Improvements in the terrestrial carbon cycle in CMIP models evaluated with satellite observations

Doctoral Dissertation of

Bettina Katharina Gier

Universität Bremen

iup

2024

Improvements in the terrestrial carbon cycle in CMIP models evaluated with satellite observations

Am Institut für Umweltphysik
vom Fachbereich für Physik und Elektrotechnik
der Universität Bremen

zur Erlangung des akademisches Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
genehmigte Dissertation

von

Bettina Katharina Gier

wohnhaft in Bremen

Erste Gutachterin:	Prof. Dr. Veronika Eyring
Zweiter Gutachter:	Prof. Dr. Gerrit Lohmann
Eingereicht am:	13. Februar 2024
Tag des Promotionskolloquiums:	13. Mai 2024

Note: This published version of the thesis contains slight modifications compared to the submitted version. These are the result of linguistic optimizations, as well as an update of references which changed since the submission of the thesis. No results or figures have been modified.

Abstract

With impacts of climate change already noticeable in every region of the world, understanding and accurately simulating the drivers of climate change is crucial. In particular, the global carbon cycle and its responses to changing carbon dioxide (CO_2) emissions plays an important role. State-of-the-art Earth System Models (ESMs) are able to simulate the complex interactions and feedbacks of physical and biogeochemical processes involved in the carbon cycle, and are routinely used to project climate change. A comprehensive evaluation of their ability to reproduce observed present-day climate is a prerequisite for reliable climate projections that are used as guidelines for climate policy. The goal of this thesis is to identify important improvements and key processes relevant to accurately simulate the carbon cycle under climate change and to provide recommendations for further model developments. This is done by a comprehensive evaluation of historical simulations from ESMs participating in the last two phases of the Coupled Model Intercomparison Project (CMIP) with satellite observations.

In a first study of this thesis, column-average CO_2 mole fraction (XCO_2) from CMIP5 and CMIP6 emission-driven ESM simulations are compared to satellite observations in the time period 2003–2014. The satellite data are a combined data product based on the SCIAMACHY/Envisat (2003-–2012) and TANSO-FTS/GOSAT (2009-–2014) instruments. The observational data are limited in their data coverage, for example due to clouds, surface reflectivity or sun elevation. Model simulations on the contrary provide data on every single grid cell around the globe. To accurately compare these simulations with gridded satellite data, the models are therefore sampled the same way as the observations in this thesis. Previous studies on CMIP data found the seasonal cycle amplitude (SCA) of atmospheric CO_2, arising due to carbon uptake by photosynthesis and release through respiration, to be increasing with increasing atmospheric CO_2 in the northern midlatitudes. In contrast, they found that the SCA based on XCO_2 satellite data shows a clear negative trend. This thesis could resolve this discrepancy and attribute it to different spatial coverage of the two satellite instruments contributing to the data set, with SCIAMACHY operating at earlier times from 2003—2012 with a higher mean SCA than the GOSAT satellite operating since 2009. Additionally, differences in spatial coverage can influence the SCA of CO_2 when missing data occur at peaks or troughs, lowering the computed SCA. This thesis could therefore show that XCO_2 data from satellite instruments provide a promising new way to evaluate the performance of ESMs on a global scale, also allowing investigation of regional effects, such as the increasing SCA at high latitudes.

Taking this observational sampling into account, the same satellite XCO_2 data are used to further evaluate climate model simulations. CMIP models are found to overestimate the carbon content of the atmosphere compared to the satellite observations. CMIP5 models have a mean bias of +10 ppm while in CMIP6 models the mean bias is reduced to +2 ppm. The spread remains large in both ensembles. The bias can be explained through the overestimation of the observed mean atmospheric growth rate (GR) of 2.0 ppm yr^{-1} that is overestimated by 0.4 ppm yr^{-1} in CMIP5 and 0.3 ppm yr^{-1} in CMIP6. Overall, this thesis showed that the CMIP6 ensemble shows better agreement with the satellite data than the CMIP5 ensemble in all considered quantities (XCO_2, GR, SCA and trend in SCA).

After investigating the atmospheric CO_2 concentration itself, the second study of this thesis evaluates present-day land carbon cycle variables with a focus on differences between concentration- and emission-driven simulations, as well as the inclusion of an interactive nitrogen cycle coupled to the carbon cycle in the models. Weaknesses in the land carbon cycle of CMIP5 have been previously found in the form of an overestimation of photosynthesis (GPP) as well as a large range in carbon stocks. Generally, CMIP6 models show an improvement over CMIP5 models in all analysed variables selected for this study (Leaf Area Index (LAI), GPP, land carbon uptake (NBP), and carbon stocks) to differing degrees. Overestimations of GPP in CMIP5 simulations are resolved in CMIP6 for models including a nitrogen cycle, but remain for models without, while the simulation of LAI remains challenging with large model spreads in both CMIP phases. While global mean land carbon uptake is well reproduced in the multi-model means, this is the result of an overestimation of NBP in the northern hemisphere, compensated by an underestimation in the southern hemisphere and the tropics. Carbon stocks remain a large uncertainty in models, and should be an area of further model improvements. Despite the increased process complexity in emission-driven simulations that fully account for the influence of climate-carbon feedbacks on atmospheric CO_2, they perform just as well as CO_2 concentration-driven simulations. Therefore, this thesis recommends the use of emission-driven over concentration-driven simulations as a default setting for future CMIP phases to represent important future Earth system feedbacks. In addition, the large improvements of photosynthesis through the inclusion of an interactive nitrogen cycle, with no adverse effects found in other areas, suggests the need to view the nitrogen cycle as a necessary part of all future carbon cycle models and raises the question about the importance of including further nutrient limitations, such as the phosphorus cycle.

All diagnostics of this thesis have been made publicly-available as part of the community-developed open source ESMValTool. This facilitates future analysis of the carbon cycle in upcoming CMIP phases and climate reports, and can easily be extended for analysing further specific model characteristics, such as dynamic vegetation. This thesis added to the understanding of the representation of the carbon and nitrogen cycles in CMIP models and leads to clear recommendations for further model improvement.

Integrated Author's References

Parts of this thesis (text, figures, and tables) are already published in several peer-reviewed studies. These studies are listed below and grouped according to the author's contributions. More information is given in Section 1.3 and in the corresponding chapters.

Publications as first author

Gier, B. K., Buchwitz, M., Reuter, M., Cox, P. M., Friedlingstein, P., & Eyring, V. (2020). Spatially resolved evaluation of Earth system models with satellite column-averaged CO$_2$. *Biogeosciences*, 17(23), 6115–6144. https://doi.org/10.5194/bg-17-6115-2020

Gier, B. K., Schlund, M., Friedlingstein, P., Jones, C. D., Jones, C., Zaehle, S., & Eyring, V. (2024). Representation of the Terrestrial Carbon Cycle in CMIP6. *Biogeosciences*, *in review*, https://doi.org/10.5194/egusphere-2024-277

Publications as co-author

Buchwitz, M., Reuter, M., Schneising, O., Noël, S., **Gier, B.**, Bovensmann, H., Burrows, J. P., Boesch, H., Anand, J., Parker, R. J., Somkuti, P., Detmers, R. G., Hasekamp, O. P., Aben, I., Butz, A., Kuze, A., Suto, H., Yoshida, Y., Crisp, D., & O'Dell, C. (2018). Computation and analysis of atmospheric carbon dioxide annual mean growth rates from satellite observations during 2003–2016. *Atmospheric Chemistry and Physics*, 18(23), 17355–17370. https://doi.org/10.5194/acp-18-17355-2018

Eyring, V., Bock, L., Lauer, A., Righi, M., Schlund, M., Andela, B., Arnone, E., Bellprat, O., Brötz, B., Caron, L.-P., Carvalhais, N., Cionni, I., Cortesi, N., Crezee, B., Davin, E. L., Davini, P., Debeire, K., de Mora, L., Deser, C., . . . **Gier, B.K.** . . . Zimmermann, K. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 – an extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of Earth system models in CMIP. *Geoscientific Model Development*, 13(7), 3383–3438. https://doi.org/10.5194/gmd-13-3383-2020

Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., Collins, W. D., **Gier, B. K.**, Hall, A. D., Hoffman, F. M., Hurtt, G. C., Jahn, A., Jones, C. D., Klein, S. A., Krasting, J. P., Kwiatkowski, L., Lorenz, R., Maloney, E., Meehl, G. A., . . . Williamson, M. S. (2019). Taking climate model evaluation to the next level. *Nature Climate Change*, 9(2), 102–110. https://doi.org/10.1038/s41558-018-0355-y

Lauer, A., Eyring, V., Bellprat, O., Bock, L., **Gier, B. K.**, Hunter, A., Lorenz, R., Pérez-Zanón, N., Righi, M., Schlund, M., Senftleben, D., Weigel, K., & Zechlau, S. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 – diagnostics for emergent constraints and future projections from Earth system models in CMIP. *Geoscientific Model Development*, 13(9), 4205–4228. https://doi.org/10.5194/gmd-13-4205-2020

Weigel, K., Bock, L., **Gier, B. K.**, Lauer, A., Righi, M., Schlund, M., Adeniyi, K., Andela, B., Arnone, E., Berg, P., Caron, L.-P., Cionni, I., Corti, S., Drost, N., Hunter, A., Lledó, L., Mohr, C. W., Paçal, A., Pérez-Zanón, N., . . . Eyring, V. (2021). Earth System Model Evaluation Tool (ESMValTool) v2.0 – diagnostics for extreme events, regional and impact evaluation, and analysis of Earth system models in CMIP. *Geoscientific Model Development*, 14(6), 3159–3184. https://doi.org/10.5194/gmd-14-3159-2021

Contributions to the IPCC Sixth Assessment Report

The author also had the honor of contributing to two chapters of the latest IPCC Working Group I Sixth Assessment Report (IPCC 2021a). First, as an acknowledged graphic developer for Chapter 3 (Eyring et al. 2021), and second as a contributing author for Chapter 5 (Canadell et al. 2021).

Canadell, J., Monteiro, P., Costa, M., Cotrim da Cunha, L., Cox, P., Eliseev, A., Henson, S., Ishii, M., Jaccard, S., Koven, C., Lohila, A., Patra, P., Piao, S., Rogelj, J., Syampungani, S., Zaehle, S., & Zickfeld, K. (2021). Global Carbon and other Biogeochemical Cycles and Feedbacks. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 673–816). Cambridge University Press. https://doi.org/10.1017/9781009157896.007

Eyring, V., Gillett, N., Achuta Rao, K., Barimalala, R., Barreiro Parrillo, M., Bellouin, N., Cassou, C., Durack, P., Kosaka, Y., McGregor, S., Min, S., Morgenstern, O., & Sun, Y. (2021). Human Influence on the Climate System. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 423–552). Cambridge University Press. https://doi.org/10.1017/9781009157896.005

Contents

1. Introduction

1.1. Motivation

Climate change is one of the most impactful topics in the world today, with many of its consequences posing a great challenge which affects the lives of people all over the world. The Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report (AR6) has found unequivocal evidence that the warming of the atmosphere, ocean and land is due to human influence (anthropogenic) and that "human-induced climate change is already affecting many weather and climate extremes in every region across the globe" (IPCC 2021b). It is further stated that the global surface temperature has increased by 1.1 °C in 2011–2020 compared to 1850–1900, which is commonly referred to as pre-industrial times. Further impacts of climate change include sea level rise, ocean acidification and an increased frequency of extreme events such as droughts. Changes brought about by past and future greenhouse gas emissions are found to be irreversible for centuries to millennia (IPCC 2021b). The $1\,°C$ warming since the pre-industrial time is a combined effect of greenhouse gases (GHGs) warming ($+1.5\,°C$), of which about half was contributed by carbon dioxide (CO_2), and cooling from anthropogenic aerosols ($-0.5\,°C$) (IPCC 2021b). The amount of CO_2 in the atmosphere has drastically increased since pre-industrial times with an average CO_2 concentration in 2023 of 419 parts per million (ppm) compared to 278 ppm in 1750 (Friedlingstein et al. 2023). Monitoring the long-term changes in atmospheric CO_2 concentrations is important to understand sources and sinks of carbon. These observations can be used to evaluate climate models, with better performance of present-day sources, sinks, and concentrations leading to a higher confidence in the models' ability to reliably simulate future projections of CO_2 concentrations under various scenarios.

The Earths' warming through GHGs is explained through the greenhouse effect: The incoming solar (shortwave) radiation passes through the atmosphere, but the outgoing infrared (longwave) radiation from the planet's surface is absorbed by GHG molecules by exciting the corresponding vibrational modes. When re-emitted, this energy is distributed isotropically, and thus some of this energy is trapped in the lower atmosphere (the troposphere) leading to a warming of the Earth's surface. The higher the concentration of GHG molecules, the stronger this heating effect becomes. CO_2 is a long-lived GHG, but not all CO_2 emitted stays in the atmosphere. Instead, land and ocean are currently acting as carbon sinks, together taking up 56 % of the emitted carbon (Friedlingstein et al. 2023; IPCC 2021b). However, while the natural land and ocean carbon sinks are projected to take up more carbon in the future under scenarios with increasing CO_2 emissions, this increase is smaller than the increase in CO_2

emissions and thus results in a smaller relative uptake (IPCC 2021b). To better understand the temporal evolution of carbon sources and sinks, the carbon cycle needs to be analyzed. On land, CO_2 in the atmosphere is absorbed by plants through photosynthesis. During the dormant season there is a net release of CO_2 due to respiration. This results in a seasonal cycle of the atmospheric CO_2 concentration.

To be able to anticipate the larger changes in the climate and its impact on a variety of environmental factors, such as temperature and the frequency of extreme events, as well as to understand the drivers of climate change, simulations by climate models are vital. These can reproduce the past observed climate, as well as project the evolution of climate change for given future scenarios. Climate models are based on a set of fundamental equations that represent the climate system with different degree of complexity (Gettelman and Rood 2016). Early General Circulation Models have been used in the community since the 1960s and simulate the physical conditions in the ocean or atmosphere. Since the 1980s, coupled climate models emerged which combine several of the individual components of ocean, atmosphere, land, and sea ice (Gettelman and Rood 2016). Climate models going beyond the representation of physical processes and additionally including biological and chemical processes such as the carbon cycle are called Earth System Models (ESMs). A large number of these models are being developed all over the world, often with a focus on different processes. To facilitate the comparison of climate model output with other climate models, an initiative was needed to standardize these modeling efforts. The objective of the Coupled Model Intercomparison Project (CMIP), which was initiated in 1995 by the Working Group on Coupled Modelling (WGCM) of the World Climate Research Programme (WCRP), aims to "better understand past, present and future climate changes" (WCRP 2024). It is a successful initiative allowing for a meaningful comparison of different model outputs by setting a framework of model experiments with common forcings and setups. An important part of CMIP is the availability of model output in a standardized and specified output format and for it to be distributed publicly through the Earth System Grid Federation (ESGF) (Eyring et al. 2016c). The most recent completed phase of CMIP is the 6th Phase (CMIP6) (Eyring et al. 2016c) with more than 100 registered contributors (PCMDI 2024). The CMIP model output is used for providing crucial input for international climate assessments, with the latest AR6 supported by data from CMIP6.

With this large amount of model data, there is a need for making the evaluation of the models more routine (Eyring et al. 2016a). The Earth System Model Evaluation Tool (ESMValTool) (Eyring et al. 2020; Lauer et al. 2020; Righi et al. 2020; Weigel et al. 2021), which is an open source community tool developed for the routine evaluation of CMIP models, fills this need. With the ability to easily include any observational datasets necessary for analysis, it facilitates the inclusion of newer observations for analysis, such as satellite observations. These satellite observations bring an opportunity for different analysis types, especially with respect to spatially resolved comparisons as opposed to requiring upscaling of in-situ data or subsampling models to areas with observational sites.

1.2. Key Science Questions

The goal of this thesis is to investigate the terrestrial carbon cycle in CMIP models with the help of satellite observations, identifying improvements and weaknesses in the new model generation and attributing these to specific advancements. This thesis therefore addresses the following three key science questions:

1. How can recent satellite observations be used to evaluate climate model simulations, and how does sparse data coverage affect results?
2. How do CMIP6 models compare to CMIP5 models in simulating atmospheric CO_2 and terrestrial carbon cycle variables, on the basis of observational and reanalysis data?
3. What effect does increased process complexity, through the inclusion of an interactively coupled nitrogen cycle and the forcing of carbon emissions instead of carbon concentrations, have on model performance?

1.3. Structure of the Thesis

Some parts of this thesis have already been published in peer-reviewed studies, listed on page vii. Chapters based on these publications are prefaced by a detailed description of the author's contributions to those publications. While the pronoun "we" is kept the same as in the published papers for consistency, all content used in this thesis from these publications including text, figures and tables were created by the author of this thesis unless specifically stated otherwise.

The structure of the thesis is as follows: Chapter 2 introduces the scientific background to frame this thesis and references literature upon which the studies of the following chapters are built upon. Chapter 3 introduces ESMValTool, a community-developed open-source software package for the analysis of Earth system models (ESMs). ESMValTool has been used for all studies in this thesis, and contributions to its code base led to co-authorship in several peer-reviewed studies (Eyring et al. 2020; Lauer et al. 2020; Weigel et al. 2021). Furthermore, this chapter also discusses contributions as contributing author to Chapter 5 (Canadell et al. 2021) and acknowledged graphic developer for Chapter 3 (Eyring et al. 2021) of the IPCC AR6, as the graphical and numerical input to these chapters was done with ESMValTool. Chapters 4 and 5 present the main results of this thesis. In Chapter 4 of this thesis, satellite column-averaged CO_2 data is used for a spatially resolved evaluation of output from CMIP5 and CMIP6 simulations, following **Gier** et al. 2020. With the knowledge from this study, the representation of the terrestrial carbon cycle in CMIP6 is analyzed in Chapter 5 based on **Gier** et al. 2024, focusing on the differences between models with and without interactive nitrogen cycle. Finally, Chapter 6 summarizes the results of this thesis and outlines possible future studies in an outlook.

2. Scientific Background

In this chapter, the scientific background relating to the thesis is introduced. This consists of the carbon cycle (Section 2.1) with a focus on the terrestrial part, an introduction to climate models (Section 2.2), and the implementation of the carbon cycle in Earth system models (ESMs) in general (Section 2.3), as well as specifics for each ESM used. Furthermore, as many different kinds of reference data sets are used throughout the thesis, Section 2.4 covers the available types of data procurement for carbon cycle variables.

2.1. The Carbon Cycle

The global carbon cycle describes how carbon is exchanged between components of the earth system, which includes the carbon emitted by humanity to the atmosphere, of which 56 % is currently absorbed by carbon sinks of the land and the ocean (Friedlingstein et al. 2023). When emitting more carbon into the atmosphere, the carbon content in the ocean and on land changes, which can in turn affect the mechanisms responsible for taking up carbon from the atmosphere, leading to feedback processes. As only carbon staying in the atmosphere is relevant for the greenhouse effect (in the form of atmospheric carbon dioxide (CO_2) acting as a greenhouse gas (GHG)), the land and ocean sinks slow down climate change. Thus, investigating the exchange processes between the different reservoirs is vital to better understand the rate of climate change.

While CO_2 is the most important GHG, it only makes up a small part of the air. The global mean atmospheric CO_2 concentration was 278 ppm in 1750 (pre-industrial level), while in 2023 it has risen to 419.3 ppm (Friedlingstein et al. 2023). ppm stands for "parts per million"; 1 ppm would correspond to one CO_2 molecule per one million particles of air. The global exchange fluxes between the atmosphere and land or ocean are usually given in GtC yr^{-1}, with 3.7 Gt of CO_2 constituting 1 GtC, or alternatively 1 PgC.

The following sections discuss the scientific background of the global carbon cycle, as well as additional details of the terrestrial carbon cycle. Unless indicated otherwise through a specific reference, this section is mainly derived from the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) Chapter 6 (Ciais et al. 2013) and references therein.

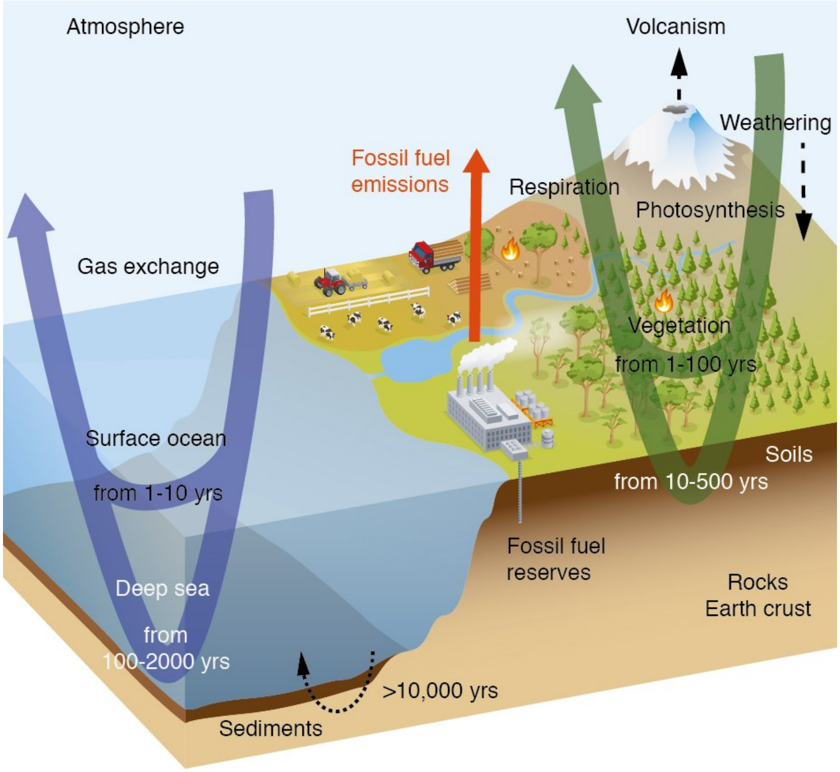


Figure 2.1.: Simplified schematic of the global carbon cycle showing the typical turnover time scales for carbon transfers through the major reservoirs. Reproduced with permission from Ciais et al. 2013 (their FAQ 6.2, Figure 1).

2.1.1. Global Carbon Cycle

Figure 2.1 shows a schematic overview of the global carbon cycle. Carbon exchanges between reservoirs in the ocean, land, and atmosphere occur on widely differing timescales, which can be split into a fast and a slow domain. Turnover times, defined as reservoir mass of carbon divided by exchange flux, of more than 10,000 years belong to reservoirs in the slow domain. This slow domain mainly consists of large carbon stores in rocks and sediments and is connected to the fast domain with turnover times up to millennia through chemical weathering, volcanic emissions, erosion, as well as sediment formation on the sea floor. The natural exchange fluxes between the slow and fast domain are relatively small ($< 0.3\,\mathrm{GtC\,yr^{-1}}$) and have been found to be relatively constant in time over the past few centuries.

The fast domain consists of the carbon in the atmosphere, the ocean, as well as the terrestrial biosphere. Carbon in the atmosphere is mainly stored in trace gases, predominantly CO_2, but also methane (CH_4), carbon monoxide (CO), further hydrocarbons, black carbon aerosols and organic compounds. In the ocean reservoir, dissolved organic carbon (DIC), made up of carbonic acid, bicarbonate and carbonate ions, dominates the carbon storage with 38,000 GtC, followed by ocean and floor sediments (1,750 GtC), dissolved organic carbon (DOC) with 700 GtC, and a small organic carbon pool of 3 GtC made predominantly of phytoplankton. Gas exchange driven by the partial CO_2 pressure difference between the atmosphere and ocean drives the air-sea CO_2 exchange, with the ocean currently taking up 2.8 GtC yr^{-1} (Friedlingstein et al. 2023). Within the ocean, carbon is transported through three mechanisms - the solubility pump, the biological pump, and the marine carbonate pump. In the terrestrial biosphere, carbon is mainly stored in organic compounds, namely 450-650 GtC in vegetation living biomass, 1500-2400 GtC in dead organic matter in litter and soils, as well as 300-700 GtC in wetland soils and approximately 1700 GtC in permafrost soils. The processes of carbon exchange between land and atmosphere, as well as the carbon exchange fluxes within the terrestrial biosphere are discussed in detail in Section 2.1.2.

As inferred from ice core measurements, the global carbon cycle was in a dynamic equilibrium before the industrial era, with a constant CO_2 concentration in the atmosphere. Since the industrial revolution, humans have been releasing large amounts of CO_2 into the atmosphere through the burning of fossil fuels and changes in land use such as deforestation. These emissions are called "anthropogenic emissions", contrasting to natural emissions (e.g., from natural fires or volcanic eruptions) and lead to an increased amount of CO_2 in the atmosphere, almost doubling the value of pre-industrial times (278 pm in 1750) at 419.3 ppm in 2023 (Friedlingstein et al. 2023). The increasing CO_2 amount in the atmosphere also impacts the carbon sinks of the ocean and land. On the ocean side, the rising CO_2 in the atmosphere leads to an increased partial pressure difference and thus a stronger ocean sink, while on land the CO_2 fertilisation effect increases photosynthesis and resulting in a larger carbon uptake on land. These are the so-called concentration-carbon feedbacks (Friedlingstein et al. 2006). Increased CO_2 also leads to global warming and other changes to the climate, giving rise to climate-carbon feedbacks (Friedlingstein et al. 2006). The rising temperature is generally found to decrease the uptake of both the land and ocean, such as through higher respiration fluxes and a lower solubility of CO_2 in a warmer ocean, with analysis on these feedbacks finding larger uncertainties on land than in the ocean (Arora et al. 2020; Friedlingstein et al. 2014; Hajima et al. 2014).

