient sensitivity to the CO_2 fertilization effect. A small, steady prescribed replacement of natural grasslands and forests by crops and pasture also contributes to losses, but if these changes did occur, then other factors are still required to compensate for them.

Combined, the above described effects can explain our land model's being an excessive sink of CO_2 during the early 20th century and an insufficient sink in the 2000's, hence the difference in atmospheric CO_2 concentration between the observed and the esm-hist experiment (Figure 2).

Now we examine spatial patterns of simulated atmospheric CO_2 concentration and fluxes. Figure 6a compares the continental scale mean 2000-2014 air-to-land fluxes between the model and CarbonTracker. Both the concentration-driven and emissions-driven historical experiments are fairly consistent with CarbonTracker data. The medians of the modeled flux agree well with CarbonTracker, however, for land, the model interquartile ranges have smaller range with underestimated maxima, indicating that our model does not capture extremes.

The spatial patterns of terrestrial biosphere CO_2 flux for the mean of 2000-2014 period are assessed in Figure 7a-c. The historical and esm-historical experiments demonstrate similar spatial patterns and magnitude of flux. Current satellite observations are not sufficient to resolve local or regional patterns of fluxes but are appropriate at the continental to hemispheric scales. Nonetheless, when the two model experiments are compared roughly to CarbonTracker data, the spatial pattern is fairly consistent; however, the magnitude is damped with overall less uptake than observed, as indicated previously in Figures 4a and 6a; this is consistent with the damped simulated seasonal atmospheric CO_2 concentration described earlier. Some small regions that show negative air-to-land flux in CarbonTracker, i.e., Indonesia, Ukraine, and the Pacific Northwest, instead have positive uptake in both historical and esm-hist experiments. These differences may be due to lack of deforestation and biomass burning in our model. The seasonal maps in Figure 7d-o provide some explanation for pronounced differences in JJA in the boreal and subarctic zones, where land CO_2 uptake is too low, and the Amazon Basin, where there is a net loss. During 2000-2014, global LAI was fairly flat in trend (Zhu et al., 2016), so the discrepancy cannot be due to different LAI. These discrepancies appear due to biases in the GISS ModelE2.1 simulated climate (see Kelley et al., 2020). The NH high latitudes (boreal, subarctic) are too cold during

the JJA growing season, causing reduced photosynthetic capacity from frost hardening in our model. The Amazon Basin has overall low precipitation and excess temperatures, leading to higher soil respiration and higher water stress, particularly in JJA during a time that should have peak productivity, resulting in net carbon losses rather than uptake and an appearance of reverse seasonality in fluxes.

Uncertainties in CarbonTracker terrestrial biosphere CO_2 flux are large in some regions, particularly in high latitudes in the Northern Hemisphere, and may reach the same magnitude as the signal itself. This is partially due to scarcity of ground observations, in particular, over Siberia and in the Amazon region. For this reason, although the comparison of model to CarbonTracker flux data provides a decent assessment of the model's skill, such a comparison of magnitudes in regional scale is not entirely conclusive, and the relative patterns should be considered taking into account the regional uncertainties of both datasets.

4.1.3 Ocean CO_2 uptake

The air-to-sea annual mean CO_2 flux and cumulative uptake for the historical and esm-hist experiments during the 1850-2014 period are shown in Figure 4b,d. The CO_2 flux into the ocean in the concentration-driven experiment is slightly higher than in its emissions-driven counterpart until about 1950, but this behavior reverses after 1950, and by 2014 both simulations estimate similar air-to-sea CO_2 flux. This difference in uptake between the historical and esm-hist experiments is arising from the difference in atmospheric CO_2 concentration between the two experiments where atmospheric CO_2 concentration in esm-hist experiment is lower than in the historical experiment until about the mid-20th century by about 10-20 ppm (Figure 2a). As explained in section 4.1.2, the low atmospheric CO_2 concentration during the initial years (1850-1900) of the esm-hist simulation is driven by the land being an excessive carbon sink. Since the atmospheric CO_2 concentration is low, air-to-sea flux is also low as the CO_2 gradient between the atmosphere and the ocean is small, leading to smaller amounts of CO_2 uptake by the ocean from the atmosphere in the esm-hist experiment compared to the historical experiment. For this reason, cumulative ocean uptake in the esm-hist experiment is lower than the historical experiment (Figure 4d), and thus, the ocean is not strongly driving the evolution of atmospheric CO_2 concentration in this period. Once the land becomes a carbon source in 1900, atmospheric CO_2 concentration begins to increase (Figure 2a), and accordingly, ocean CO_2 uptake begins to increase as well (Figure 4b,d). The land becomes a sink again in 1960, however, atmospheric CO_2 keeps increasing until the end of the simulation (Figure 2a) mainly due to sharply increasing anthropogenic CO_2 emissions, and the ocean uptake keeps increasing in response to the increasing atmospheric CO_2 concentration. The deviation in ocean uptake between the historical and esm-hist decreases towards the end of the simulation, and at the end, the cumulative ocean carbon uptake in the historical experiment is only slightly higher (~6 GtC or about 4%) than the esm-hist experiment. The cumulative uptake by the ocean in our model is consistent with studies that have utilized observation-based data synthesized by the GLODAPv2 (Olsen et al., 2016). Net cumulative ocean uptake of both anthropogenic and natural CO_2 during the 1994-2007 period in the historical and esm-hist experiments, which are 26.2 GtC and 29.2 GtC, respectively, are within the uncertainty range of 29±5 GtC reported in Table 3 of Gruber et al. (2019).

Figure 6b compares the mean 2000-2014 air-to-sea flux between the model and CarbonTracker regionally. Both the concentration-driven and emissions-driven historical experiments are generally consistent with CarbonTracker data for the northern and southern mid-latitude regions, and the subtropics regions are not far off. Modeled fluxes in the tropics

are biased showing decreased outgassing and in some places even uptake CO_2 due to biases in Ekman induced upwelling that produce warmer surface temperatures and shallower mixing for alkalinity and DIC (as discussed in detail in Lerner et al, 2020). Polar regions are noticeably different between the model and CarbonTracker. Although the model does simulate slightly stronger uptake than observed (Lerner et al, 2020), CarbonTracker uncertainties in the region are also large.

The spatial patterns of air-to-sea CO_2 flux for the mean of 2000-2014 period in the model and observations are compared in Figure 7a-c. The historical (Figure 7a) and esm-hist (Figure 7b) experiments demonstrate very similar spatial patterns and magnitude of air-to-sea CO_2 flux when compared against each other, which was also observed in Figure 6b. When the two model experiments are compared to CarbonTracker data, the spatial pattern is broadly consistent where the regions close to the poles are mostly CO_2 sinks and tropical regions are CO_2 sources. Differences are most apparent in the Southern Ocean, equatorial Pacific, and subpolar regions. Although both the model and CarbonTracker indicate Subpolar North Atlantic Ocean as having high positive air-to-sea CO_2 flux, both historical and esm-hist experiments indicate even higher positive fluxes in this region. For the equatorial oceans, both simulations show less outgassing than the CarbonTracker data. Below, we briefly discuss the mechanisms underlying the CO_2 flux bias; a more detailed discussion can be found in Lerner et al. (2020).

Subpolar Regions: The model overestimates the CO_2 flux from the atmosphere to the ocean in both the subpolar North Atlantic and Pacific due to the model's underestimate of pCO_2 in both regions. However, the underlying cause of the pCO_2 underestimation differs between these two regions. In the subpolar North Atlantic, the model overestimates wintertime mixing (Figure S1a,b), which results in the model overestimating surface DIC and alkalinity. If the biases in alkalinity and DIC were equivalent, then one would expect no impact on the bias in pCO_2 since alkalinity acts to decrease pCO_2 while DIC contributes positively to pCO_2. However, the alkalinity bias in intermediate and deep waters is greater than that for DIC in this region (Figure S2). Thus convective mixing contributes to the surface alkalinity bias more strongly than it does to the DIC bias, causing the model to have higher excess alkalinity (the difference between alkalinity and DIC) in the surface waters of this region (Figure S1c,d).

In the Subpolar North Pacific, however, the bias appears to be temperature driven, as the model underestimates temperature by ~2 °C (Figure S1e,f), which lowers the solubility of CO_2 and increases the flux of CO_2 into the ocean from the atmosphere in this region. Additionally, in the Subpolar North Pacific, the model systematically overestimates DIC and alkalinity (as in the Subpolar North Atlantic), but since excess alkalinity is not systematically larger than observations (Figure S1c,d), temperature likely plays a stronger role than alkalinity and DIC biases in driving the air-to-sea CO_2 flux bias in this region.

Southern Ocean: The model overestimates alkalinity and the CO_2 flux from the atmosphere to the ocean in the Southern Ocean as well. However, the reason behind the excess alkalinity bias in the Southern Ocean is unclear, as the mixed layer depth in the model is consistent with observations, and primary production (PP) is only slightly overestimated by the model in austral winter. While the model predicts a mixed layer depth consistent with observations, it also displays a positive excess alkalinity bias in the Atlantic Sector of the Southern Ocean below 200 m (Figure S3). Since the model's mixed layer depth can exceed 200 m in some regions of the Southern Ocean (Figure S1a,b), particularly during austral winter, the bias in surface excess alkalinity may stem from regions where the model transports subsurface waters that are depleted in DIC and enriched in alkalinity compared to observations.

Equatorial Pacific: In the Equatorial Pacific, two factors contribute to the model's underestimation of outgassing of CO_2. The first is the model's underestimation of wind speed (Figure S1g,h), which reduces the gas transfer velocity as well as Ekman induced upwelling of deep waters enriched in DIC, and hence the air-to-sea CO_2 flux. The second is the model's high excess alkalinity, particularly along the coast of South America (Figure S1c,d). This excess alkalinity bias likely stems from the upwelling of waters within and below the thermocline that also contain excess alkalinity that is higher than observed (Figure S4).

In three of the four regions of largest CO_2 flux bias, the bias seems at least partially to be due to the transport of subsurface and intermediate waters that have positive biases in excess alkalinity. While the cause of these biases is unclear, a possible explanation is the model's representation of calcium carbonate dissolution below the surface layer. The model does not contain a prognostic calcium carbonate pool; rather, the model assumes that the export flux of calcium carbonate below the compensation depth (75 m in the model) is exponentially decreasing with depth, with the maximum flux proportional to the integrated primary production above the compensation depth. The dissolution of calcium carbonate, and hence the alkalinity source below the compensation depth, is then calculated as the vertical divergence of the aforementioned export flux. Notably, the calcium carbonate dissolution rate does not depend on the saturation state of seawater. Such a treatment may overestimate the alkalinity source in the mesopelagic and bathypelagic zones, since in the model the calcium carbonate flux is tightly linked to primary production, whereas in the real ocean this flux stems from production of only calcium carbonate producing organisms (e.g., coccolithophores, foraminifera, etc.).

Biases in the air-to-sea CO_2 flux may also stem from biases in nutrient distributions and their impact on productivity. Interestingly, the regions of the largest negative surface nitrate biases, in the Equatorial Pacific and Southern Ocean (Figure S1i,j), are associated with positive biases in primary production (Figure S1k,l). This suggests that low nitrate concentrations in these regions are at least partly due to increased nutrient utilization by phytoplankton, though in the equatorial Pacific they are likely also a consequence of reduced upwelling of nutrient rich sub-thermocline waters due to weaker surface wind than observed. High productivity tends to decrease DIC and contributes to both regions being a weaker CO_2 source (equatorial Pacific) or stronger CO_2 sink (Southern Ocean) than observed (Figure 7a,c).

Physical ocean biases are documented in Kelley et al. (2020) and in Lerner et al. (2020): there is still a double ITCZ (Intertropical Convergence Zone) in the Pacific and excessive precipitation patterns in the tropics as well as larger than observed snowfall in the mid latitudes. Sea surface temperature positive biases are found mostly in the Eastern Boundary Upwelling and in the storm tracks regions and are associated with radiation and cloud precipitable water deficiencies. Salinity positive biases are mostly found in marginal seas but also the broader tropical open ocean regions and are associated with precipitation biases. Deeper than observed mixing is found in the North Atlantic deep water formation regions and the Southern Ocean south of the Antarctic Polar Front.

4.2   Biogeochemically coupled historical simulations

The global annual mean cumulative uptake of CO_2 by the land and ocean reservoirs in both concentration-driven and emissions-driven biogeochemically coupled historical scenarios are shown in Figure 8c-f. As expected, the biogeochemically coupled scenarios for both concentration-driven and emissions-driven experiments for the land have higher uptake

of CO_2 than their fully coupled counterparts towards the end of the simulation (Figure 8c,e) since there is no enhancement to autotrophic and soil respiration in the absence of warming from the radiative effect. The historical and hist-bgc experiments have identical cumulative uptake until about 1900 while the hist-bgc experiment accumulates more carbon after 1900, with total land carbon stocks in 2014 enhanced by 30.2 GtC in the hist-bgc experiment compared to the historical, stored in soil carbon (Table B1). For the emission-driven counterpart, this identical accumulation period lasts until approximately 1960, as there is less of a CO_2 fertilization effect due to lower atmospheric CO_2 concentration during this period compared to the historical atmospheric CO_2 concentration. At the end of the simulation, the esm-hist-bgc experiment stores 25.5 GtC more on land than the esm-hist (Table B4). The esm-hist-bgc experiment has negligibly lower cumulative CO_2 uptake than the hist-bgc experiment by 4.6 GtC (Figure 8c,e). Soil carbon stocks are higher by ~3% in the biogeochemically coupled experiments compared to the preindustrial, and by ~1% when the radiative effect is included.

Our model releases any enhanced carbon uptake to the soil, but we also encountered a model error in which there is excess accumulation of labile carbon (non-structural carbohydrates, NSC) in grasses (cOther in Table B1), which lack calculation of potential height growth or simulated full mortality to move this biomass as litter to the soil. This plant labile carbon pool does not respire in our model, whereas it would more realistically be transformed to higher plant density or leaf area, litterfall, and more soil carbon storage, some of which would be respired back to the atmosphere. The labile carbon accumulation in all plants is 223 GtC, a third of the total vegetation biomass of 563 GtC. Given the estimate of Erb et al. (2018) of ~380-536 (mean 450) GtC in global vegetation biomass, and our woody stem biomass of 310 GtC, this accumulated labile carbon is perhaps excessive by ~60 GtC, but not excessive enough to adversely impact our simulated carbon cycle given uncertainties in both soil carbon and vegetation biomass.

The cumulative ocean uptake of CO_2 shows very small variations between the fully and biogeochemically coupled configurations of both concentration-driven and emissions-driven simulations (Figure 8d,f). For the concentration-driven simulations, the biogeochemically coupled case has a slightly higher cumulative uptake at the end of the simulation (historical: 136.3 GtC, hist-bgc: 150.3 GtC in 2014), whereas the fully and biogeochemically coupled versions are almost identical in the emissions-driven simulations (esm-hist: 130.2 GtC, esm-hist-bgc: 129.8 GtC in 2014). In the concentration-driven historical simulations, the atmospheric CO_2 concentration is prescribed, and hence, it is exactly the same in the fully and biogeochemically coupled simulations. However, in the hist-bgc experiment, there is no coupling of the atmospheric CO_2 evolution to the radiation, hence the surface ocean does not warm as much as in the fully coupled experiment (Figure S5b in the Supporting Information), resulting in the uptake being higher in hist-bgc due to the solubility pump being stronger (Figure S5a). The role of the biological pump manifests itself through primary production (PP) and carbon export (cexp, computed in the model at 75 m depth). Initially (earlier than 1950) both PP and cexp are weaker in the biogeochemically coupled experiment than in the fully coupled experiment (Figure S5d,e) because warmer water in the fully coupled experiment stimulates primary production. During this period, PP and cexp contribute to a tendency towards reduced ocean carbon uptake in the biogeochemically coupled vs. the fully coupled experiment. However, the difference in the solubility pump is greater than the difference in the biological pump between these two experiments, so that carbon uptake is actually greater in the biogeochemically coupled experiment than the fully coupled experiment. However, after 1970's, both PP and cexp get stronger in the biogeochemically coupled experiment than in the fully coupled experiment

(thus now contributing to stronger uptake than in the fully coupled experiment) due to the fact that mixed layer depth (MLD) in the fully coupled experiment is shoaling (surface warming leads to more stable stratification hence shallower mixing) and leads to enhanced nutrient depletion at the surface, limiting PP in the fully coupled experiment (Figure S5c). Therefore, in the concentration-driven simulations, it is the solubility pump that drives the uptake differences with minor contributions from the biological pump in the early part of the simulations (before 1970), but the biological pump plays a more important role in driving the uptake differences in the later part of the simulations (after 1970).

However, the ocean response in the emissions-driven simulations is more complicated than in the concentration-driven simulations because the ocean adjusts interactively to the atmospheric CO_2 concentrations. Unlike in the concentration-driven simulations, there is an evolving atmospheric CO_2 concentration difference between the esm-hist and ems-hist-bgc experiments, such that before 1910, atmospheric CO_2 concentration in the two experiments is essentially the same (+/- 2 µatm), and hence, differences in ocean uptake are small (Figure S6a).

In the period 1910-1950, atmospheric CO_2 concentration in the esm-hist-bgc experiment is greater than in esm-hist experiment, but the difference is decreasing to zero by 1950. Thus, during this period the ocean uptake in the biogeochemically coupled experiment is greater than in the fully coupled experiment. After 1950, the atmospheric CO_2 concentration in the esm-hist experiment increases faster than in the esm-hist-bgc experiment, leading to more uptake in the fully coupled experiment which overcomes the effect of surface warming. Therefore, the difference between the uptake in these two experiments decreases with time, switches sign around 1980, and remains small thereafter. Primary production and carbon export differences increase (Figure S6d,e). However, whereas PP difference is similar to the one in the concentration-driven simulations (Figure S5d), cexp difference is smaller than in the concentration-driven simulations (Figure S5e), hence the very small uptake difference at the end of the simulations, possibly indicating a different role of the biological pump in the emissions-driven simulations. The different behavior in the emissions-driven and the concentration-driven simulations is going to be the subject of subsequent studies as it requires more in-depth, quantitative analysis than the one provided here.

Overall, the ocean takes up a much greater amount of carbon than the land in either concentration-driven or emissions-driven simulations, as was the case for the fully coupled historical and esm-hist experiments. This trend is similar to many CMIP5 models where the ocean uptake was mainly in the range of 100-150 GtC in the 2000's and was greater than the land uptake which showed considerable disagreements but nevertheless had mostly net zero or negative cumulative uptake in the 2000's (Friedlingstein et al., 2014).

The spatial distributions of the difference between the fully and biogeochemically coupled simulations (2000-2014 mean) are shown in Figure 9. On land, the CO_2 fertilization effect without the radiative effect does not result in uniform enhancement of the uptake, but there are regional variations for both concentration-driven and emissions-driven simulations. However, these regional differences are not all the same for the concentration-driven versus emissions-driven experiments. Both concentration-driven and emissions-driven simulations show greater uptake in the fully coupled experiments (historical, esm-hist) in the northeastern U.S., Russia, south-central Africa, and eastern China (Figure 9a,b). However, only the concentration-driven simulations show greater uptake with the fully coupled experiment than the biogeochemically coupled experiment in a crop-dominated area of South America, at the boundary between North and South America, and in a large portion of continental Europe

(Figure 9a,b). Alaska, India, and southeastern Asia show larger uptake in the fully coupled experiment in emission-driven simulations (Figure 9a,b). These are reflected in seasonal maps (Figure 9c-j). The fully coupled esm-hist experiment shows substantially lower uptake than its biogeochemically coupled counterpart during winter months in Europe (DJF) and crop-dominated area of South America (JJA) which is not the case for concentration-driven simulations (Figure 9c,d). In DJF, southeastern Asia shows larger uptake in the fully coupled experiment in the emission-driven mode than in the concentration-driven mode (Figure 9c,d). India and Alaska show larger uptake in the fully coupled experiment in the emission-driven mode in SON and JJA, respectively (Figure 9i,j). For the ocean, the air-to-sea CO_2 flux difference between the fully and biogeochemically coupled experiments is smaller (Figure 9a,b). Generally, the ocean shows weaker uptake in the fully coupled simulations over most of the ocean due to warmer surface temperatures, shallower mixed layers, and weaker mixing (Figure S5,S6). There are two notable exceptions where the uptake is stronger in the fully coupled than the biogeochemically coupled simulations: i) the Norwegian-Greenland Seas and ii) the Weddell Sea. In both regions, sea ice melting due to warmer temperatures leads to more ocean surface available for exchanges with the atmosphere and hence uptake there. The concentration-driven simulations are characterized by less sea ice melting than the emissions-driven simulations, hence the effects in the emissions-driven simulations are more pronounced in the Norwegian-Greenland Seas. The equatorial Pacific and Atlantic Oceans show larger uptake in the fully coupled simulations than the biogeochemically coupled simulations in emission-driven experiments which is noticeably different from the concentration-driven simulations. The role of the biological pump is smaller than the solubility pump in the concentration-driven simulations, however the emissions-driven simulations merit more in-depth analysis in future work.