2.1.2. Land Carbon Cycle

Plant photosynthesis causes a net uptake of CO_2 from the atmosphere by land of about 120 GtC yr^{-1}. This process is also called gross primary production (GPP). Photosynthesis refers to the process of plants using energy from solar radiation to turn CO_2 from the atmosphere and water into Glucose ($C_6H_{12}O_6$) and oxygen (O_2). Carbon fixed into plants through photosynthesis is either directly released through autotrophic respiration, or further cycled through the plant to grow leaves, stems and roots. When herbivores eat plants, they also take up the carbon stored within and the CO_2 generated from cellular respiration is emitted back into the atmosphere. After the death of a plant or parts of it, litter is produced and the carbon therein enters the litter pools. Decomposing litter releases carbon to the atmosphere through hetereotrophic respiration, or is incorporated into soil and becomes soil organic matter. While carbon released to the atmosphere remains constant throughout the whole year, photosynthesis only occurs during the growing season, resulting in a net release of CO_2 from the land and thus rising atmospheric CO_2 in the dormant season, which eventually leads to a seasonal cycle of atmospheric CO_2 (Keeling et al. 1989). Due to the greater land mass in the Northern Hemisphere (NH), the global atmospheric CO_2 concentration follows the seasonal cycle of the NH, with peaks during northern winter and troughs during northern summer. The amplitude of that seasonal CO_2 cycle (hereafter referred to as SCA - seasonal cycle amplitude), has increased over the last 50 years, with larger increases in higher latitudes (Barnes et al. 2016; Graven et al. 2013; Keeling et al. 1995; Keeling et al. 1996; Myneni et al. 1997; Piao et al. 2018; Yin et al. 2018). Multiple drivers of this process have been explored in various studies, like CO_2 fertilization, land-use change and climate warming (Bastos et al. 2019; Fernández-Martínez et al. 2019; Zhao et al. 2016). All processes combined, land currently takes up 3.8 GtC yr^{-1} from the atmosphere (Friedlingstein et al. 2023).

The land carbon cycle is tightly coupled to the nitrogen cycle, as plants require reactive nitrogen (every nitrogen species other than N_2) to grow. Reactive nitrogen can be created through two natural processes (lightning and biological nitrogen fixation (BNF)), as well as three anthropogenic methods: the Haber-Bosch process, agricultural BNF, and fossil fuel combustion. The anthropogenic production has been larger than the natural production in the last decades. Plants in nitrogen-poor ecosystems will be limited in their uptake of carbon as plant growth cannot occur without nitrogen. The availability of reactive nitrogen changes in response to climate change, generally increasing with warmer temperatures and increased precipitation. Similarly, phosphorus also plays a role in the terrestrial carbon uptake, with a phosphorus limitation being more likely in high latitudes, while nitrogen limitation dominates in the tropics. While the land carbon cycle has been extensively studied, its fundamental properties have not been fully understood yet (Luo and Smith 2022).

2.2. Climate Models

Due to the sheer size and complexity of the earth system, it is not possible to research and fully understand it by solely performing laboratory experiments under controlled circumstances. While we can measure some aspects of its past and current state (see Section 2.4), for predictions related to the earth system, numerical climate models are indispensable. These models come in different forms, from simple 1D energy balance models calculating the temperature of the earth's surface using the Stefan-Boltzmann equation to ESMs with hundreds of thousands of lines of code (Alexander and Easterbrook 2015) simulating the entire earth system beyond just physical processes. No matter their complexity, they are based on fundamental laws of physics, such as the primitive equations describing the conservation of mass, momentum, energy and moisture which are solved numerically in the models. Due to the large scale and complexity of the earth system and the finite computational resources, climate models split the earth into three-dimensional boxes called grid cells, whose size can vary substantially

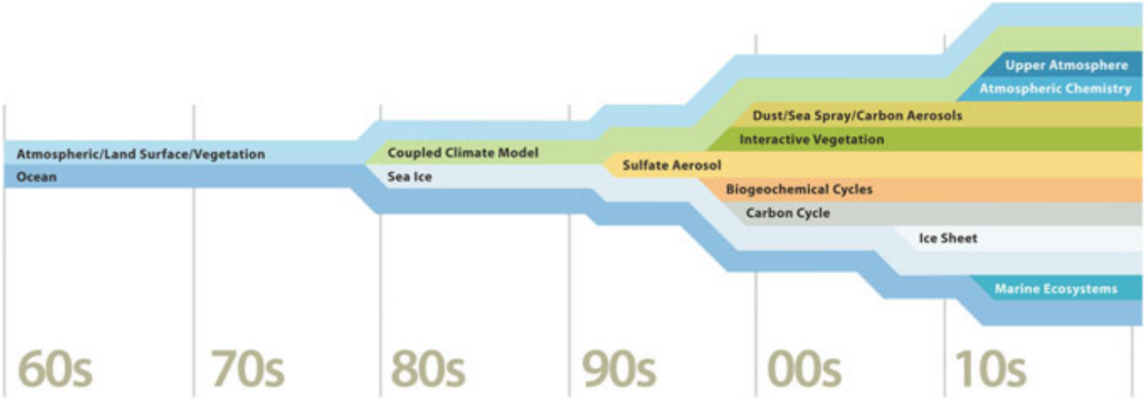


Figure 2.2.: Evolution of earth system components treated in climate models over time. Adapted with permission from Gettelman and Rood 2016.

across models. A typical ESM has a horizontal resolution, corresponding to latitudes and longitudes, of about 100 km and a vertical resolution, typically measured in pressure level or height, of approximately 1 km (Eyring et al. 2021). There are many processes, such as the formation of clouds or photosynthesis in plants, which occur on scales much smaller than the size of the grid cells, yet have a significant impact on the overall climate system. These processes are parameterized to approximate their effect at the scale of a grid cell by generating additional terms for the rest of the system (Gettelman and Rood 2016). Forcings are external inputs as boundary conditions to a climate model, often split into natural processes such as the solar radiation, or anthropogenic drivers such as GHG emissions and concentrations.

2.2.1. History of Climate Models

First numerical climate models were created in the 1960s (Gettelman and Rood 2016) and have been developed with increasing complexity (Figure 2.2) ever since. The first numerical climate models originated from weather prediction models and simulated purely physical aspects of individual components of the climate system like atmosphere, land, and ocean. While for in-depth analysis of specific parts the simulation of individual components of the climate system may be sufficient, for a general overview of the whole climate system the interactions between these components need to be considered in a consistent fashion. In 1975, Manabe et al. 1975 published the first paper detailing the results of a global Atmosphere-Ocean General Circulation Model (AOGCM), which is a model that couples the atmosphere and ocean. AOGCMs were the standard climate models assessed in the climate reports up to and including the IPCC Fourth Assessment Report (AR4) (Flato et al. 2013). Throughout the years, the resolution of climate models has increased. A typical gridbox for the first AOGCM had a width of 500 km, while those used in AR4 were finer at a scale of 110 km (Le Treut et al. 2007). The Sixth Assessment Report (AR6) in 2021 included some high resolution models with scales smaller than 25 km (Eyring et al. 2021). Besides increasing the resolution, the complexity of the models has increased with time by adding additional components or considering a wider range of processes within existing components. Cox et al. 2000 show results of coupling a dynamic global vegetation model to an AOGCM, highlighting the effect of carbon-cycle feedbacks on climate change. Climate models coupling AOGCMs with atmospheric chemistry and biogeochemical modules such as the carbon cycle are called ESMs (Jones 2020), and are the basis for model analysis in AR5 and the current AR6 (Canadell et al. 2021). More details about the carbon cycle in ESMs are discussed in Section 2.3. Throughout this history of climate models, modeling groups have shared their code or schemes, and some modeling groups built on top of existing models to supplement their own components. This results in relationships between models and statistically dependent projections to be considered in multi model ensembles (Kuma et al. 2023).

2.2.2. The Coupled Model Intercomparison Project (CMIP)

With the rising number of climate models, an initiative was needed to set up a common framework that would help facilitate the comparison of model output. In 1995, the Working Group on Coupled Modelling (WGCM) of the World Climate Research Programme (WCRP) initiated the Coupled Model Intercomparison Project (CMIP) (Meehl et al. 2000) with aims to "better understand past, present and future climate changes arising from natural, unforced variability or in response to changes in radiative forcing in a multi-model context" (WCRP 2024). CMIP model output provides crucial input for international climate assessments, with the latest AR6 supported by data from CMIP6. CMIP6 is the last completed phase of CMIP (Eyring et al. 2016c) with 114 registered contributors (PCMDI 2024). By setting up a framework of model experiments with common forcings and a uniform output structure, CMIP allows direct comparisons of the output from a large ensemble of climate model simulations.

CMIP6 consists of three elements (Eyring et al. 2016c). The first element is a handful of common experiments which include a *historical* simulation from 1850–2014 (the *historical* setup of the predecessor model generation CMIP5 covered 1850–2005) as well as the so-called DECK (Diagnostic, Evaluation and Characterization of Klima) experiments. These DECK experiments are four baseline experiments, a historical atmospheric Model Intercomparison Project (MIP) simulation (*amip*) with prescribed sea surface temperature and sea ice concentrations, a pre-industrial control run (*pi-control*), a simulation forced with an abrupt quadrupling of CO_2 (*abrupt-4xCO2*), as well as a simulation forced with a $1\,\%\,\mathrm{yr}^{-1}$ increase of CO_2 (*1pctCO2*). These constitute the core of figure 2.3, which shows the experimental design of CMIP6. All models participating in CMIP6 are required to run these experiments with the aim to document basic characteristics of models and maintain continuity to facilitate evaluations across different CMIP phases.

The second element of CMIP6 is to ensure "common standards, coordination, infrastructure, and documentation that will facilitate the distribution of model outputs and the characterization of the model ensemble" (Eyring et al. 2016c), and is mandatory for all models participating

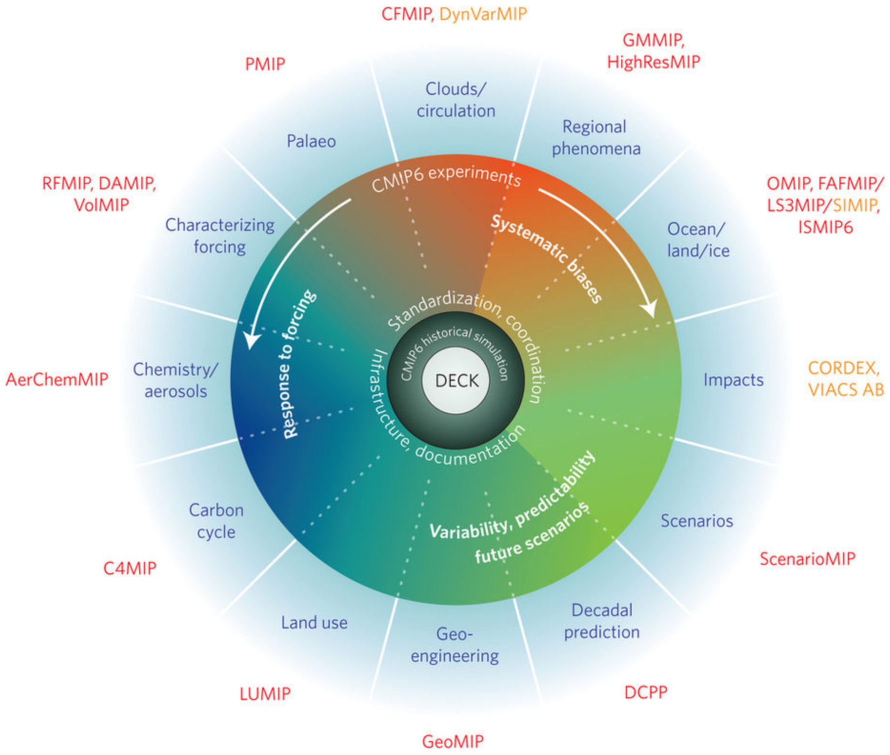


Figure 2.3.: Schematic of the overall design of CMIP/CMIP6 and the 23 CMIP6-Endorsed MIPs. Adapted with permission from Simpkins 2017.

in CMIP. It manifests in the use of common controlled vocabularys (CVs) which describe models and simulations. This results in consistent variables names and abbreviations, as well as a commonly defined output type. Model output is made publicly available on the Earth System Grid Federation (ESGF) nodes, allowing for easy access to directly compare model simulations.

The third and last part of CMIP6 is an ensemble of endorsed MIPs which build on the DECK experiments to target specific research areas and scientific questions. Participation in these MIPs is purely voluntary for the modeling groups. Of particulate note for this thesis is the Scenario Model Intercomparison Project (ScenarioMIP) (O'Neill et al. 2016), which provides forcings for runs simulating the future, and the Coupled Climate-Carbon Cycle Model Intercomparison Project (C4MIP) (Jones et al. 2016) with specialized runs for carbon cycle analysis. The ScenarioMIP experiments are based on Shared Socioeconomic Pathways (SSPs) (Riahi et al. 2017), a set of pathways based on theoretical assumptions of societal development (O'Neill et al. 2016). Five SSPs are considered, ranging from a sustainability scenario (SSP1) to a future of fossil-fueled development (SSP5), and coupled with associated global average forcing levels in the year 2100. For example, the SSP5-8.5 run follows the SSP5 scenario with a radiative forcing of 8.5 $\mathrm{W\,m^{-2}}$ in 2100. The CMIP5 equivalent to these SSPs are the so-called Representative Concentration Pathways (RCPs) (van Vuuren et al. 2011), which were only based on the radiative forcing in 2100, with RCP8.5 being the equivalent to SSP5-8.5.

C4MIP focuses on carbon cycle feedbacks and interactions in climate simulations (Jones et al. 2016). The core CMIP experiments also include simulations forced with prescribed CO_2 emissions instead of prescribed CO_2 concentrations, such as a historical emission-driven simulation (*esm-hist*). While these were not mandatory for models participating in CMIP, they were considered prerequisites for all models participating in C4MIP, as the only two simulations required for participation in C4MIP were a future extension of the emission-driven historical simulation using the SSP5-8.5 scenario (*esm-ssp585*), as well as a biogeochemically coupled version of the 1 % per year increasing CO_2 simulation (*1pctCO2-bgc*). This refers to a simulation where the radiation code of a model sees a constant, preindustrial concentration of CO_2, while the carbon cycle components of the model are allowed to respond to the increasing CO_2 concentration. Voluntary further simulations were divided into the categories of idealized simulations with different couplings of the biogeochemistry module and nitrogen deposition, as well as various scenario simulations of biogeochemically coupled SSP5 future scenarios. C4MIP also existed for CMIP5, thus making emission-driven simulations available for both CMIP phases.

With each new CMIP phase, additional models and improved model components are introduced, highlighting the need for comparison between phases to see if model weaknesses were identified and overcome, or to investigate potential reductions in systematic biases (Eyring et al. 2019) or introductions of new biases through increasing model complexity with the inclusion of additional processes. Even if the increase of process-realism of models does not lead to a reduction of present-day biases, a neutral impact can still be considered a successful step in model improvement and likely leads to more realistic future projections (**Gier** et al. 2024).

This thesis uses climate model data from CMIP5 and CMIP6 to evaluate the terrestrial carbon cycle, with the variables and models specific to each study discussed in the corresponding chapters.

2.3. Carbon Cycle in Earth System Models

Due to the key role of anthropogenic emissions of CO_2 as driver of climate change, the carbon cycle is an important process to be simulated in ESMs. Previous studies analysing the carbon cycle in CMIP5 historical model simulations found that while models reproduce the main climate variables and their seasonal evolution correctly, more carbon cycle–specific fields showed weaknesses, like a general overestimation of photosynthesis (Anav et al. 2013a). Over time, the process-realism has been increased to address weaknesses and biases, as well as to be able to more accurately simulate the future. In this section, the historical evolution of the representation of the land carbon cycle in models is briefly explained, followed by a general overview of the processes to be considered in land models. A matrix approach to land carbon cycle modeling is shown as a comprehensive way of representing the carbon cycle in models.

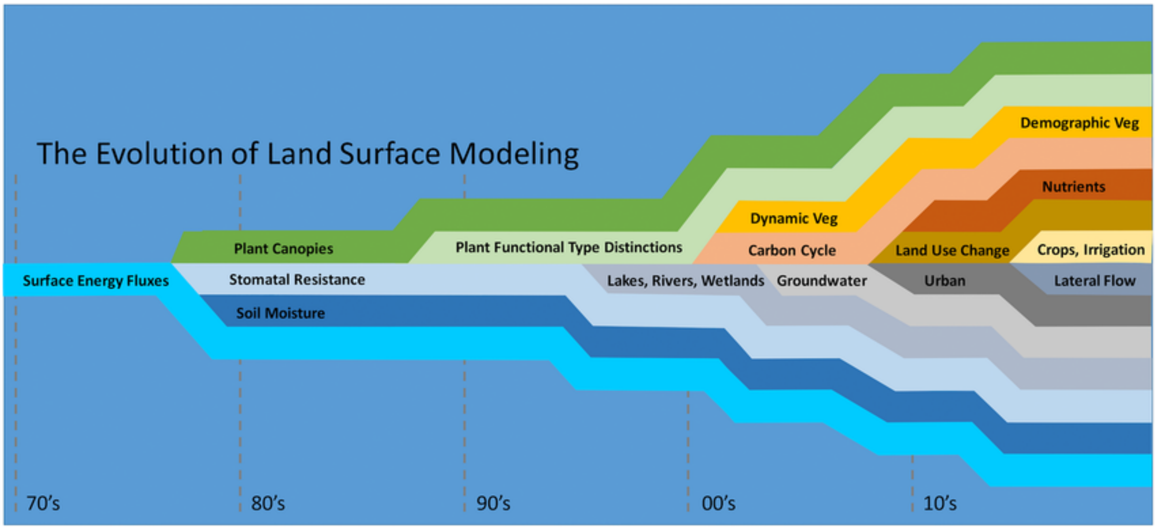


Figure 2.4.: Schematic of the evolution of land surface model processes, depicting the timing for new processes to be commonly included in ESMs. Individual models may differ from this series due to a different focus or evolution pathway. Adapted with permission from Rosie and Koven 2020.

Finally, special features and details of each of the land models which are part of the CMIP models used throughout this thesis are described.

2.3.1. Representation of the Land Carbon Cycle in Models over the Years

The carbon cycle in ESMs is split into different components. This thesis places a focus on the land carbon cycle and its interaction with the atmosphere. The land carbon cycle is modeled in land surface models (LSMs), which are numerical models calculating water, energy and carbon fluxes between the land surface and the atmosphere. Figure 2.4 gives an approximate overview on when certain processes became prevalent in LSMs. LSMs started with the need for physical boundary terms to represent the influence of land on atmospheric processes using surface energy fluxes. This was followed by basic modeling of soil moisture, stomatal resistance and plant canopies, which were later separated into different plant functional types (PFTs). In the early 2000's, model results that showed that the carbon cycle coupled to the climate system can substantially impact the rate of global warming (Cox et al. 2000) emphasized the need to include biogeophysical and biogeochemical feedbacks to the climate system. Thus, these processes became important targets for model development. One of the newest broadly introduced processes is the coupling of the nitrogen cycle to the carbon cycle to consider nutrient limitation. Modern LSMs encompass a large amount of fields: physics, biochemistry, physiology, ecology, hydrology, geography, statistics, mathematics, and high-performance computing (Fisher and Koven 2020). Figure 2.5 shows an example schematic of the representation of the carbon, nitrogen and phosphorus cycles in the state-of-the-art ORCHIDEE LSM.

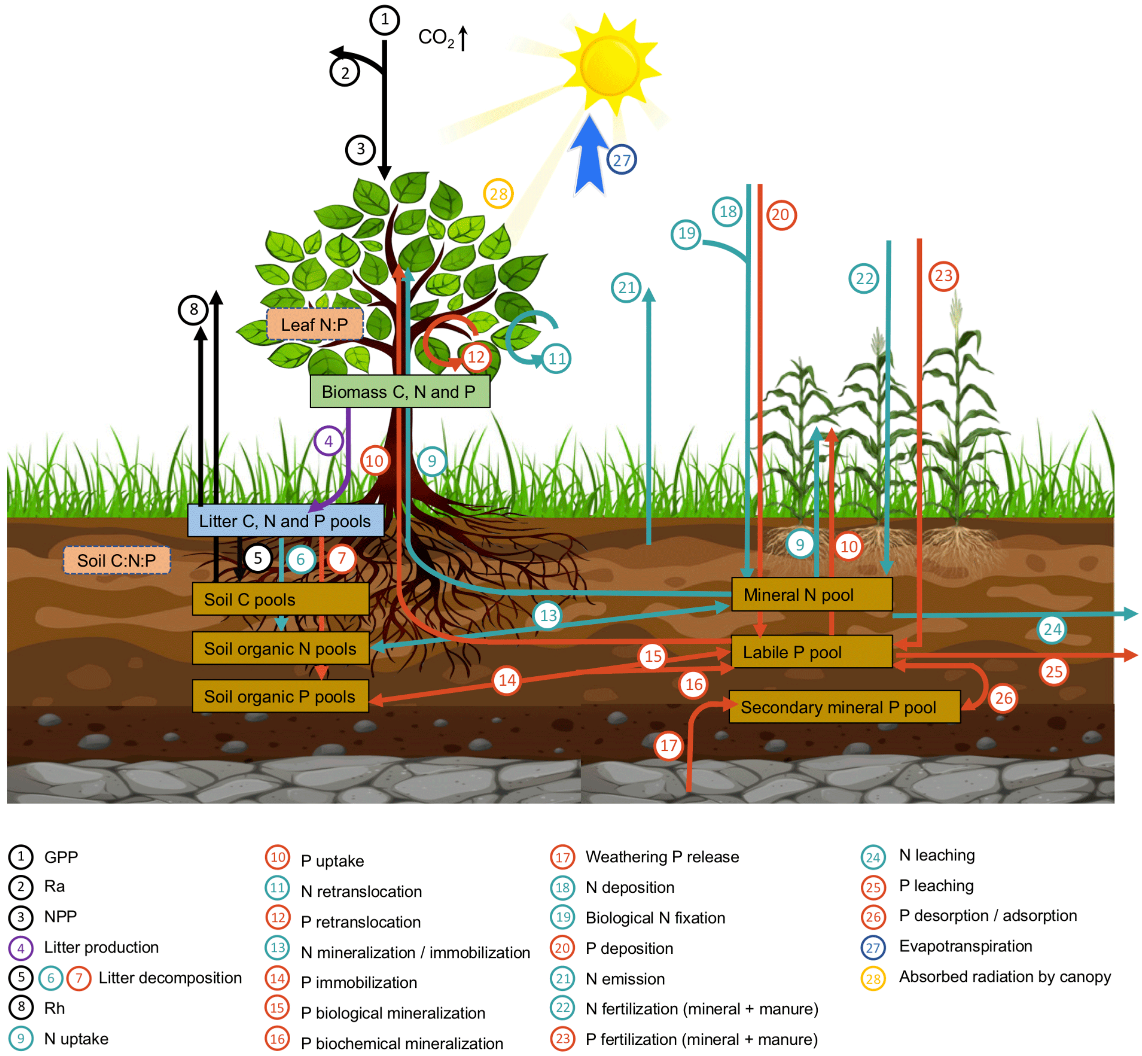


Figure 2.5.: Schematic of C, N, and P cycles considered in ORCHIDEE-CNP Adapted with permission from Sun et al. 2021.

2.3.2. Basic Processes Modeled in Land Surface Models

The following brief overview of the processes considered in LSMs is largely based on Fisher et al. 2014. Processes modeled in LSMs are split into four categories: leaf scale, canopy scale, plant scale and ecosystem scale.

On the leaf scale, stomatal conductance is modeled to regulate the uptake of CO_2 from and the release of water to the atmosphere through the stomata, which are small pores on leaves. These pores are not always open, nor do all plant species follow the same opening times. The rate of this CO_2 uptake (= the rate at which plants produce usable carbon through photosynthesis) is called gross primary production (GPP). In modern LSMs, there are three widely used approaches to parameterize GPP. The most common one is the biochemical approach through enzyme kinetics which unifies carbon, water and energy fluxes through stomatal conductance with common parameterizations following Farquhar et al. 1980 and Collatz et al. 1992. Further possibilities to model photosynthesis include the light-use efficiency (light used in photosynthesis is a fraction of absorbed photosynthetically active radiation, which can be measured), or the carbon assimilation which calculates a maximum possible GPP which is downscaled by taking into account environmental properties.

Canopy scaling describes how to scale leaf properties to an entire canopy. This often employs scaling an average leaf over the total leaf area index (LAI). LAI measures the total leaf area per ground area. Further canopy scale processes include the canopy temperature, interception of precipitation, aerodynamics and phenology, the latter describing the cycle of plant life. Phenology is largely determined using temperature as a threshold for leaf photosynthesis. Some models have moved from modeling phenology to prescribing it using observational products. The most important canopy scale components are plant functional types (PFTs). Each PFT combines a group of plants with similar behaviour, such as deciduous trees which drop their leaves, or evergreen trees which do not. The most common PFTs used are broadleaf evergreen trees, broadleaf deciduous trees, needleleaf deciduous trees, grasses and shrubs. Some models may further split these classes into multiple subclasses.

On the plant scale, autotrophic respiration (R_a), i.e., carbon returning to the atmosphere through maintenance and growth of plants, is typically modeled to depend on temperature or the maximum potential of plants to take up carbon. Net primary production (NPP) denotes the carbon taken up by photosynthesis after accounting for the autotrophic respiration, and is generally calculated as GPP minus R_a. The net carbon taken up by plants is typically allocated for growth and/or used for further survival and reproduction of the plant. LSMs generally allocate carbon to wood, roots and leaves.

Most processes modeled by LSMs take place on the ecosystem scale. LSMs have different soil layers and soil pools to store and cycle carbon, water and nutrients. The number of different layers varies largely between models. Plant growth can be limited by nutrient availability such as nitrogen or phosphorus. In CMIP5, only the Community Land Model (CLM) LSM included an interactively coupled nitrogen cycle, while this has become more common in CMIP6, with the ORCHIDEE LSM additionally including a coupled phosphorus cycle (Figure 2.5). As plants turn into litter, litter decomposition occurs which releases CO_2 into the atmosphere through heterotrophic respiration (R_h). This is often modeled based on temperature, but also moisture and carbon content. The total amount of carbon available to the ecosystem after accounting for respiration is called net ecosystem production (NEP), which is calculated as GPP minus R_a and R_h (a positive value denotes carbon uptake by land). Finally, the overall land carbon uptake, often referred to as net biome productivity (NBP), additionally includes carbon emissions from fires and other disturbances and is calculated as NEP minus disturbance fluxes. Further processes to be considered include snowmelt, permafrost thawing, infiltration and percolation, evapotranspiration, water balance, competition among PFTs for space, the establishment of how plants occupy new spaces, and plant mortality. Disturbances like fires or volcanic eruptions affect vegetation in random and punctuated manners. Beside CO_2, O_2 and H_2O, further trace gases, such as methane, are emitted which can affect weather and climate.