4.3   1% per year increasing CO_2 simulations

CO_2 exchanges between the atmosphere and the land and ocean for the fully, biogeochemically, and radiatively coupled 1pctCO2 experiments are shown in Figure 10a,b. At the time of quadrupling of CO_2, i.e. after 140 years, we obtain 66% higher uptake of CO_2 on land and 30% higher uptake in the ocean (Table B1,2) in the 1pctCO2-bgc experiment compared to the 1pctCO2 experiment. In the radiatively coupled experiment, 1pctCO2-rad, surface warming, while the land and ocean carbon reservoirs only see preindustrial atmospheric CO_2 concentrations leads to both land and ocean becoming CO_2 sources. In all three 1pctCO2 experiments, the global uptake of the ocean is higher than that of the terrestrial biosphere (Figure 10a-d).

Since our model does not simulate changes in LAI, height growth, or cover change that might result from either fertilization or warming, the changes in land carbon stocks are due to changes in soil carbon (CO_2 fertilization enhances litterfall; warming enhances soil respiration) and in the labile carbon pool in vegetation (CO_2 fertilization increases GPP; warming can both increase carboxylation capacity of CO_2, $V_{c,max}$, and increase autotrophic respiration). In Figure 10e, higher GPP in the 1pctCO2 experiment than in the 1pctCO2-bgc experiment shows that there is a thermal fertilization effect with the increase of $V_{c,max}$ with temperature. However, for the net land carbon uptake trends, warming-induced soil respiration losses are in excess of this enhancement to GPP. Thus, comparing Figures 10c and 10g, trends in land carbon are almost entirely accounted for by changes in soil carbon, with a loss of ~600 GtC in the 1pctCO2-rad experiment compared to 1pctCO2.

The spatial distributions of the change in stored carbon in the terrestrial biosphere for the duration of the 140 year period of the simulations are shown in Figure 11a-c. The fully

and biogeochemically coupled experiments show similar patterns. On land, the biogeochemically coupled experiment stores more carbon at locations of high productivity, e.g., tropical and temperate regions. Increase in productivity is greater in the biogeochemically coupled experiment where respiration is not increased in the absence of radiatively induced climate warming. Most notable differences between the fully and biogeochemically coupled experiments are the net negative regions in northern Eurasia and northeastern North America found in the 1pctCO2 experiment and the uniform lack of any regions of carbon loss in the 1pctCO2-bgc experiment. The 1pctCO2-rad experiment shows largely net negative change in stored carbon with small patches of positive change in the northern high latitudes. As shown in Figure 10g, the predominant high latitude losses are due to increased soil respiration with climate warming, since in our experimental configuration, the vegetation cover and structure are fixed. The plant labile carbon pool of vegetation also will be more depleted from enhanced vegetation respiration at higher temperatures (not shown), but soil respiration accounts for almost all losses to the atmosphere here. The small patches of increased net uptake are due to warming in these locations boosting GPP to a greater extent than increasing soil respiration.

The spatial distributions of the change in stored carbon in the ocean over the 140 years of the simulations are shown in Figure 11. The fully coupled and biogeochemically coupled experiments demonstrate very similar patterns, implying that the carbon disequilibrium between the atmosphere and the ocean surface, i.e. the solubility pump, drives most of the ocean uptake, especially in the region of intense uptake in subpolar North Atlantic Ocean and along the pathway of the North Atlantic western boundary current. In this region, the biogeochemically coupled experiment shows greater increase in stored carbon than the fully coupled experiment because in the latter the effect of surface warming and enhancement of outgassing are included. An important characteristic of the ocean response in the 1pctCO2 and 1pctCO2-rad simulations with the GISS coupled model is the drastic decline of the Atlantic Meridional Overturning Circulation (AMOC) driven primarily by intense freshening in the polar/subpolar regions through melting of land and ocean ice. At the beginning of the simulation, the AMOC at 26 °N is about 25 Sv, and by the time of doubling of atmospheric CO_2 concentration, it has reached 18 Sv, and at the time of quadrupling, it is at 10 Sv. The AMOC slowdown leads to decreased uptake in the North Atlantic via the weakening of the solubility pump in the 1pctCO2-rad simulation but less so in the fully coupled simulation. While primary production is stimulated by surface warming in both the fully and radiatively coupled simulations, it is also impacted by the shallowing of the mixed layer and the reduced availability of nutrients from below. As a result, PP declines with time in both the fully and radiatively coupled simulations but remains stable in the biogeochemically coupled simulation. The Southern Ocean plays a critical role in the solubility pump because, while it remains unchanged in the 1pctCO2-bgc experiment, it becomes the dominant carbon sink in the fully coupled simulation due to the weakening of the North Atlantic sink with the AMOC strength reduction and the relative increase in importance of the Southern Ocean sink. The biogeochemically coupled simulation has no surface warming due to climate change, thus ocean circulation and mixing are stronger than in the fully coupled simulation which explains the stronger uptake.

The terrestrial biosphere uptake of carbon is also more sensitive to climate change than the ocean uptake. The uptake of carbon by the terrestrial biosphere is reduced to a greater extent in the radiatively coupled 1pctCO2-rad experiment (Figure 10c,d, 11c).

4.3.1 Carbon cycle feedbacks

A set of metrics known as carbon-concentration and carbon-climate feedback parameters are frequently used to quantify the response of carbon cycle storage on land and in the ocean to changes in atmospheric CO_2 concentration and climate (Arora et al., 2013; Boer & Arora, 2009, 2010, 2013; Friedlingstein et al., 2003, 2006; Gregory et al., 2009; Jones et al., 2016; Plattner et al., 2008; Roy et al., 2011; Schwinger et al., 2014; Zickfeld et al., 2011). In this section we discuss the feedback parameters estimated from our simulations.

We follow the integrated flux method (Friedlingstein et al., 2006; Arora et al., 2013) which expresses the changes in stored carbon in land and ocean components of the carbon cycle as an approximate linear sum of effects arising from changes in atmospheric CO_2 concentration and climate, which is written as

$$\Delta C_{L,O}^{COU} = \beta_{L,O} \Delta C_A^{COU} + \gamma_{L,O} \Delta T^{COU} \quad\text{...} \quad 3$$

where $\Delta C_{L,O}$, ΔC_A, β, γ, and ΔT represent the change in total land or ocean carbon storage (GtC), change in atmospheric CO_2 concentration (ppm), the carbon-concentration feedback parameter (GtC/ppm), the carbon-climate feedback parameter (GtC/K), and change in global mean surface air temperature (K), respectively. The subscripts A, L and O signify atmosphere, land and ocean components of the carbon cycle, respectively, and the superscript COU indicates a fully coupled simulation. The change in total carbon storage in the land or the ocean ($\Delta C_{L,O}$) is computed as the change in cumulative flux at a given year of simulation with respect to the cumulative flux of the initial year. Change in atmospheric CO_2 concentration (ΔC_A) is also with respect to the initial year. The change in surface air temperature is computed with respect to the mean of the initial ten years and the mean of the ten years prior to the desired year, e.g., for the year 1990, ΔT is computed as the difference between the mean temperatures of 1980-1990 and 1850-1860. Ten-year means are used in order to reduce large variabilities in surface air temperature. Here, change in global mean surface air temperature, ΔT, is regarded as an acceptable proxy to climate change (Friedlingstein et al., 2006; Schwinger et al., 2014). The β and γ parameters are computed as

$$\beta_{L,O} = \Delta C_{L,O}^{BGC} / \Delta C_A^{BGC} \quad\text{...} \quad 4$$

$$\gamma_{L,O} = \left(\Delta C_{L,O}^{COU} - \Delta C_{L,O}^{BGC} \right) / \Delta T^{COU} \quad\text{...} \quad 5$$

where the superscripts BGC indicate biogeochemically coupled simulations. Alternatively, γ parameter can be computed from the radiatively coupled experiment as

$$\gamma_{L,O} = \Delta C_{L,O}^{RAD} / \Delta T^{RAD} \quad\text{...} \quad 6$$

(e.g., Arora et al., 2013) though Schwinger et al. (2014) concluded that carbon-climate feedback computed this way underestimates the reduction of ocean carbon uptake (superscripts RAD indicate radiatively coupled simulations). Positive β parameter indicates negative feedback of land and ocean to the atmosphere because the growth of atmospheric CO_2 concentration is hindered from enhanced uptake in response to higher levels of atmospheric CO_2 concentration. On the other hand, negative β parameter indicates positive feedback of land and ocean to the atmosphere. Similarly, for the γ parameter, the feedback of land and ocean to climate change (warming surface air temperature) is negative and positive

when γ is positive and negative, respectively. For the computation of γ parameter, we follow the recommended COU‑BGC difference method (equation 5) from a CMIP5 study (Schwinger et al., 2014) because this method captures the response of the ocean to climate change under rising atmospheric CO_2 concentration. We only use equation 6 to compute γ when comparing our results to CMIP5 global multi-model-mean results reported in Arora et al. (2013).

The feedback parameters from the annual mean global cumulative uptake in the 1pctCO2 experiments are shown in Figure 12a,c. Throughout the simulation, the β for both land and ocean remains positive, meaning that the land and ocean take up more carbon with increasing atmospheric CO_2 concentration, and the feedback to the atmosphere is negative, because carbon uptake by land and ocean tends to reduce atmospheric concentrations and therefore climate change. Over the first few decades, β increases with the rising atmospheric CO_2 concentrations, but after approximately 80 years in the land and 70 years for the ocean, β decreases as we increase CO_2 in the atmosphere, meaning that the land and ocean sinks weaken. For the ocean, this behavior arises from the fact that the uptake increases almost as a cube root function (Figure 10b) where the rate of general change of uptake becomes almost constant at ~0.1 GtC/yr, indicating a saturation of carbon storage reservoirs. The atmospheric CO_2 concentration, however, keeps increasing exponentially at 1%/yr according to the definition of the 1pctCO2 experiments (Section 3), and the rate of change in atmospheric CO_2 concentration surpasses the rate of general change in ocean uptake near year 70. Therefore, equation 4 dictates that β decreases sharply. The land also experiences a similar effect where, in our model, the uptake rate decreases because of decreasing fertilization effect due to saturation of fixed leaf photosynthetic capacities for carboxylation of CO_2. At any given year, β for the ocean is larger than that for the land which is consistent with the larger cumulative CO_2 uptake of ocean compared to the land in the 1pctCO2 experiment (Figure 10c,d).

Unlike the β parameter, the γ parameter remains negative for both land and ocean at any given time and maintains an almost constant decreasing trend for the duration of the simulation (Figure 12a; computed using the $COU - BGC$ approach in equation 5). γ is negative because the land and ocean release more carbon to the atmosphere (increased outgassing) and therefore their carbon storage reservoirs decrease with increasing global temperatures. The feedback to the atmosphere is positive because it tends to enhance global warming. The magnitude of the land γ parameter is larger than that of the ocean, consistent with the larger cumulative loss of carbon from the land than the ocean in the 1pctCO2-rad experiment (Figure 10c,d).

Relative contributions of the different carbon cycle feedbacks and the atmospheric carbon burden are quantified in Figure 12d. The components are determined as $\beta_L \Delta C_A^{COU}$: land carbon-concentration, $\beta_O \Delta C_A^{COU}$: ocean carbon-concentration, $\gamma_L \Delta T^{COU}$: land carbon-climate, $\gamma_O \Delta T^{COU}$: ocean carbon-climate, and $m \Delta C_A^{COU}$: airborne where m is 2.12 GtC/ppm (Appendix 1; Arora et al., 2013; Gregory et al., 2009), and the five components are expressed as decimal fractions of the total sum. The magnitude of carbon-concentration feedback contribution is significantly larger than the carbon-climate feedback, meaning that the response of land and ocean is greater to increased atmospheric CO_2 concentration than to increased surface air temperature leading to negative overall carbon cycle feedback to CO_2 forcing. Throughout the simulation, both ocean and land carbon-concentration feedback contributions remain mostly constant with slight increases in the beginning. The land carbon-climate feedback contribution also remains mostly constant except with an opposite sign and on a much smaller scale. The ocean carbon-climate feedback contribution experiences a steady decrease (increase in magnitude) until about 100 years into the simulation, however,

its magnitude is smaller than the land counterpart at the time of quadrupling of atmospheric CO_2 concentration. The airborne fraction decreases in the first half of the simulation, reaching to about 0.5, then the second half experiences a slight increase.

The evolution of feedback parameters from the GISS ModelE2.1 is mostly comparable to those from some CMIP5 models reported in Arora et al. (2013). Here, the GISS ModelE2.1 γ parameters are computed using the $COU - BGC$ method (equation 5, Figure 12a) but also using the RAD method (equation 6; Figure 12b) for consistency with Arora et al (2013). At the time of quadrupling of atmospheric CO_2 concentration, the GISS ModelE2.1 ocean carbon-concentration feedback, 0.85 GtC/ppm, is similar to the value reported for the CMIP5 model mean (0.80±0.07 GtC/ppm). The land carbon-concentration parameter for the GISS ModelE2.1 model is 0.68 GtC/ppm which is smaller than the CMIP5 model mean value (0.92±0.44 GtC/ppm). The ocean carbon-climate parameter, γ_O, for GISS ModelE2.1 is larger (in the negative sense; −17.7 GtC/K) than the ensemble mean for some of the CMIP5 models (−7.8±2.9 GtC/K), indicating that the ocean in GISS ModelE2.1 is a stronger source of CO_2 with warming which relates to the large warm biases of GISS ModelE2.1 as well as the slowdown of the ocean overturning circulation with a warming climate. γ_L of the model (−54.8 GtC/K) is close to the CMIP5 estimate (−58.4±28.5 GtC/K).

Lastly, estimating the geographic distribution of the feedback parameters in order to highlight the regions which contribute most to their magnitude is informative. To do so, we computed the feedback parameters at each grid point based on equations 3 - 5 for the year of quadrupling of atmospheric CO_2 concentration (140th year; Figure 13). The β and γ parameters are now expressed per unit area (GtC/m^2/ppm and GtC/m^2/K) similar to Boer & Arora (2010), Boer & Yu (2003), Ciais et al. (2013), and Roy et al. (2010) in order to normalize by grid cell area. In order to avoid singularities in β and γ locally when the denominators become very small, i.e. for changes in atmospheric CO_2 concentration (ΔC_A^{BGC}) and temperature (ΔT^{COU}) locally which are very small, global averages in these quantities are used (Boer & Yu, 2003; Ciais et al., 2013).

The carbon-concentration feedback parameter is almost entirely positive throughout the globe. Regions with the largest contribution to the carbon-concentration feedback over land are found at equatorial tropical and subtropical regions, eastern North America, Europe, East Asia, and mid-latitude South America which is consistent with the fact that these regions are not limited by water or temperature and thus have high response to elevated atmospheric CO_2 concentration. This pattern is mostly consistent with CMIP5 multi-modal-mean β distribution in Ciais et al. (2013). For the ocean, areas with largest contributions are found in subpolar North Atlantic Ocean, which is a region of intense uptake of CO_2, and along the pathway of the North Atlantic Deep Water along the western boundary in the North Atlantic, as mentioned in Section 4.3. However, this differs from the CMIP5 multi-modal-mean (Ciais et al., 2013) which only shows moderately large β values at subpolar North Atlantic Ocean and small β values in the North Atlantic western boundary current. The discrepancy could be attributed to the fact that GISS ModelE2.1 has stronger overturning circulation in the North Atlantic than the CMIP5 model mean.

The carbon-climate feedback parameter is mostly negative with the locations of (negatively) largest γ occurring at similar locations as the largest β, e.g., tropical land and northwestern Atlantic Ocean. However, unlike β, γ switches sign over some distinct regions. For the terrestrial biosphere, patches of positive γ are found in northern North America and Eurasia. These patches are characterized by increases in GPP due to warming in these high latitude areas boosting temperature-sensitive photosynthetic capacity, as noted earlier. For the

ocean, positive γ values occupy the subtropical gyres, though the magnitude is generally lower than the positive land γ values. The Arctic Ocean also shows positive γ. Negative γ implies outgassing whereas positive implies uptake, hence in the subtropical gyres and in the Arctic Ocean, GISS ModelE2.1 takes up CO_2, due to carbon transport divergence.

4.4   SSP5-8.5 scenarios

4.4.1   Atmospheric CO_2 concentration

Figure 14a shows the global mean atmospheric CO_2 concentrations in the emissions-driven esm-ssp585 experiment and in the concentration-driven ssp585-bgc experiment. Atmospheric CO_2 concentration in the esm-ssp585 and the ssp585-bgc increase at nearly the same pace while the esm-ssp585 experiment maintains a slightly higher atmospheric CO_2 concentration throughout the simulation. In the final year of the simulation (year 2100) in the esm-ssp585 experiment, the atmospheric CO_2 concentration is 1160 ppm, whereas it is 1136 ppm in the concentration-driven experiment. We note, however, that the esm-ssp585 experiment started from the end of the esm-hist experiment which was at 7.2 ppm higher atmospheric CO_2 concentration than the concentration-driven historical experiment (Figure 2a).

Simulated global mean surface air temperatures in the esm-ssp585 and ssp585-bgc experiments are shown in Figure 14b. Temperature increases almost linearly in the esm-ssp585 experiment, reaching approximately 18.9 °C by 2100. The ssp585-bgc experiment, despite an absence of the radiative warming effect, still exhibits a small increase due to forcings in the SSP8-8.5 scenario other than CO_2 (e.g., aerosols, ozone, and land use change; O'Neal et al., 2016). This is likely due also to reduction in stomatal conductance at higher surface CO_2 concentration causing increases in surface temperature.

Figure 15 shows the distribution of column-averaged atmospheric CO_2 concentration; differences between the 2015-2024 mean and 2045-2054 mean are shown in Figure 15a,b, and differences between 2015-2024 mean and 2091-2100 mean are shown in Figure 15c,d. In the esm-ssp585 experiment, extensive warming is evident over the globe already by mid-21st century, except a small region in the North Atlantic which exhibits cooling of about 6 °C (the "cold blob"). Northern high latitudes (mainly the Arctic Ocean) warm by more than 6 °C while surface air temperatures over the Southern Ocean remain unchanged and the warming over Antarctica does not exceed 2 °C. Tropical oceans warm by about 4 °C (Figure 15c). Surface air warming is significantly reduced in the ssp585-bgc experiment (Figure 15b) compared to the esm-ssp585 experiment; the cold blob in the North Atlantic is only 2 °C colder than the beginning of the simulation, and the air over the Southern Ocean exhibits another cold "blob" in the Weddell Sea, but also other regions around Antarctica.

By the end of the 21st century, surface warming has increased everywhere, although the Southern Hemisphere has warmed at a faster rate (Figure 15c). The North Atlantic cold spot is still cooling at almost the same rate, however, its geographic location has shifted slightly towards the northeast. Southern Hemisphere high latitudes experience now analogous warming as their Northern Hemisphere counterparts. Interestingly, in the ssp585-bgc experiment, the North Atlantic cold blob is broader and significantly colder than in the esm-ssp585 experiment, and the Southern Ocean is also cooling.