Most ESMs in the current iteration of CMIP are run using prescribed CO_2 concentrations. Some models can additionally perform simulations driven by CO_2 emissions, which requires an interactive carbon cycle to model the carbon fluxes (Friedlingstein et al. 2014). We call these concentration- and emission-driven simulations, respectively. Only emission-driven simulations have fully active climate-carbon cycle feedbacks for future projections, which are essential for the simulations to be self-consistent. Consequently, there are recommendations to prioritize the use of emission-driven simulations in CMIP7 (Sanderson et al. 2023). This adds further relevance and importance to the evaluation of the carbon cycle with more models required to implement an interactive carbon cycle to be able to compute CO_2 concentration from prescribed CO_2 emissions. This thesis uses output from both concentration- and emission-driven simulations.

2.3.3. Matrix Approach to Land Carbon Cycle Modeling

Luo and Smith 2022 describe a matrix approach to land carbon cycle modeling, which can be used to unify the different models. It has already been applied to some land models of ESMs, such as CLM (Huang et al. 2018a; Lu et al. 2020) and ORCHIDEE (Huang et al. 2018b). In this section we explain the matrix approach following Luo and Smith 2022 as a way to model the carbon cycle in models in a general form, while later sections discuss specific features of the land surface models used throughout this study.

Considering the example of litter, two pools can be introduced: a donor pool donating the litter, as well as a recipient pool which receives the litter. The transfer rate, or litterfall, is proportional to the amount of litter in the donor pool, and independent of the recipient pool. In mathematical terms, with X(t) denoting the donor pool size, ξ(t) an environmental scalar describing the effects of phenology, wind, and other environmental factors, as well as the litter decay constant k varying with litter type and location, the rate of litterfall can be expressed as:

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = \xi(t)kX(t) \tag{2.1}$$

Figure 2.6 shows how to translate the basic processes of the land carbon cycle into a model using carbon pools and fluxes. To describe this structure, equations need to consider both the carbon input into the pools, as well as the carbon leaving the pools. The first carbon pool, X_1 describes the carbon content of leaves. It receives carbon from photosynthesis and loses carbon through litterfall and autotrophic respiration. This can be represented mathematically by

$$\frac{\mathrm{d}X_l(t)}{\mathrm{d}t} = b_1\mu(t) - \xi(t)k_lX_l(t) \tag{2.2}$$

with $\mu(t)$ denoting the amount of total incoming carbon from NPP (= photosynthesis minus autotrophic respiration), b_l the fraction of incoming carbon attributed to leaves, k_l the rate of

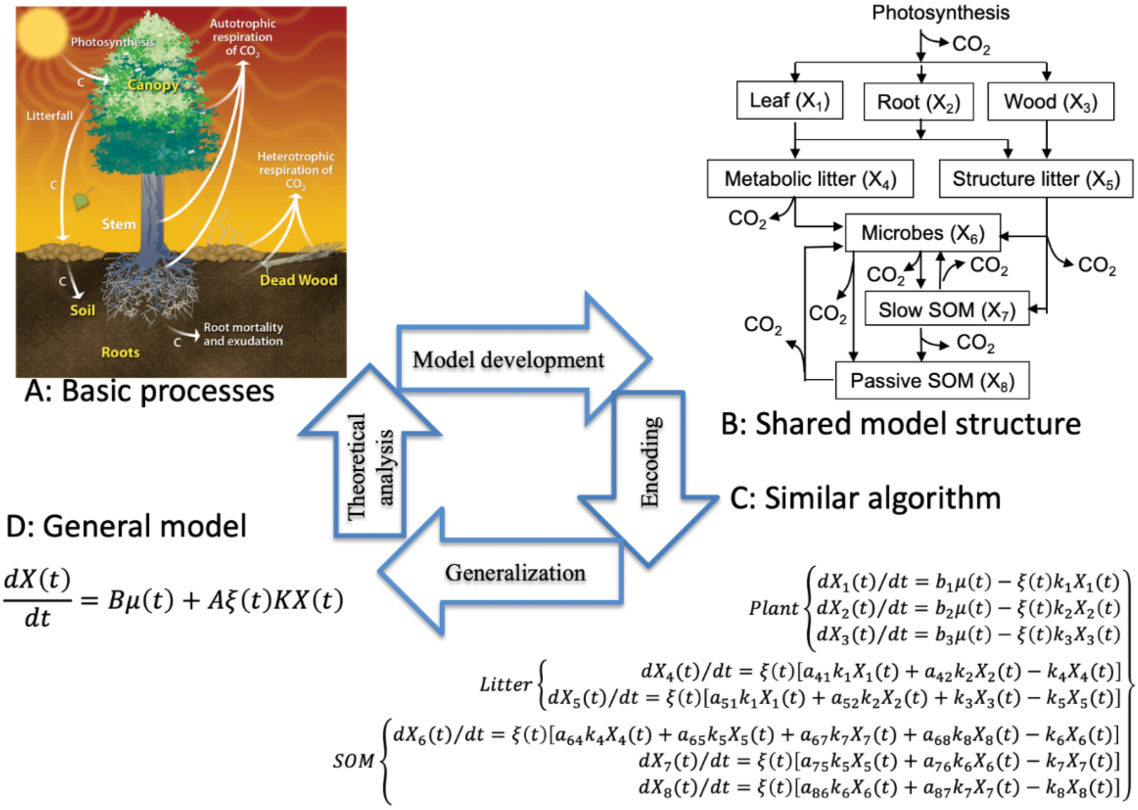


Figure 2.6.: A generalized matrix model of the terrestrial carbon cycle. (A) The basic carbon cycle processes are represented by four fundamental properties for all terrestrial ecosystems. (B) The four properties have been incorporated into terrestrial carbon cycle models with a pool-and-flux structure. (C) The structure is typically encoded using a set of balance equations with carbon input into and output from each pool. (D) The balance equations of terrestrial carbon cycle models can be converted to a matrix equation. Thus, the matrix equation can be considered as a general system equation (or a dynamical equation) for the terrestrial carbon cycle Adapted with permission from Luo and Smith 2022.

senescene, describing the aging of leaves, as well as the environmental modifier $\xi(t)$. Equations for the carbon balance of the other pools can be derived in the same way, with some pools receiving CO_2 from several other pools, and also removing CO_2 from the system through respiration. b_i is the coefficient for incoming carbon to pool i, a_{ij} the transfer coefficient from pool i to pool j and k_i the coefficient for carbon loss. All these single equations can be combined into a matrix equation

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = B\mu(t) + A\xi(t)KX(t) \tag{2.3}$$

where $X(t)$ is a vector denoting the pool sites, B a vector of partitioning coefficients b_i for carbon input, $\mu(t)$ the total carbon input rate, A a matrix with diagonal values of -1 and transfer coefficients a_{ij} for the pools in the off-diagonal, K a diagonal matrix for the carbon loss coefficients k_i signifying mortality for plant pools and decomposition for litter and soil pools. $\xi(t)$ as an environmental modifier can either be a scalar or a diagonal matrix accounting for carbon cycle responses to changes in temperature, moisture and oxygen.

This equation can be used for any land carbon cycle model which follows first-order kinetics with a variable number of carbon pools. Several land models also include vertical layers of soil pools. These models generally only consider carbon vertical transfers between adjacent layers. To include the 10 vertical soil layers of CLM4.5, Huang et al. 2018a extended the matrix equation by an additional term

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = B\mu(t) + A\xi(t)KX(t) + V(t)X(t) \tag{2.4}$$

which explains the vertical movement with the matrix V quantifying the transfers to the layers above and below the current layer. For CLM5, Lu et al. 2020 used four matrix equations to describe the carbon and nitrogen cycles, with two each for vegetation and soil. This included the vegetation carbon dynamic equation

$$\frac{\mathrm{d}X_\mathrm{v}(t)}{\mathrm{d}t} = B\mu(t) + \left(A_\mathrm{phc}(t)K_\mathrm{phc} + A_\mathrm{gmc}(t)K_\mathrm{gmc} + A_\mathrm{fic}(t)K_\mathrm{fic}\right) X_\mathrm{v}(t) \tag{2.5}$$

where the transfer between pools from matrices $A\xi(t)$ and K is split into time dependent contributions $A_j(t)K_j$ with j subscripts *phc* for phenology, *gmc* for gap mortality from harvest and natural causes, and *fic* for fire processes. This split leads to a more intuitive equation which is also computationally more efficient for spin-up.

2.3.4. Specifics of the Carbon Cycle in individual Land Surface Models

The special characteristics, such as the number of PFTs or the inclusion of a nitrogen cycle, of the LSMs which are part of the ESMs used throughout this thesis are listed below, taken from the Appendix of **Gier** et al. 2024. The sections are listed by LSM instead of ESMs as several models share the same LSM.

CABLE + CASA-CNP

The Commmunity Atmosphere-Biosphere Land Exchange model (CABLE, Kowalczyk et al. 2013) version 2.4 is a LSM coupled to the biogeochemistry module Carnegie-Ames-Stanfoard Approach carbon cycle model with nitrogen and phosphorus cycles (CASA-CNP, Wang et al. 2010) used in ACCESS-ESM1-5 (Ziehn et al. 2020). CASA-CNP and thus ACCESS-ESM1.5 is the only model in this study to include a phosphorus cycle coupled to the land carbon-nitrogen cycle. A sensitivity study of allowable emissions to nutrient limitation found a reduction of the land carbon uptake by 35-40 % with nitrogen limitation and a further 20-30 % reduction with nitrogen and phosphorus limitation on the carbon cycle (Zhang et al. 2013), showing the importance of nutrient limitation.

A simple land-use scheme accounts for annual net change in the vegetation tile fractions of each grid-cell which consider 10 vegetated and three non-vegetated surfaces. Three live and six dead carbon pools are modelled. LAI in ACCESS-ESM1-5 is calculated from specific leaf area and the size of the leaf carbon pool, while phenology is prescribed. In the CMIP5 model ACCESS-ESM1, which is not considered in this paper due to a lack of variables on the ESGF, LAI was significantly higher than observations, mainly due to an overestimation of LAI in the NH, despite a significant underestimation of LAI in the tropics. To better match the observations, two parameters were adjusted for ACCESS-ESM1-5: one PFT specific parameter used in the parametrisation for the maximum carboxylation rate and thus related to the nitrogen cycle, as well as one global parameter related to the daytime leaf respiration rate. Further changes to the model since CMIP5 include the conservation of land carbon, which was not conserved in CMIP5, as well as the inclusion of wetland tiles in the biogeochemistry calculation and the removal of a spin-up condition which ensured a minimum nitrogen and phosphorus level in soil pools.

CLASS + CTEM

The land component in the Canadian Earth System Models (CanESM) is divided into the physical part represented by the Canadian Land Surface Scheme (CLASS, Verseghy et al. 1993; Verseghy 1991, 2000) and the biogeochemical processes as simulated by the Canadian Terrestrial Ecosystem Model (CTEM, Arora and Boer 2003, 2005; Arora 2003). In the CMIP5 model (CanESM2, Arora et al. 2011) version 2.7 of CLASS was used, while the CMIP6 models CanESM and CanESM-CanOE (Swart et al. 2019a) employ CLASS v3.6. While neither version includes a nitrogen cycle, a parameter representing terrestrial photosynthesis downregulation is included to simulate the effect of nutrient constraints. This parameter is increased in CanESM5 compared to the previous version CanESM2, resulting in a higher land carbon uptake in CanESM5. Four PFTs are considered in CLASS, while CTEM increases the number to nine PFTs so that phenology can be simulated prognostically. LAI is dynamically simulated and three live and two dead carbon pools are considered. Added features since CanESM2 include dynamic wetlands and their diagnostic methane emissions.

CLM

The Community Land Model (CLM, UCAR 2020) is the most commonly used land model in this thesis, with 11 models across CMIP5 and CMIP6 using 3 different versions of it. For the CMIP5 models, CLM3.5 (Oleson et al. 2008) was used by FIO-ESM and CLM4 (Lawrence et al. 2011) was used in CESM1-BGC and NorESM1-ME. In CMIP6 CLM4 is used for SAM0-UNICON and TaiESM who mainly adapted the CESM1 configuration (Lee et al. 2020). CMCC-CM2-SR5 and CMCC-ESM2 use CLM4.5 (Koven et al. 2013), while the newest version CLM5 (Lawrence et al. 2019) is a part of CESM2, CESM2-WACCM, NorESM2-LM and NorESM2-MM.

While in CLM3.5 nitrogen limitation was merely represented by a downregulation factor, CLM4 introduced the coupled carbon-nitrogen cycle. Further improvements in CLM4 included transient land cover change modeling, changes to the PFT distribution and more realistic modeling of permafrost regions. To reduce biases found in CLM4 such as low soil carbon stocks and unrealistic values for GPP and LAI in several regions, such as a stark overestimation in the tropics, several parametrisations were changed in CLM4.5. Modifications were made to the canopy processes, including co-limitations on photosynthesis and photosynthetic parameters. Newly introduced features included a vertically resolved soil biogeochemistry with vertical mixing of soil carbon and nitrogen and a more realistic distribution of biological fixation over the year. The structure of the litter and soil carbon and nitrogen pools was adapted to the Century model and ^{13}C and ^{14}C carbon isotopes were introduced.

Finally, in CLM5 many major components of the land model were updated, with a focus on a better representation of land use and land-cover change as well as a more mechanistic treatment of key processes. Changes included a stronger soil moisture control on decomposition, the use of ^{13}C and ^{14}C isotopes for crops, and several changes to the nitrogen cycle and its impact on photosynthesis. Flexible plant C:N ratios were introduced to eliminate instantaneous down-regulation of photosynthesis, leaf nitrogen was optimized in the form of the leaf use of nitrogen for assimilation (LUNA, Ali et al. 2016) model and a model handling the fixation and uptake of nitrogen (FUN, Shi et al. 2016) was included. With respect to the land use and land cover aspect of the model, land unit weights are no longer fixed during the simulation and the transient PFT distribution was updated. CLM5 considers 22 live and 7 dead carbon pools as well as 22 PFTs.

CoLM+BNU-DGVM

The Common Land Model (CoLM, Dai et al. 2003) which shares an initial version with CLM but was then developed separately, is the land model component for the CMIP5 model BNU-ESM in the CoLM2005 version. CoLM includes a photosynthesis-stomatal conductance model for sunlit and shaded leaves separately. While carbon-nitrogen cycle interactions were included in the model, they were turned off for the CMIP5 simulations due to not being fully evaluated at the time (Ji et al. 2014).

HAL

The land model for the MRI models in both its CMIP5 version MRI-ESM1 and the CMIP6 version MRI-ESM2-0 is the Hydrology, Atmosphere, and Landsurface model (HAL, Hosaka 2011). It consists of three submodels called SiByl (vegetation) with grass and canopy vegetation layers, SNOWA (snow), and SOILA (soil) with 14 soil layers in the CMIP5 experiments.

ISBA-CTRIP

The land component for the CNRM-ESM2-1 model is presented by the Interaction Soil-Biosphere-Atmosphere (ISBA) LSM and the total runoff integrating pathways (CTRIP) river routing model (Decharme et al. 2019; Delire et al. 2020). ISBA-CTRIP simulates plant physiology, leaf phenology, carbon allocation and turnover, wild fires and carbon cycling through litter and soil (Séférian et al. 2019). Land use processes are prescribed instead of simulated, while land cover changes are used to represent anthropogenic disturbances. While the model does not include an interactive nitrogen cycle, its effects are included through an artificial downregulation of photosynthesis and a reduced specific leaf area with increasing CO_2 concentration. Six live and seven dead carbon pools are considered, along with 16 PFTs (Gibelin et al. 2008). Changes since the previous version used in CNRM-ESM-1 include improvements to the photosynthetic and autotrophic respiration schemes.

CTRIP includes carbon leaching through the soil and subsequent transport of dissolved organic carbon to the ocean. As chemical species such as dissolved inorganic carbon are not included, the air-water carbon exchange in the river routing model CTRIP cannot be computed. This leads to a carbon cycle which is not fully bounded.

JSBACH

JSBACH is the land component of the MPI-ESM model, with version 3.2 used for MPI-ESM1.2 (Mauritsen et al. 2019). In the previous version, parameters in the model for decomposition of dead organic matter were tuned to reproduce the historical atmospheric CO_2 concentrations, with soil and litter carbon stocks merely being the result of this tuning. In version 3.2 decomposition is handled by the YASSO model (Tuomi et al. 2011) based on litter and soil data resulting in no unconstrained parameters. YASSO simulates four fast soil carbon pools and one slow pool. A total of 18 dead carbon pools are considered due to a different application based on the woody and non-woody origins, as well as above and below ground decomposition. Additionally, three live carbon pools (natural vegetation, crops, pasture) and 13 PFTs are simulated by JSBACH, while permafrost carbon is not considered. The dynamical vegetation component interacts with the land use changes, modifying the land use data set to conform to the JSBACH setup. JSBACH3.2 includes an interactive terrestrial nitrogen cycle (Goll et al. 2017) driven by the nitrogen demand of the carbon cycle. Further adjustments in v3.2 include the change for carbon timescales in wood pools to be PFT specific.

JULES

The joint UK land environment simulator (JULES, Best et al. 2011; Clark et al. 2011) is the land model for the CMIP5 models HadGEM2-CC and HadGEM2-ES with the terrestrial carbon cycle following the Top-down Representation of Interaction of Foliage and Flora Including Dynamics (TRIFFID, Cox 2001) dynamic vegetation scheme, and the CMIP6 model UKESM1-0-LL (Sellar et al. 2019) employs JULES version 5.0.

Improvements in the version used for UKESM1-0-LL include the introduction of nitrogen cycling, as well as developments to plant physiology and functional types and land use. In this model, nitrogen controls biomass and LAI within TRIFFID, thus only indirectly affecting photosynthetic capacity, as well as limiting the decomposition of litter into soil carbon. For better agreement with observations, global total GPP was tuned down through a reduction of the quantum efficiency of photosynthesis. Furthermore, crop and pasture areas were separated and a harvest carbon flux was introduced. UKESM1 has four soil carbon pools, nine natural PFTs - increased from five in prior versions - and four PFTs for crop and pasture.

LM

The GFDL Land Model (Anderson et al. 2004, p. LM) is used in the GFDL Earth System Models, with the CMIP5 models GFDL-ESM2G and GFDL-ESM2M (Dunne et al. 2012; Dunne et al. 2013) using version 2.0 while the CMIP6 model GFDL-ESM4 employs version 4.1 (Dunne et al. 2020). Neither version includes an interactive nitrogen cycle.

Improvements since CMIP5 include updated soil types in the CORPSE model (Sulman et al. 2014; Sulman et al. 2019), hydrology, radiation, as well as the inclusion of a new fire model FINAL (Rabin et al. 2018) with daily computations instead of previously annual figures and a new model for vegetation dynamics through the Perfect Plasticity Approximation (Weng et al. 2015, PPA). LM4.1 includes six live carbon pools for leaves, fine roots, heartwood, sapwood, seeds and nonstructural carbon, 20 vertical soil levels split into separate fast and slow pools and pools for soil microbes and microbial products. Six PFTs are included representing C3 grass, C4 grass, tropical trees, temperate deciduous trees, and cold evergreen trees. Land use is accounted for through annual wood harvesting, crop planting and harvesting, pasture grazing, and newly included rangelands.

LPJ-GUESS

The Lund-Potsdam-Jena General Ecosystem Simulator (LPJ-GUESS, Smith et al. 2014) in combination with the Hydrology Tiled ECMWF Scheme of Surface Exchanges over Land (HTESSEL, Balsamo et al. 2009) is the land model used in the EC-Earth models EC-Earth3-CC and EC-Earth3-Veg (Döscher et al. 2022), the difference between the models consisting of CC having additional ocean biogeochemistry (PISCES) and atmospheric composition for CO_2 (TM5), letting it perform CO_2 emission-driven simulations.

HTESSEL solves the energy and water balance at the land surface, while vegetation types and vegetation coverage is interactively provided by the coupled LPJ-GUESS, which includes an interactive nitrogen cycle. Compared to the common area-based vegetation schemes, the interactive coupling of LPJ-GUESS to an atmospheric model should improve realism on longer timescales (Döscher et al. 2022). LPJ-GUESS includes 10 litter pools, seven vegetation carbon pools, as well as five soil carbon pools. Wildfires, disturbances and land use change are simulated on a yearly time step and distributed evenly throughout the year to conserve carbon mass. Land-use change dynamics are considered together with a crop module (Lindeskog et al. 2013) including five crop functional types. Three types of plant phenology - evergreen, seasonal-deciduous, and stress-deciduous - are considered, with only the latter two being simulated with an explicit phenological cycle. Seasonal-deciduous PFTs have a fixed growing season length of 210 days, while the growing season for stress-deciduous PFTs is determined by a threshold for the water stress.

MATSIRO + SEIB-DGVM/VISIT-e

The Minimal Advanced Treatments of Surface Interaction and RunOff (MATSIRO, Takata et al. 2003) is the physical land model for the MIROC-ESM family — MIROC-ESM and MIROC-ESM-CHEM (with coupled atmospheric chemistry) for CMIP5 (Watanabe et al. 2011) and MIROC-ES2L for CMIP6 (Hajima et al. 2020a)—, which consists of a single layer canopy, three snow layers and six soil layers down to a depth of 14 m. For the CMIP6 version a physically based parametrization for snow distribution and snow-derived wetlands was added.

Biogeochemistry in MIROC-ESM and MIROC-ESM-CHEM is simulated by the Spatially Explicit Individual-Based Dynamic Global Vegetation Model (SEIB-DGVM, Sato et al. 2007). It includes 13 PFTs split into two for grass and eleven for trees, as well as two organic carbon pools. Light capture competition among trees is explicitly modeled instead of parametrized.

MIROC-ES2L for CMIP6 uses the Vegetation integrative SImulator for Trace gases model (VISIT, Ito and Inatomi 2012), with changes for coupling to the ESM (adding the -e suffix), such as including leaf-nitrogen concentrations and thus limitations to enable fully coupled climate-carbon-nitrogen projections, and land-use change processes to get more use out of new LUC forcing data sets, such as using five types of land cover. The model does not simulate explicit dynamic vegetation. Three Vegetation carbon pools (leaf, stem, and root) are dynamically regulated and have constant turnover rates to three litter and three soil pools. 12 vegetation types are considered. A daily timestep is used for the land ecosystem and land biogeochemistry.

ORCHIDEE

The ORganizing Carbon and Hydrology in Dynamic EcosystEms (ORCHIDEE, Boucher et al. 2020; Krinner et al. 2005) land model is used in the IPSL models, version 1 for the CMIP5 models IPSL-CM5A-LR, IPSL-CM5A-MR, and version 2 in the CMIP6 model IPSL-CM5B-LR. The model considers 15 PFTs, as well as 8 vegetation carbon, 4 litter carbon and 3 soil carbon pools. Plant and soil carbon fluxes are computed every 15 min, the same as the atmospheric physics timestep, while slow processes like soil and litter carbon dynamics are computed daily instead. The CMIP5 model used a two-layer bucket model for its soil hydrology, while in CMIP6 an 11-layer soil hydrology scheme is employed. Photosynthesis is parametrized based on the common Farquhar and Collatz schemes for C3 and C4 respectively. Nutrient limitation in CMIP6 is introduced through downregulation using a logarithmic function of the CO_2 concentration.

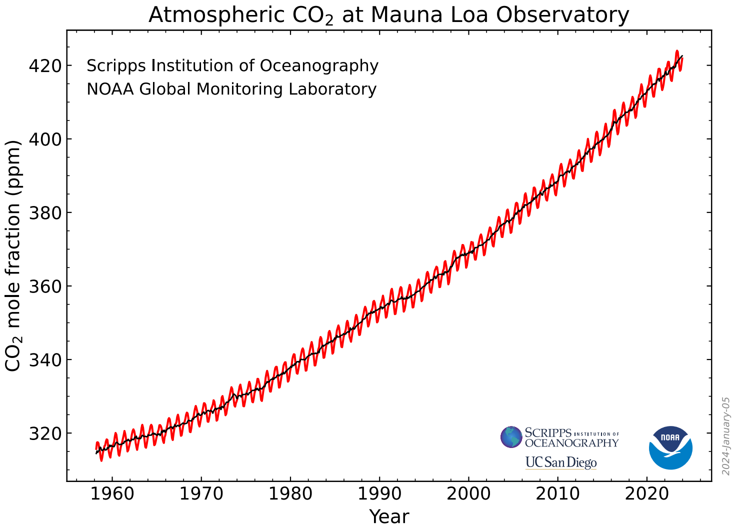


Figure 2.7.: Monthly mean atmospheric CO_2 concentration at Mauna Loa, Hawaii. Adapted with permission from NOAA 2005.

INMCM

The carbon cycle module for INMCM (Volodin 2007) includes a single soil carbon pool. The most important changes with respect to INMCM4 for the CMIP6 models lie in the atmospheric component of the model, as well as some upgrades to the oceanic component, but no changes to the carbon cycle (Volodin et al. 2017b).

2.4. Observations for carbon cycle variables

To evaluate the present-day performance of climate models, observations of the current climate are indispensable. In situ ground-based measurements of CO_2 first started in 1958 in Mauna Loa (Hawaii, USA). They provided the first evidence of an increase in atmospheric CO_2 concentrations due to fossil fuel combustion (Keeling et al. 1976), and these measurements persist to this day. This is the so-called Keeling Curve shown in Figure 2.7, which illustrates the rising CO_2 concentration in the atmosphere since the beginning of the measurements superimposed by the seasonal cycle due to carbon uptake and release by plants. Since then many further observatories around the globe have started measuring atmospheric CO_2 and other gases. However, most of these observatories are concentrated in the US and Europe, with only sparse coverage in the rest of the world, such as the tropics.

While these ground-based measurements now cover a relatively long time period, they are only point based and cannot accurately show regional and spatial distributions of CO_2. Satellite measurements of CO_2 complementing these ground-based measurements started in

2002 with near infrared (NIR)/short-wave infrared (SWIR) nadir-based (downward-looking) satellite retrievals (Buchwitz et al. 2005). Unlike in situ measurements, satellite instruments measure the column-average CO_2 mole fraction (XCO_2), a dimensionless quantity obtained by dividing the vertical column of CO_2 by the vertical column of dry air. Details on the calculation of XCO_2 are given in Section 2.4.1. To obtain surface fluxes from XCO_2 measurements, they are used as inputs for inverse modeling of atmospheric transports (Basu et al. 2013; Chevallier et al. 2014; Houweling et al. 2015; Reuter et al. 2014). Alternatively, some models assimilate observational data to constrain process parameters, e.g. CCDAS (Kaminski et al. 2013), while some skepticism about the reliability of satellite observations remains, as XCO_2 estimates can differ vastly from inversion experiments in the tropics (Crisp et al. 2022).

Satellite measurements are also important for other carbon cycle variables. The only direct way to measure LAI is to harvest leaves and measure the area of each leaf. This is not feasible for tall forest canopies or for large scale measurements. As such, the only reliable means for a global spatiotemporally continuous LAI data set is through remote sensing (Cao et al. 2023), estimating LAI through the amount of light transmitted by a plant canopy. Commonly this data is used alongside some regional ground-based measurements to feed machine learning algorithms to arrive at a final data product. As such, while many LAI reference data sets use the same underlying satellite measurements, their resulting data products can look different. Another example is GPP, which can be measured by Solar-induced fluorescence (SIF), which is directly correlated to GPP. SIF can be obtained through the oxygen-A band observed by GHG satellites, enabling them to measure GPP alongside GHG concentrations (Sellers et al. 2018).

The specific reference datasets used in this study to analyze CMIP model output are discussed in the relevant chapters in more detail.

2.4.1. Computation of XCO_2

Model simulations provide CO_2 as a 3D variable on a latitude-longitude grid with a vertical pressure coordinate. The calculation from this 3D CO_2 to the 2D XCO_2 coordinate for comparison with satellite XCO_2 data as published in the Appendix of **Gier** et al. 2020, following the description in Buchwitz and Reuter 2016, is reproduced here:

$$XCO_2 = \frac{\sum n_\mathrm{d} \cdot c_{CO_2}}{\sum n_\mathrm{d}} \tag{2.6}$$

Here, c_{CO_2} represents the modeled CO_2 dry-air mole fraction on model layers (i.e., layer centers or full levels) and n_d the number of dry-air particles (air molecules excluding water vapor) within these levels. The summations are performed over all model layers. The number of dry-air particles can be computed as follows:

$$n_\mathrm{d} = \frac{N_\mathrm{a} \cdot \Delta p \cdot (1 - q)}{m_\mathrm{d} \cdot g}. \tag{2.7}$$

N_a is the Avogadro constant ($6.022140857 \times 10^{23}\,\mathrm{mol}^{-1}$) and m_d the molar mass of dry air ($28.9644 \times 10^{-3}\,\mathrm{kg\,mol}^{-1}$). Δp is the pressure difference (in hPa) computed from the model's pressure levels (i.e., layer boundaries or half levels) surrounding the model layers, q is the modeled specific humidity (in kg/kg), and g is the gravitational acceleration approximated by

$$g = \sqrt{g_0^2 - 2 \cdot f \cdot \phi}. \tag{2.8}$$

This includes the model's geopotential ϕ (in $\mathrm{m}^2\,\mathrm{s}^{-2}$) on layers, the free air correction constant $f = 3.0825959 \times 10^{-6}\,\mathrm{s}^{-2}$ and the gravitational acceleration g_0 on the geoid approximated by the international gravity formula depending only on the latitude φ:

$$\begin{aligned} g_0 = 9.780327 \cdot \Big[& 1 + 0.0053024 \cdot \sin^2(\varphi) \\ & - 0.0000058 \cdot \sin^2(2\varphi) \Big]. \end{aligned} \tag{2.9}$$

3. Routine evaluation of climate models

With the growing number of climate models and climate model output made available within CMIP, with an expected total data volume of 20 PB for CMIP6 compared to the 2 PB of CMIP5 (Petrie et al. 2021), routine evaluation of these models becomes increasingly important for baseline analysis. This was also discussed in a workshop on *Earth System Model Evaluation to Improve Process Understanding* hosted by the Aspen Global Change Institute the author participated in, which resulted in a perspective paper (Eyring et al. 2019) to summarize the discussions and findings the participants arrived at. It stresses the importance of readily available climate model evaluation tools to facilitate a more rapid and comprehensive evaluation of model simulations. A software facilitating such routine analysis is the Earth System Model Evaluation Tool (ESMValTool) (https://www.esmvaltool.org), an open-source community-developed diagnostics and performance metrics tool. This chapter introduces the structure of ESMValTool in Section 3.1, followed by the author's contributions to its development not documented in other chapters in Section 3.2. The author also used ESMValTool for contributions to the latest IPCC AR6, which are detailed in Section 3.3.

3.1. Earth System Model Evaluation Tool (ESMValTool)

After its first release in 2016 (Eyring et al. 2016b), ESMValTool has undergone constant development for new diagnostics, new preprocessing operations, and enhanced performance capabilities. By attaching detailed provenance information to all output, all results produced by the tool are fully traceable and reproducible. A schematic representation of the structure of ESMValTool is shown in Figure 3.1. With the introduction of version 2, ESMValTool has been split into two parts. The first part is the core functionalities of the ESMValTool (ESMValCore) (Righi et al. 2020), which contains Python-based core functionalities for preprocessing the data according to the users' needs, as well as ensuring the technical standards of input data. After loading the data, checks and fixes are performed to ensure that the input data are compliant to the Climate Model Output Rewriter (CMOR) format, a common standard expected from all models participating in CMIP. ESMValTool uses this standard for all its data processing. Optional preprocessing operations provided by ESMValCore include spatial and temporal subsetting or regridding, masking, multi-model statistics, temporal or spatial statistics, or unit conversions. Furthermore, ESMValCore can derive custom non-CMOR variables from the input data. Since ESMValCore conveniently collects commonly used preprocessing operations in one location, it can also be used as a standalone package for preprocessing data.

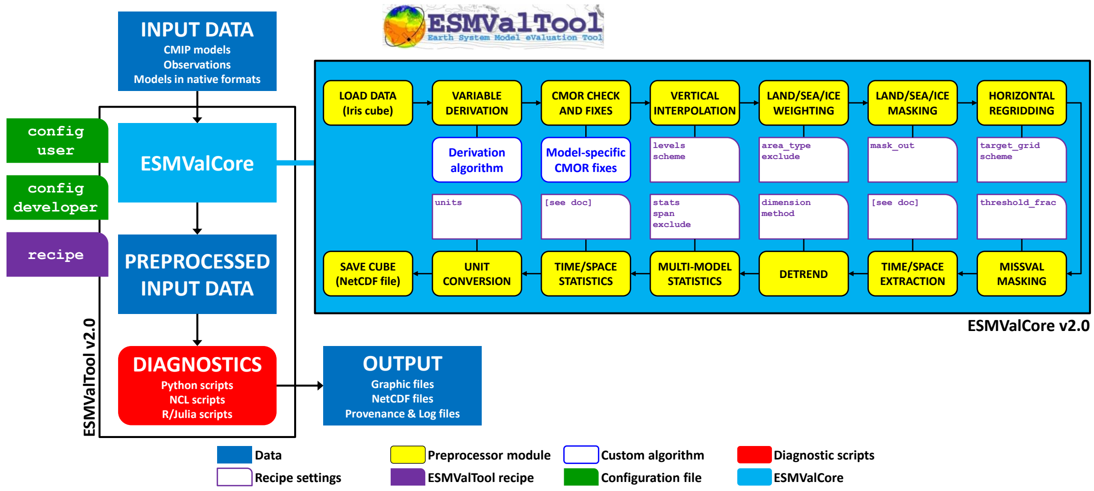


Figure 3.1.: Schematic representation of the system architecture of Earth System Model Evaluation Tool (ESMValTool) version 2. Input Data are preprocessed by core functionalities of the ESMValTool (ESMValCore), which includes preprocessor operations for regridding, interpolation, data fixing, multi-model statistics and more. These preprocessed data are fed into diagnostic scripts, which create the final output files, most commonly plots and data in a Network Common Data Form (NetCDF) format, for which full provenance are given. The configuration is done through the recipe, where the user specifies the input datasets, the preprocessing functions and the diagnostic scripts to be used. Adapted with permission from Righi et al. 2020.

The second part is ESMValTool, which uses the preprocessed data from ESMValCore to run so-called diagnostics defined in recipes. These recipes are the main control files of ESMValTool, and include the input datasets, the preprocessing functions and the diagnostic scripts used. Diagnostic scripts, which may be written in Python, NCL, R or Julia, are responsible for running scientific evaluations on the preprocessed data sets. There are many different diagnostics covering a wide range of scientific topics. These include large-scale diagnostics for quasi-operational and comprehensive evaluation of ESMs (Eyring et al. 2020), diagnostics for emergent constraints and analysis of future projections (Lauer et al. 2020), and diagnostics for extreme events, regional model and impact evaluation and analysis (Weigel et al. 2021). Outputs of these diagnostics are usually figures and corresponding NetCDF files, as well as log files containing provenance information. To use observations and reanalysis data in ESMValTool which do not adhere to the CMOR standards, cmorizer scripts are included in the software package to reformat these data into the CMOR standard. ESMValTool is developed open-source on the GitHub repositories at https://github.com/ESMValGroup. It is released under the Apache License, version 2.0, with the latest release publicly available on Zenodo (Andela et al. 2023b; Andela et al. 2023a).

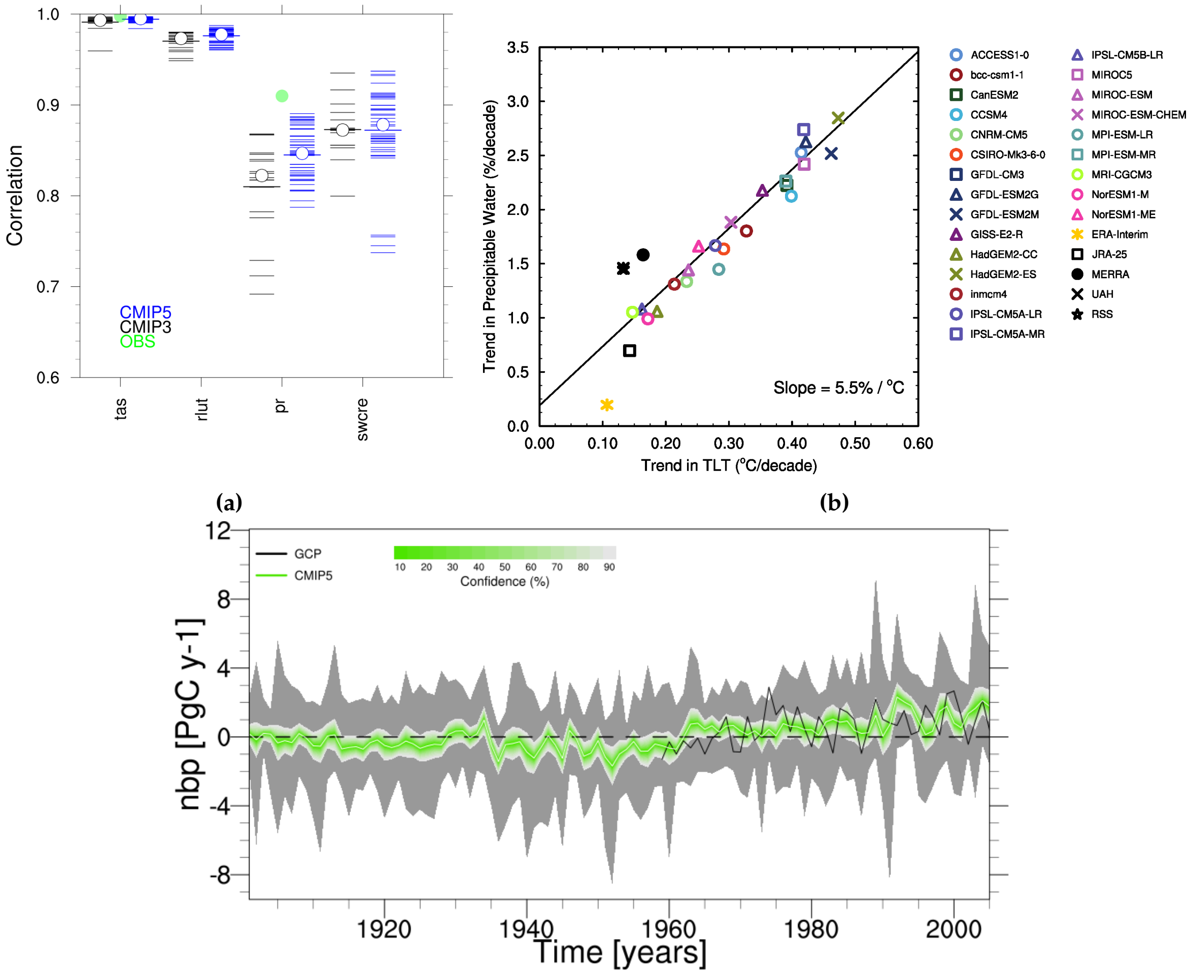
3.2. Contributions to ESMValTool

Apart from the author's main studies presented in Chapters 4 and 5, several contributions have been made to ESMValTool, partly covered in scientific documentation papers (Eyring et al. 2020; Lauer et al. 2020; Weigel et al. 2021). Additionally, the author of this thesis is part of the scientific lead development team of ESMValTool.

ESMValTool includes recipes to reproduce analyses from international climate assessments. One of these is the group of *recipe_flato_13ipcc_figure_*.yml* recipes in the *ipccwg1ar5ch9* folder, covering a large part of the figures from the climate model evaluation of Chapter 9 of the IPCC's AR5 (Flato et al. 2013). The author of this thesis implemented the three figures depicted in Figure 3.2.

Firstly, Figure 9.6 of Flato et al. 2013 shows the centred pattern correlations, which measure the similarity of two patterns between models and observations for the annual mean climatology over the period 1980–1999 (after removing the global mean). While the figure was coded to be able to be applied for several different projects and variables, the original figure analyzed the CMIP3 and CMIP5 ensemble for the near-surface air temperature (*tas*), top of the atmosphere outgoing longwave radiation (*rlut*), precipitation (*pr*), and top of the atmosphere shortwave cloud radioactive effect (*swcre*). Improvements in all variables can be seen for the CMIP5 ensemble, with *tas* and *rlut* showing the highest correlation, while *pr* and *swcre* show a large model spread in both ensembles. This figure has also been used for the new IPCC AR6 Chapter 3 (Eyring et al. 2021) Figure 3.43 with small modifications. Secondly, Figure 9.9 of Flato et al. 2013 was implemented in version 1 of ESMValTool. It shows a scatterplot for the decadal trends of CMIP5 models for precipitable water as a function of the decadal trend in lower tropospheric temperature (TLT), with a high correlation between the two variables. Due to complications of porting the variable derivation of TLT in the initial version of ESMValTool version 2, as well as the non-widespread use of this variable, this figure has not been ported to the version 2 of ESMValTool. Lastly, Figure 9.26 of Flato et al. 2013 was added by the author to this recipe, as well as *recipe_anav13jclim.yml*, which is the original source of this figure. It depicts the global land carbon uptake for CMIP5 models between 1900–2005 in comparison with data from the Global Carbon Project (GCP; Le Quéré et al. 2009). Models and observations agree well with each other and show large interannual variations with an increasingly higher land carbon uptake since the mid 20th century. This is discussed in more detail in Chapter 5.

Another contribution included work on the implementation of the *recipe_smpi.yml*, as well as its porting to version 2 of ESMValTool (Figure 3.3). This figure is based on the Single Model Performance Index (SMPI) introduced in Reichler and Kim 2008. The SMPI combines different climate variables of atmospheric, surface, and oceanic origin to compute a single normalized index signifying the overall performance of a model. Smaller SMPI values correspond to better-performing models, while a value of 1 marks the average performance. The radius of the circle shown in the plot is the 95 % confidence interval of the index. The SMPI allows a quick estimation the best-performing models across all considered variables.



(a) **(b)**

(c)

Figure 3.2.: Figures contributed to *recipe_flato13ipcc.yml* which reproduces select figures from Flato et al. 2013. **(a)** Figure 9.6: Centred pattern correlations between models and observations calculated for the annual mean climatology over the period 1980–1999. The colors indicate the different CMIP projects and observations, with individual models shown as a dash. The thicker dash represents the ensemble mean and the open circle the ensemble median. Depicted are the four variables global mean near-surface temperature at 2m (*tas*), top of the atmosphere (TOA) outgoing longwave radiation (*rlut*), precipitation (*pr*) and TOA shortwave cloud radiative effect (*swcre*). **(b)** Figure 9.9: Scatterplot showing the decadal trends of CMIP5 models for precipitable water as a function of the decadal trend in lower tropospheric temperature (TLT) over the world's oceans in the tropics (20 °S to 20 °N). The trends are calculated over the 1988–2012 period, with CMIP5 models extended with the RCP8.5 simulations. This figure was only included in version 1 of ESMValTool in the ESMValTool-private repository. **(c)** Figure 9.26: Global Land carbon uptake for CMIP5 models for 1900–2005 against observations from the Global Carbon Project (GCP; Le Quéré et al. 2009). The white line represents the multi-model mean, with the colored shading corresponding to the confidence limits of the ensemble mean assuming a *t*-distribution, and the grey area representing the total range of the model ensemble.

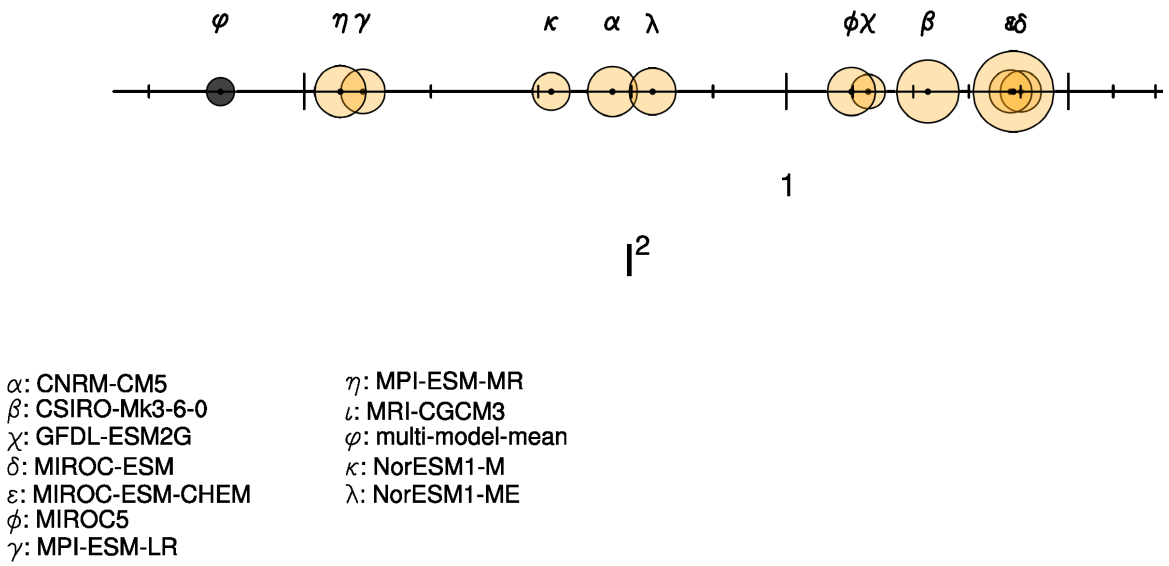


Figure 3.3.: Single Model Performance Index (SMPI) I^2 for individual models (circles), and the multi-model mean (black circle). The size of the circles indicate the length of the 95 % confidence intervals. Implemented in ESMValTool (*recipe_smpi.yml*) based on the SMPI introduced by Reichler and Kim 2008.

3.3. Contributions to the IPCC AR6

The author of this thesis had the honor of contributing to two chapters of the IPCC AR6. The contribution to Chapter 3 (Eyring et al. 2021) is in the form of an acknowledged graphic developer for Figure 3.31 (Figure 3.4) and 3.32 (Figure 3.5). Figure 3.31 (Figure 3.4) shows the evaluation of historical emission-driven CMIP6 simulations for 1850–2014 with observational data from the GCP (Friedlingstein et al. 2019). Panel (a) shows the atmospheric CO_2 concentration of models, which are very similar to observations from NOAA (Dlugokencky and Tans 2020). Panel (b) shows the near-surface air temperature anomaly with respect to the 1850–1900 mean compared to observations from HadCRUT4 (Morice et al. 2012). This panel additionally shows the concentration driven simulations of the same models in dashed lines. Models in both concentration- and emission-driven simulations show similar performance, and both models and observations show an increasing temperature, reaching an average of 1 °C increase in 2014. Panels (c) and (d) show the global land and ocean carbon uptake, respectively, against observations from the GCP (Friedlingstein et al. 2019). These panels employ a 10-year running mean for better visibility. The ocean uptake is further offset to be 0 in 1850 to correct for pre-industrial riverine-induced carbon fluxes. These panels show an increasing carbon uptake by both land and ocean in response to the rising atmospheric CO_2 concentration in panel (a). Both are in agreement with the observations, while the model spread and interannual variability are larger over land than over the ocean.

Figure 3.32 (Figure 3.5) shows the relative change in the seasonal cycle amplitude (SCA) of global land carbon uptake for CMIP6 simulations from 1961–2014 compared to CO_2 observations and NBP estimations from JMA-TRANSCOM, as well as the seasonal cycle of the

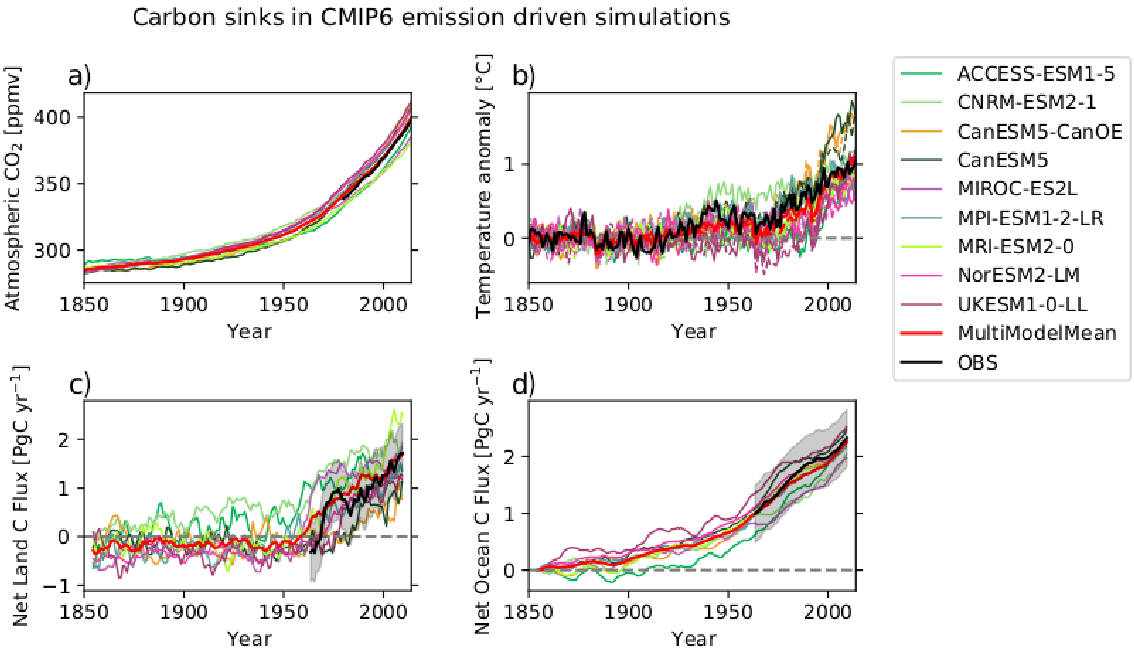


Figure 3.4.: *IPCC AR6 Figure 3.31*: Evaluation of historical emission-driven CMIP6 simulations for 1850-–2014. Observations (black) are compared to simulations of global mean **(a)** carbon dioxide concentration (ppmv), with observations from the National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratories (ESRL) (Dlugokencky and Tans 2020); **(b)** surface air temperature anomaly (°C) with respect to the 1850-–1900 mean, with observations from HadCRUT4 (Morice et al. 2012); **(c)** land carbon uptake (PgC yr^{-1}); and **(d)** ocean carbon uptake (PgC yr^{-1}), both with observations from the Global Carbon Project (GCP; Friedlingstein et al. 2019) and grey shading indicating the observational uncertainty. Land and ocean carbon uptakes are plotted using a 10-year running mean for better visibility. The ocean uptake is offset to 0 in 1850 to correct for pre-industrial riverine-induced carbon fluxes. Reproduced with permission from Eyring et al. 2021.

models for the first (orange) and last (green) ten years in the inset. For comparison between the NBP and CO$_2$ data, the relative change is offset to the 1961–1970 mean for long timeseries and the last ten years for shorter timeseries. It is evident that the SCA has increased over time, but the specific reasons for this increase are still debated. The code for the aforementioned ESMValTool IPCC AR6 Chapter 3 figures is publicly available on GitHub in this branch: https://github.com/ESMValGroup/ESMValTool-AR6-OriginalCode-FinalFigures/tree/ar6_chapter_3_tina.

Furthermore, the author of this thesis contributed to Chapter 5 of the IPCC AR6 (Canadell et al. 2021) as contributing author. In this chapter, the author worked in close collaboration with Peter Cox (University of Exeter, UK) Prabir Patra (JAMSTEC, Japan) on Figures 5.23 (Section 3.3 top), 5.24 (Section 3.3 bottom) and 5.25 (Figure 3.7) by preprocessing CMIP6 model data with ESMValTool, and formatting the output in a requested format ready for plotting. This included checking the data for errors and viability, as well as giving feedback on the final figures produced using these data.