4.4.3 Terrestrial biosphere and ocean CO$_2$ uptake

Air-to-land and air-to-sea annual mean CO$_2$ flux in the ssp585-bgc and esm-ssp585 experiments are shown in Figure 14c,d. As the ssp585-bgc experiment experiences only a small amount of surface air temperature increase, the air-to-land flux in this experiment is higher due to less respiration than in the fully coupled esm-ssp585 experiment (Figure 14c). Both experiments have rising trends in both fluxes after 2014 (Figure 14c,d), but warmer temperatures in esm-ssp585 lead to lower air-to-land and air-to-sea fluxes than in the ssp585-bgc experiment, resulting in lower cumulative uptake for esm-ssp585 (Figure 14e,f). For land, the rising trend peaks at 2070; thereafter in ssp585-bgc, the trend remains flat, while in esm-ssp585 it reverses slightly then rises again at 2100. The air-to-sea CO$_2$ flux peaks higher than land at 2080; after the peak, in ssp585-bgc the trend plateaus, whereas in esm-ssp585 it reverses slightly. At the end of the simulations, the air-to-land fluxes of the ssp585-bgc and esm-ssp585 experiments reach 6.0 GtC/yr and 3.8 GtC/yr, respectively (Figure 14c), and air-to-sea flux in the ssp585-bgc and esm-ssp585 experiments reach 7.2 GtC/yr and 5.2 GtC/yr, respectively, consistent with the esm-ssp585 result at the mean of the CMIP5 models. The partitioning of CO$_2$ into the atmosphere, land, and ocean can be seen in Figure 14e,f. In both experiments, CO$_2$ remaining in the atmosphere is roughly twice as much as CO$_2$ accounted for by the land and ocean combined uptake. The land uptake is noticeably smaller in the esm-ssp585 experiment in comparison to ssp585-bgc experiment owing to lower air-to-land flux throughout the simulation. The esm-ssp585 trends and magnitudes are in accord with the behavior of CMIP5 models that do not include a terrestrial nitrogen cycle but have land use change, as compared in Friedlingstein et al. (2014). However, the slight rise in the last ~3-5 years of the GISS ModelE2.1 esm-ssp585 simulated flux is not seen in the CMIP5 models.

Figure 16 shows the distribution of change in stored carbon; differences between 2015-2024 mean and 2045-2054 mean are shown in Figure 16a,b, and differences between 2015-2024 mean and 2091-2100 mean are shown in Figure 16c,d. In the ssp585-bgc experiment, most of the globe undergoes a positive change where the productive regions on land (e.g., southeast Asia) and North Atlantic Ocean (extending into the Arctic Ocean) show the highest increase in stored carbon (Figure 16b,d). The Southern Ocean also shows a noticeably large increase in stored carbon (Figure 16b,d). In the esm-ssp585 experiment, the extent of high uptake region in the North Atlantic Ocean is markedly smaller compared to the ssp585-bgc experiment, and especially in the high latitudes of this region, the increase in stored carbon is not different from other parts of the ocean (Figure 16a,c). Changes in carbon storage in the Southern Ocean in the esm-ssp585 experiment are similar to those in the ssp585-bgc experiment. There are parts of the land (the boreal zone in northern Russia) that show net decrease in stored carbon in the ssp585-bgc experiment (Figure 16b,d), and this is even more pronounced in the esm-ssp585 experiment (Figure 16a,c). The variabilities in change in stored carbon on land can be largely attributed to soil carbon, which increased by 9.3% from 1567 GtC at the end of the esm-hist experiment to 1714 GtC in the esm-ssp585 experiment by 2100 (Table B1). Net primary productivity nearly doubles for land by the end of the esm-ssp585 experiment (~45 GtC/yr) compared to the esm-hist experiment in 2014 (85 GtC/yr) (Table B1); since our model configuration releases carbon that would be allocated to plant height growth instead as litter to the soil, esm-ssp585 soil respiration in 2100 also nearly doubles, due to both warming and more substrate, such that the net increase in soil carbon reflects the CO$_2$ fertilization effect. The majority of the globe experiences a net increase in stored carbon; however, the northern high latitudes (boreal zone and subarctic) losses, i.e., northern Russia and northern North America, show more pronounced and more widespread net decrease in stored carbon (Figure 16a,c). These losses are due to increased

soil respiration from a shorter frozen season for the soil, with the spring thaw earlier and warmer by 2-5°C in May.

Figure 17 shows the evolution of the components of the carbon budget in the 1850-2100 period from esm-hist and esm-ssp585 experiments. Though the land CO_2 uptake (0.4 GtC/yr excluding land use change) and ocean CO_2 uptake (2.5 GtC/yr) combined accounted for about half of airborne CO_2 (6.3 GtC/yr) at the end of the historical period (2014), this proportion decreases in the SSP8-8.5 scenario. By 2100, the land uptake (4.1 GtC/yr excluding land use change) and ocean uptake (5.2 GtC/yr) accounts for about a third of airborne CO_2 (26.6 GtC/yr). The land and ocean are net sinks, slightly less at the end of the historical period compared to the inventory of Friedlingstein et al. (2019), which also has a slightly higher fossil fuel emissions estimate than the CMIP6 prescribed emissions. Their future increases as net sinks are in accord with the majority of CMIP5 models (Friedlingstein et al. 2014).

4.5   Climate biases and future regional climates

As a way to interpret our simulated climates, we classify the esm-hist and esm-ssp585 climates here with the Köppen-Geiger system, a well-known, long-developed geographer's collation of the correlation between climate (seasonality of monthly surface temperature and precipitation) and land biome types (Köppen, 1884). This system is widely used to assess the ecological suitability of climates for vegetation and to predict potential biome changes with climate change (e.g. Peel et al., 2007; Poulter et al., 2011; Rawal et al., 2014; Rubel & Kottek, 2010; Rubel & Kottek, 2011; Sanderson, 1999). Figure 18 shows the Köppen-Geiger classification for observed climate and the emissions-driven experiments. These figures use the 30-class scheme of Kottek et al. (2006) and Rubel & Kottek (2011) and color table of Beck et al. (2018). Observed climate in Figure 18a uses the 2001-2010 mean surface temperature from the Climate Research Unit (CRU TS 3.22; Harris et al. 2014), and the same decade mean precipitation from the Global Precipitation Climatology Centre (GPCC v6; Schneider et al. 2011).

As there was negligible difference in Köppen-Geiger classes between the concentration-driven historical experiment climate at the end of the period compared to the emissions-driven experiment, esm-hist, here we compare only esm-hist to the observed climate. The comparison of esm-hist (average of 2005-2014) to the CRU/GPCC observed climate shows that GISS ModelE2.1 captures the statistics of surface temperature and precipitation that distinguish biomes as classified by Köppen-Geiger. However, there are some important biases: 1) GISS ModelE2.1 is too dry in the tropics, particularly the Amazon Basin (Kelley et al., 2020), as is a well-known issue for many ESMs (Yin et al., 2012; Koutroulis, 2016), but the Köppen-Geiger classification makes this clearly significant for vegetation, with insufficient precipitation for Af, equatorial rainforest, and Am, equatorial monsoon, and some areas more appropriate for arid hot steppe, BSh, than for forest (in the Amazon basin); 2) GISS ModelE2.1 does not quite capture arid cold desert climates (BWk) as in central Eurasia and in general produces climates with slightly more precipitation; 3) the temperate/boreal transition (Dfa/Dfb) and tundra/subarctic (Dfc) are at lower latitudes in Eurasia in the esm-hist experiment, and Alaska is colder (ET instead of Dsa and Dsb), which reflect GISS ModelE2.1's cold bias in higher latitudes; 4) GISS ModelE2.1 does not capture the area of Siberia that is classified as Dwb (where there is extensive larch forest), instead simulating this as having colder winters (Dwc); 5) ModelE2.1 also does not capture the diversity of climate types (snow with dry hot or warm summers, Dsa, Dsb) on the Siberian peninsula due to being too cold, and meanwhile it simulates these climates in Central Eurasia

where there otherwise should be boreal forest and tundra/subarctic climates (Dfb, Dfc); 6) in the esm-hist experiment, Australia is slightly more humid than observed (more arid hot steppe, BSh, than arid hot desert, BWh); 7) in the Central-Western U.S., where observed climate is classified as arid cold steppe (BSk), as well as in Central Asia (BSk, and also Bwk, arid cold desert), GISS ModelE2.1 esm-hist climate simulates these regions as more humid types (Dsa, Dsb) or colder (Dfc).

The comparison between esm-hist and esm-ssp585 experiments (Figure 18b,c) shows major changes at both continental and finer regional scales: 1) warming climate shifts the temperate/boreal transition (Dfa, Dfb) to higher latitudes, with temperate forest climate nearly replacing boreal, and a broad band of Dfa, extending along ~60 °N; northward shift of tundra climate (Dfc) nearly eradicates polar (tundra or subarctic) (ET) environments; 2) reduction of the precipitation and higher temperatures in the Western U.S. shift the biomes to their hotter versions (Dsb to Dsa, and Csb to Csa, due to hotter summer temperature); 3) in continental Europe, all biomes shift toward those with hotter summers, and most of Spain changes from warm temperate with dry hot summers (Csa) to arid hot steppe (BSh); 4) the Tibetan Plateau changes polar tundra (ET) to still dry though warmer (Dwc), but with some more humid areas (Dfc); and 5) Alaska warms and Australia dries in the esm-ssp585 experiment relative to the esm-hist experiment, to be more similar to the 2001-2010 observed climate, which implies that they may in reality change commensurate to their shifts relative to the esm-hist experiment.

Some differences in biome class for the observations and experiments are between biomes with fine distinctions and climate/biome correlation uncertainties that can be due to imperfections in the Köppen-Geiger classification scheme, which is subject to many revisions and refinements in the literature (e.g. Cannon, 2012; Guetter & Kutzbach, 1990; Peel et al., 2007; Poulter et al., 2011; Trewartha, 1968; Trewartha & Horn, 1980; Zscheischler et al., 2012). Beck et al. (2018) produced a classification of future climate based on combining projections of CMIP5 models for 2071-2100. In their scheme, they changed the threshold between temperate (C) and cold (D) climates from 0 °C to 3 °C, and they did not include the class As. However, the broad patterns of climate change can be compared to our results. As with GISS ModelE2.1's esm-ssp585 experiment, the CMIP5 models also predict: 1) the broad swath of "snow fully humid with hot summer" (Dfa), corresponding to temperate broadleaf forests, across ~60 °N, and GISS ModelE2.1 marginalizes boreal needleleaf forest climate (Dfb) whereas the CMIP5 models maintain more of this type in its migration northward; 2) the warming of Alaska and the Tibetan Plateau; and 3) the dramatic regional changes in continental Europe. The CMIP5 models do not predict the warming of Australia as in GISS ModelE2.1, and they capture the region of Siberia where our model does not, and are on average not as dry-biased in the Amazon as GISS ModelE2.1.

6   Summary and conclusions

The performance of the NASA GISS ModelE2.1 Earth System Model (GISSE2.1-G-CC) in simulating a fully coupled carbon cycle is presented here for the first time, following the experimental protocol for C4MIP. The GISS ModelE2.1 carbon cycle incorporated land and ocean components coupled interactively to the atmosphere and ice components. In this work, we assessed the CMIP6 C4MIP historical simulations, in both concentration-driven and emissions-driven modes, against observation-based data; described three variations of the concentration-driven 1%/yr increasing atmospheric CO_2 concentration simulations that differed in the degree of coupling of the carbon cycle to radiation (1pctCO2, 1pctCO2-bgc,

1pctCO2-rad); and described the SSP5-8.5 fully coupled emissions-driven and biogeochemically coupled concentration-driven scenario experiments.

The historical simulations demonstrated that, within the uncertainties, our model was largely able to reproduce the observed atmospheric CO_2 concentration and land and ocean fluxes. The identified departures from observed atmospheric CO_2 concentration were tied to differences seen in the fluxes. The behavior of land carbon uptake had a stronger impact than the ocean uptake in influencing the agreement of modeled versus observed atmospheric CO_2 concentration. In the first part of the simulation through the mid 1900's, erroneously represented land use change as well as fixed 2004 LAI (higher than during earlier in the century) led to higher land carbon uptake and hence lower modeled atmospheric CO_2 concentration than the observed. In the 2000's, the major discrepancy was attributed to low land uptake of carbon in the northern high latitudes and in the Amazon in the summer, due to GISS ModelE2.1 biases in simulated climate. CO_2 fertilization as well as warming-enhanced photosynthetic capacity overall lead toward convergence to observed atmospheric CO_2 by the end of the historical period, and a trend of land as a continued land sink at the end of the historical period. Though the land played a larger role in influencing the variability of carbon uptake, the ocean uptake mostly dominated total carbon uptake, particularly after mid 1900's.

The biogeochemically coupled historical simulations, where radiation is not coupled to the climate and thus has no climate change due to radiative forcing, showed that more carbon is taken up than in the fully coupled historical simulations. The ocean uptake in emission-driven simulation, however, experienced an equal amount of carbon uptake as the fully coupled case. The land is more sensitive to climate change, leading to a difference of about 40 GtC more accumulated at the end of the simulation whereas the difference for the ocean was at most 10 GtC. There were hardly any differences observed in the spatial distribution of fluxes compared to those of fully coupled cases. Our experimental framework with fixed observed 2004 LAI serves as a type of control experiment, as it enabled straightforward interpretation of why simulated fluxes and atmospheric CO_2 concentration depart from observations without being subject to the complexity of, for example, poorly simulated phenology or uncertain simulated distributions of natural vegetation types. This enabled us to separate the roles of land cover change and LAI, to identify deficiencies in our model implementation of land cover change, and to discern an insufficient CO_2 fertilization effect that can be improved with future model simulations that include prognostic growth and globally changing LAI. We are not yet able to quantify where lack of fire could be a significant aspect of our model's departure from observed atmospheric CO_2 concentration.

The SSP5-8.5 scenario experiments indicated that the land and ocean uptake increases rapidly at the beginning of the simulations due to the high rate of atmospheric CO_2 concentration increase, however, the uptake rate decreased over time. By the end of the simulation in the fully coupled esm-ssp585 experiment, the proportion of CO_2 remaining in the atmosphere was larger than that at the end of the esm-hist experiment where it accounted for roughly 3 times the combined land (excluding land use change) and ocean uptake. Total ocean uptake at the end of the esm-ssp585 experiment was nearly 3 times the total land uptake (including land use change). Most regions of the globe in the esm-ssp585 experiment showed an increase in stored carbon, however, notable regions of net negative change were identified (e.g., North Atlantic Ocean, northern Russia, northern North America). The GISS ModelE2.1 results for the esm-ssp585 experiment fall squarely within the range of trends simulated by the CMIP5 models that do not include a terrestrial nitrogen cycle and that include land cover change (Friedlingstein et al, 2014).

According to Köppen-Geiger climate classifications, the predicted climate at the end of the esm-ssp585 experiment reproduces climate changes like those predicted by CMIP5 models (Beck et al. 2018), notably the northward migration of temperate deciduous forest climate to a temperate/boreal boundary at ~60 degrees °N, spreading along this latitude across Eurasia. The GISS ModelE2.1 esm-ssp595 projected climate also reproduces the several changes in climate types in continental Europe to drier and warmer types. Major climate type changes occur largely for temperate to subarctic climates. That the boreal zone and subarctic specifically show net carbon losses in the esm-ssp585 scenario whereas the rest of the world increases as a net sink is due to greater soil respiration from a shorter frozen season and warmer and more widespread thawed period for soils.

The 1%/yr increasing atmospheric CO_2 concentration simulations demonstrated the behavior of carbon uptake by the different components of our model in an idealized scenario where CO_2 is the only forcing. The biogeochemically coupled simulation experienced the greatest positive change in storage, and the radiatively coupled simulation experienced a net loss of carbon from the land and ocean reservoirs. As with the historical case, the land was more sensitive to climate change than the ocean. Tropical land biosphere and northern Atlantic Ocean were among the key regions identified to have had the largest change in stored carbon.

Carbon-concentration (β) and carbon-climate (γ) feedback parameters were computed from the three variations of 1%/yr CO_2 simulations. The ocean had larger magnitude of β parameter while the land had larger magnitude of γ parameter. The total carbon-concentration feedback dominated the total carbon-climate feedback, however, the magnitude of the rate of increase decreased for the carbon-concentration feedback while that for the carbon-climate feedback did not. The carbon-concentration feedback was almost completely positive throughout the globe, however, the carbon-climate feedback contained some regions with positive values while the majority of the globe was negatively valued. Compared to CMIP5 model ensemble, the feedback parameters from our model were consistent to within one standard deviation except the ocean γ parameter which was larger in magnitude.

Future improvements in how GISS ModelE simulates the carbon cycle will certainly be seen with prescription of more accurate land cover change and land use state transitions, as well as prognostic LAI and growth that can capture global trends like the greening in the 2000's and fire dynamics. The land model's geographically height-varying vegetation structure, biophysics and soil biogeochemistry otherwise perform well in providing terrestrial biomass and soil carbon stocks and fluxes, but require better boundary conditions in historical land cover and LAI, so implementation of the Hurtt et al. (2011) land use state-transitions is planned. Prognostic phenology is a weakness of land models, and may introduce other errors, as seen in most C4MIP models (Friedlingstein et al. 2006), but sensitivities of interannual variability and growth trends with climate would be possible to explore. Whether the addition of nitrogen and disturbance dynamics can lend significant improvement compared to the role of land use change remains to be seen, but these additional dynamics are known to reduce model sensitivity to CO_2 fertilization and climate forcings (Friedlingstein et al. 2014). Meanwhile, the northern high latitude cold bias of the climate model emerges as significant in impacting the simulated carbon cycle on land, but it has yet to be explained. GISS ModelE2.1 also simulates arid or semi-arid regions in the western U.S. and central Eurasia as slightly more humid climate types for both present and future climate. Uncovering the causes of these model climate biases will be important to improving predictions of the carbon cycle, as the Köppen-Geiger classifications show they are significantly different enough to impact the type of biome that can be supported by the simulated climate.

Future development in the ModelE ocean carbon cycle will seek to address current limitations that may give rise to biases in ocean biogeochemical properties. First, alkalinity sources and sinks are based on integrated primary production (alkalinity sink) and an assumed calcium carbonate dissolution rate that decreases exponentially with depth (alkalinity source) and does not depend on the calcium carbonate saturation state of seawater. In addition, riverine inflow does not contribute to alkalinity. Second, while the model includes N_2 fixation, denitrification is not explicitly represented in the model. Rather, nitrate is lost in each grid cell in proportion to the concentration of nitrate in the grid cell and the vertical integral of N_2 fixation. While this treatment forces the nitrogen budget to be closed, it can also result in unrealistic spatial variability in nitrate concentration. Lerner et al. (2020) found that in the GISS ModelE2.1 historical simulations, nitrate is generally underestimated in the surface ocean compared to observations, which may partially be a consequence of our treatment of denitrification, though it may also stem from a lack of iron limitation in observed high nutrient low chlorophyll areas such as the Southern Ocean. Therefore, future model development will focus on the inclusion of a prognostic PIC tracer, which will improve ocean model's representation of alkalinity by allowing for a relaxation of the current strong dependence of alkalinity on net primary production, and by having a dissolution rate that depends on the saturation state. It will also focus on the inclusion of the explicit representation of denitrification, nitrification, and remineralization that depend on oxygen concentration. This will allow for an improved representation of the ocean's carbon and nitrogen cycles, which should improve the model's distribution of biomass, productivity, DIC, and nutrients.

Besides the enhancements in the dynamical levels and details of the land and ocean models described above, additional coupled carbon cycle processes to be developed in GISS ModelE2.1 include natural fire dynamics and biomass burning, and the effect of ash from land on ocean nutrient inputs and on aerosol and dust feedbacks to radiation. Such additions will couple both disturbance and nutrient dynamics, as well as add further coupling between land, ocean, and atmospheric chemistry.

Following this study's introduction of the NASA GISS ESM ModelE as a contributor to C4MIP, the GISS carbon cycle group will continue to conduct further analyses of these and additional experiments in more coupled model intercomparison projects. Since all models have structural uncertainty and also currently no modeling groups are able to cover the full complexity of the Earth system's carbon cycle, the unique aspects of the NASA GISS ESM's implementation, and ongoing developments, should further contribute to understanding of carbon cycle-climate feedbacks and impacts of different human greenhouse gas emissions trajectories on dynamics of the system.