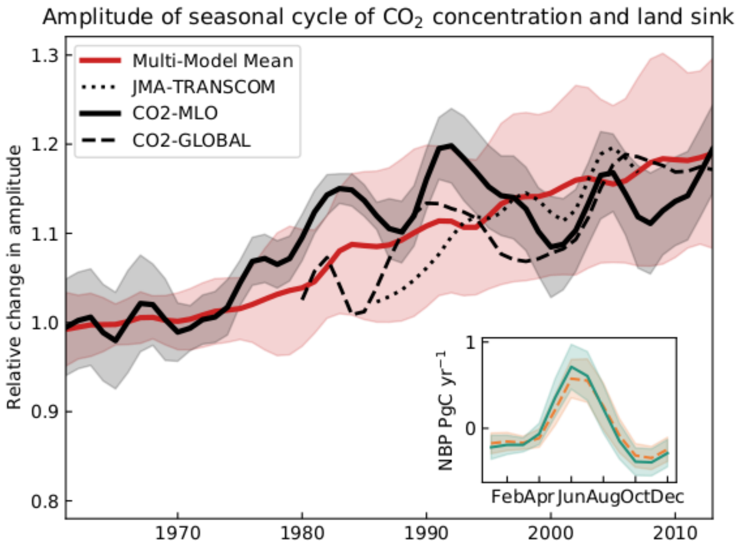


Figure 3.5.: *IPCC AR6 Figure 3.32*: Relative change in the amplitude of the seasonal cycle of global land carbon uptake in the historical CMIP6 simulations from 1961-–2014. Net biosphere production estimates from 19 CMIP6 models (red), the data-led reconstruction JMA-TRANSCOM (Maki et al. 2010; dotted) and atmospheric CO_2 seasonal cycle amplitude changes from observations (global as dashed line, Mauna Loa Observatory (MLO) (Dlugokencky et al. 2018) in bold black). Seasonal cycle amplitude is calculated using the curve fit algorithm package from the National Oceanic and Atmospheric Administration Earth System Research Laboratory (NOAA ESRL). Relative changes are referenced to the 1961—1970 mean and for short time series adjusted to have the same mean as the model ensemble in the last 10 years. Interannual variation was removed with a nine-year Gaussian smoothing. Shaded areas show the one sigma model spread (grey) for the CMIP6 ensemble and the one sigma standard deviation of the smoothing (red) for the CO_2 MLO observations. Inset: average seasonal cycle of ensemble mean net biosphere production and its one sigma model spread for 1961—1970 (orange dashed line, light orange shading) and 2005-–2014 (solid green line, green shading). Reproduced with permission from Eyring et al. 2021.

Figure 5.23 (Section 3.3 top) shows the changes in ocean (a) and land (b) carbon storage against data from the GCP (Friedlingstein et al. 2019) similar to Figure 3.31 but for the cumulative global sum and for concentration driven simulations. Similar to Figure 3.31, the models fit well to the observations, with a larger range in the land carbon storage than the ocean. In the report this range in the land carbon uptake is attributed in part to land-use change. Figure 5.24 (Section 3.3 bottom) shows the mean uptake 2000–2009 in a zonal distribution for both CMIP5 and CMIP6 models, compared to atmospheric inversion estimates. While the ocean carbon sink is well simulated by the CMIP models, the land carbon uptake shows large underestimations by the models in high latitudes and large uncertainties in the tropics. This is also found in the analysis in Chapter 5 and analyzed in more detail there.

Figure 5.25 (Figure 3.7) shows the evolution of the land and ocean carbon sinks for several different future projections in concentration driven simulations. As the land and ocean carbon uptakes are primarily influenced by atmospheric CO_2, the evolution of the sinks varies greatly

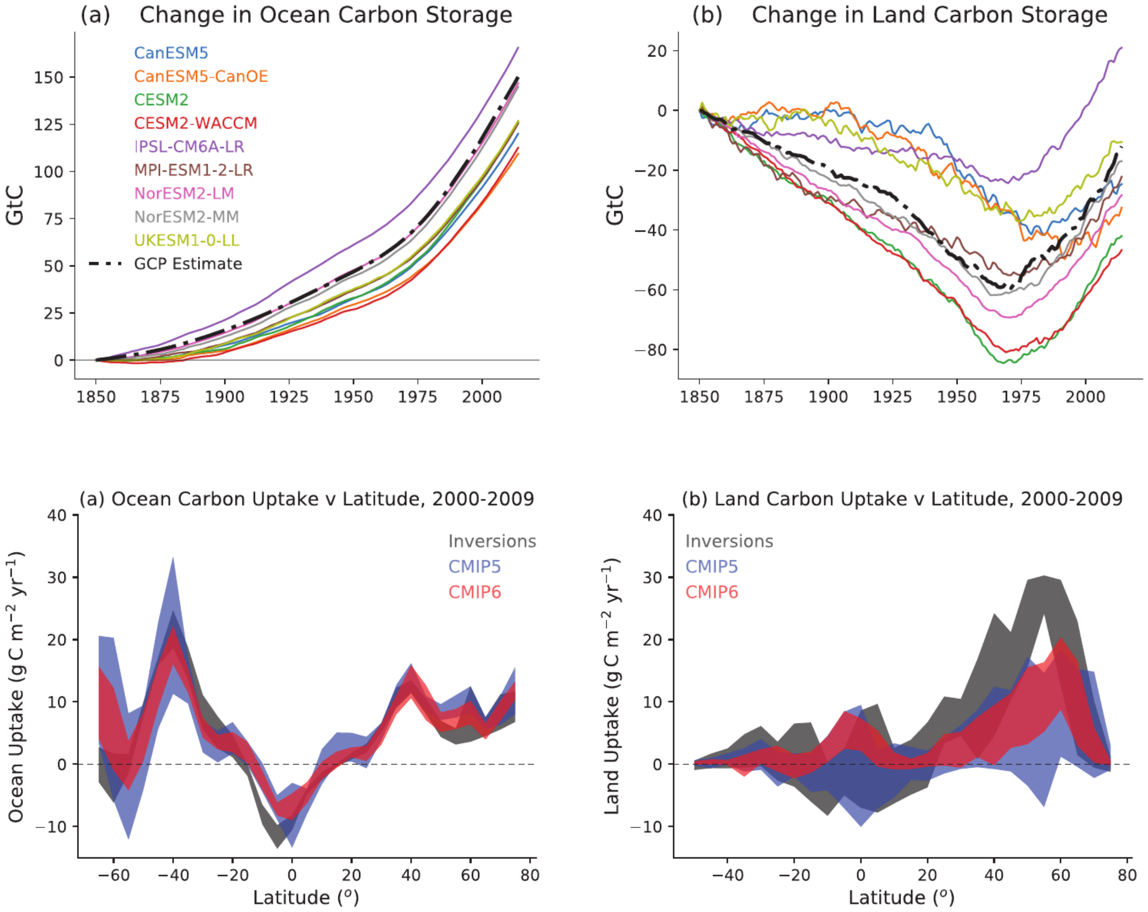


Figure 3.6.: Figures from IPCC AR6 Chapter 5 for which historical data were provided after preprocessing with Earth System Model Evaluation Tool (ESMValTool). **Top**: *IPCC AR6 Figure 5.23*. CMIP6 Earth system model concentration-driven historical simulations for 1850 to 2014, compared to observation-based estimates from the global carbon project (GCP). **(a)** Cumulative ocean carbon uptake from 1850 (PgC); **(b)** cumulative land carbon uptake from 1850 (PgC). Only models that simulate both land and ocean carbon fluxes are shown here. **Bottom**: *IPCC AR6 Figure 5.24*. Comparison of modelled zonal distribution of contemporary carbon sinks against atmospheric inversion estimates for 2000-–2009: **(a)** ocean carbon uptake; **(b)** net land uptake. Latitude runs from 90 °S (i.e., –90 °N) to 90 °N. Positive uptake represents a carbon sink to ocean/land while negative uptake represents a carbon source. The land uptake is taken as net biome productivity (NBP) and so includes net land-use change emissions. The bands show the mean ± 1 standard deviation across the available inversions (black bands, 3 models), CMIP5 ESMs (blue bands, 12 models for the ocean, 12 models for the land), and CMIP6 ESMs (red bands, 11 models for ocean, 10 models for land). Reproduced with permission from Canadell et al. 2021.

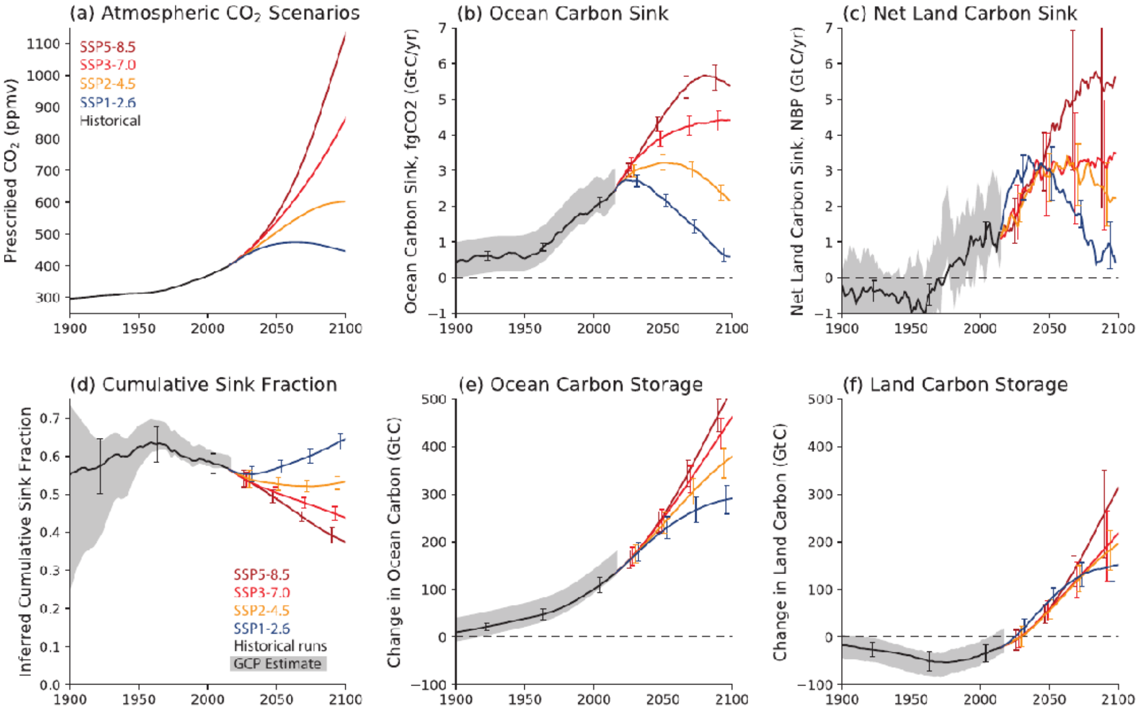


Figure 3.7.: *IPCC AR6 Figure 5.25*: Modelled evolution of the global land and ocean carbon sinks from 1900 to 2100 in concentration-driven CMIP6 Earth system model (ESM) scenario runs. (SSP1-2.6: blue; SSP2-4.5: orange; SSP3-7.0: red; SSP5-8.5: brown): **(a)** prescribed atmospheric CO_2 concentrations; **(b)** five-year running mean ocean carbon sink (GtC yr^{-1}); **(c)** five-year running mean net land carbon sink (GtC yr^{-1}); **(d)** inferred cumulative sink fraction of emissions from 1850; **(e)** change in ocean carbon storage from 1850 (GtC); **(f)** change in land carbon storage from 1850 (GtC). Thick lines represent the ensemble mean of the listed ESM runs, and the error bars represents ± 1 standard deviation about that mean. The grey wedges represent estimates from the Global Carbon Project (GCP), assuming uncertainties in the annual mean ocean and net land carbon sinks of 0.5 GtC yr^{-1} and 1 GtC yr^{-1} respectively, and uncertainties in the changes in carbon stores (ocean, land and cumulative total emissions) of 25 GtC. The net land carbon sink is taken as net biome productivity (NBP) and so includes any modelled net land-use change emissions. Reproduced with permission from Canadell et al. 2021.

across the different SSP scenarios. Simulations with larger CO_2 increases lead to larger carbon and land sinks, but lower sink fractions. Again the land shows a much larger uncertainty in the models than the ocean, evident in the larger bars representing the standard deviation.

4. Spatially resolved evaluation of ESMs with satellite column-averaged carbon dioxide

CO_2 is the most important anthropogenic GHG. The uncertain remaining budget of CO_2 emissions to reach specific temperature targets under global warming stresses the need for models to correctly simulate atmospheric CO_2 concentrations. As model projections cannot be tested for their reliability through experiments, they are instead analyzed for their present-day performance using observations of the earth system. While most long-term CO_2 observations are from ground-based stations which report an increase of about 45 % since pre-industrial times (Ciais et al. 2013), they are spatially sparse. Satellite measurements instead can enable spatially resolved evaluations of ESMs. This chapter addresses the research questions of how recent satellite observations can be used to evaluate climate model simulations, and how sparse data coverage can affect the results by comparing unsampled model data to model data sampled in the same way as the satellite observations. Furthermore, on the basis of the satellite XCO_2 data, output from CMIP5 and CMIP6 model simulations is compared to analyze the influence of the increased process complexity of the newer model generation. The satellite data used in this chapter are introduced in Buchwitz et al. 2018, to which the author contributed in discussions about computation methods. These discussions helped the author familiarize with the data set.

This chapter is based on an already published study (**Gier** et al. 2020). The author of this thesis led the writing and analysis of the paper, and also developed the diagnostics which have since been included in the ESMValTool in the recipe *recipe_gier2020bg.yml* for full reproducibility of all figures. Section 4.1 introduces the data products, models and methods used in this study. Satellite XCO_2 measurements are briefly compared to CO_2 flask measurements and model simulations in Section 4.2, while Section 4.3 evaluates CMIP5 and CMIP6 simulations with XCO_2 satellite data, split into individual sections focusing on the time series, growth rate (GR), and SCA, respectively. The section on SCA additionally includes a part detailing the influence of the observational sampling. Finally, Section 4.4 summarizes and concludes this study.

4.1. Data and Methods

4.1.1. Observational datasets

Satellite XCO_2

We use the Observations for Model Intercomparisons Project (obs4MIPs) version 3 (O4Mv3) XCO_2 satellite data (Buchwitz et al. 2017b; Buchwitz et al. 2018). obs4MIPs hosts observationally based datasets which have been formatted according to the CMIP model output requirements (e.g., variable definitions, coordinates, frequencies) in order to facilitate an easier comparison between observations and models (Ferraro et al. 2015; Teixeira et al. 2014; Waliser et al. 2020). The satellite product used here is a gridded (level-3) monthly data product with a $5° \times 5°$ spatial resolution following the obs4MIPs format, produced as part of the Copernicus Climate Change Service (C3S). The O4Mv3 product is retrieved from two satellite instruments: Scanning Imaging Absorption Spectrometer for Atmospheric CHartographY (SCIAMACHY)/Envisat (Bovensmann et al. 1999; Burrows et al. 1995) and the Thermal And Near infrared Sensor for carbon Observation Fourier transform spectrometer (TANSO-FTS)/Greenhouse Gases Observing Satellite (GOSAT) (Kuze et al. 2009).

This monthly mean XCO_2 satellite dataset covers a 14-year time span (2003–2016). It is obtained by gridding the level-2 product (individual soundings) generated with the ensemble median algorithm (EMMA) (Reuter et al. 2013), in this case EMMA version 3.0 (EMMAv3; Reuter et al. 2017). EMMA combines several different XCO_2 level-2 satellite data products from SCIAMACHY/Envisat (2003–2012) and TANSO-FTS/GOSAT (2009–2016) and includes a bias correction to all products during overlap phases, resulting in a good agreement during the overlap period. This product was validated against Total Carbon Column Observing Network (TCCON) (Wunch et al. 2011) ground-based observations of XCO_2, revealing a +0.23 ppmv global bias, a relative accuracy (defined as standard deviation of the station-to-station biases) of 0.3 ppmv and a very good stability in terms of a linear bias trend ($-0.02 \pm 0.04\,\mathrm{ppmv\,yr^{-1}}$) (Buchwitz et al. 2017a). While the dataset ends in 2016, our evaluation only goes up to the year 2014 because the historical simulations for CMIP6 end in 2014 and scenarios from the emission-driven simulations that could be used to extend the runs were not yet available at the time of the study for all considered models.

The number of observations depends significantly on the location with most points over locations with low cloud cover, high surface reflectivity and (at least) moderate to high Sun elevation. Coverage over ocean is sparse as ocean retrievals are only included from GOSAT Sun-glint mode observations – outside of glint conditions, the reflectivity of water is very low in the NIR/SWIR spectral region. Figure 4.1 shows the mean monthly coverage of the dataset for 2003–2014. In Section 4.3, we will show that taking into account this sampling in the evaluation of ESMs is essential for a proper comparison.

The dataset also contains uncertainty estimates for each grid cell, with a mean value of 0.92 ppmv, accounting for both statistical uncertainties from the individual soundings and

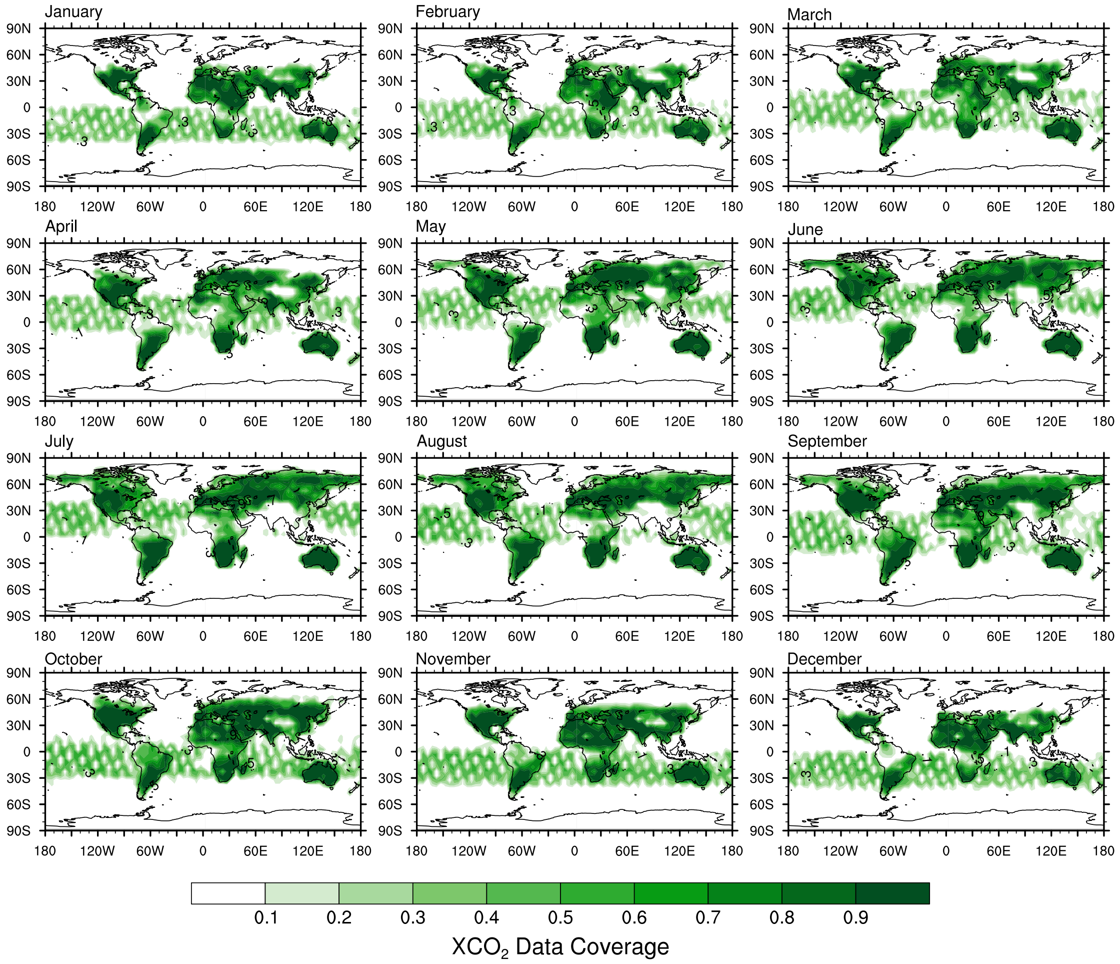


Figure 4.1.: Mean fractional coverage of monthly XCO_2 satellite data for 2003—2014. A value of 0 (white) signifies no available data, while a value of 1 (dark green) means that this grid cell contains data for all years of this month. Adapted with permission from **Gier** et al. 2020.

uncertainties from potential regional and temporal biases (Buchwitz et al. 2017b). However, the overall uncertainties are small compared to inter-model differences (see Section 4.1.3) and are therefore neglected in our analysis.

Surface CO_2 measurements

For the comparison of satellite XCO_2 and surface CO_2 data in Section 4.2, we have obtained surface flask measurements from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network (Dlugokencky et al. 2020). Measurement sites at locations with no available satellite data were excluded from the analysis, which ruled out the four baseline observatories in Mauna Loa and Samoa, as well as the South Pole and Point Barrow sites. Furthermore, sites which did not collect data during the period from 2003–2014 were discarded. From the remaining sites, a sample of five sites was chosen which had the best coverage of different latitudes, and when latitudes were similar, different longitudes were selected for increased spatial coverage. The selected sites are listed in Table 4.1.

Code	Location	Latitude [°]	Longitude [°]	Altitude [m]	Start year
ASK	Assekrem, Algeria	23.2625	5.6322	2710	1995
CGO	Cape Grim, Australia	−40.6800	144.6800	94	1984
LEF	Park Falls, United States	45-945	269.7300	868	1994
HUN	Hegyhátsál, Hungary	46.950	16.650	248	1993
WIS	Ketura, Israel	30.8595	34.7809	482	1995

Table 4.1.: List of active NOAA surface flask measurement sites used in this study.

4.1.2. Model simulations

We use monthly mean output data from 10 CMIP5 and 10 CMIP6 models which performed emission-driven simulations, with two of the CMIP5 and five of the CMIP6 models including a nitrogen cycle. Table 4.3 and Table 4.2 list all the CMIP5 and CMIP6 models used in this paper along with their atmosphere, land and ocean model component, respectively. Only models with an interactive carbon cycle are able to perform the emission-driven simulations, in which the emissions rather than the concentrations of the greenhouse gases are prescribed (Eyring et al. 2016c; Taylor et al. 2012). This allows the carbon cycle in the models to react to changes in climate and atmospheric CO_2 by adjusting their carbon fluxes to the new climate conditions and providing the atmospheric CO_2 concentration as an output (Friedlingstein et al. 2014). In order to facilitate the comparison between the satellite data and the CMIP5 emission-driven simulations, the historical simulations (1850–2005) were extended beyond 2005 with simulations from Representative Concentration Pathway (RCP) 8.5 (2006–2100), for which most ESM simulations are available. Since the period of observations only extends a decade beyond the historical runs, the choice of emissions scenario has a negligible impact on the results that we present below. For CMIP6, only the historical simulations are used, which end in 2014. For CMIP5, only one model had more than one ensemble member performing the emission-driven RCP8.5 simulation, and thus only one ensemble member for each model has been used. In CMIP6, several models have three or more ensemble members. We consider all of them in Figure 4.3 for the time series to show the models' intrinsic variability but then proceed with the analysis with only the first ensemble member for each model. The different initial value ensemble members perform similarly to each other for the analysis presented in this paper, and using an ensemble mean would reduce the interannual variability found in each individual member.

Table 4.2.: CMIP5 models analyzed in this study. D marks models including dynamic vegetation, and N marks models including nitrogen cycles.

Model	Institute	Atmosphere model	Land model	Ocean model	Comment	Main reference
BNU-ESM	College of Global Change and Earth System Science, China	CAM3.5	CoLM + BNU-DGVM	MOM4p1 + IBGC	D	Ji et al. 2014
CanESM2	Canadian Center for Climate Modeling and Analysis, BC, Canada	CanAM4	CLASS2.7 + CTEM1	CMOC		Arora et al. 2011
CESM1-BGC	National Center for Atmospheric Research Boulder, CO, USA	CAM4	CLM4	POP2 + BEC	N	Gent et al. 2011; Lindsay et al. 2014
FIO-ESM	The First Institute of Oceanography, SOA, China	CAM3.0	CLM3.5 + CASA	POP2.0 + OCMIP-2		Bao et al. 2012; Qiao et al. 2013
GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory, United States	AM2	LM3.0	GOLD + TOPAZ2	D	Dunne et al. 2012; Dunne et al. 2013
GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory, United States	AM2	LM3.0	MOM4.1 + TOPAZ2	D	Dunne et al. 2012; Dunne et al. 2013
MIROC-ESM	MIROC, Japan	MIROC-AGCM + SPRINTARS	MATSIRO + SEIB-DGVM	COCO3.4 + NPZD	D	Watanabe et al. 2011
MPI-ESM-LR	Max Planck Institute for Meteorology, Hamburg, Germany	ECHAM6	JSBACH + BETHY	MPIOM + HAMOCC5	D	Giorgetta et al. 2013

MRI-ESM1	Meteorological Research Institute, Japan	MRI-AGCM3.3 + MASINGAR mk-2 + MRI-CCM2	HAL	MRI.COM3	D	Adachi et al. 2013; Yukimoto et al. 2011; Yukimoto et al. 2012
NorESM1-ME	Norwegian Climate Center, Norway	CAM4-Oslo	CLM4	HAMOCC5	N	Tjiputra et al. 2013

Table 4.3.: CMIP6 models analyzed in this study. D marks models including dynamic vegetation, and N marks models including nitrogen cycles.

Model	Institute	Atmosphere model	Land model	Ocean model	Comment	Main reference and data DOI
ACCESS-ESM1-5	Commonwealth Scientific and Industrial Research Organisation, Australia	UM7.3	CABLE2.4 with CASA-CNP	MOM5 + WOMBAT	N	Law et al. 2017; Ziehn et al. 2017, Data: Ziehn et al. 2019
CanESM5	Canadian Center for Climate Modeling and Analysis, BC, Canada	CanAM5	CLASS-CTEM	NEMO 3.4.1. + CMOC		Swart et al. 2019a, Data: Swart et al. 2019b
CanESM5-CanOE	Canadian Center for Climate Modeling and Analysis, BC, Canada	CanAM5	CLASS-CTEM	NEMO 3.4.1. + CanOE		Swart et al. 2019a, Data: Swart et al. 2019c
CNRM-ESM2-1	CNRM-CERFACS, France	ARPEGE-Climat v6.3 + SURFEX v8.0	ISBA + CTRIP	NEMO v3.6 + GELATO + PISCESv2		Séférian et al. 2019, Data: Seferian 2019
GFDL-ESM4	Geophysical Fluid Dynamics Laboratory, United States	AM4.1	LM4.1	OM4 MOM6 + COBALTv2	D	Dunne et al. 2020, Data: Krasting et al. 2018

MIROC-ES2L	MIROC, Japan	MIROC-AGCM + SPRINTARS	VISIT-e + MATSIRO6	COCO + OECO v2	N	Hajima et al. 2020a, Data: Hajima et al. 2020b
MPI-ESM1-2-LR	Max Planck Institute for Meteorology, Hamburg, Germany	ECHAM6.3	JSBACH3.2	MPIOM1.6 + HAMOCC6	N, D	Mauritsen and Roeckner 2020, Data: Wieners et al. 2019
MRI-ESM2-0	Meteorological Research Institute, Japan	MRI-AGCM3.5 + MASINGAR mk-2r4c + MRI-CCM2.1	HAL	MRI.COMv4		Yukimoto et al. 2019a, Data: Yukimoto et al. 2019b
NorESM2-LM	Norwegian Climate Center, Norway	Modified CAM6	CLM5	HAMOCC	N	Seland et al. 2020, Data: Seland et al. 2019
UKESM1-0-LL	Met Office Hadley Centre, United Kingdom	Unified Model	JULES-ES-1.0	NEMO + MEDUSA-2	N, D	Sellar et al. 2019, Data: Tang et al. 2019

4.1.3. Methods

Sampling of observations and models

For comparison of model and satellite data, first the CO_2 data of the models were converted to XCO_2 data. The model data were interpolated to the grid of the satellite dataset using a bilinear interpolation scheme and grid cells with missing values in the satellite data were also set to missing values in the model fields. Further sampling considerations are discussed in more detail in Section 4.3.3 and in Appendix A.

Most analysis is carried out with regional averages covering several grid cells. Unless specifically stated otherwise, these are calculated by taking the arithmetic averages over all grid cells weighted by their area for each month.

Calculation of growth rate, seasonal cycle amplitude and growing season temperature anomaly

We compute the GR following the method described in Buchwitz et al. 2018. Monthly resolved annual GRs are calculated by subtracting the XCO_2 value 6 months in the future from the one 6 months in the past. Then these monthly resolved GRs are averaged to a yearly GR for a calendar year, and any year with less than 7 months of data is flagged as missing. The addition of the 7-month data availability was introduced to be consistent with the constraint on SCA as explained below.

We define the SCA as the peak-to-trough amplitude in a calendar year of the detrended time series. The time series is detrended with the cumulative sum of monthly GRs, using the annual mean GRs as substitution for missing values where necessary. The SCA is calculated by subtracting the minimum from the maximum value for each year with a minimum data availability of 7 months. When investigating the seasonal cycle of observationally sampled simulations at higher latitudes, the maximum value of the time series was generally only accounted for if more than 7 months of data were available. We therefore introduce the cutoff of 7-month data availability to preserve as many peaks as possible without restricting the data too much. However, as peak preservation cannot be guaranteed when any missing values are present, we can only speak of an effective SCA. The absolute SCA is not as important in our comparison, because we use the same sampling for both the model and observations.

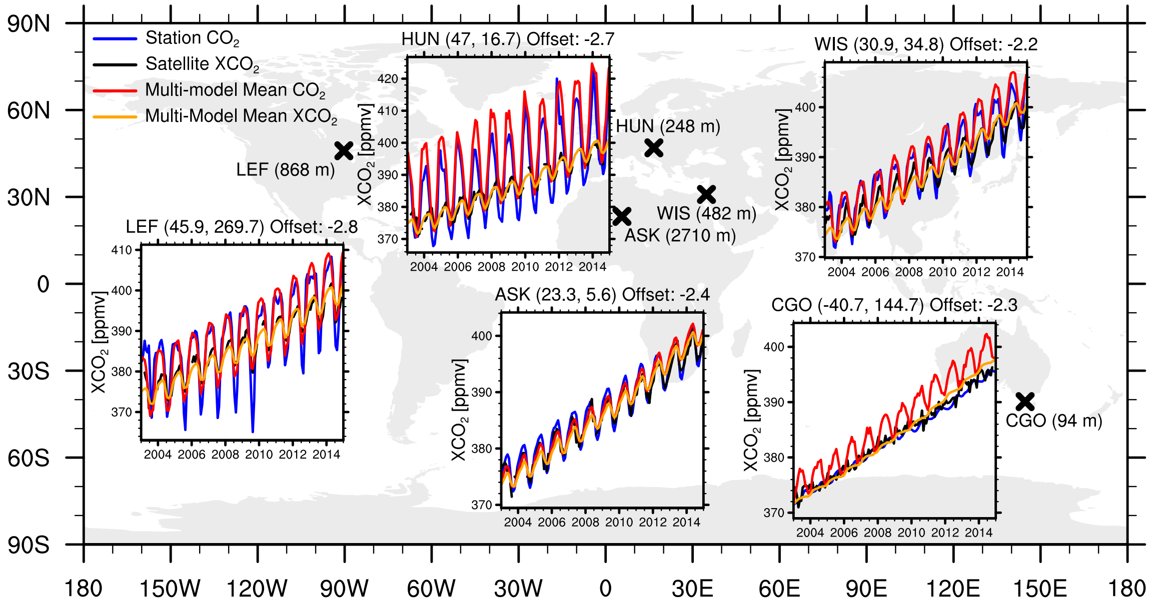
The growing season temperature anomaly ΔT is calculated from the NASA Goddard Institute for Space Studies (GISS) Surface Temperature Analysis (GISTEMP) version 4 (Hansen et al. 2010) temperature anomaly map following Schneising et al. 2014. The data are masked to include only vegetated areas, using the MODIS land cover classification (Friedl et al. 2010). Surface temperature anomalies are calculated with respect to their monthly climatologies. The data are averaged over the growing season if they cover only one hemisphere (April–September for the NH; December to May for the Southern Hemisphere (SH)). Additionally, if the data cover both hemispheres, the whole year is taken into account. The growing season averages are taken because the temperature has a large influence on the plant growth and the resulting biospheric CO_2 fluxes, which in turn drive both the SCA and interannual variability of the GR (Schneising et al. 2014).

4.2. Comparison of XCO_2 and surface CO_2

Until recent years, most model–observation comparisons have been carried out using in situ surface CO_2 data (e.g., Wenzel et al. 2016). As such, it is interesting to compare the differences between XCO_2 and surface CO_2 at different locations. Figure 4.2 shows a comparison of time series between both kinds of observations and the multi-model mean (MMM) for both XCO_2 and surface CO_2 for CMIP6 (top) and CMIP5 (bottom) models. The MMM for both XCO_2 and surface CO_2 is offset to have the same mean value as the satellite data, and this offset is noted above each time series panel. It is interesting to note that the offset appears to be larger at higher latitudes, thus showing a different latitudinal gradient between the models and the satellite data, indicating potential issues with surface fluxes or transport in the models. The MMM and satellite data are averaged between all grid cells covering a $5° \times 5°$ radius around the stations, which results in a maximum of four grid cells to be considered. The mean and GR of XCO_2 and surface CO_2 are in very good agreement, while the MMM overestimates both variables at all sites, with the overestimated mean XCO_2 arising from the effect of higher GRs over time. Furthermore, the offset from the modeled surface CO_2 is higher than that of XCO_2, while this difference is smaller for CMIP5. This might be due to the fact that the CMIP5 offset for MMM XCO_2 was larger overall with approximately 10 ppmv compared to the CMIP6 offset of approximately 2 ppmv.

SCA is higher at higher latitudes and also generally higher at the surface compared to the column average. This is to be expected as the processes dominating the seasonal cycle, respiration and photosynthesis, take place at the surface, leading to the higher SCA for station data and surface $CO_2$from models compared to the XCO_2 values. Mixing of air coming from lower latitudes with lower SCA dampens the SCA in the column compared to surface SCA. This is evident in the increasing SCA difference between XCO_2 and surface CO_2 going from low-latitude to high-latitude stations, with no discernible seasonal cycle in the SH due to the lack of land and vegetation. The MMM follows this trend in the observations, although it underestimates the higher-latitude SCA, with a larger underestimation at the surface while capturing the XCO_2 SCA relatively well. As this study aims at evaluating model simulations with satellite data, further analysis is restricted to XCO_2.

(a) CMIP6



(b) CMIP5

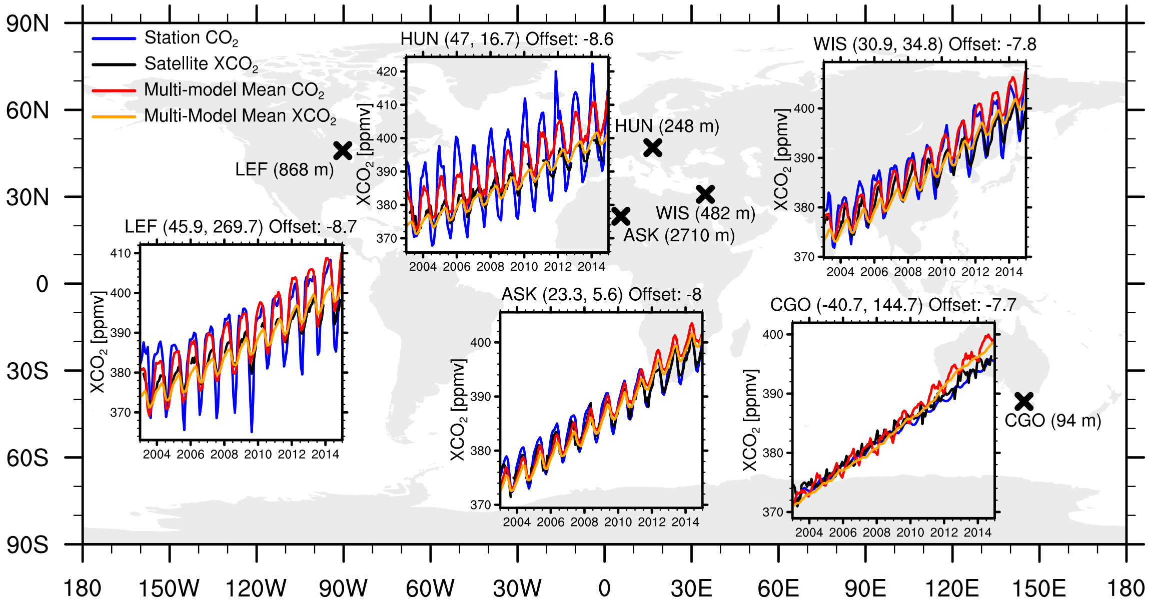


Figure 4.2.: Comparison of time series from satelliteXCO_2 (black), MMM XCO_2 (orange) and surface CO_2 (red), and NOAA surface CO_2 station data (blue) at selected sites, with the coordinates noted in brackets above the time series and the altitudes shown in the map plot (see Table 4.1). The MMM for both XCO_2 and surface CO_2 was offset to have the same average value as the satellite XCO_2 for better comparison, and this offset is noted above each time series. CMIP6 and CMIP5 MMMs are shown in the top **(a)** and bottom **(b)** panels, respectively. Adapted with permission from **Gier** et al. 2020.

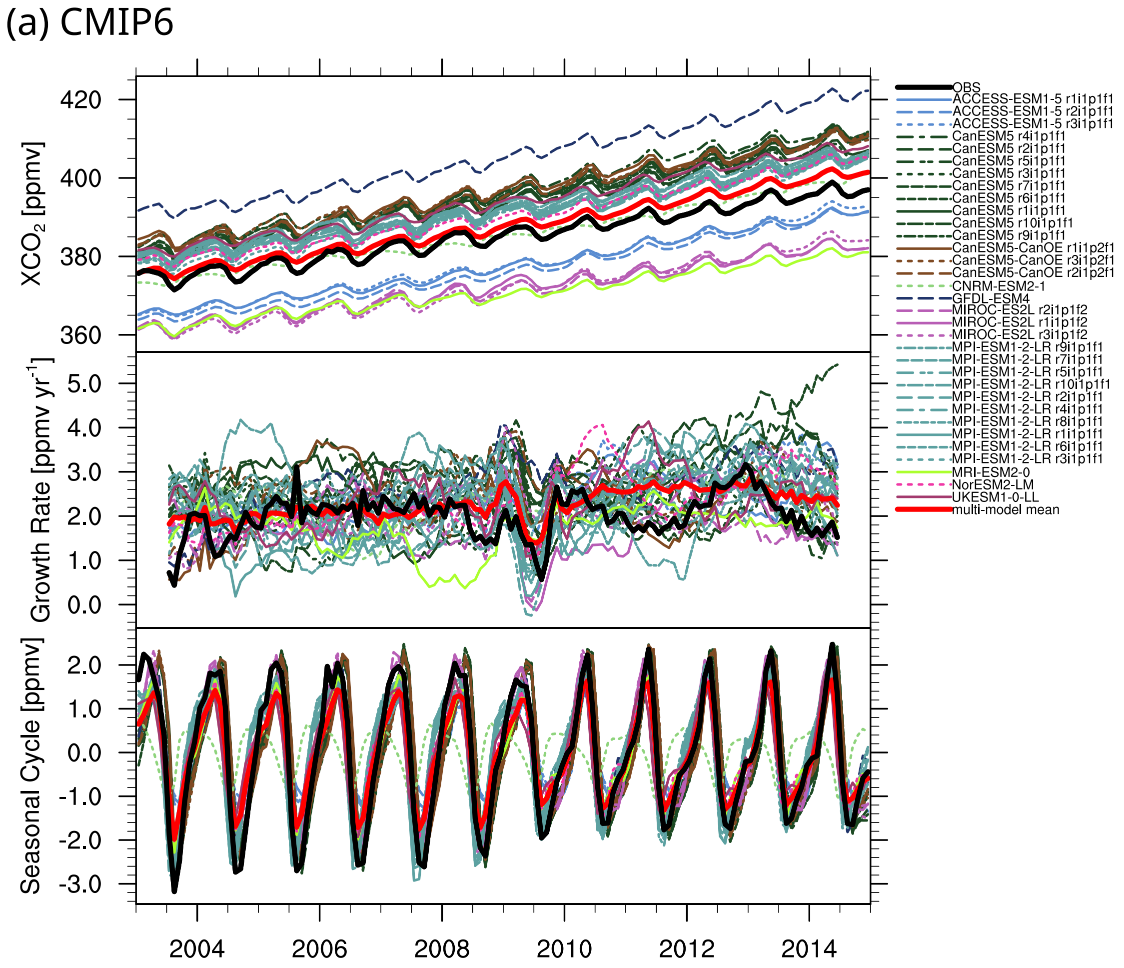


Figure 4.3.: (a) Global time series of monthly mean XCO_2 from 2003 to 2014 for the emission-driven CMIP6 model simulations in comparison to satellite XCO_2 data (bold black line). The model output is sampled as the satellite data. The top panels show the time series, while the middle panels show the computed monthly GR, which has been used to detrend the data to obtain the seasonal cycle shown in the bottom panel. All available ensemble members for each model are shown to demonstrate the intrinsic variability of the models. Adapted with permission from **Gier** et al. 2020.

4.3. Evaluation of CMIP simulations with satellite data

4.3.1. XCO_2 time series

The globally averaged time series of XCO_2 is shown in Figure 4.3 in the top panel, with CMIP6 (a) and CMIP5 (b) models sampled as the satellite observations (see Section 4.1.3). The observational uncertainty is too small to be seen in this plot and is therefore neglected. The middle panel shows the monthly resolved annual GR and the bottom panel the detrended seasonal cycle. All available ensemble members for CMIP6 models are used to show their internal variability. All ensemble members perform similar to one another. The MMM is computed using the first ensemble member, which is also used in the further analysis. As in Figure 4.2, an increase of XCO_2 with time and a pronounced seasonal cycle for all

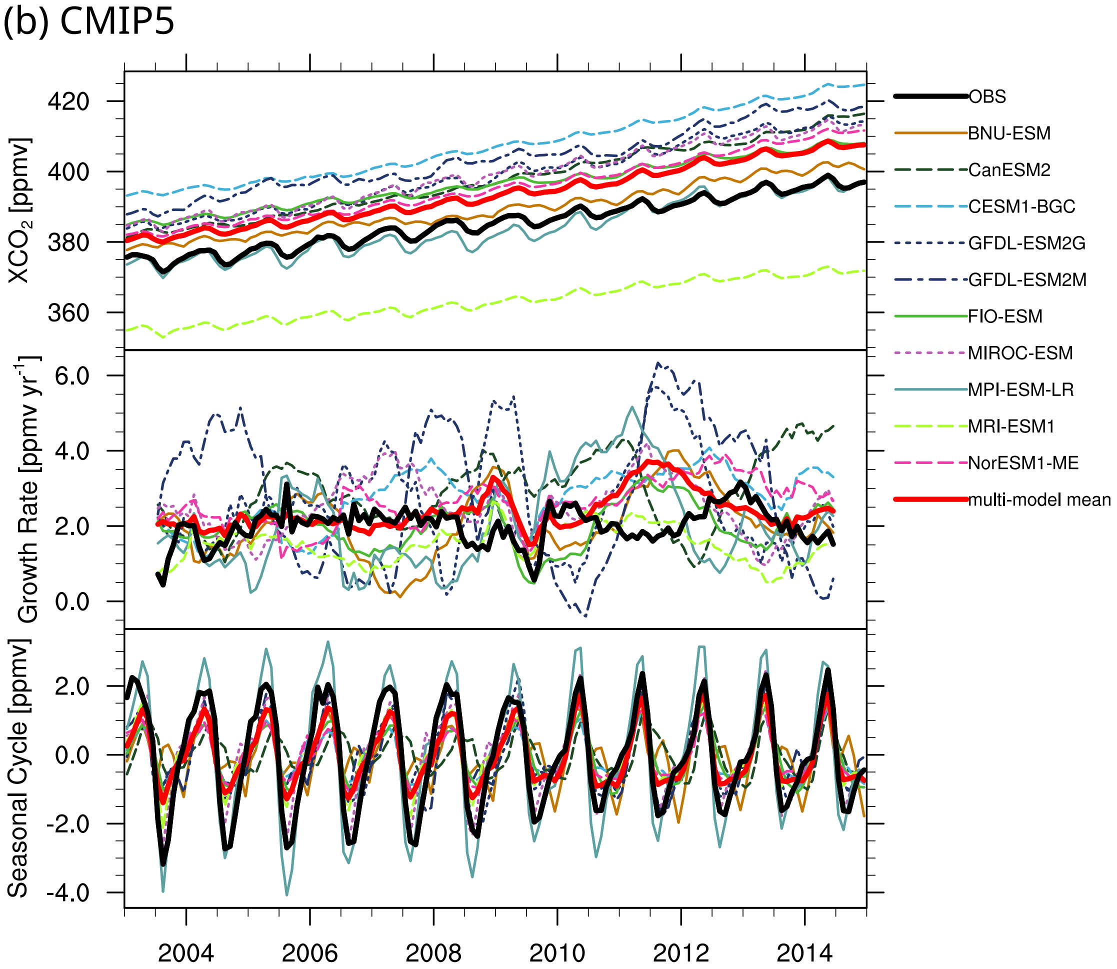


Figure 4.3: (b) Same as Figure 4.3a but for CMIP5. Only one ensemble member is shown. Adapted with permission from **Gier** et al. 2020.

models can be seen. The focus here is on the absolute values, as the trend (GR) and SCA are discussed in dedicated sections below. The CMIP6 models display a large range of absolute XCO_2 values, ranging from an underestimation by 15 ppmv (MRI-ESM2.0 and MIROC-ES2L) to an overestimation by 20 ppmv (GFDL-ESM4). The model closest to the observations is CNRM-ESM2-1, which reproduces the mean value well, with the next closest models being NorESM2-LM and MPI-ESM1.2-LR, both overestimating XCO_2 by about 5 ppmv. The MMM shows an overestimation by approximately 2 ppmv or equivalently a time shift of 1 year. While the spread in the models has not decreased much since CMIP5, the overestimation of the MMM has decreased from 10 to 2 ppmv. Furthermore, CMIP6 models which have predecessors in CMIP5 show similar biases as their predecessors, besides the MIROC models, which overestimated the mean by 15 ppmv in CMIP5 and underestimates it by that much in CMIP6. Both MRI models underestimate XCO_2 significantly, while GFDL-ESM4 overestimates the atmospheric content even more. The MRI-ESM1 model was the only model in CMIP5 to underestimate XCO_2 with respect to the observations, and this was by about 20 ppmv. This model underestimates the historical warming, causing plant and soil respiration to be too low, which leads to a larger land sink and a reduced atmospheric CO_2 concentration (Adachi et al. 2013). This underestimation has been reduced by about 5 ppmv in CMIP6. The GFDL models show an overestimation of about 15 ppmv in both ensembles, and both CanESM models are 10 ppmv too high. A minor improvement can be seen for NorESM-LM over NorESM1-ME, with a decrease of the overestimation from 15 to 10 ppmv.

4.3.2. Growth rate

The middle panel of Figure 4.3 shows that while models capture the interannual variability of the GR quite well, they overestimate the mean GR compared to the observations. The correlation coefficient for the MMM is at 0.48 in CMIP6 and 0.07 in CMIP5 which shows a large improvement in this area. The pronounced feature in 2009 is due to the introduction of the GOSAT data which changed the shape of the seasonal cycle and thus due to its calculation the monthly resolved annual GR. Fortunately, this feature gets averaged out when computing the annual GR and does not tangibly affect our conclusions. Figure 4.4 shows the global mean GR of XCO_2 for 2003–2014 and its standard deviation over all years depicted as error bars, with the observations shown in black and the MMM in red. The annual mean GR of the satellite data is 1.9 ± 0.4 ppmv yr^{-1}, while the CMIP5 models (right) range from 1.5 ± 0.4 (MRI-ESM1) to 3.0 ± 0.9 ppmv yr^{-1} (CanESM2) with a MMM of 2.4 ± 0.4 ppmv yr^{-1}. In CMIP6 (left), the MMM is slightly lower at 2.3 ± 0.3 ppmv yr^{-1}, and the spread has decreased by 0.6 ppmv yr^{-1}, with a range from 1.7 ± 0.4 (MRI-ESM2.0) to 2.6 ± 0.7 ppmv yr^{-1} (GFDL-ESM4). As expected from Figure 4.3, the models – with the exception of MRI-ESM1, MRI-ESM2.0 and MIROC-ES2L – overestimate the GR, leading to an increased XCO_2 level in the present-day atmosphere compared to observations. The interannual variability of the GR for the models is generally higher than that of the observations but well reproduced in the MMM.

The spatial variability of the GR is small, as CO_2 is long lived and well mixed in the atmosphere with a 1-year mean interhemispheric crossing time. Thus, there should be no significant regional changes on an annual level. Buchwitz et al. 2018 found the GR of the satellite dataset to be in agreement with those quoted by NOAA ESRL's global and Mauna Loa time series, as well as robust over several latitude bands. Our own analysis also shows only very small regional differences in the GR (not shown). No significant changes to the annual GRs due to the satellite spatial coverage were found.

Trend of growing season temperature and interannual variability (IAV) of CO_2 GR

Emergent constraint (EC)s are relationships defined using an ensemble of models, between a measurable aspect of current or past climate and an aspect of Earth system feedback in the future, which can be constrained using observational data (Eyring et al. 2019). Cox et al. 2013 developed an EC on the sensitivity of tropical land carbon to climate change using the sensitivity of the IAV of CO_2 GR to the IAV of tropical temperature, which was later

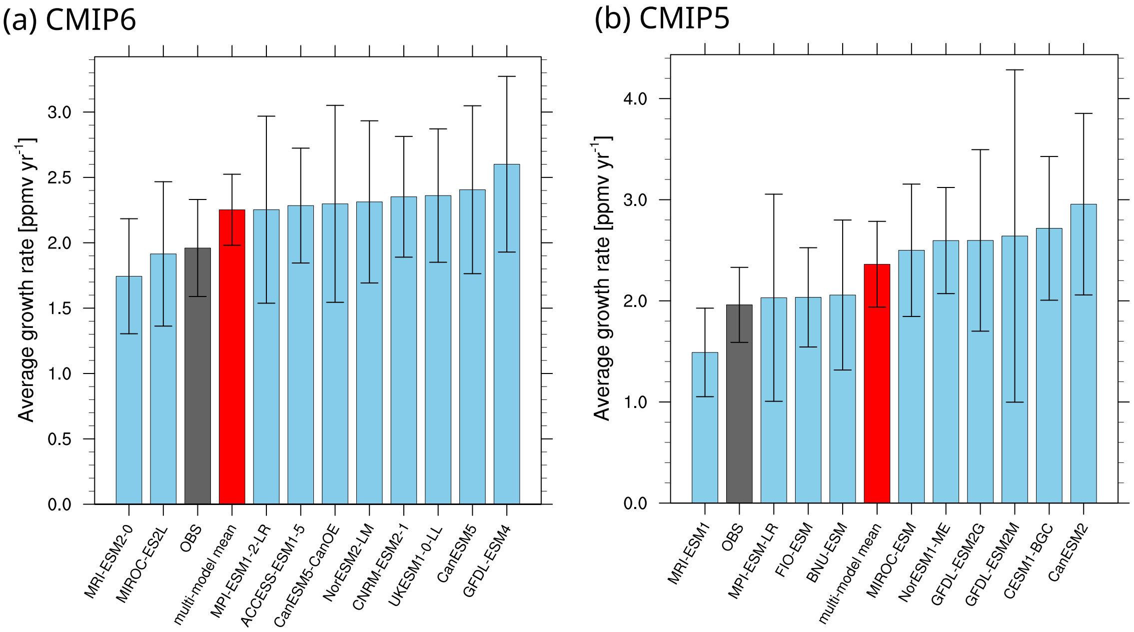


Figure 4.4.: Global mean and standard deviation over all years of annual GR of XCO_2 during 2003–2014 for CMIP6 models **(a)** and CMIP5 models **(b)**. The black bar represents the satellite observations, while the red bar depicts the MMM. Adapted with permission from **Gier** et al. 2020.

adapted by Wenzel et al. 2014 to CMIP5 models. Figure 4.5 shows the sensitivity of the IAV of XCO_2 GR to the tropical growing season temperature IAV for CMIP6 (left) and CMIP5 (right) models, both compared with observations. The observational temperature is taken from the GISTEMP v4 dataset (Lenssen et al. 2019) and the models use their own modeled temperature. We find an observational value of $-0.23 \pm 0.70\,\mathrm{ppmv\,yr^{-1}\,K^{-1}}$ for the 2003–2014 period. However, when using the full span of the satellite data until 2016, the slope increases to $0.75 \pm 0.6\,\mathrm{ppmv\,yr^{-1}\,K^{-1}}$ (not shown), as the additional years show both a high growing season temperature and GR IAV, coinciding with a strong El Niño. This shows that the time period 2003–2014 is not sufficiently long to reproduce the EC, although this may become feasible once CMIP6 emission-driven future simulations are available for a longer time overlap between models and observations. CMIP5 model values for the timeframe 2003–2014 range from 0.53 ± 0.51 (NorESM1-ME) to $3.14 \pm 0.63\,\mathrm{ppmv\,yr^{-1}\,K^{-1}}$ (MRI-ESM1), with only CESM1-BGC showing a negative trend of $-0.64 \pm 0.55\,\mathrm{ppmv\,yr^{-1}\,K^{-1}}$. The MMM has a value of $1.79 \pm 0.80\,\mathrm{ppmv\,yr^{-1}\,K^{-1}}$. In CMIP6, the range is significantly decreased with a minimum of $1.14 \pm 0.56\,\mathrm{ppmv\,yr^{-1}\,K^{-1}}$ (ACCESS-ESM1.5) to a maximum of $3.37 \pm 0.71\,\mathrm{ppmv\,yr^{-1}\,K^{-1}}$ (CanESM5-CanOE) and a MMM of $1.14 \pm 0.37\,\mathrm{ppmv\,yr^{-1}\,K^{-1}}$.