Appendix A

Atmospheric CO₂ mass/concentration ratio

The atmospheric CO_2 mass to concentration conversion factor, m, used in Section 4.3.1 is derived as follows: the product of mass of atmosphere (5.1×10^{18} kg) and the quotient of molecular weights of carbon (12.01 g/mol) and air (28.93 g/mol) gives $5.1 \times 10^{18} \times (12.01/28.93) \cong 2.117 \times 10^{18}$ kg. Volume mixing ratio of CO_2 in parts per million (ppm) is 1×10^{-6}/ppm which then gives $m = 2.117 \times 10^{18} \times 1 \times 10^{-6} = 2.117 \times 10^{12}$ kgC/ppm, or approximately 2.12 GtC/ppm.

Appendix B

The tables in this section summarize the C4MIP data request variables for tier-1 and a subset of tier-2, as described in Jones et al. (2016) for 9 experiments discussed in this study and the preindustrial control experiment. Updates since Jones et al. (2016) include the carbon flux to the atmosphere from land use change.

Appendix C

Köppen-Geiger classes

References

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., & Wu, T. (2013). 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

Beck, H.E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*, 180214. https://doi.org/10.1038/sdata.2018.214

Ball, T. & Berry, J. (1985). A Simple Empirical Model of Stomatal Control. *Plant Physiology*, 77 (n. Supplement 4). 91.

Boer, G. J. & Yu, B. (2003). Climate sensitivity and response. *Climate Dynamics*, *20*, 415–429.

Boer, G. J. & Arora, V. K. (2009). Temperature and concentration feedbacks in the carbon cycle. *Geophysical Research Letters*, *36*, L02704. https://doi.org/10.1029/2008GL036220

Boer, G. J. & Arora, V. K. (2010). Geographic aspects of temperature and concentration feedbacks in the carbon cycle. *Journal of Climate*, *23*, 775–784. https://doi.org/10.1175/2009JCLI3161.1

Boer, G. J. & Arora, V. K. (2013). Feedbacks in emission-driven and concentration-driven global carbon budgets. *Journal of Climate*, *26*, 3326–3341. https://doi.org/10.1175/JCLI-D-12-00365.1

Bosilovich, M. G., Lucchesi, R., & Suarez, M. (2015). *MERRA-2: File Specification* [GMAO Office Note No. (Version 1.0)]. http://gmao.gsfc.nasa.gov/pubs/office notes

Cannon A. J. (2012). Köppen versus the computer: comparing Köppen-Geiger and multivariate regression tree climate classifications in terms of climate homogeneity. *Hydrology and Earth System Sciences*, *16*, 217-229.

Guetter P. J., and Kutzbach J. E. (1990) A modified Köppen classification applied to model simulations of glacial and interglacial climates. Climatic Change, 16: 193-215.

Carlson, C. A., Bates, N. R., Hansell, D. A., and Steinberg, D., K. (2001). Carbon cycle. *Encyclopedia of Ocean Sciences* (2nd ed.), *1*, 390–400. https://doi.org/10.1016/B978-012374473-9.00272-1

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R. B., Piao, S., & Thornton, P. (2013). *Carbon and Other Biogeochemical Cycles*. In: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Collatz, G. J., Ball, J. T., Grivet, C. & Berry, J. A. (1991). Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. *Agricultural and Forest Meteorology*, *54*. 107–136.

da Cunha, L. C., Buitenhuis, E. T., Le Quéré, G. X., C., & Ludwig, W. (2007). Potential impact of changes in river nutrient supply on global ocean biogeochemistry. Global Biogeochemical Cycles, 21(4), 1-15.

Erb, K. H., Kastner, T., Plutzar, C., Bais, A. L. S., Carvalhais, N., Fetzel, T., Gingrich, S., Haberl, H., Lauk, C., Niedertscheider, M., Pongratz, J., Thurner, M. & Luyssaert, S. (2018). Unexpectedly large impact of forest management and grazing on global vegetation biomass. *Nature*, *553*(7686). 73–76.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*, 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Farquhar, G. D. & von Caemmerer, S. (1982). Modelling photosynthetic response to environmental conditions. In O. L. Lange, C. B. Osmond, H. Ziegler. Berlin, *Encyclopedia of Plant Physiology (NS)* (12B. pp. 549–587). Berlin: Springer.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., & Schaaf, C. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, *83*(1-2). 287–302.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A. & Huang, X. M. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*(1). 168–182.

Friedlingstein, P., Dufresne, J. -L., Cox, P. M., & Rayner, P. (2003). How positive is the feedback between climate change and the carbon cycle? *Tellus*, *55B*, 692–700. https://doi.org/10.1034/j.1600-0889.2003.01461.x

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., & Zeng, N. (2006). Climate–carbon cycle feedback analysis: Results from the C^{4}MIP model intercomparison. *Journal of Climate*, *19*, 3337–3353.

Friedlingstein, P., Meinshausen, M., Arora, V., K., Jones, C. D., Anav, A., Liddicoat, S., K., & Knutti, R. (2014). Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, *26*, 511–526. https://doi.org/10.1175/JCLI-D-12-00579.1

Friedlingstein, P., et al. (2019). Global carbon budget 2019. *Earth System Science Data*, *11*, 1783–1838. https://doi.org/10.5194/essd-11-1783-2019

García-Pando, C., Miller, R., Perlwitz, J. P., Rodríguez, S., Prospero, J. M. (2016). Predicting the mineral composition of dust aerosols: Insights from elemental composition measured at the Izaña Observatory. *Geophysical Research Letters*, *43*(19), 10,520–10,529. https://doi.org/10.1002/2016GL069873

GISTEMP Team (2020). GISS Surface Temperature Analysis (GISTEMP), version 4. NASA Goddard Institute for Space Studies. Dataset accessed 2020-03-31 at https://data.giss.nasa.gov/gistemp/

Gregg, W., Conkright, M. E., Ginoux, P., O'Reilly, J. E., & Casey, N. W., (2003). Ocean primary production and climate: global decadal changes. *Geophysical Research Letters*, *30*(15), 1809. https://doi.org/10.1029/2003GL016889

Gregg, W., & Casey, N. W. (2007). Modeling coccolithophores in the global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography*, *54*, 447–477. https://doi.org/10.1016/j.dsr2.2006.12.007

Gregory, J., M., Jones, C. D., Cadule, P., & Friedlingstein, P. (2009). Quantifying carbon Cycle Feedbacks. *Journal of Climate*, *22*, 5232–5250. https://doi.org/10.1175/2009JCLI2949.1

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., Monaco, C. L., Mathis, J. T., Murata, A., Olsen, A., Perez, F. F., Sabine, C. L., Tanhua, T., & Wanninkhof, R. (2019). The oceanic sink for anthropogenic CO_2 from 1994 to 2007. *Science*, *363*, 1193–1199. https://doi.org/10.1126/science.aau5153

Hansen, J. et al. (2007). Dangerous human-made interference with climate: A GISS modelE study. *Atmospheric Chemistry and Physics*, *7*, 2287–2312.

Harris, I., Jones, P. D., Osborn, T. J. & Lister, D.H. (2014). Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. International Journal of Climatology, *34*. 623–642. https://doi.org/10.1002/joc.3711

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., & Zhang, Q. (2018). Historical (1750-2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS). *Geoscientific Model Development*, *11*, 369–408. https://doi.org/10.5194/gmd-11-369-2018

Hoffman, F. M., Randerson, J. T., Arora, V. K., Bao, Q., Cadule, P., Ji, D., Jones, C. D., Kawamiya, M., Khatiwala, S., Lindsay, K., Obata, A., Shevliakova, E., Six, K., D., Tjiputra, J., F., Volodin, E., M., & Wu, T. (2014). Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models. *Journal of Geophysical Research: Biogeosciences*, *119*, 141–162. https://doi.org/10.1002/2013JG002381

Holte, J., Talley, L. D., Gilson, J., & Roemmich, D. (2017). An Argo mixed layer climatology and database. *Geophysical Research Letters*, *44* (11), 5618–5626.

Hougton, R. A. (2003). The contemporary carbon cycle. *Treatise on Geochemistry*, *8*, 473–513. https://doi.org/10.1016/B0-08-043751-6/08168-8

Houghton, R. A., Butman, D., Bunn, A. G., Krankina, O. N., Schlesinger, P., & Stone, T. A. (2007). Mapping Russian forest biomass with data from satellites and forest inventorie., *Environmental Research Letters*, *2*, 045032. https://doi.org/10.1088/1748-9326/2/4/045032

Hurtt, G. C., Frolking, S., Fearon, M. G., Moore, B., Shevliakova, E., Malyshev, S., Pacala, S., W., & Houghton, R. A. (2006). The underpinnings of land-use history: three

centuries of global gridded land-use transitions, wood harvest activity, and resulting secondary lands. *Global Change Biology*, *12*, 1208–1229.

Hurtt, G. C. et al. (2011). Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic Change*, *109*, 117. https://doi.org/10.1007/s10584-011-0153-2

IPCC (2013). Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, NY, USA.

Jones, C. D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., Graven, H., Hoffman, F., Ilyina, T., John, J. G., Jung, M., Kawamiya, M., Koven, C., Pongratz, J., Raddatz, T., Randerson, J. T., & Zaehle, S. (2016). C4MIP – The Coupled Climate–Carbon Cycle Model Intercomparison Project: experimental protocol for CMIP6. *Geoscientific Model Development*, *9*, 2853–2880. https://doi.org/10.5194/gmd-9-2853-2016

Jones, C. D., Frölicher, T. L., Koven, C., Bopp, L., MacDougall, A. H., Matthews, H. D.., Zickfeld, K., Rogelj, J., Tokarska, K. B., Gillett, N., Ilyina, T., Meinshausen, M., Mengis, N., Séférian, R., & Eby, M. (2019). The Zero Emission Commitment Model Intercomparison Project (ZECMIP) contribution to CMIP6: Quantifying committed climate changes following zero carbon emissions. *Geoscientific Model Development*, *12*, 4375–4385. https://doi.org/10.5194/gmd-12-4375-2019

Kaushik, A., Graham, J., Dorheim, K., Kramer, R., Wang, J., & Byrne, B. (2020), The future of the carbon cycle in a changing climate, *Eos, 101*. https://doi.org/10.1029/2020EO140276

Kelley, M. et al. (2020). GISS-E2.1: Configurations and Climatology. *Journal of Advances in Modeling Earth System Models*, Accepted online 10 July 2020: e2019MS002025, DOI 10.1029/2019MS002025

Kiang, N. Y., Koster, R. D., Moorcroft, P. R., Ni-Meister, W. & Rind, D. (2006). Ent: A Dynamic Global Terrestrial Ecosystem Model for Coupling with GCMs. *American Geophysical Union Fall Meeting*. San Francisco, CA.

Kim, D., Lee, M., & Seo, E. (2019). Improvement of Soil Respiration Parameterization in a Dynamic Global Vegetation Model and Its Impact on the Simulation of Terrestrial Carbon Fluxes. *Journal of Climate*. https://doi.org/10.1175/JCLI-D-18-0018.1

Kottek, M., Grieser, J., Beck, C., Rudolf, R. & Rubel, F. (2006). World map of the Koppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, *15*(3). 259–263.

Koutroulis, A. G., Grillakis, M. G., Tsanis, I. K., & Papadimitriou, L. (2016). Evaluation of precipitation and temperature simulation performance of the CMIP3 and CMIP5 historical experiments. *Climate Dynamics*, *47*(5-6). 1881–1898.

Köppen, W. P. (1884). Die Wärmezonen der Erde, nach der Dauer der heissen, gemässigten und kalten Zeit, und nach der Wirkung der Wärme auf die organische Welt betrachtet. Meteorologische Zeitschrift.

Kriegler, E., Bauer, N., Popp, A., Hupenöder, F., Leimbach, M., Strefler, J., Baumstark, L., Leon Bodirsky, B., Hilaire, J., Klein, D., Mouratiadou, I., Weindl, I., Bertram, C., Dietrich, J., Luderer, G., Pehl, M., Pietzcker, R., Piontek, F., Lotze-Campen, H.,

Biewald, A., Bonsch, M., Giannousakis, A., Kreidenweis, R., Müller, C., Rolinski, S., Schultes, A., Schwanitz, J., Stevanovic, M., Calvin, K., Emmerling, J., Fujimori, S., Edenhofer, O. (2017). Fossil-fueled development (SSP5): An energy and resource intensive scenario for the 21st century. *Global Environmental Change*, *42*, 297–315. https://dx.doi.org/10.1016/j.gloenvcha.2016.05.015

Latto, R., & Romanou, A. (2018). The "Ocean Carbon States" database: A proof-of-concept application of cluster analysis in the ocean carbon cycle. *Earth Systems Science Data*, 10, 609–626. https://doi.org/10.5194/essd-10-609-2018

Lauvset, S., 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., & Watelet, S. (2016). 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

Lawrence, D. M., Hurtt, G. C., Arneth, A., Brovkin, V., Calvin, K. V., Jones, A. D., Jones, C. D., Lawrence, P. J., de Noblet-Ducoudré, N., Pongratz, J., Seneviratne, S. I., & Shevliakova, E. (2016). The Land Use Model Intercomparison Project (LUMIP) contribution to CMIP6: rationale and experimental design. *Geoscientific Model Development*, *9*, 2973–2998. https://doi.org/10.5194/gmd-9-2973-2016

Lenssen, N., Schmidt, G., Hansen, J., Menne, M., Persin, A., Ruedy, R., and Zyss, D. (2019). Improvements in the GISTEMP uncertainty model. *Journal of Geophysical Research: Atmospheres*, *124*(12), 6307–6326. https://doi.org/10.1029/2018JD029522

Le Quéré et al. (2018). Global Carbon Budget 2017. *Earth Systems Science Data*, 10, 405–448.

Lerner, P. Romanou, A., Kelley, M., Romanski, J., Ruedy, R., & Russel, G. (2020). Drivers of air-sea CO_2 flux seasonality and its long-term changes in the NASA-GISS model CMIP6 submission. *Journal of Advances in Modeling Earth Systems*, submitted.

Marshall, J., Scott, J., Romanou, A., Kelley, M., & Leboissetier, A. (2017). The dependence of the ocean's MOC on mesoscale eddy diffusivities: A model study. *Ocean Modelling*, *111*, 1– 8. https://doi.org/10.1016/j.ocemod.2017.01.001

Matthews, E. (1983). Global vegetation and land use: New high-resolution data bases for climate studies. *Journal of Climate Applied Meteorology*, *22*, 474–487

Meehl, G. A., Covey, C., McAvaney, B., Latif, M., & Ronald, J. (2005). Overview of the Coupled Model Intercomparison Project. *Bulletin of the American Meteorological Society*, *86* (1), 89–93. 10.1175/BAMS-86-1-89

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P., B., Beyerle, R., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Melders, G. J. M., Vollmer, M. K., Wang, R. H. J., Weiss, R. (2017). Historical greenhouse gas concentrations for climate modelling (CMIP6). *Geoscientific Model Development*, *10*, 2057–2116. https://doi.org/10.5194/gmd-10-2057-2017

Miller, R. L., Cakmur, R. V., Perlwitz, J. P., Geogdzhayev, I. V., Ginoux, P., Kohfeld, K. E., Koch, D., Prigent, C., Ruedy, R., Schmidt, G. A., & Tegen, I. (2006). Mineral dust aerosols in the NASA Goddard Institute for Space Sciences ModelE atmospheric

general circulation model. *Journal of Geophysical Research*, *111*, D06208. https://doi.org/10.1029/2005JD005796

O'ishi, R., Abe-Ouchi, A., Prentice, I. C., & Sitch, S. (2009). Vegetation dynamics and plant CO_2 responses as positive feedbacks in a greenhouse world. Geophysical Research Letters, 36, L11706. https://doi.org/10.1029/2009GL038217

Olsen, A., Key, R. M., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Pérez, F. F. & Suzuki, T. (2016). The Global Ocean Data Analysis Project version 2 (GLODAPv2) – an internally consistent data product for the world ocean. *Earth System Science Data*, *8*, 297–323. https://doi.org10.5194/essd-8-297-2016

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., & Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*, 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

Orr, J. C., Najjar, R. G., Aumount, O., Bopp, L., Bullister, J. L., Danabasoglu, G., Doney, S. C., Dunne, J. P., Dutay, J.-C., Graven, H., Griffies, S. M., John, J. G., Joos, F., Levin, I., Lindsay, K., Matear, R. J., McKinley, G. A., Mouchet, A., Oschlies, A., Romanou, A., Schlitzer, R., Tagliabue, A., Tanhua, T., & Yool, A. (2017). Biogeochemical protocols and diagnostics for the CMIP6 Ocean Model Intercomparison Project (OMIP). *Geoscientific Model Development, 10,* 2169-2199. https://doi.org/10.5194/gmd-10-2169-2017

Peel, M. C., Finlayso, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, *11*. 1633–1644.

Perlwitz, J. P., Pérez García-Pando, C., & Miller, R. L. (2015a). Predicting the mineral composition of dust aerosols – Part 1: Representing key processes. Atmospheric Chemistry and Physics. 15, 11593–11627. https://doi.org/10.5194/acp-15-11593-2015

Perlwitz, J. P., Pérez García-Pando, C., & Miller, R. L. (2015b). Predicting the mineral composition of dust aerosols – Part 2: Model evaluation and identification of key processes with observations. *Atmospheric Chemistry and Physics*, *15*, 11629–11652. https://doi.org/10.5194/acp-15-11629-2015

Piao, S., et al. (2013). Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO_2 trends, *Global Change Biology*, *19*, 2117–2132. https://doi.org10.1111/gcb.12187

Plattner, G. -K., Knutti, R., Joos, F., Stocker, T. F., von Bloh, W., Brovkin, V., Cameron, D., Driesschaert, E., Dutkiewicz, S., Eby, M., Edwards, N. R., Fichefet, T., Hargreaves, J. C., Jones, C. D., Loutre, F., Matthews, H. D., Mouchet, A., Müller, S. A., Nawrath, S., Price, A., Solkolov, A., Strassmann, K. M., & Weaver, A. J. (2008). Long-term climate commitments projected with climate–carbon cycle models. *Journal of Climate*, *21*, 2721–2751. https://doi.org/10.1175/2007JCLI1905.1

Pongratz, J., Reick, C., Raddatz, T., & Claussen, M. (2008). A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles*, *22*, GB3018. https://doi.org/10.1029/2007GB003153

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., & Klooster, S. A. (1993). Terrestrial ecosystem production: a process model based on global satellite and surface data. *Global Biogeochemical Cycles*, *7*(4). 811–841.

Poulter, B., Ciais, P., Hodson, E., Lischke, H., Maignan, F., Plummer, S., & Zimmermann, N. E. (2011). Plant functional type mapping for earth system models. *Geoscientific Model Development*, *4*. 993–1010.

Prather, M. J. (1986), Numerical advection by conservation of second-order moments. *Journal of Geophysical Research*, *91*, 6671– 6681. https://doi.org/10.1029/JD091iD06p06671

Ramankutty, N. & Foley, J. A. (1999). Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles*, *13*(4), 997-1027. https://doi.org/10.1029/1999GB900046

Randerson, J. T., Hoffman, F. M., Thornton, P. E., Mahowald, N. M., Lindsay, K., Lee, Y., Nevison, C. D., Doney, S. C., Bonan, G., Stöckli, R., Covey, C., Running, S. W., & Fung, I. Y. (2009). Systematic assessment of terrestrial biogeochemistry in coupled climate–carbon models. Global Change Biology, 15(10), 2462–2484. https://doi.org/10.1111/j.1365-2486.2009.01912.x

Roemmich, D., & Gilson, J. (2009). The 2004–2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program. *Progress in Oceanography*, *82*(2), 81–100. https://doi.org/10.1016/j.pocean.2009.03.004

Romanou, A., Gregg, W. W., Romanski, J., Kelley, M., Bleck, R., Healy, R., Nazarenko, L., Russell, G., Schmidt, G. A., Sun, S., & Tausnev, N. (2013). Natural air-sea flux of CO_2 in simulations of the NASA-GISS climate model: Sensitivity to the physical ocean model formulation. *Ocean Modelling*, *66*, 26–44. https://doi.org/10.1016/j.ocemod.2013.01.008

Romanou, A., Romanski, J., & Gregg, W. W. (2014). Natural ocean carbon cycle sensitivity to parameterizations of the recycling in a climate model. *Biogeosciences*, *11*, 1137– 1154, https://doi.org/10.5194/bg-11-1137-2014

Romanou, A., Marshall, J., Kelley, & M., Scott, J. (2017). Role of the ocean's AMOC in setting the uptake efficiency of transient tracers. *Geophysical Research Letters*, *44*(11), 5590–5598. https://doi.org/10.1002/2017GL072972

Rosenzweig, C., & Abramopoulos, F. (1997). Land surface model development for the GISS GCM. *Journal of Climate*, *10*, 2040–2054.