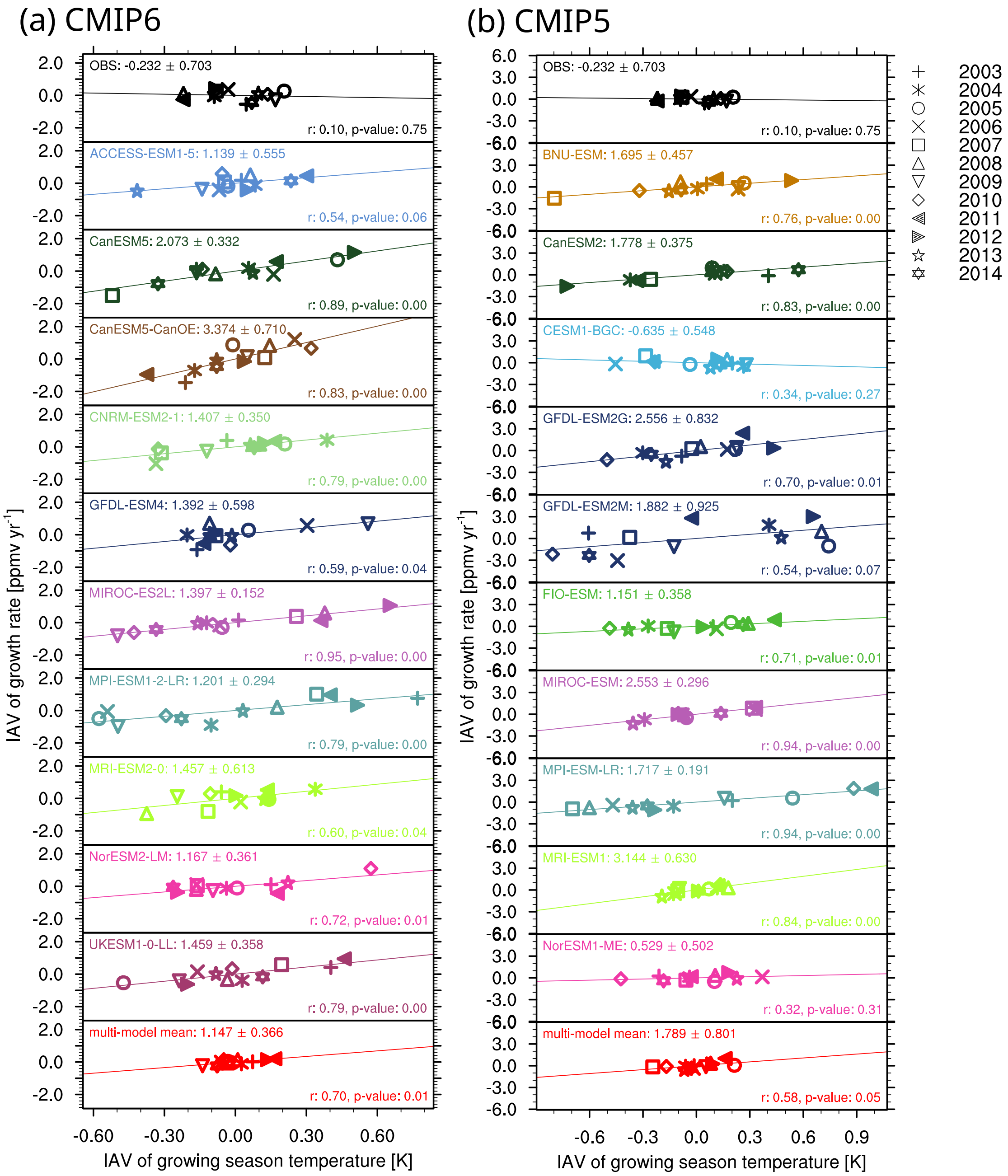


Figure 4.5.: Sensitivity of the IAV of the XCO_2 GR in the tropics (30° S–30° N) to the IAV of tropical growing season temperature for CMIP6 models **(a)** and CMIP5 models **(b)**. The observational temperature taken from the GISTEMP temperature anomaly map, while the models use their own simulated temperature. A linear regression is performed on the data for each dataset. Model colors are the same as in Figure 4.3, and symbols denote the years. In the top left of each panel, the regression coefficient and its uncertainty is shown, while the bottom right states the Pearson correlation coefficient and p value. Adapted with permission from **Gier** et al. 2020.

4.3.3. Seasonal cycle amplitude

This section about the SCA is divided into two subsections, with the first one taking a closer look at inter-model differences, while the second subsection is devoted to the impact of observational sampling.

Model differences

The lower panel in Figure 4.3 shows the detrended global seasonal cycle for all models. Models in CMIP6 (a) show a strong improvement in their ability to capture both the SCA, as well as its phase compared to CMIP5 (b) but still underestimate the SCA. The correlation coefficient to the observed seasonal cycle is 0.93 in CMIP5 and 0.98 in CMIP6 for the MMM. The only model in CMIP6 to significantly underestimate the SCA is CNRM-ESM2-1. Two errors have been identified causing this dampened seasonal cycle: ocean carbon fluxes are lagged in time, and in the emission-driven simulations, CO_2 is considered as an active tracer and coupled with atmospheric chemistry. These chemical fields are restored to global mean climatological concentrations at the model surface, acting as a damping component to the CO_2 concentrations (Séférian et al. 2019). Figure 4.6 shows maps of the climatological mean SCA (2003–2014) for all models, with the global mean given in the top right and the zonal averages shown in the panel attached to the right of the maps. All CMIP6 models (Figure 4.6a) underestimate the SCA compared to satellite observations (Figure 4.7, middle) in the global mean, with the closest mean SCA being MIROC-ES2L. This underestimation was already present in CMIP5 (Figure 4.6b), with several studies discussing it for surface CO_2 SCA (Graven et al. 2013; Wenzel et al. 2016). In CMIP6, the MMM has a globally averaged mean SCA of 3.25 ppmv, compared to 2.92 ppmv for CMIP5, while the observations show an effective SCA of 5.89 ppmv (Figure 4.7).

Both models and observations show the well-known increasing SCA with increasing latitude, due to the more pronounced seasonal cycle of the climate at higher latitudes. Most models show a decreased growth from 0 to 30° N, with higher SCA increases in the lower tropics and northern midlatitudes. Overall, the zonal distribution is quite similar throughout the models, with UK-ESM1-0-LL showing increased SCA at 30–90° S. Tropical land regions in northern South America, Africa and southeast Asia show increased SCA values compared to the ocean SCA at this latitude for the same model. While in the GFDL CMIP5 models this was so pronounced that these regions showed a higher CMIP5 than the higher latitudes (Dunne et al. 2013), this is no longer the case for GFDL-ESM4 in CMIP6. Dunne et al. 2013 attributed the GFDL problem in CMIP5 to the seasonality of respiration in the northern latitudes and an Amazonian low-precipitation bias which reversed the seasonal cycle synchronizing it with the African and Oceanian rain forests. The improvement in CMIP6 is due to a reduced ocean–atmosphere CO_2 flux in the Southern Hemisphere, as well as the reduction of the high tropical land–atmosphere fluxes expressed over the ocean (Dunne et al. 2020).

The SCA in the CMIP5 MPI-ESM-LR model is on average twice as large as the observed one. The high SCA has been discussed in Giorgetta et al. 2013, where it was attributed to a

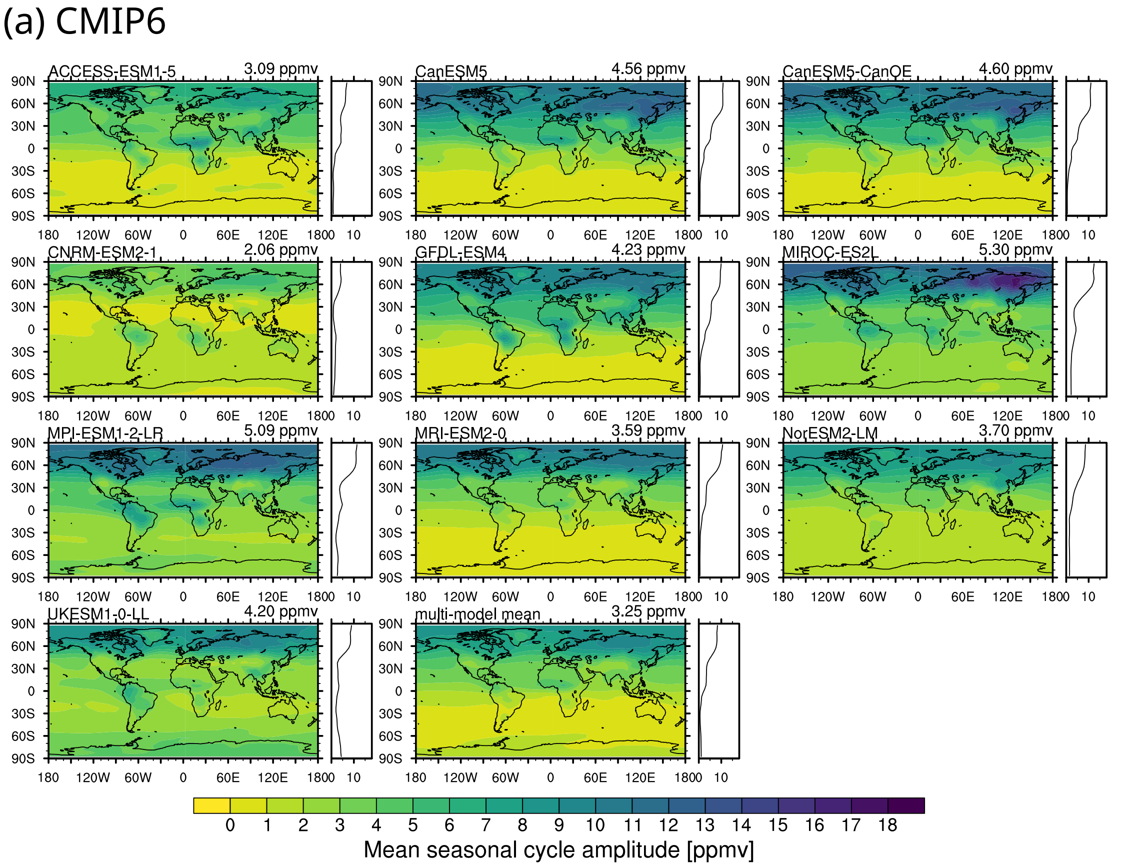


Figure 4.6.: (a) Maps of mean annual SCA for 2003–2014 for the CMIP6 models. The model name is given in the top left of each panel, and the top right shows the global average of the mean annual seasonal cycle. The panel to the right of the maps shows the same zonal mean SCA. Adapted with permission from **Gier** et al. 2020.

combination of an overestimation of net primary productivity in ocean and land biology and uncertainties in atmospheric tracer transport. A particularly severe overestimation was seen in the SH when comparing to station data. As shown in Figure 4.7, we additionally find a large overestimation in XCO_2 SCA in the NH, in particular in the extratropics. For the CMIP6 successor model, MPI-ESM1.2-LR, the SCA is still the highest of the model ensemble but is no longer twice as high as the other models. However, it now shows a more pronounced SCA over the tropical land regions mentioned above, which was not as dominant in CMIP5.

It is known that nitrogen limitations tend to suppress CO_2 fertilization (Reich et al. 2006). Of the four models with the lowest overall SCA in CMIP5 (CESM1-BGC, FIO-ESM, NorESM1-ME and BNU-ESM), two of them – CESM1-BGC and NorESM1-ME – include a nitrogen cycle. The SCAs of NorESM1-ME and CESM1-BGC are very similar, which can be attributed to sharing the same land model (CLM4). FIO-ESM uses the predecessor CLM3.5 and is also comparable to the other two models. It was found that CLM4 had an unrealistically strong nitrogen limitation, which has been reduced in CLM5 (Wieder et al. 2019). In CMIP6, ACCESS-ESM1.5, MPI-ESM1.2-LR, MIROC-ES2L, NorESM2-LM and UKESM1.0-LL include a nitrogen model

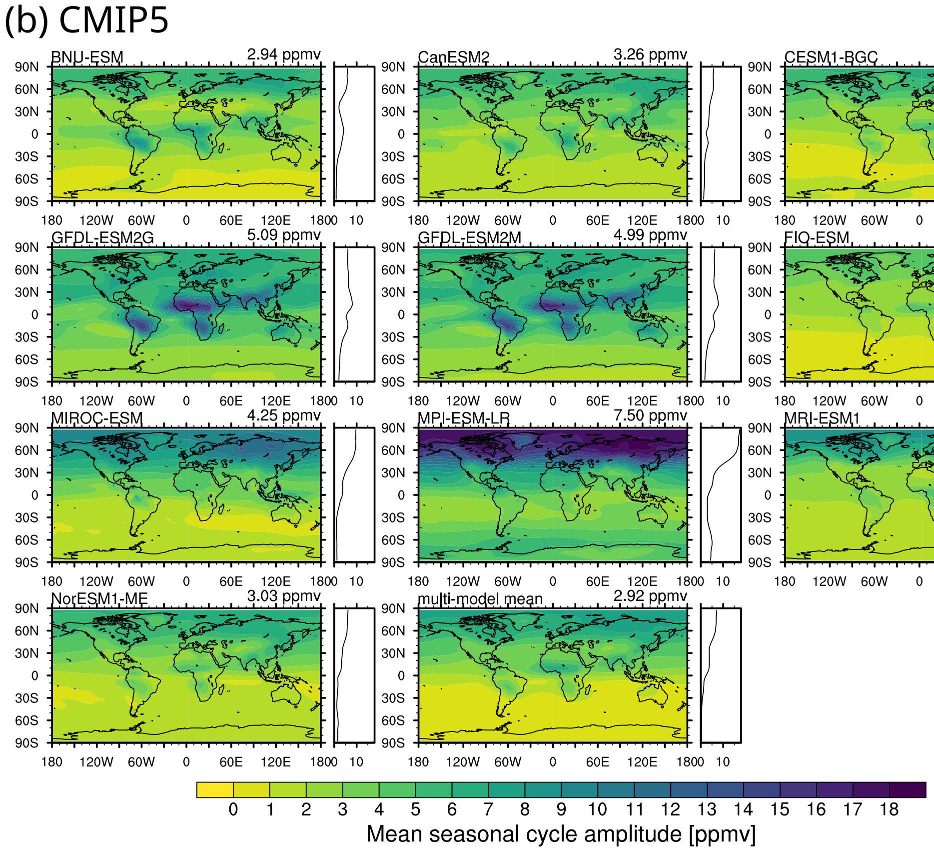


Figure 4.6: (b) Same as Figure 4.6a but for CMIP5.Adapted with permission from **Gier** et al. 2020.

but none of them share the same land model. While ACCESS-ESM1.5 has the second lowest overall SCA, MPI-ESM1.2-LR and MIROC-ES2L have the highest, and thus the observation from CMIP5 models that N-cycle models feature a lower SCA no longer stands for CMIP6.

Influence of sampling

There are a number of ways to compare model SCA to observational SCA, beginning with a grid box comparison. Figure 4.7 shows a comparison of the MMM of CMIP6 (Figure 4.7a) and CMIP5 (Figure 4.7b) to observations. The top right shows the unsampled SCA. The top left panel shows the effective SCA when using observational sampling and the middle panel the satellite data's effective SCA. All numbers are given in ppmv. For an easier comparison, the bottom panels show the absolute difference plots, with the left panel depicting the difference between sampled model and observations, and the right panel the difference between the sampled and unsampled model. Observational sampling slightly lowers the SCA, which is to be expected, as it could lead to masking out the peaks or troughs. While this effect was minimized by imposing the restriction of only computing the SCA of a year when at least 7 months of data are available, it is still a possibility. We therefore classify this SCA as an

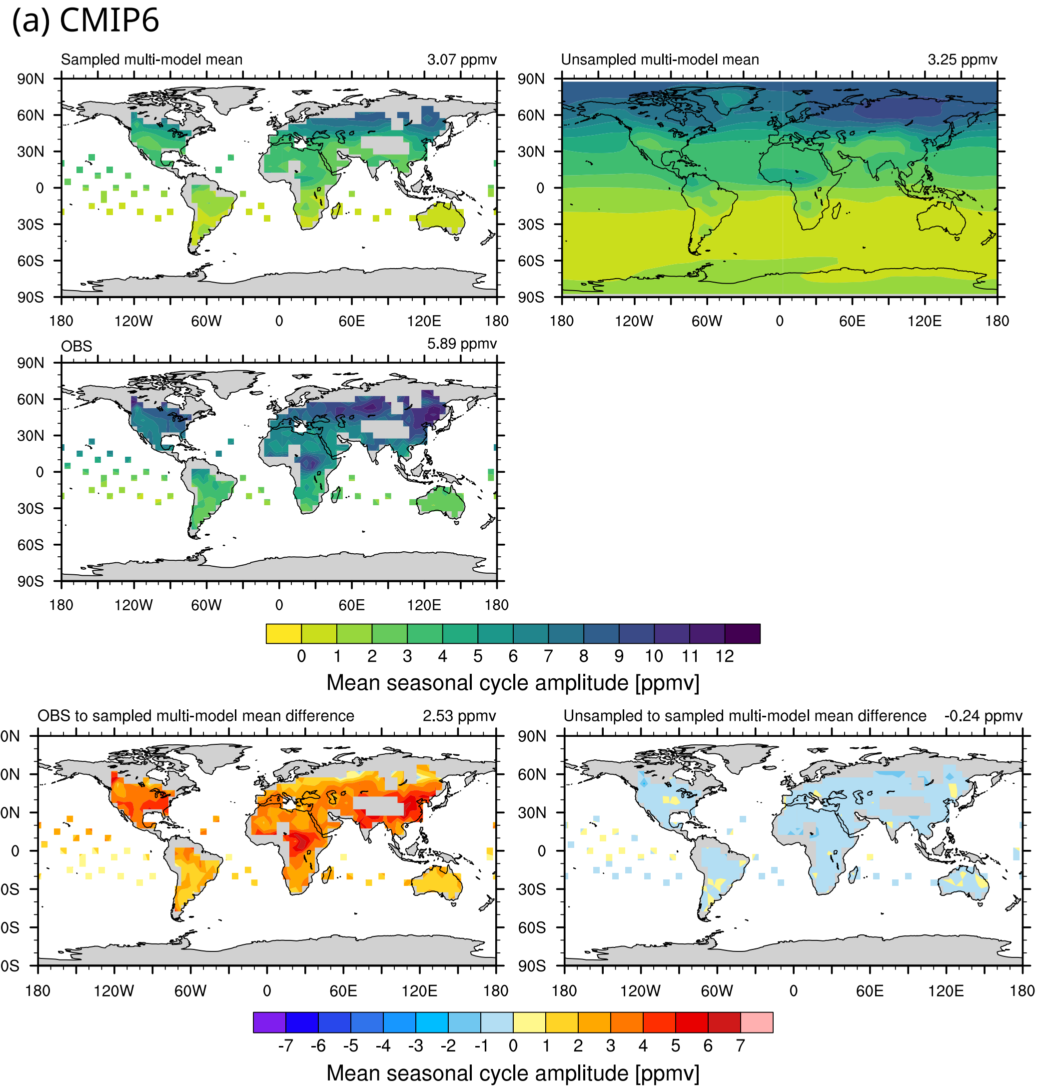


Figure 4.7.: (a) Maps of mean SCA of the CMIP6 MMM for 2003–2014. Top: SCA of MMM with observational sampling (left) and without sampling (right). Middle: SCA of the satellite observations. Bottom: difference between observations and sampled model data (left) and sampled and unsampled model (right). Adapted with permission from **Gier** et al. 2020.

"effective SCA". However, the SCA does not seem to decrease significantly through sampling and the difference does not follow a trend in latitude, so a grid box comparison seems feasible. This paves the way for more comprehensive spatial investigations, which previously relied on data from ground-based stations with sparse spatial coverage. While the stations provide data in higher latitudes that the satellite dataset does not cover, in the tropics and midlatitudes the spatial coverage of the satellite is superior to the ground-based stations. A downside with this approach is the sparsity of the data when using observational sampling. Furthermore, this becomes a computationally expensive operation, as the SCA will need to be calculated for each grid box.

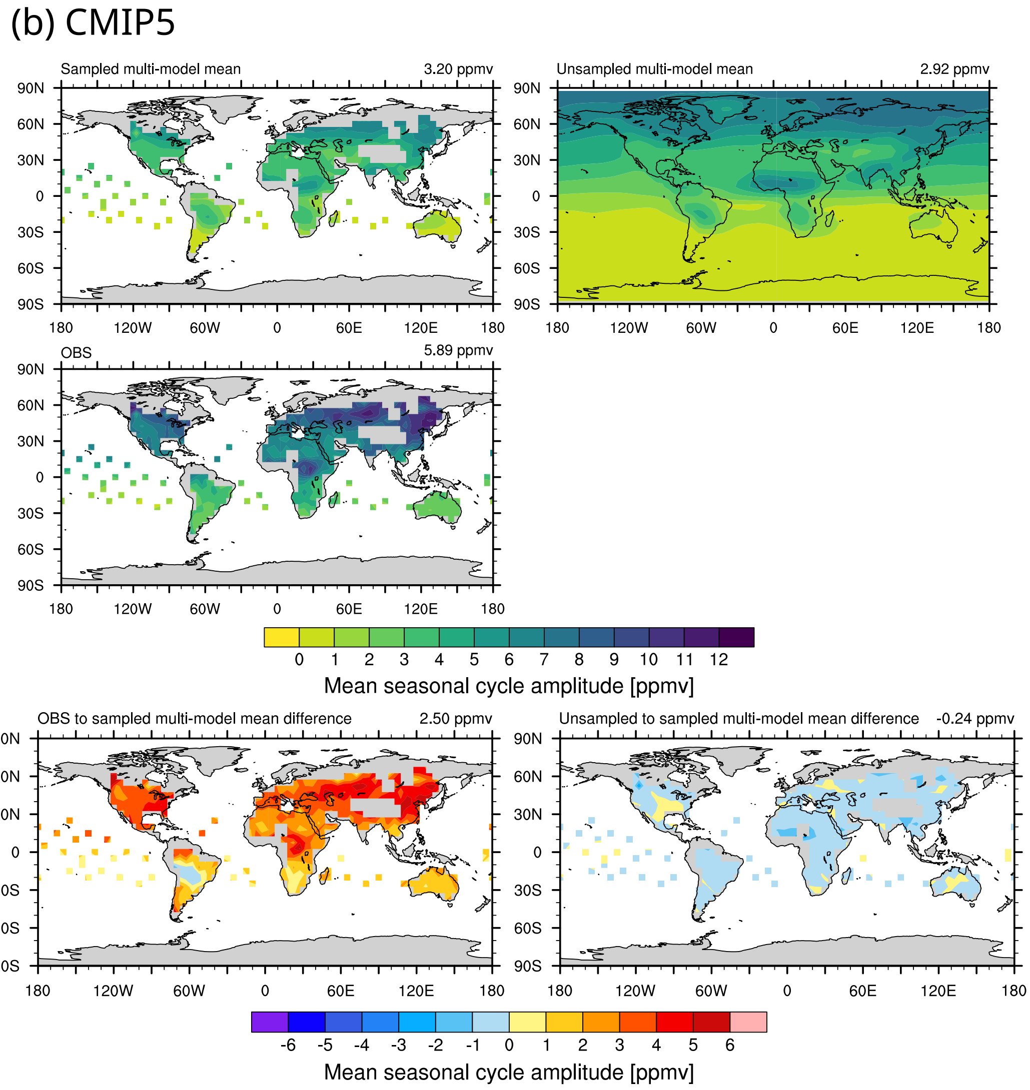


Figure 4.7: (b) Same as Figure 4.7a but for the CMIP5 MMM. Adapted with permission from **Gier** et al. 2020

Another approach often used in model analysis is area averaging, e.g., over different latitude bands like the tropics or the northern midlatitudes. Using surface flask measurements, Wenzel et al. 2016 found an increased SCA with rising CO_2 concentration for CMIP5 using model data from the full historical simulation (1850–2005) – CO_2 fertilization – and used this to establish an EC on the fertilization of terrestrial GPP. Figure 4.8 shows the SCA trend of CMIP6 models versus XCO_2 for 2003–2014 in the northern midlatitudes (30–60° N), including a linear regression including slope, mean SCA, Pearson correlation coefficient and p value. The left panel shows the trend in the unsampled models, while the right one shows the trend when following the sampling of the satellite data. The SCA was computed after a weighted area average was determined on the XCO_2 time series. While some of the unsampled models show

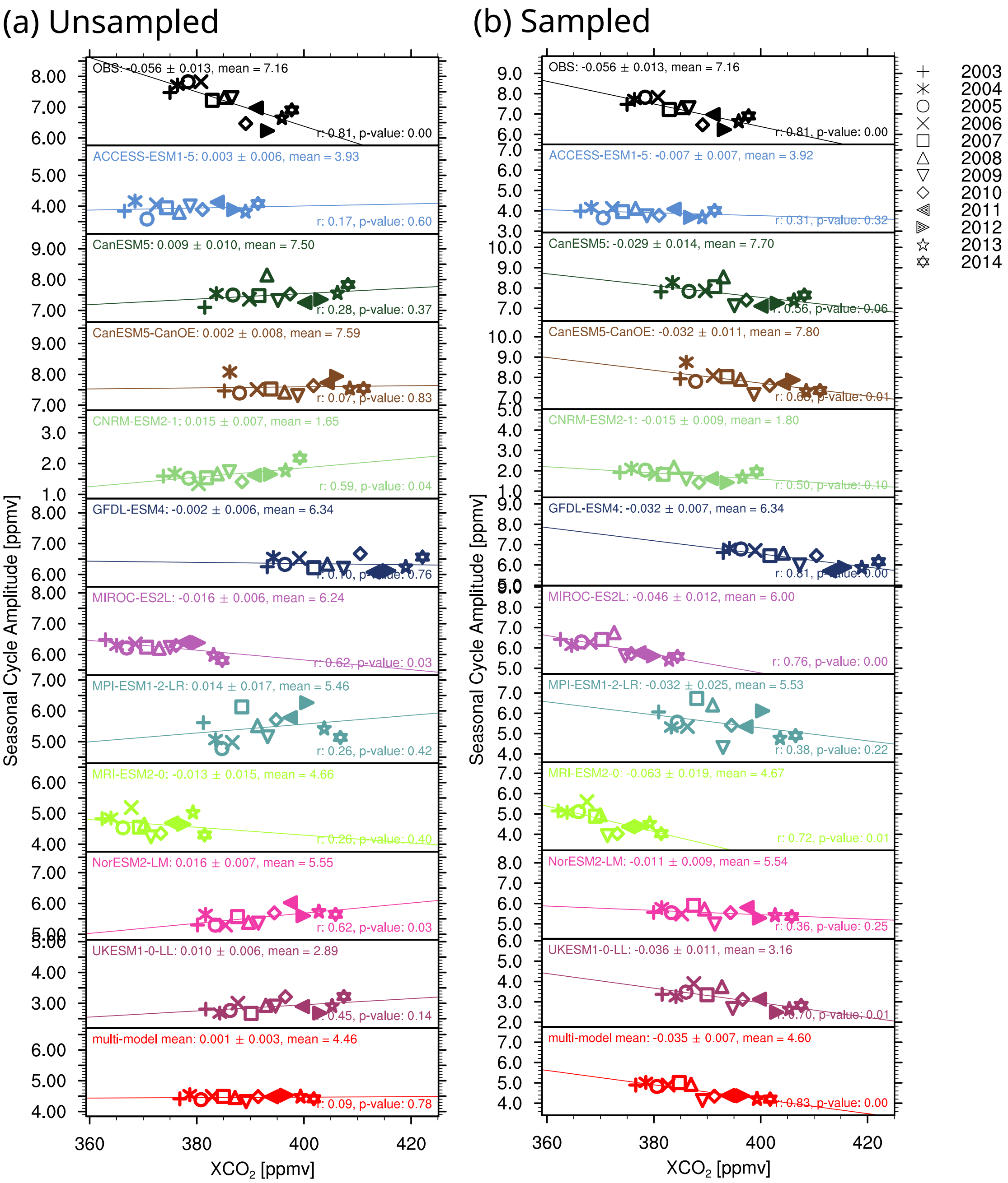


Figure 4.8.: Trend of SCA with XCO_2 for 2003–2014 for the northern midlatitudes (30–60° N), including a linear regression with slope and mean SCA given in the top left of each panel and the Pearson correlation coefficient as well as the p value in the bottom right. Symbols denote the different years and model colors are consistent with previous figures. The left panels **(a)** show unsampled CMIP6 models, while CMIP6 models sampled according to the satellite data are shown on the right **(b)**. Note that the y-axis range for each plot is the same and only differs by a shift. Adapted with permission from Gier et al. 2020.