Roy, T., Bopp, L., Gehlen, M., Schneider, B., Cadule, P., Frolicher, T. L., Segschneider, J., Tjiputra, J., Heinze, C., & Joos, F. (2011). Regional impacts of climate change and atmospheric CO_2 on future ocean carbon uptake: A multimodel linear feedback analysis. *Journal of Climate*, *24*, 2300–2318. https://doi.org/10.1175/2010JCLI3787.1

Russell, G., Miller, J., & Rind, D. (1995). A coupled atmosphere-ocean model for transient climate change studies. *Atmosphere-Ocean*, *33*, 683–730. https://doi.org/10.1080/07055900.1995.9649550

Scharlemann, J. P. W., E. V. J. Tanner, R. Hiederer & V. Kapos (2014). Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Management*, *5*(1), 81–91.

Schmidt, G. A., et al. (2006). Present day atmospheric simulations using GISS ModelE: Comparison to in-situ, satellite and reanalysis data. *Journal of Climate*, *19*, 153–192. https://doi.org10.1175/JCLI3612.1

Schmidt, G. A., et al. (2014). Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *Journal of Advances in Modeling Earth Systems*, *6*, 141–184. https://doi.org/10.1002/2013MS000265

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B. & Ziese, M. (2011). GPCC Full Data Reanalysis Version 6.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. Global Precipitation Climatology Centre, National Oceanic and Atmospheric Administration.

Schwinger, J., Tjputra, J., F., Heinze, C., Bopp, L., Christian, J., R., Gehlen, M., Ilyina, T., Jones, C. D., Salas-Mélia, D., Segschneider, J., Séférian, R., & Totterdell, I. (2014). 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

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research: Biogeosciences*, *116*, G04021. https://doi.org/10.1029/2011JG001708

Sitch, S., et al. (2008), Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using 5 DynamicGlobal Vegetation Models (DGVMs). *Global Change Biology*, *14*, 1–25.

Spitters, C. J. T. (1986). Separating the diffuse and direct component of global radiation and its implications for modeling canopy photosynthesis. Part II. Calculation of canopy photosynthesis. *Agricultural and Forest Meteorology*, 38. 231–242.

Tian, Y., Wang, Y., Zhang, Y., Knyazikhin, Y., Bogaert, J., & Myneni, R. B. (2002a). Radiative transfer based scaling of LAI retrievals from reflectance data of different resolutions, *Remote Sensing of the Environment*, *84*, 143–159.

Tian, Y., Woodcock, C. E., Wang, Y., Privette, J. L., Shabanov, N. V., Zhou, L., Zhang, Y., Buermann, W., Dong, J., Veikkanen, B.,Häme, T., Andersson, K., Ozdogan, M., Knyazikhin, Y., Myneni, R. B. (2002b). Multiscale analysis and validation of the MODIS LAI productII. Sampling strategy. *Remote Sensing of the Environment*, *83*, 431–441.

Todd-Brown, K. E., Randerson, J. T., Hopkins, F., Arora, V., Hajima, T., Jones, C., Shevliakova, E., Tjiputra, J., Volodin, E., Wu, T., Zhang, Q., & Allison, S. D. (2014). Changes in soil organic carbon storage predicted by Earth system models during the 21st century. *Biogeosciences*, *11*, 2341–2356. https://doi.org/10.5194/bg-11-2341-2014

Trewartha, G. T. (1968). *An introduction to climate*. McGraw-Hill, New York.

Trewartha, G. T., & Horn, L. H. (1980). *An introduction to climate*. McGraw Hill, New York, NY.

Welp, L. R., et al. (2011). Interannual variability in the oxygen isotopes of atmospheric CO_2 driven by El Niño. *Nature*, *477*(7366), 579582. https://doi.org.10.1038/nature10421

Wenzel, S., Cox, P. M., Eyring, V., & Friedlingstein, P. (2014). 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

Westberry, T., Behrenfeld, M. J., Siegel, D. A., & Boss, E. (2008). Carbon-based primary productivity modeling with vertically resolved photoacclimation. *Global Biogeochemical Cycles*, *22*(2), 1-18. https://doi.org/10.1029/2007GB003078

Yin, L., Fu, R., Shevliakova, E., & Dickinson, R. E. (2013). How well can CMIP5 simulate precipitation and its controlling processes over tropical South America? *Climate Dynamics*, *41*(11), 3127–3143.

Yuan, W. et al. (2019). Increased atmospheric vapor pressure deficit reduces global vegetation growth. *Science Advances*, *5*(8), eaax1396. https://DOI.org/10.1126/sciadv.aax1396

Zeebe, R. E. (2012). History of seawater carbonate chemistry, atmospheric CO2, and ocean acidification. *Annual Review of Earth and Planetary Sciences*, *40*, 141–165

Zhu, Z. et al. (2016). Greening of the Earth and its drivers. *Nature Climate Change*, *6*, 791–795. https://doi.org/10.1038/nclimate3004

Zickfeld, K., Eby, M., Matthews, H. D., Schmittner A., & Weaver, A. J. (2011). Nonlinearity of carbon cycle feedbacks. *Journal of Climate*, *24*, 4255–4275. https://doi.org/10.1175/2010JCLI3898.1

Zscheischler, J., Mahecha, M. D., & Harmeling, S. (2012). Climate classifications: the value of unsupervised clustering. *Proceedings of the International Conference on Computational Science*, Iccs 2012, 9. 897–906.

Table 1. Summary of all experiments conducted by GISS ModelE2.1 for C4MIP of CMIP6. Experiments used in this study are indicated in bold letters. Dashed lines in the category column indicate experiments not defined in the protocols, but still conducted by our group. Biogeochemically coupled experiments have land and ocean carbon cycle components that experience changing levels of CO_2, but there is no radiative warming due to emissions. Similarly, in a radiatively coupled experiment, the climate is changing due to radiative effects of increased emissions, but the land and ocean components experience constant levels of CO_2 in the atmosphere similar to the preindustrial era. Direct climate change due to rising surface air temperatures is only present in the full and the radiatively coupled experiments.

Category	Experiment ID	Driving Type	Period	Coupling
DECK	piControl	Concentration	1850	Full*
	esm-piControl	Emission	1850–2100	Full
	1pctCO2	**Concentration**	**1850–2000**	**Full***
Historical	**historical**	**Concentration**	**1850–2015**	**Full***
	esm-hist	**Emission**	**1850–2015**	**Full**
Tier-1	**1pctCO2-bgc**	**Concentration**	**1850–2000**	**Biogeochemical**
	esm-ssp585	**Emission**	**2015–2100**	**Full**
Tier-2	**1pctCO2-rad**	**Concentration**	**1850–2000**	**Radiative**
	hist-bgc	**Concentration**	**1850–2015**	**Biogeochemical**
	ssp585-bgc	**Concentration**	**2015–2100**	**Biogeochemical**
----------	**esm-hist-bgc**	**Emission**	**1850–2015**	**Biogeochemical**
----------	hist-rad	Concentration	1850–2015	Radiative

*Concentration driven runs are not "fully coupled" strictly speaking, since the land/ocean carbon sinks do not feed back to climate.

Table B1. Summary of tier-1 land carbon budgets for the GISS ModelE2.1 C4MIP experiments. A subset of tier-2 carbon pools for land are also listed.

	piControl*	historical Year 2014	hist-bgc Year 2014	1pctCO2 Year 140	1pctCO2-bgc Year 140	1pctCO2-rad Year 140	esm-hist Year 2014	esm-hist-bgc Year 140	esm-ssp585 Year 2100	ssp585-bgc Year 2100
Fluxes (GtC/year)	GtC/year	GtC/year	GtC/year	GtC/year	GtC/year	GtC/year	GtC/year	GtC/year	GtC/year	GtC/year
Gross primary production (GPP)	99.7	116.7	114.5	202.5	192.1	102.3	117.6	114.2	198.9	185.7
Autotrophic respiration (RA)	65.9	72.9	69.4	118.4	105.2	72.5	72.9	69.9	113.5	100.3
Net primary productivity (NPP)	33.8	43.8	45.1	84.0	86.9	29.8	44.7	44.3	85.4	85.4
Heterotrophic respiration (RH)	33.8	44.1	42.7	79.9	80.1	31.8	44.3	43.1	81.3	79.0
Fire natural emission to atmosphere (fFireNet)	Not simulated	Not simulated	Not simulated	Not simulated	Not simulated	Not simulated	Not simulated	Not simulated	Not simulated	Not simulated
Net flux from land use change to atmosphere (cProduct)	NA	$< 10^{-6}$	$< 10^{-6}$	NA	NA	NA	$< 10^{-6}$	$< 10^{-6}$	0.4	0.4
Net flux to atmosphere (netAtmosLandCO2flux)	0.0	0.3	-2.4	- 4.1	-6.8	1.9	-0.4	-1.2	-3.8	-6.0
Stocks (GtC)	GtC	GtC	GtC	GtC	GtC	GtC	GtC	GtC	GtC	GtC
Organic										
cVeg total	571.1	562.7	563.0	585.8	591	567.3	562.7	563.2	572.6	572.0
cLeaf	7.6	6.8	6.8	7.6	7.6	7.6	6.8	6.8	6.8	6.8
cStem	341.9	310.3	309.9	341.3	341.6	342.3	310.6	310.5	295.0	295.12
cRoot	25.6	23.3	23.3	25.5	25.6	25.6	23.3	23.3	23.0	23.1
cOther	196.0	222.3	223.0	219.0	216.2	191.8	222.0	222.6	250.1	249.4
cSoil	1550.3	1566.7	1596.6	1894.3	2113.7	1362.3	1566.7	1591.8	1713.9	1837.3

Inorganic	Not simulated	Not simulated	Not simulated	Not simulated	Not simulated	Not simulated	Not simulated	Not simulated	Not simulated	Not simulated

*The piControl value is the average of the last 30 years of the spin-up, as initial conditions for other experiments.

Table B2. Summary of tier-1 ocean carbon budgets for the GISS ModelE2.1 C4MIP experiments.

	piControl*	historical Year 2014	hist-bgc Year 2014	1pctCO2 Year 140	1pctCO2-bgc Year 140	1pctCO2-rad Year 140	esm-hist Year 2014	esm-hist-bgc Year 2014	esm-ssp585 Year 2100	ssp585-bgc Year 2100
Fluxes (GtC/year)	GtC/year	GtC/year	GtC/year	GtC/year	GtC/year	GtC/year	GtC/year	GtC/year	GtC/year	GtC/year
Riverflux of carbon	0.64	0.67	0.63	0.71	0.68	0.66	0.69	0.64	0.74	0.70
Air-to-sea flux of carbon	0.10	1.96	2.64	5.02	6.55	-1.18	2.54	2.50	5.17	7.20
Grazing	12.95	12.52	14.49	11.43	15.53	11.18	12.44	12.84	9.28	12.07
Phytoplankton growth	21.63	21.07	23.42	19.24	24.76	18.96	20.88	21.47	16.46	20.27
Phytoplankton + zooplankton respiration	2.95	2.92	3.21	2.76	3.37	2.73	2.89	2.95	2.37	2.86
Remin. of organic carbon	18.70	18.53	20.11	16.49	21.13	16.35	18.39	19.05	14.23	16.86
Ocean net primary production	19.5	19.0	21.1	17.3	21.8	17.1	18.8	19.3	14.82	18.24
Carbon export	3.82	3.76	4.12	3.25	4.26	3.23	3.77	3.91	2.83	3.41
Calcite flux	0.44	0.43	0.48	0.39	0.50	0.39	0.23	0.44	0.34	0.42
Stocks (GtC)	GtC	GtC	GtC	GtC	GtC	GtC	GtC	GtC	GtC	GtC
Phytoplankton	0.46	0.42	0.48	0.34	0.47	0.34	0.40	0.47	0.31	0.37
Zooplankton	0.11	0.10	0.13	0.08	0.12	0.08	0.10	0.12	0.07	0.09
Detritus organic carbon	0.089	0.081	0.102	0.057	0.097	0.057	0.078	0.098	0.049	0.065
Dissolved organic carbon	45.2	44.1	46.6	37.4	46.9	37.2	43.7	45.8	34.3	39.2

Dissolved inorganic carbon	35258.4	35395.0	35405.7	35860.9	35981.7	35206.1	35389.9	35385.9	35811.1	35854.7

*The piControl value is the average of the last 30 years of the spin-up, as initial conditions for other experiments.

Table B3. Summary of tier-1 ocean nitrate, ammonium, iron, silicate, and alkalinity budgets for the GISS ModelE2.1 C4MIP experiments.

	piControl*	historical Year 2014	hist-bgc Year 2014	1pctCO2 Year 140	1pctCO2-bgc Year 140	1pctCO2-rad Year 140	esm-hist Year 2014	esm-hist-bgc Year 2014	esm-ssp585 Year 2100	ssp585-bgc Year 2100
Stocks (Gt)	Gt	Gt	Gt	Gt	Gt	Gt	Gt	Gt	Gt	Gt
Nitrate	421.7	447.3	447.0	459.5	456.0	459.4	447.4	446.9	420.5	419.0
Ammonium	6.34	6.74	6.74	6.95	6.88	6.95	6.75	6.74	6.34	6.30
Iron	1.06	1.06	1.06	1.05	1.06	1.05	1.06	1.06	1.05	1.05
Silicate	3012.4	3030.6	3029.5	3046.2	3036.3	3046.2	3031.2	3029.4	3021.9	3017.9
Alkalinity	3211.00	3210.58	3210.59	3210.37	3210.44	3210.37	3210.58	3210.60	3211.01	3211.03

*The piControl value is the average of the last 30 years of the spin-up, as initial conditions for other experiments.

Table B4. Total carbon stocks

Total Carbon (GtC)	piControl*	historical Year 2014	hist-bgc Year 2014	1pctCO2 Year 140	1pctCO2-bgc Year 140	1pctCO2-rad Year 140	esm-hist Year 2014	esm-hist-bgc Year 2014	esm-ssp585 Year 2100	ssp585-bgc year 2100
Ocean	35304.3	35440.0	35453.0	36029.4	35898.9	35243.9	35434.3	35432.5	35845.8	35894.5
Land	2121.4	2129.4	2159.6	2487.9	2704.6	1919.6	2129.5	2155.0	2286.9	2409.6
Atmosphere	603.2	842.4	842.4	2428.0	2428.5	603.1	858.1	833.4	2458.9	2407.76

*The piControl value is the average of the last 30 years of the spin-up, as initial conditions for other experiments.

Table C1. Köppen-Geiger biome/climate classes, with color scheme from http://koeppen-geiger.vu-wien.ac.at/

	Class	Description
1	Af	Equatorial rainforest
2	Am	Equatorial monsoon
3	As	Equatorial savannah with dry summer
4	Aw	Equatorial savannah with dry winter
5	BWk	Arid desert cold
6	BWh	Arid desert hot
7	BSk	Arid steppe cold
8	BSh	Arid steppe hot
9	Csa	Warm temperate with dry hot summer
10	Csb	Warm temperate with dry warm summer
11	Csc	Warm temperate with dry cool summer and cold winter
12	Cwa	Warm temperate with dry winter and hot summer
13	Cwb	Warm temperate with dry winter and warm summer
14	Cwc	Warm temperate with dry cold winter and cool summer
15	Cfa	Warm temperate fully humid with hot summer
16	Cfb	Warm temperate fully humid with warm summer
17	Cfc	Warm temperate fully humid with cool summer and cold winter
18	Dsa	Snow with dry hot summer
19	Dsb	Snow with dry warm summer
20	Dsc	Snow with dry cool summer and cold winter
21	Dsd	Snow with dry summer extremely continental
22	Dwa	Snow with dry winter and hot summer
23	Dwb	Snow with dry winter and warm summer
24	Dwc	Snow with dry cold winter and cool summer
25	Dwd	Snow with dry winter extremely continental
26	Dfa	Snow fully humid with hot summer
27	Dfb	Snow fully humid with warm summer
28	Dfc	Snow fully humid with cool summer and cold winter
29	Dfd	Snow fully humid extremely continental
30	EF	Polar frost
31	ET	Polar tundra
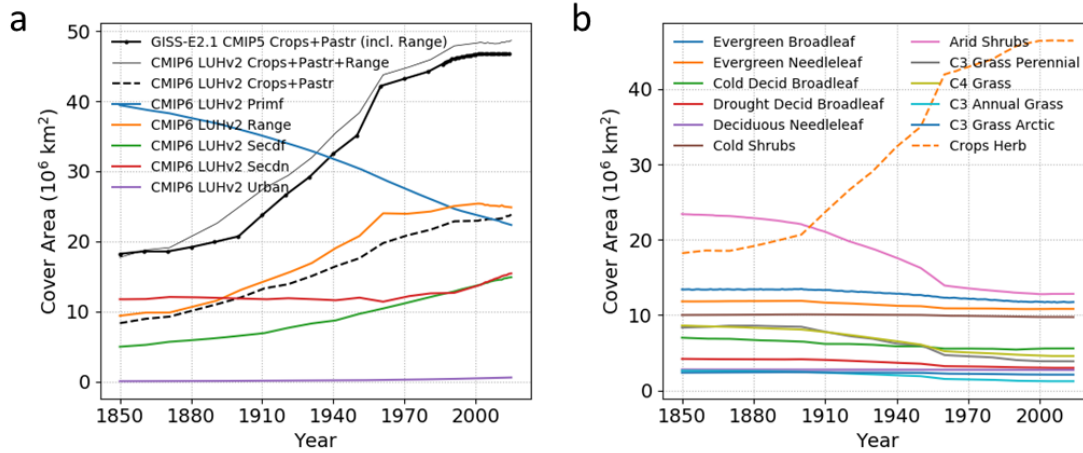
Figure 1. (a) Global land use change area prescribed in GISS ModelE2.1 for 1850-2014 (GISS-E2.1 CMIP5 crops+pastr incl. range), compared to the CMIP6 Land Use Harmonization v.2 (LUHv2) components of land use change. (b) Ent Terrestrial Biosphere Model plant functional types prescribed with CMIP5 historical cover change in GISS ModelE2.1. In the CMIP5 data set, rangeland was included in pasture, while in CMIP6 they are separated. Trend differences due to including rangeland increase after 1900, with a steep rise over 1950-1960. Legend abbreviations are crops (the sum of 5 crop types in the LUH or LUH2), pastr (pasture), range (rangeland), primf (primary forest), secdf (secondary forest), urban (urban).
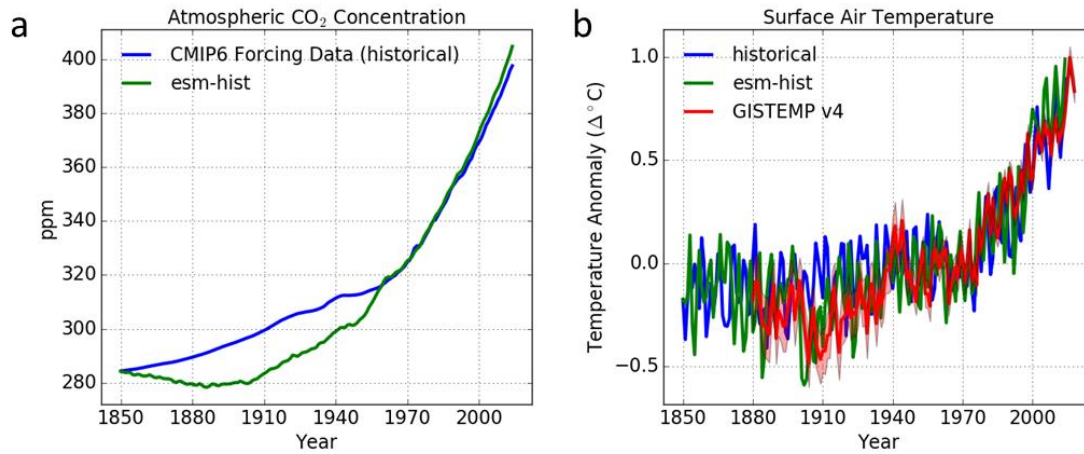
Figure 2. a) Yearly global mean atmospheric CO_2 concentration (in ppm) computed prognostically by the model in the emissions-driven experiment (esm-hist) compared to CMIP6 forcing atmospheric CO_2 data (Meinshausen et al., 2017) used in the concentration-driven historical experiment. b) Annual global mean simulated surface air temperature anomalies from the historical (hist) experiment, esm-hist experiment, and the GISTEMP v4 observation (GISTEMP Team, 2020; Lenssen et al., 2019). Anomaly is defined as the deviation from the mean of 1951-1980 period which are 13.96 °C and 13.88 °C for the historical and esm-hist experiments, respectively. The GISTEMP record is from 1880 to 2019 and is based on combined land-surface air and sea-surface water data. The red shaded areas indicate annual uncertainties at 95% confidence interval.
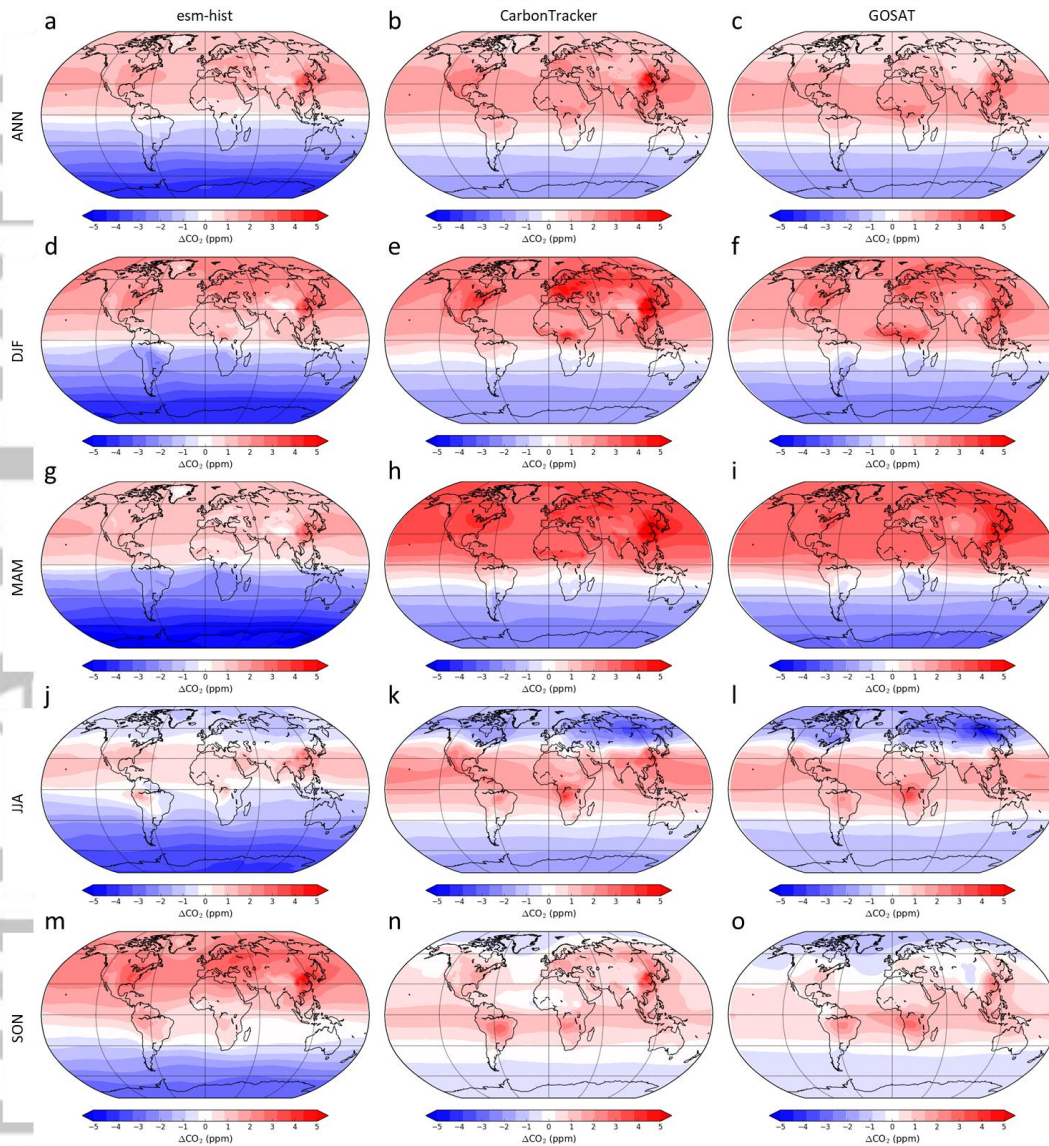
Figure 3. Distribution of 2010-2014 mean, column-averaged atmospheric CO_2 concentration differences in a) the GISS ModelE2.1 emissions-driven experiment (esm-hist) and compared to the same 5 year mean of b) CarbonTracker CT2017 atmospheric CO_2 mole fraction data (Peters et al., 2007) and c) GOSAT global atmospheric CO_2 gas concentration data (JAXA/NIES/MOE). Values shown are anomalies from the global mean for each of the distributions (400, 390, and 390 ppm for esm-hist, CarbonTracker, and GOSAT, respectively). For CarbonTracker, monthly global CO_2 mole fraction netCDF files were downloaded. For GOSAT, level L4B annual global CO_2 distribution netCDF files (each containing monthly data) were downloaded. The labels on the left indicate seasonal means as (d,e,f,) DJF: December, January, February; (g,h,i) MAM: March, April, May; (j,k,l) JJA: June, July, August; and (m,n,o) SON: September, October, November. ANN indicates annual mean. The model, CT2017, and GOSAT resolutions are 2° lat × 2.5° lon, 2° lat × 3° lon, and 2.5° lat × 2.5° lon, respectively, and they were all interpolated into a common grid 2° lat × 2.5° lon.
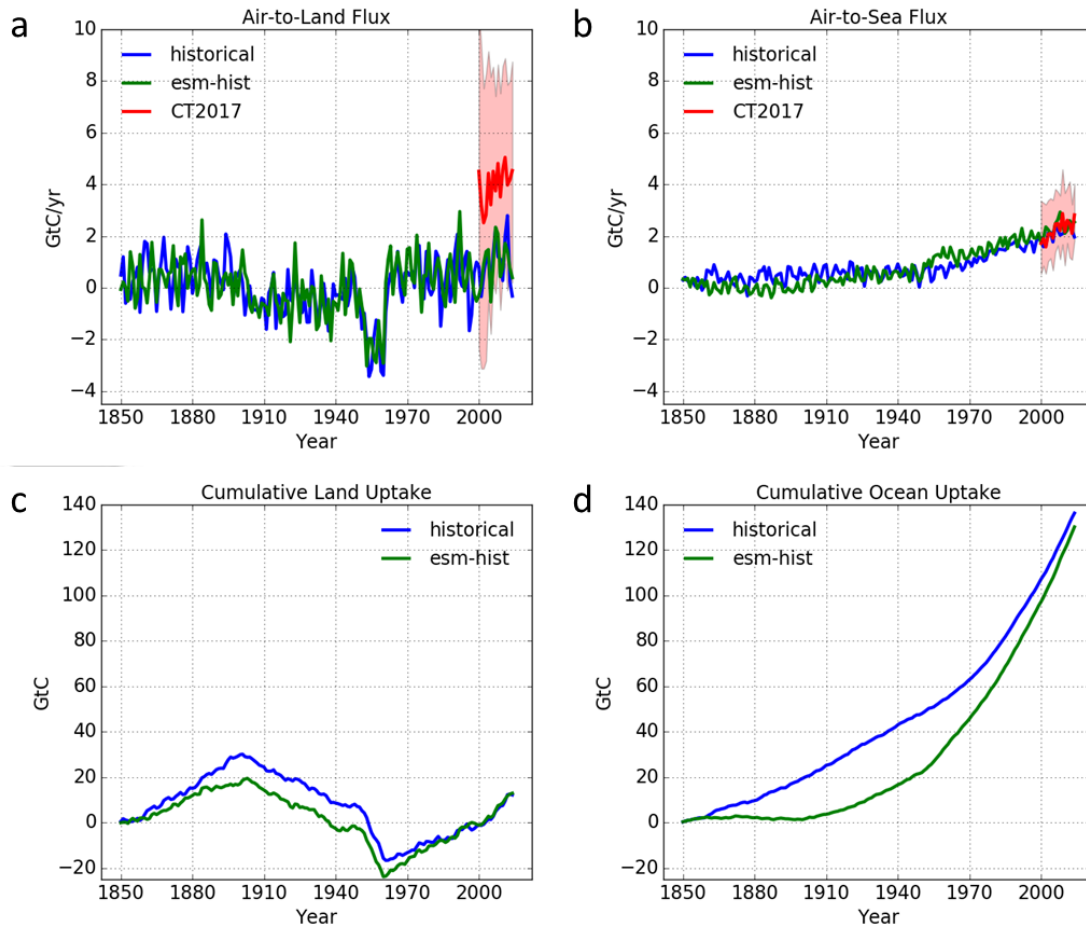
Figure 4. (a) Yearly global mean terrestrial biosphere CO_2 flux and (b) ocean CO_2 flux (positive into surface) in historical and esm-hist experiments. Red lines are global mean CO_2 flux estimates from CarbonTracker CT2017 data. The red shaded areas indicate the root-mean-square sum of one standard deviation uncertainties from CT2017 data. (c) Cumulative CO_2 uptake by the land and (d) ocean in historical and esm-hist experiments. Negative values indicate outgassing.
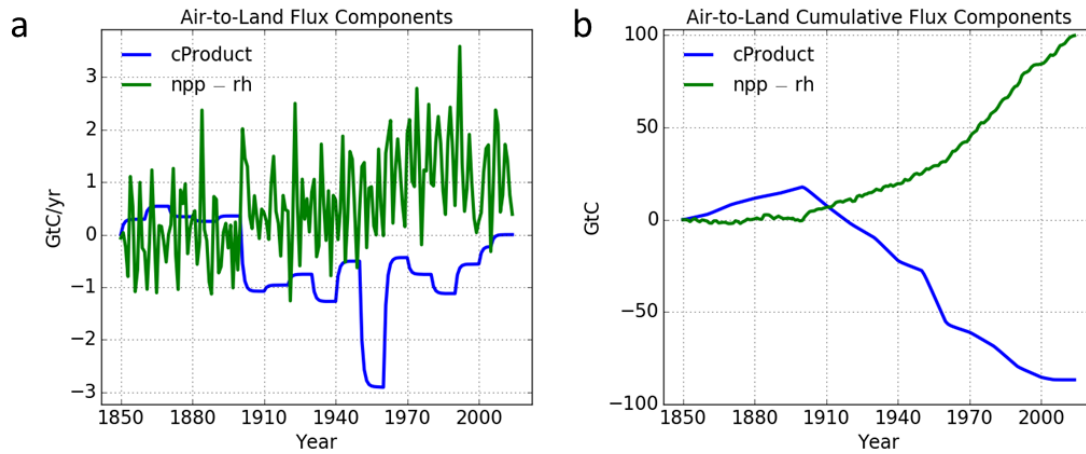
Figure 5. Comparison of the components of the total air-to-land flux for the esm-hist experiment. Positive flux is into the land surface. The difference between npp and rh equals the flux in the absence of the carbon flux in products of land use change. The terms $npp - rh$ and $cProduct$ have similar values in opposite signs which leads to relatively small uptake of carbon by the land by the end of simulation. The concentration-driven historical experiment has very similar trends.
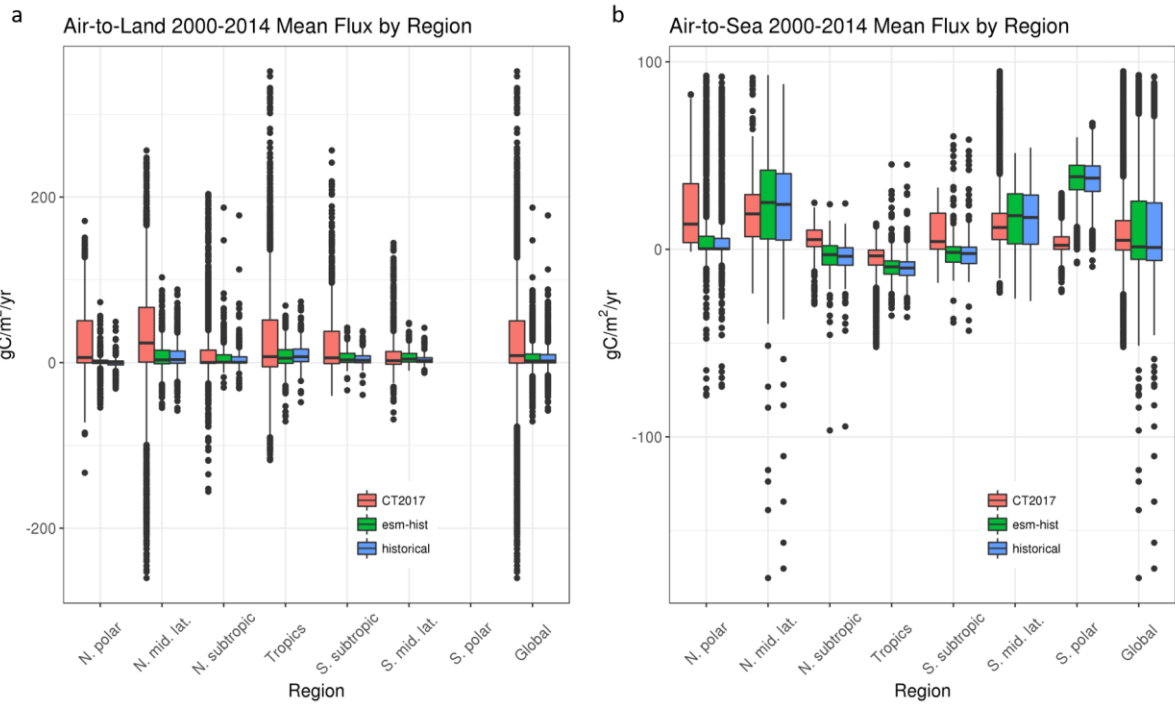
Figure 6. Boxplot summarizing the 2000-2014 mean (a) terrestrial biosphere and (b) ocean CO$_2$ flux (positive into surface) from the historical and esm-hist experiments and CarbonTracker CT2017 data (Peters et al., 2007). Boxes and middle lines inside the boxes indicate the interquartile range (IQR; 25th to 75th percentile) and median, respectively. Whiskers and points indicate ±1.5 × IQR and outliers, respectively. Land in the South polar (S. polar) region is covered with ice and has no flux. The ranges of latitudes used to define regions are North polar (N. polar): +60 to +90, N. mid. lat.: +30 to +60, N. subtropics: +15 to +30, Tropics: -15 to +15, S. subtropics: -15 to -30, S. mid. lat.: -30 to -60, S. polar: -60 to -90.
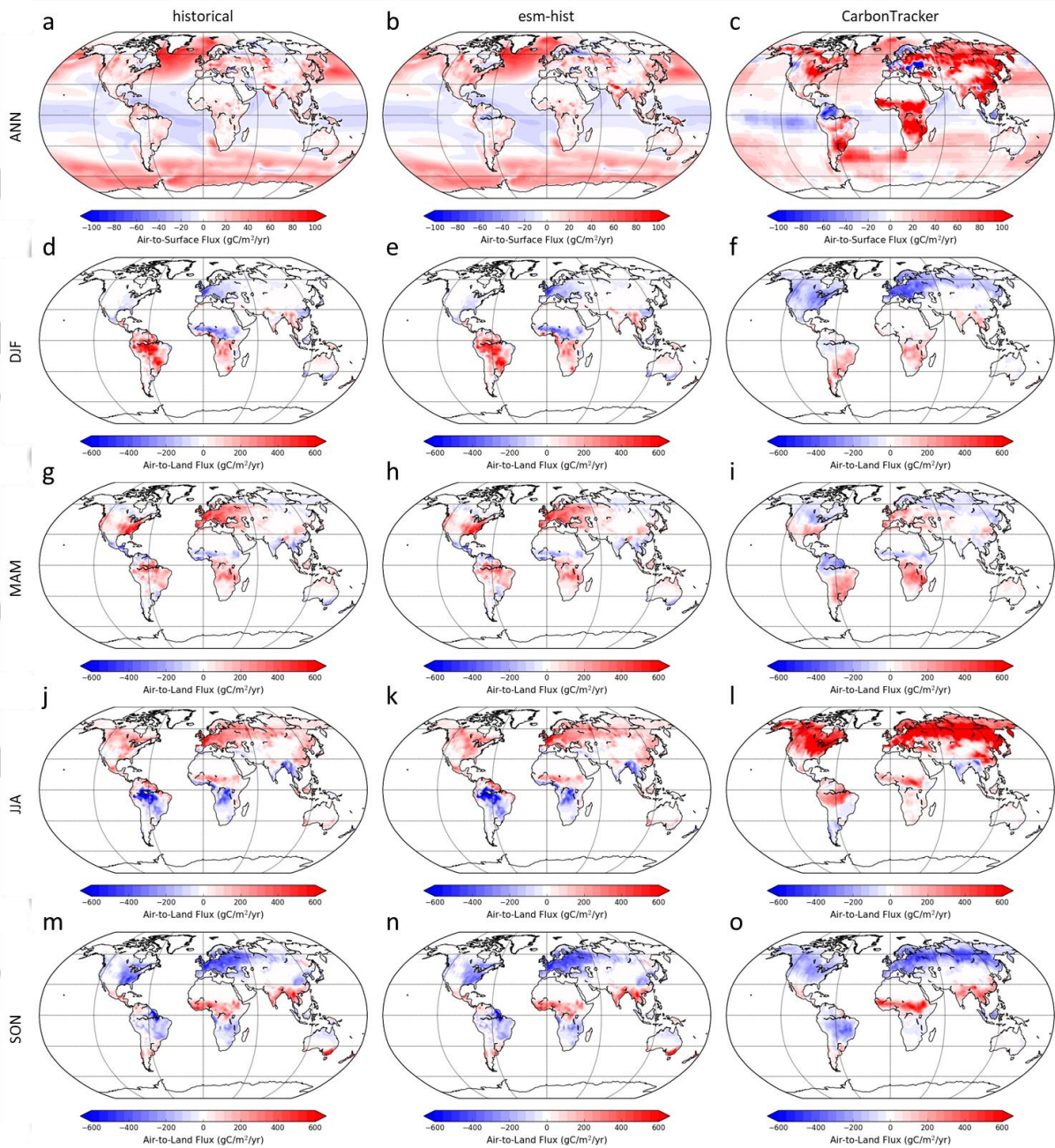
Figure 7. Distribution of 2000-2014 mean terrestrial biosphere and ocean CO_2 flux (positive into surface) in (a) historical and (b) esm-hist simulations compared to the same 15 year mean of (c) CarbonTracker CT2017 data (Peters et al., 2007). Distribution of seasonal terrestrial biosphere flux are shown. The labels on the left indicate seasonal means as (d,e,f) DJF: December, January, February; (g,h,i) MAM: March, April, May; (j,k,l) JJA: June, July, August; and (m,n,o) SON: September, October, November. ANN indicates annual mean (note different colorbar scale). The model ocean grid has been converted to 2° lat × 2.5° lon grid consistent with the land. CT2017 data has also been converted to 2° lat × 2.5° lon. For land, high latitudes in CT2017 data have especially high uncertainties. The coefficient of determination (r^2), bias, and standard error of the fluxes from historical and esm-hist experiments relative to fluxes from CarbonTracker are shown in Supplementary Information Table S1.
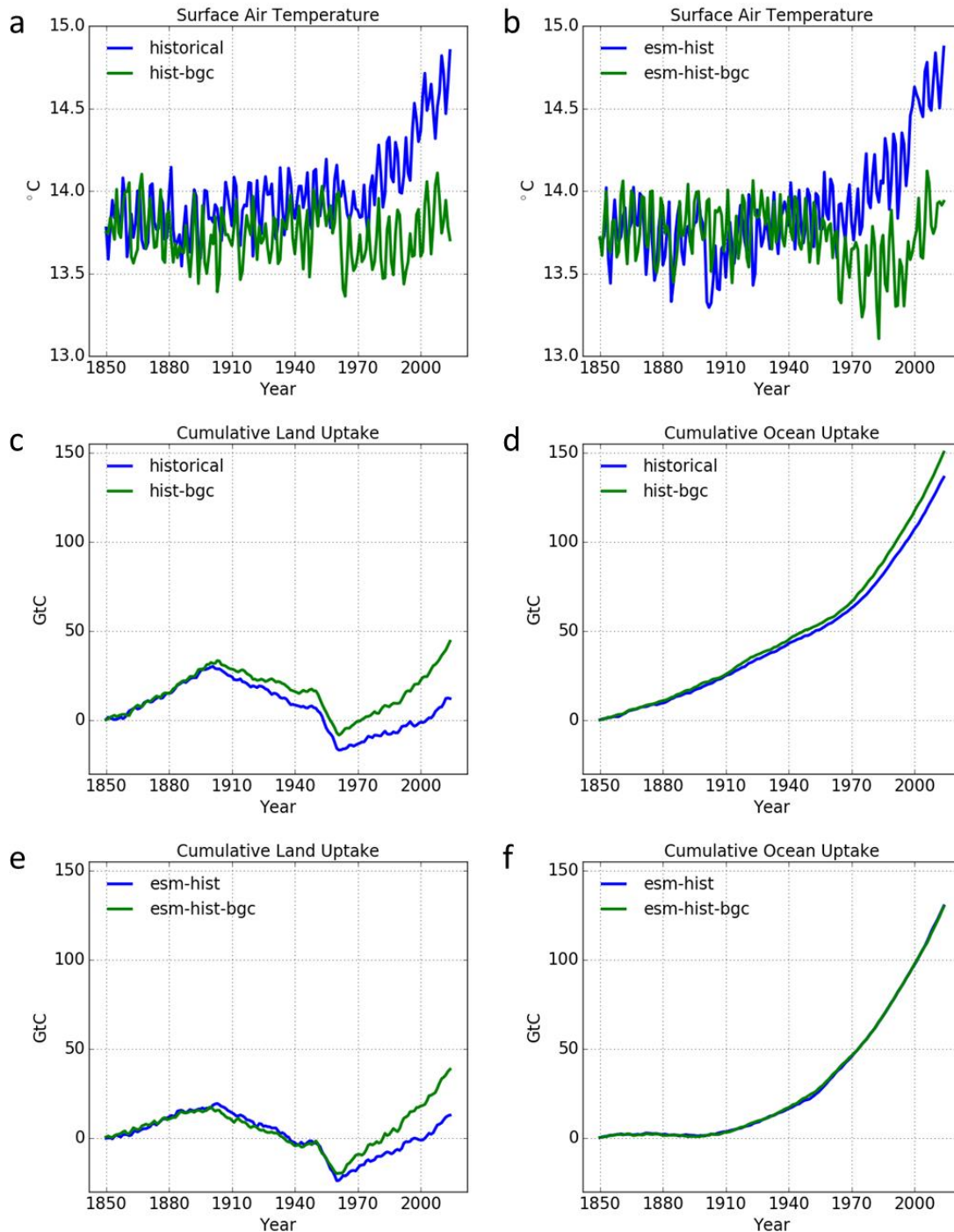
Figure 8. (a) Yearly global mean surface temperature comparing the fully and biogeochemically coupled versions of concentration-driven historical simulations. (b) Yearly global mean surface temperature comparing the fully and biogeochemically coupled versions of emissions-driven esm-hist simulations. (c,d) Yearly global mean land and ocean cumulative uptake comparing the fully and biogeochemically coupled versions of concentration-driven historical simulations. (e,f) Yearly global mean land and ocean cumulative uptake comparing the fully and biogeochemically coupled versions of emissions-driven esm-hist simulations.
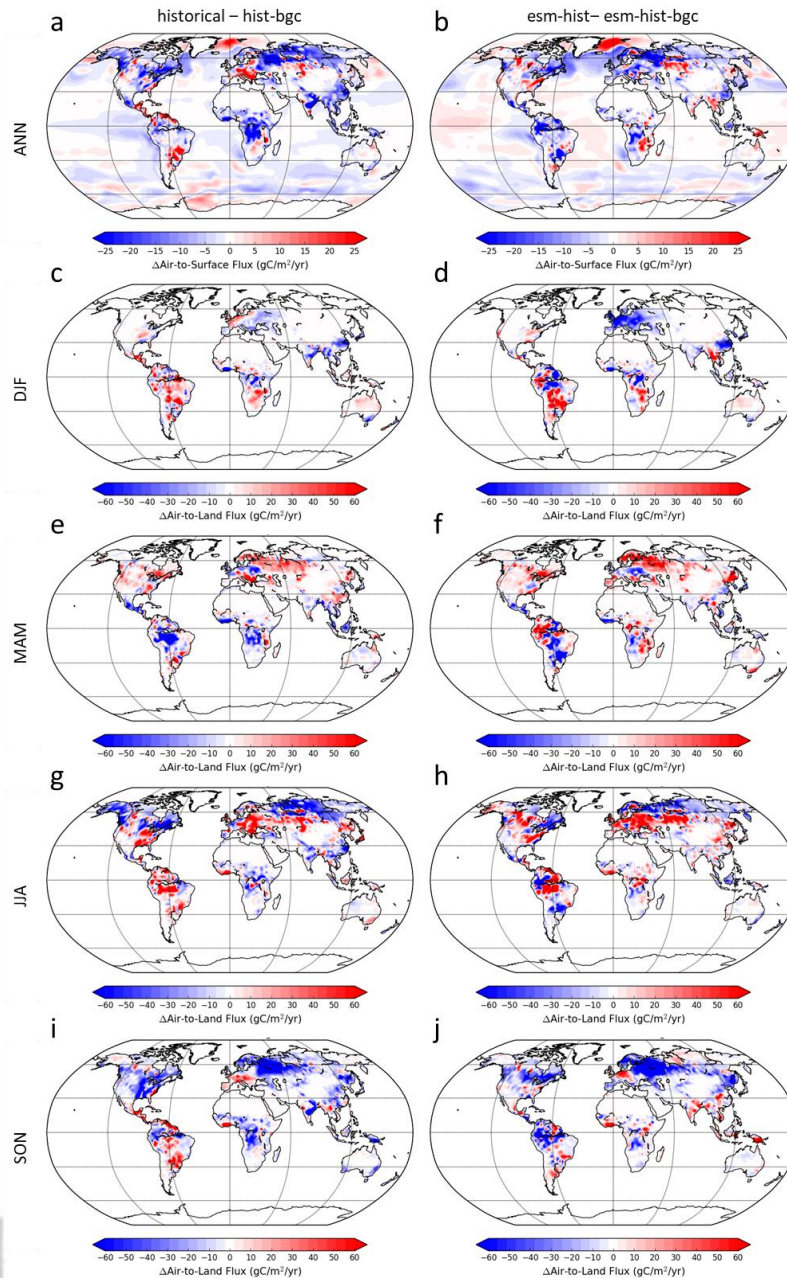
Figure 9. The distribution of differences in 2000-2014 mean air-to-land and air-to-sea flux between the fully and biogeochemically coupled historical simulations. (a) Difference in air-to-land and air-to-sea annual CO_2 flux between the historical and hist-bgc experiments. (b) Difference in air-to-land and air-to sea annual CO_2 flux between esm-hist and esm-hist-bgc experiments. (c,e,g,i) Distribution of differences in 2000-2014 seasonal air-to-land flux between the historical and hist-bgc experiments. (d,f,h,j) Distribution of differences in 2000-2014 seasonal air-to-land flux between the esm-hist and esm-hist-bgc experiments. The labels on the left indicate seasonal means as (c,d) DJF: December, January, February; (e,f) MAM: March, April, May; (g,h) JJA: June, July, August; and (i,j) SON: September, October, November. ANN indicates annual mean (note different colorbar scale). The model ocean grid has been converted to 2° lat × 2.5° lon grid consistent with the land. Positive flux is defined as CO_2 into the land or ocean surface.
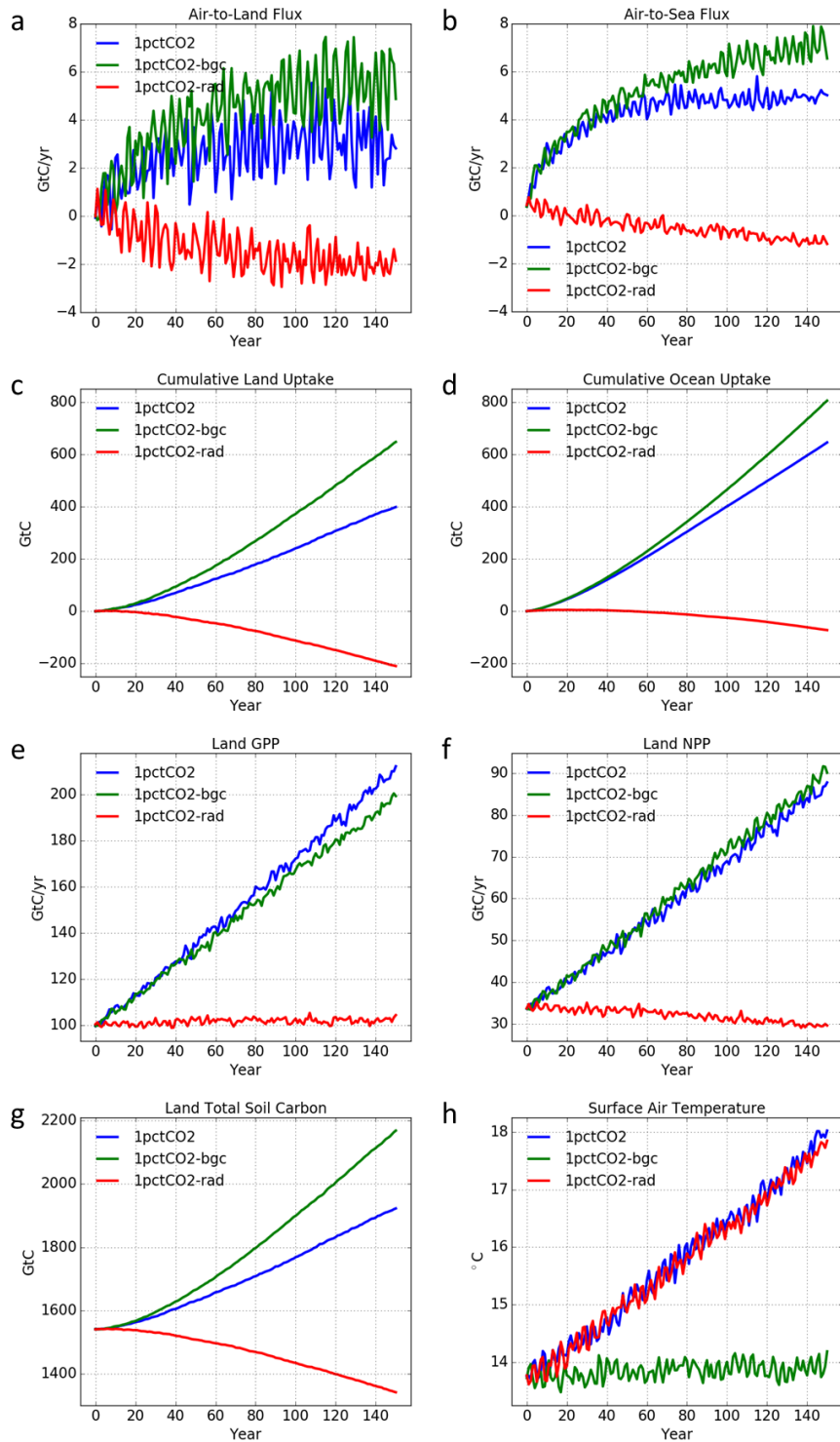
Figure 10. Selected diagnostics from the three variations of 1pctCO2 experiments. Yearly global means of: (a) air-to-land CO_2 flux, (b) air-to-sea CO_2 flux, (c) cumulative land CO_2 uptake, (d) cumulative ocean CO_2 uptake, (e) land GPP, (f) land NPP, (g) total soil carbon, and (h) surface air temperature. Flux convention is defined as positive into the land or the ocean surface.
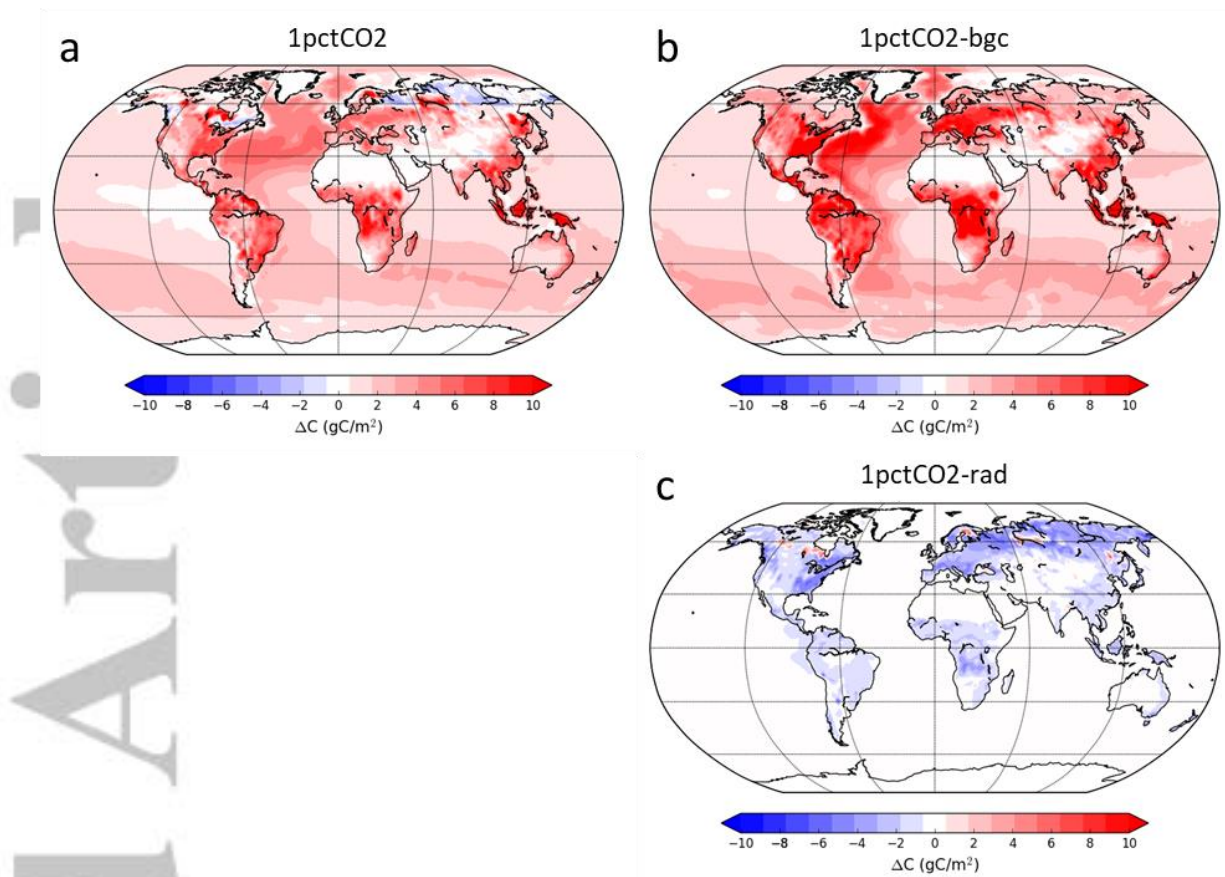
Figure 11. Spatial distribution of change in stored land and ocean carbon from year 0 to year 140 (time to quadruple CO₂) of the three 1pctCO2 simulations: (a) fully coupled 1%/yr concentration-driven experiment, (b) biogeochemically coupled 1%/yr concentration-driven experiment, and (c) radiatively coupled 1%/yr concentration-driven experiment. The ocean flux has been interpolated onto 2° lat × 2.5° lon grid consistent with the land.
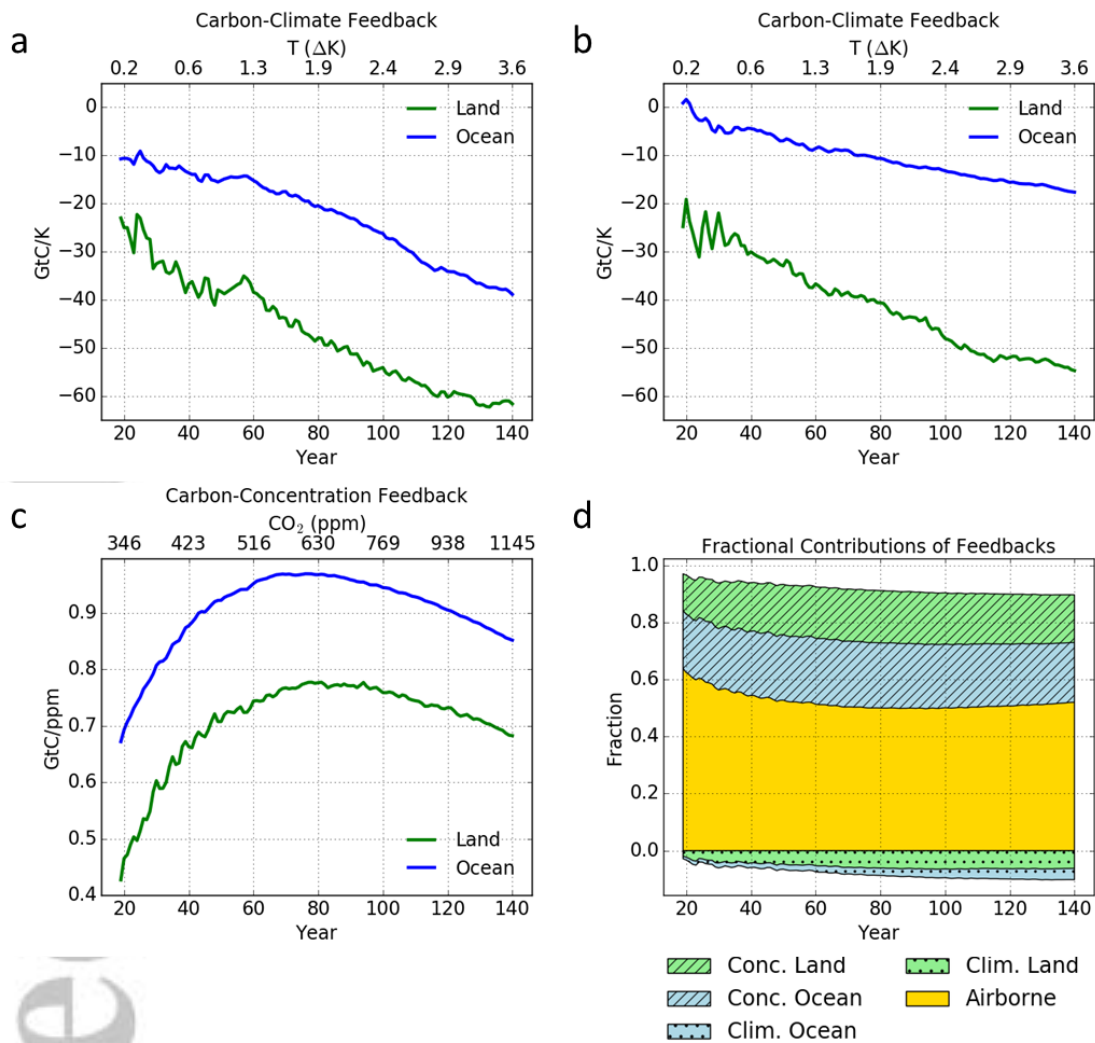
Figure 12. The yearly global feedback parameters (a) the carbon-climate feedback γ computed with $COU - BGC$ method (equation 5), (b) the carbon-climate feedback γ computed with the RAD method (equation 6), and, (c) the carbon-concentration feedback, β computed with equation 4. Units of β and γ terms are GtC/ppm and GtC/K, respectively. (d) Quantification, in fractions, of the land carbon-concentration, ocean carbon-concentration, land carbon-climate, and ocean carbon-climate feedback strengths along with the atmospheric carbon burden for each year derived from the β and γ feedback parameters. At any given year, the sum of the magnitudes of the five components equals one. Colors indicate the different carbon reservoirs. Hash and dot patterns indicate the total carbon-concentration and carbon-climate feedback contributions, respectively. γ in (d) is computed using the $COU - BGC$ method.
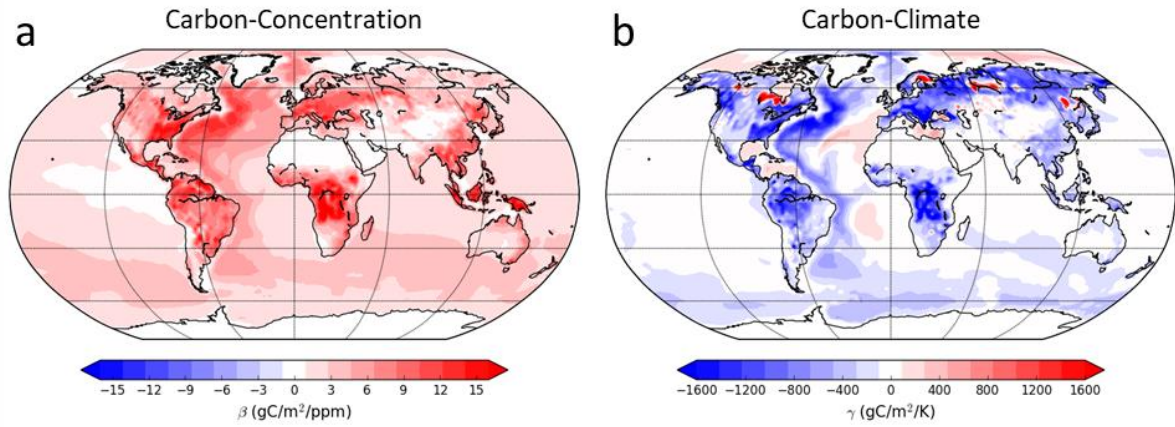
Figure 13. Distribution of (a) carbon-concentration (β) and (b) carbon-climate (γ) feedback parameters per unit area at the end of the 1pctCO2 simulation computed using method from equations 4 and 5 for each grid point with global mean ΔC_A^{BGC} and ΔT^{COU}. Ocean grid has been converted to 2° lat × 2.5° lon grid consistent with the land.
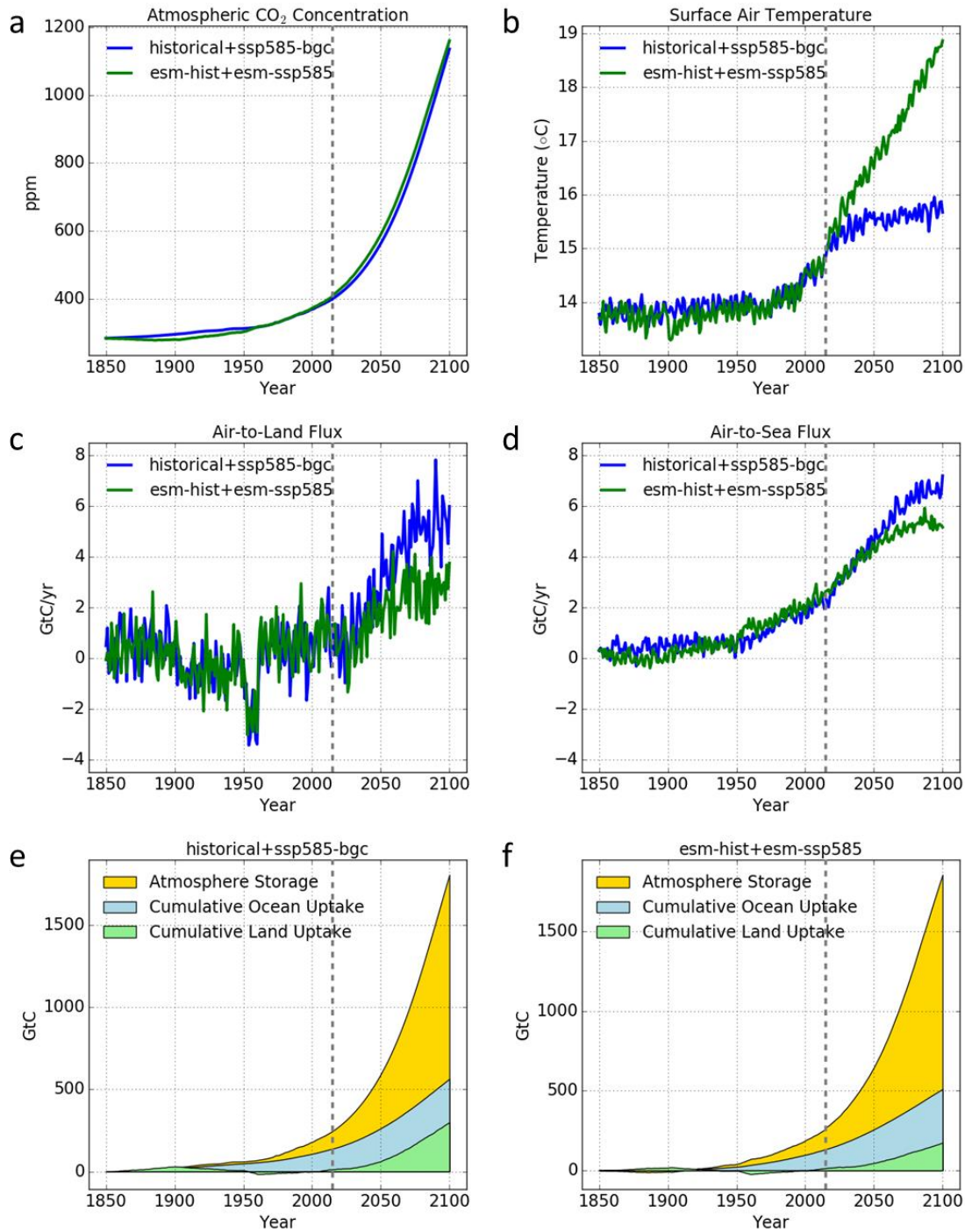
Figure 14. Combined historical and SSP5-8.5 (esm-ssp585 and ssp585-bgc experiments) yearly global mean trajectories of (a) the atmospheric CO$_2$ concentration, (b) surface air temperature, (c) air-to-land flux, (d) air-to-sea flux, and cumulative (e) land and (f) ocean uptake. Left and right of the gray dashed line corresponds to the historical periods (historical and esm-hist experiments) and future periods (ssp585-bgc and esm-ssp585 experiments), respectively. Positive flux is defined as flux into the land/ocean surface. In e and f, atmospheric storage of carbon is also shown for comparison.
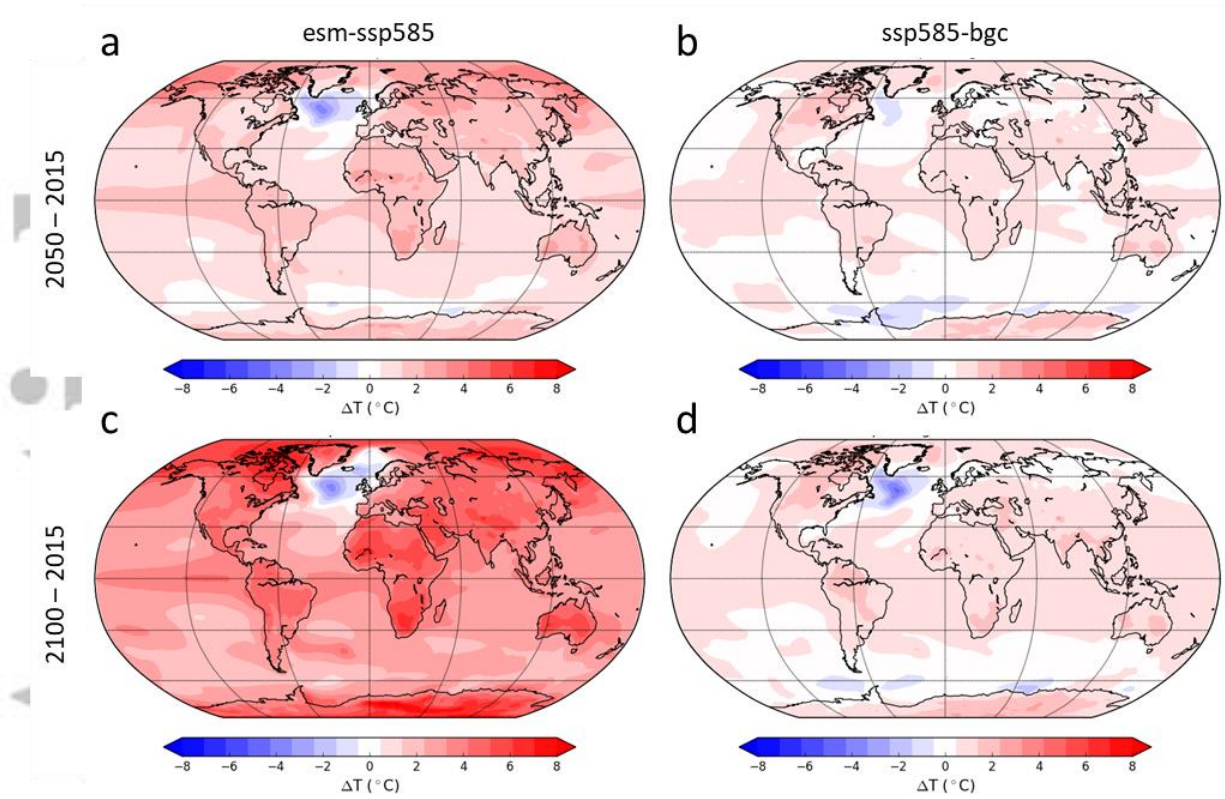
Figure 15. Distribution of change in surface air temperature for (a,c) esm-ssp585 and (b,d) ssp585-bgc experiments. (a,b) Difference between 2015-2024 mean and 2045-2054 mean. (c,d) Difference between 2015-2024 mean and 2091-2100 mean.
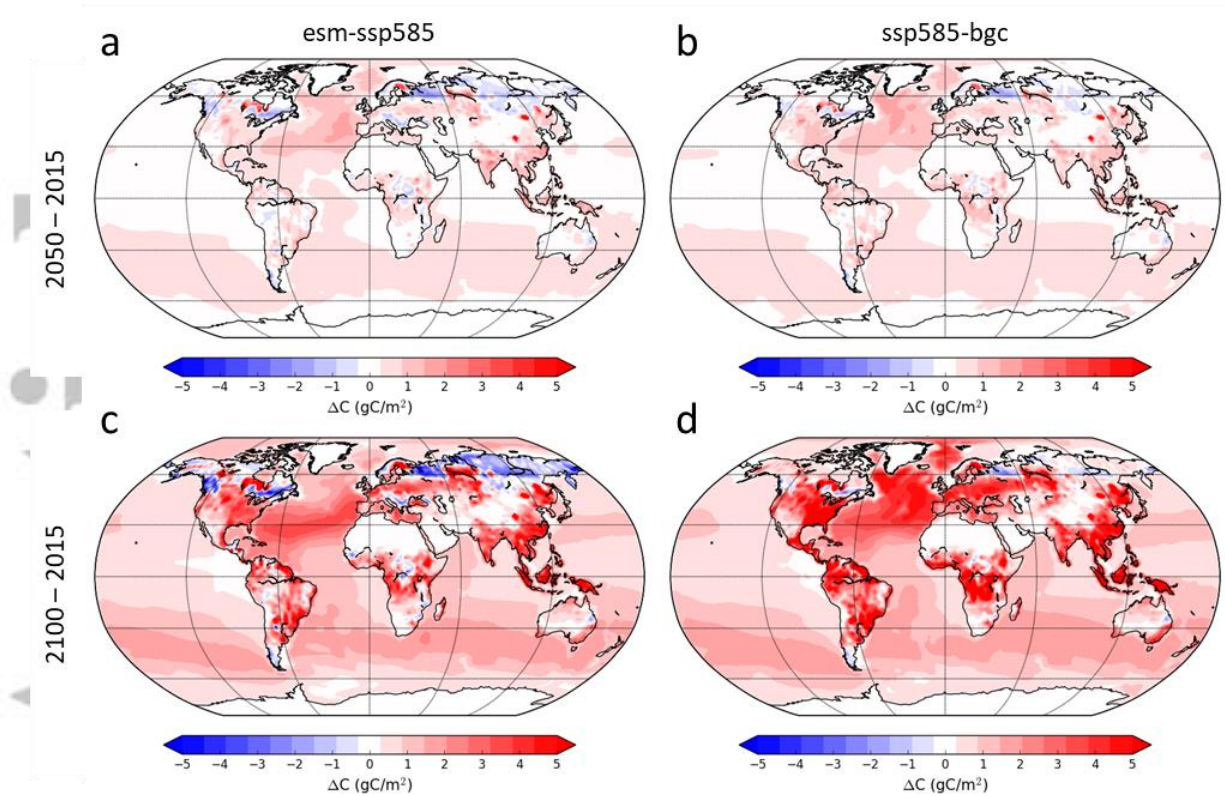
Figure 16. Distribution of change in stored land and ocean carbon for (a,c) esm-ssp585 and (b,d) ssp585-bgc experiments. (a,b) Difference between 2015-2024 mean and 2045-2054 mean. (c,d) Difference between 2015-2024 mean and 2091-2100 mean. The ocean grid has been interpolated onto 2° lat × 2.5° lon grid consistent with the land.

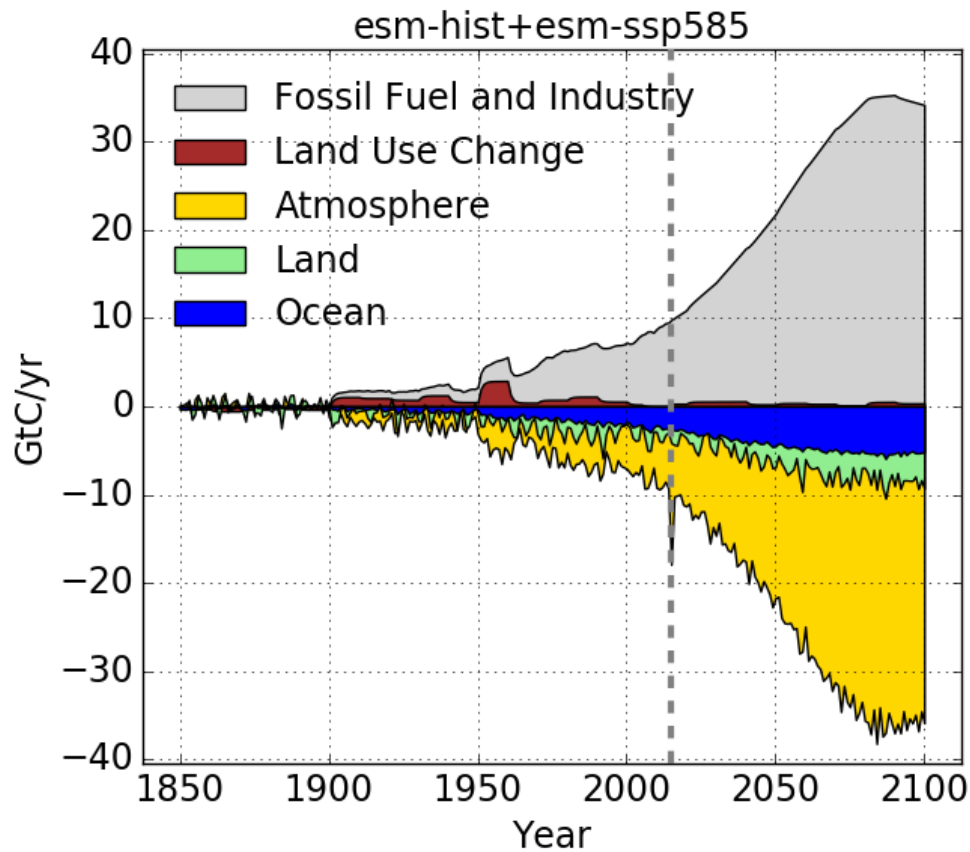*Figure 17. Components of the carbon budget in terms of CO$_2$ flux for the 1850-2100 period from esm-hist and esm-ssp585 experiments. Positive flux is defined as CO$_2$ emitted from fossil fuel/industry and land use change, and negative flux is defined as partitioning of CO$_2$ into the atmosphere, land, and ocean reservoirs. Land component excludes land use change. Left and right of the gray dashed line corresponds to the 2014 boundary between the historical periods (esm-hist experiment) and future periods (esm-ssp585 experiment), respectively.*
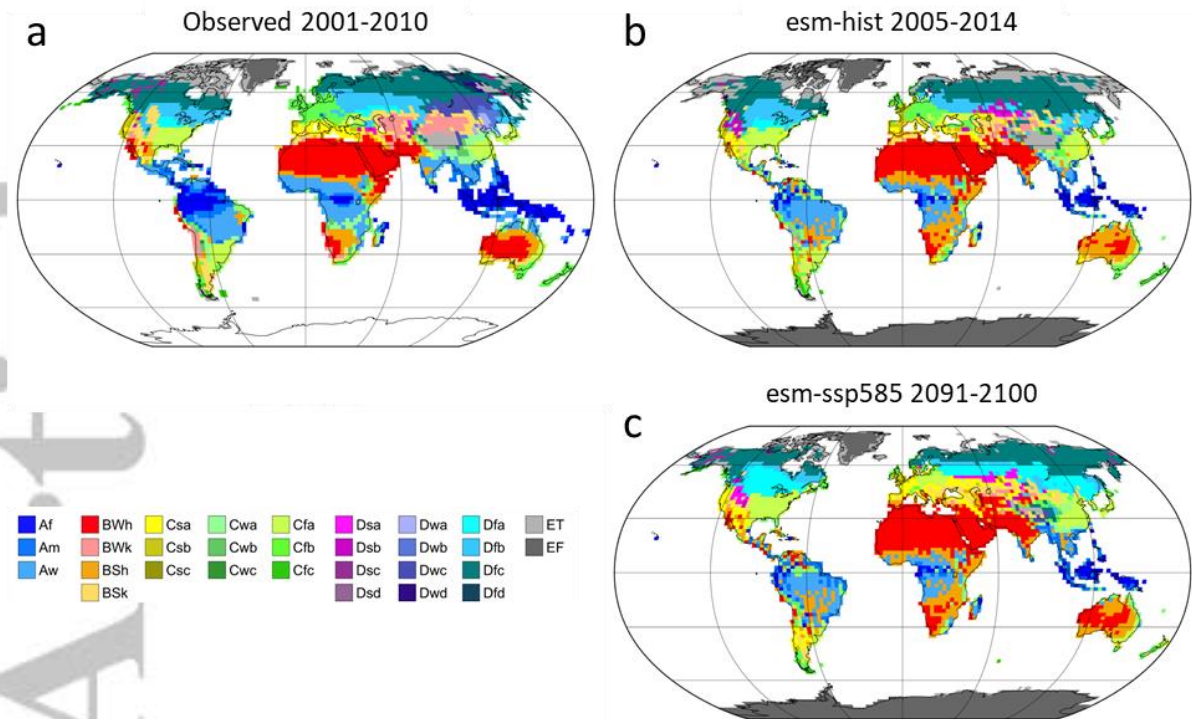
Figure 18. Köppen-Geiger classification of climates for land biomes for climates from: (a) surface temperature observations from the Climate Research Unit TS-3.22 (CRU TS-3.22) and precipitation from the Global Precipitation Climatology Centre (GPCC); (b) esm-hist 2005-2014 GISS ModelE2.1 simulation; and (c) esm-ssp585 2091-2100 GISS ModelE2.1 simulation. A detailed legend to Köppen-Geiger categories is provided in Appendix C.