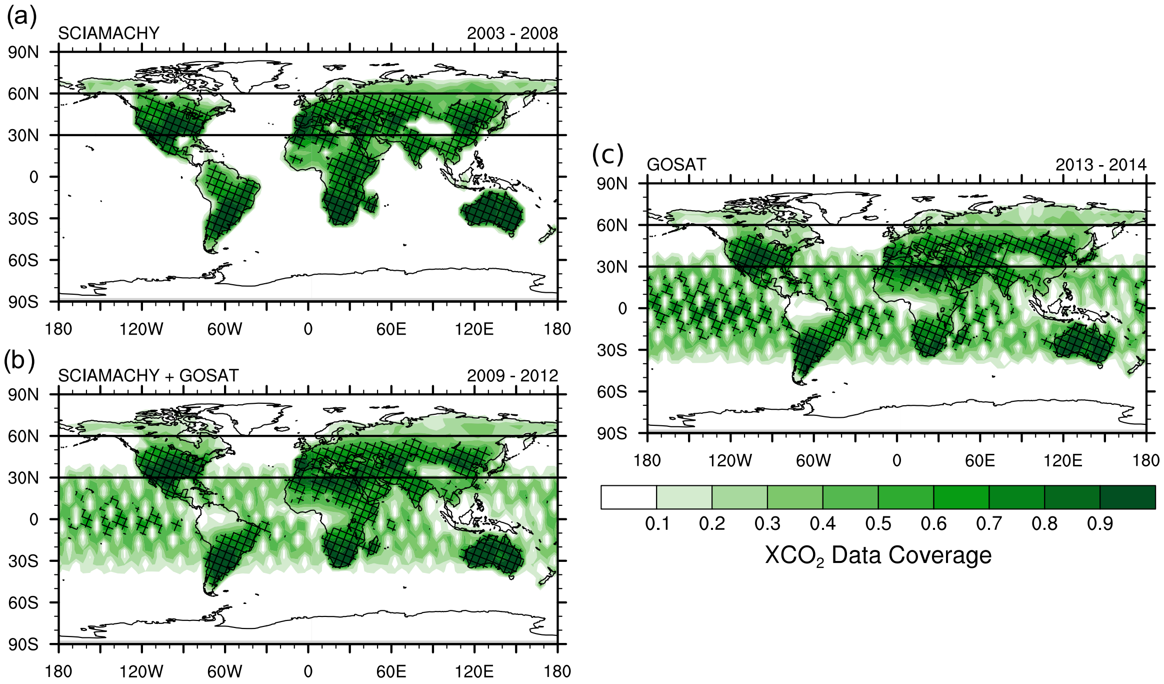


Figure 4.9.: Data coverage of the satellite observations for **(a)** 2003–2008 containing only SCIAMACHY data, **(b)** 2009–2012 containing the overlap of SCIAMACHY and GOSAT data and **(c)** 2013–2014 containing only GOSAT data. The patterned area highlights values above 0.5. Adapted with permission from **Gier** et al. 2020.

an increasing SCA trend with increasing XCO_2, which is in agreement with the findings from Wenzel et al. 2016, others do not show a statistically significant trend and the MMM shows an insignificant positive trend. The sampled model data (right) show a significant negative trend. Calculating the average with a zonal average before summing up the latitude bands does not change this result.

To investigate this change in trend due to observational sampling, Figure 4.9 shows the data coverage for different time periods, 2003–2008 for SCIAMACHY only measurements (a), the overlap between the two satellites in 2009–2012 (b) and 2013–2014 for the GOSAT satellite only (c), with the pattern marking areas with a coverage of 50 % or above. Above 50° N, SCIAMACHY measurements include more areas with 50 % or more coverage compared to GOSAT measurements. With a larger SCA in higher latitudes, it implies that SCIAMACHY would have a larger average SCA over this region compared to GOSAT, hence artificially generating a decreasing trend in the observed SCA, when moving from SCIAMACHY to GOSAT. Figure 4.10 shows the CMIP6 effective SCA trend with XCO_2 using SCIAMACHY and GOSAT masks obtained from masking out points with less than 50 % coverage. While the slopes remain largely the same, the mean effective SCA is higher in the models using the SCIAMACHY mask than when using the GOSAT mask. This mean SCA difference is larger than the total SCA difference within a model using the same sampling for the whole time period. Thus, when considering the observational time series and its full sampling,

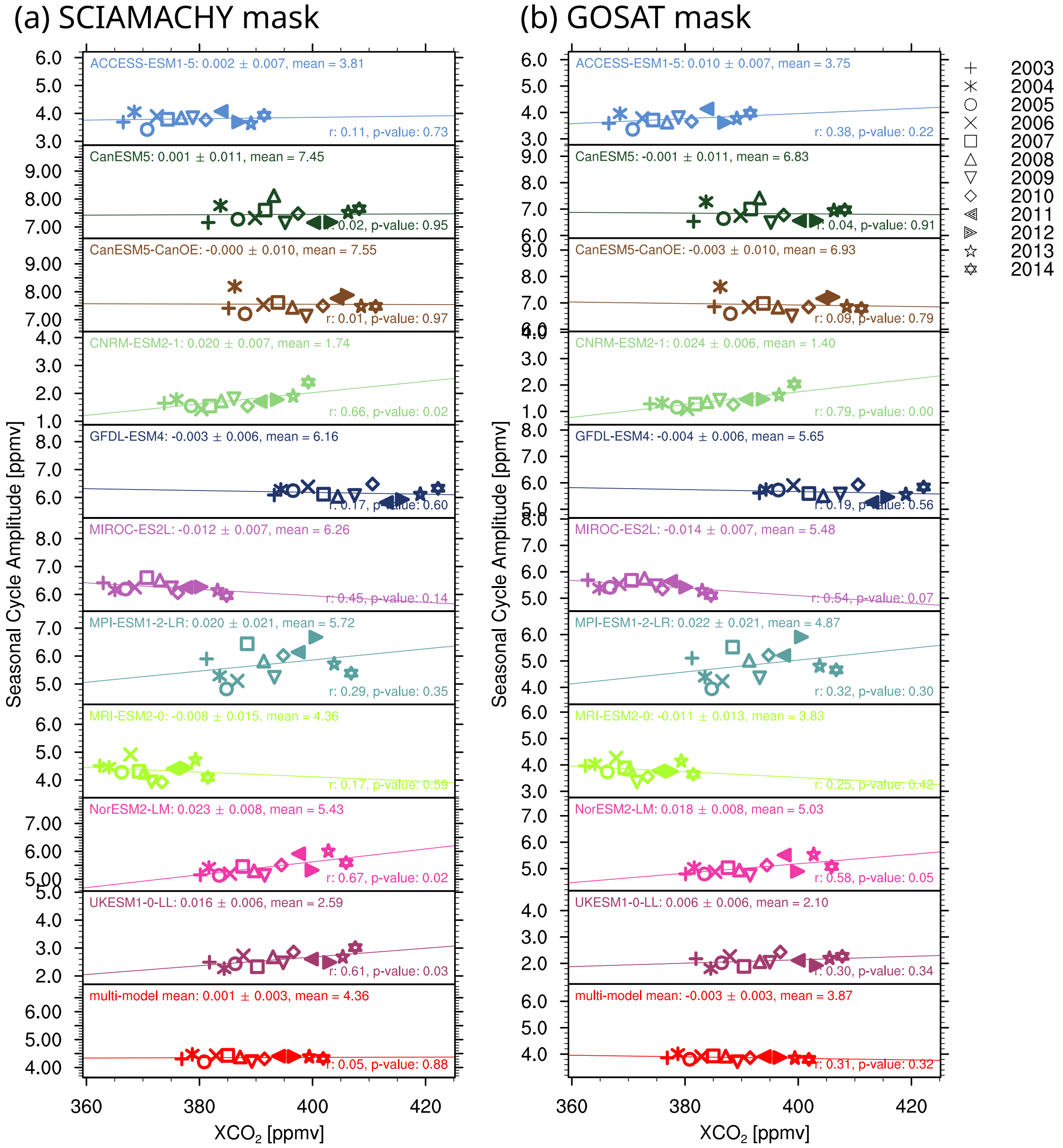


Figure 4.10.: Same as Figure 4.8 but with CMIP6 models masked using **(a)** the SCIAMACHY mask and **(b)** the GOSAT mask, with the masks derived from Figure 4.9, masking out points with less than 50 % coverage in those time periods.

the trend intrinsic to the model is dominated by the negative SCA difference going from the SCIAMACHY to the GOSAT data coverage and thus changed to the negative trend seen in the observations. We can therefore attribute at least part of the negative trend in the satellite data to the different data coverage of the two satellites. We are able to reproduce this negative trend with the models, when these are sampled consistently with the satellite data. This study on sampling also holds true for CMIP5 models, with the equivalent figures shown in Appendix A (Figure A.1 and Figure A.2).

Further impacts on CO_2 concentrations could come through temporal sampling, such as the fact that the satellite data only include measurements with low cloud cover and are limited to 13:00 Local Time (LT). While cloud cover can impact photosynthesis, the response can be fundamentally different for various ecosystems (Still et al. 2009) we expect a larger effect from the diurnal cycle in CO_2 which is included in the model means but not the satellite data. Due to a lack of CO_2 data from models with a higher temporal resolution, this effect cannot be estimated in this study.

4.4. Summary and conclusion

In this paper, we have evaluated the performance of CMIP5 and CMIP6 ESMs with interactive carbon cycle (Table 4.3 and Table 4.2) against column-integral CO_2 (XCO_2) data from satellite retrievals. Our analysis has compared ESM simulations to the 2003–2014 obs4MIPs XCO_2 satellite dataset O4Mv3 retrieved from radiance spectra measured by the SCIAMACHY/Envisat (2003–2012) and TANSO-FTS/GOSAT (2009–2014) satellite instruments. The O4Mv3 data product has a spatial resolution of $5° \times 5°$ and monthly time resolution. For CMIP5, the historical simulations covering the period 2003–2005 were combined with simulations from the RCP8.5 scenario (2006–2014), and for CMIP6 the historical simulations were used (2003–2014). The evaluation of the CMIP models with the satellite data focused on the time series, the GR and the SCA XCO_2. All SCAs computed with a masked time series are considered to be "effective" SCAs due to the possibility of masking out peaks and troughs.

The XCO_2 time series comparison shows that most models overestimate the carbon content of the atmosphere relative to the satellite observations in both model ensembles, with a lower overestimation for the CMIP6 models of 2 ppmv for the MMM and a wide range of individual model differences of −15 to +20 ppmv. The CMIP5 models overestimate by 5 to 25 ppmv with the exception of the MRI-ESM1 model, which underestimates by 20 ppmv. The CMIP5 MMM overestimates by 10 ppmv compared to the observations, which has also previously been found for surface comparisons (Friedlingstein et al. 2014; Hoffman et al. 2014). Overall, CMIP6 models follow the same trends as their CMIP5 counterparts but with reduced systematic biases.

The XCO_2 annual mean GR is typically slightly overestimated compared to the observational value of $1.9 \pm 0.4\,\mathrm{ppmv\,yr^{-1}}$. CMIP6 models range from 1.7 ± 0.4 (MRI-ESM2.0) to $2.6 \pm 0.7\,\mathrm{ppmv\,yr^{-1}}$ (GFDL-ESM4) with a MMM of $2.3 \pm 0.3\,\mathrm{ppmv\,yr^{-1}}$. CMIP5 models have a slightly higher MMM GR of $2.4 \pm 0.4\,\mathrm{ppmv\,yr^{-1}}$ and a larger spread, with the CMIP5 lowest model being MRI-ESM1 at $1.5 \pm 0.4\,\mathrm{ppmv\,yr^{-1}}$ and the highest CMIP5 GR shown by CanESM2 at $3.0 \pm 0.9\,\mathrm{ppmv\,yr^{-1}}$.

All models capture the expected increase of the SCA with increasing latitudes, but most models underestimate the SCA to differing degrees in different regions. This result is in line with previous studies (Graven et al. 2013; Wenzel et al. 2016). Models with similar model components show similar behavior, with models including a nitrogen cycle generally showing a lower SCA in CMIP5, but this influence is not clear in CMIP6. Finally, the connection between SCA and XCO_2 was investigated in the northern midlatitudes. Most models from both ensembles show a positive trend, i.e., an increase of the SCA with XCO_2, consistent with findings for surface CO_2 (Wenzel et al. 2016). However, the satellite product shows a strong negative trend in contrast to the models and surface based observations. We have attributed this trend reversal to the sampling characteristics of the satellite products. The average effective SCA is higher for models sampled according to the SCIAMACHY/Envisat as opposed to the TANSO-FTS/GOSAT mean data coverage. As the early time series is based solely on the SCIAMACHY/Envisat data and the last years only use data from TANSO-FTS/GOSAT, this introduces an artificial negative trend which dominates the positive trend shown by the unsampled models. This demonstrates the importance of equal sampling of models and observations in model evaluation studies.

There are several ways to improve on this analysis in the future. With more available future scenario simulations, the analysis can be extended to a longer time series, making use of longer observational time series, such as the one introduced in Reuter et al. 2020. Higher temporal resolution of the models would enable studies on the effect of the diurnal cycle of CO_2 on the monthly mean and also allow for the construction of a co-located time series with the level-2 satellite data. This could help highlight some of the causes of model biases by being able to pinpoint time and space where they occur more precisely. Model biases may also result from the CMIP experimental design, such as requiring the climate state to be in equilibrium in 1850 while the real world may not have been (Bronselaer et al. 2017), or the parameterizations of biological and physical processes not allowing the system to change rapidly enough (Hoffman et al. 2014). Along with a longer time series, newer satellites, such as Orbiting Carbon Observatory-2 (OCO-2) or the planned Sentinel 7 bring higher resolutions and more data, potentially helping to fill the gaps and reduce the impact of the sampling we discussed in Section 4.3.3.

Overall, the CMIP6 ensemble shows improved agreement with the satellite data in all considered quantities (mean XCO_2, GR, SCA and trend in SCA), with the biggest improvement shown in the mean XCO_2 content of the atmosphere. The paper demonstrates the great potential of satellite data for climate model evaluation as this allows us to go beyond regional means or single point observations from in situ data and also enables the investigation of regional effects on SCA, such as the increase in SCA at higher latitudes.

5. Representation of the Terrestrial Carbon Cycle in CMIP6

The study presented in Chapter 4 shows large improvements for CMIP6 models to simulate the mean XCO_2 content of the atmosphere compared to CMIP5, as well as the general ability to reproduce the seasonal cycle of XCO_2, albeit the SCA is underestimated by most models. The seasonal cycle of CO_2 in the atmosphere is predominantly the result of carbon uptake and release by plants on land. Therefore, the next logical step in analyzing the carbon cycle and understanding the reason for the underestimation of the SCA is to take a closer look at the land carbon cycle and its representation in ESMs. The uncertainty of the global carbon cycle is dominated by land (Canadell et al. 2021), further underlining the importance of understanding the sources of this uncertainty. Previously, Anav et al. 2013a investigated both the land and ocean carbon cycle for CMIP5 historical concentration-driven simulations. They found that while most models were able to reproduce the main climatic variables and their seasonal evolution correctly, some more specific fields showed weaknesses, like an overestimation of photosynthesis. This lead to updates and enhancements of the LSMs for CMIP6 ESMs, such as the inclusion of a nitrogen cycle coupled to the carbon cycle, as well as photosynthesis updates and soil hydrology improvements (e.g., Danabasoglu et al. 2020; Delire et al. 2020; Wiltshire et al. 2021).

To understand what effect the increased process complexity through the inclusion of the interactively coupled nitrogen cycle, as well as the difference between concentration and emission-driven simulations have on model performance, the study presented in this chapter groups models according to three characteristics: their CMIP phase to see a general change between the model generations, their inclusion of an interactive nitrogen cycle, and the simulation type (concentration-driven vs. emission-driven).

This chapter is based on a recently submitted study (**Gier** et al. 2024). The author of this thesis led the writing and analysis of the paper, and also produced all figures. The recipes will be published in the ESMValTool upon final publication of the paper for full reproducibility of all figures. The data used are described in Section 5.1. Section 5.2 analyzes important carbon cycle variables (LAI, GPP, NBP) regarding their trends, growth rate, seasonal cycle and spatially resolved temporal means in the different groupings. Lastly, the individual model performance is evaluated regarding their climatological seasonal cycles and pattern correlations. Section 5.3 summarizes and concludes the study.

5.1. Data

5.1.1. CMIP simulations

Model simulations from both CMIP Phase 6 (Eyring et al. 2016c), and Phase 5 (Taylor et al. 2012) are used, with Table 5.1 and Table 5.2 listing model characteristics such as their atmosphere and land model components in addition to their main references. A more comprehensive summary of the land model components of the CMIP models is given in Appendix A. Models were selected due to their availability on ESGF nodes for the considered variables.

This study focuses on historical simulations, which aim to reproduce the observed climate since the pre-industrial times. They span from 1850 to 2005 (CMIP5) and 1850 to 2014 (CMIP6). Both simulations with prescribed greenhouse gas concentrations (concentration-driven) and prescribed CO_2 emissions (emission-driven) are considered, but evaluated separately and compared to each other. Models participating in the emission-driven simulations, marked in bold in the tables, use their interactive carbon cycle to determine the distribution of natural and anthropogenic carbon fluxes across the land, marine and atmospheric reservoirs instead of relying on prescribed atmospheric CO_2 concentrations (Friedlingstein et al. 2014).

Very few CMIP5 models had a coupled carbon-nitrogen cycle. While the BNU-ESM model included carbon-nitrogen interactions, they were turned off for the CMIP5 model simulations as the nitrogen cycle had not been fully evaluated (Ji et al. 2014). Therefore, a nitrogen cycle was included in two out of 18 CMIP5 models (CESM1-BGC, NorESM1-ME) which both use the CLM4 land model and in 15 out of 23 CMIP6 models spread over six different land models - with CLM in different versions accounting for eight CMIP6 models (v5: CESM2, CESM2-WACCM, NorESM2-LM, NorESM2-MM, v4.5: CMCC-CM2-SR5, CMCC-ESM2, v4: SAM0-UNICON, TaiESM1). The other land models in CMIP6 with a coupled nitrogen cycle are LPJ-GUESS (EC-Earth3-CC, EC-Earth3-Veg), JSBACH (MPI-ESM-1-2-HAM, MPI-ESM1-2-LR), CABLE+CASA-CNP (ACCESS-ESM1-5), JULES-ES (UKESM1-0-LL) and Visit-e (MIROC-ES2L). This shows a large bias towards the CLM land model in CMIP6 which needs to be considered while analyzing the MMM.

To facilitate a direct comparison of CMIP5 and CMIP6 data in figures containing temporal means, only data up to 2005 representing the end of the CMIP5 historical simulations are considered. Unless stated otherwise, figures use mean data over the time period 1986-2005. Only one realization per model is used, as different ensemble members perform similarly to each other with respect to the carbon cycle and using an ensemble mean would lead to an under representation of the internal variability present in individual ensemble members. MMMs were computed separately for each project and experiment combination, as well as an additional distinction between models with and without interactive nitrogen models, and are computed on the monthly gridded data, for which models are regridded to a common 2°x2° grid. MMMs are neither weighted according to the interdependence of the models and model components, nor according to their performance relative to observational products.

Table 5.1.: CMIP6 models analysed in this study. Under Comments, D stands for models including dynamic vegetation and N for models including Nitrogen cycles. Models for which emission driven simulations are also analysed are marked in bold script.

Model	Institute	Atmosphere model	Land model	Comment	Main reference
ACCESS-ESM1-5	Commonwealth Scientific and Industrial Research Organisation, Australia	UM7.3	CABLE2.4, CASA-CNP	N	Law et al. 2017; Ziehn et al. 2017; Ziehn et al. 2020
CanESM5	Canadian Center for Climate Modeling and Analysis, Canada	CanAM5	CLASS3.6, CTEM1.2	D	Swart et al. 2019a
CanESM5-CanOE	Canadian Center for Climate Modeling and Analysis, Canada	CanAM5	CLASS3.6, CTEM1.2	D	Swart et al. 2019a
CESM2	National Center for Atmospheric Research, USA	CAM6	CLM5	N, D	Danabasoglu et al. 2020
CESM2-WACCM	National Center for Atmospheric Research, USA	WACCM6	CLM5	N, D	Danabasoglu et al. 2020
CMCC-CM2-SR5	The Euro-Mediterranean Centre on Climate Change, Italy	CAM5	CLM4.5	N	Cherchi et al. 2019
CMCC-ESM2	The Euro-Mediterranean Centre on Climate Change, Italy	CAM5	CLM4.5-BGC	N	Lovato et al. 2022
CNRM-ESM2-1	CNRM-CERFACS, France	ARPEGE-Climate v6.3 + SURFEX v8.0	ISBA + CTRIP		Séférian et al. 2019
EC-Earth3-CC	EC-Earth Consortium, Europe	IFS 36r4 + HTESSEL + TM5	LPJ-GUESS	N, D	Döscher et al. 2022

EC-Earth3-Veg	EC-Earth Consortium, Europe	IFS 36r4 + HTESSEL	LPJ-GUESS	N, D	Döscher et al. 2022
GFDL-ESM4	Geophysical Fluid Dynamics Laboratory, United States	AM4.1	LM4.1	D	Dunne et al. 2020
INM-CM4-8	Institute for Numerical Mathematics, Russian Academy of Science, Russia	Inbuilt	Inbuilt		Volodin et al. 2018
INM-CM5-0	Institute for Numerical Mathematics, Russian Academy of Science, Russia	Inbuilt	Inbuilt		Volodin et al. 2017a; Volodin et al. 2017b
IPSL-CM6A-LR	L'Institut Pierre-Simon Laplace, France	LMDZ6A	ORCHIDEEv2		Boucher et al. 2020
MIROC-ES2L	MIROC, Japan	MIROC-AGCM + SPRINTARS	VISIT-e & MATSIRO6	N	Hajima et al. 2020a
MPI-ESM-1-2-HAM	HAMMOZ-Consortium, Europe	ECHAM6.3– HAM2.3	JSBACH3.2		Neubauer et al. 2019; Tegen et al. 2019
MPI-ESM1-2-LR	HAMMOZ-Consortium, Europe	ECHAM6.3	JSBACH3.2	N, D	Mauritsen et al. 2019
MRI-ESM2-0	Meteorological Research Institute, Japan	MRI-AGCM3.5 + MASINGAR mk-2r4c + MRI-CCM2.1	HAL		Yukimoto et al. 2019a
NorESM2-LM	NorESM Climate Modeling Consortium, Norway	Modified CAM6	CLM5	N, D	Seland et al. 2020
NorESM2-MM	NorESM Climate Modeling Consortium, Norway	Modified CAM6	CLM5	N, D	Seland et al. 2020
UKESM1-0-LL	Met Office Hadley Centre, United Kingdom	Unified Model + UKCA	JULES-ES-1.0	N, D	Sellar et al. 2019

SAM0-UNICON	Seoul National University, Republic of Korea	CAM5 + UNICON	CLM4	N	Park et al. 2019
TaiESM1	Research Centre for Environmental Changes, Academia Sinica, Taiwan	Modified CAM5.3	Modified CLM4	N	Lee et al. 2020

Table 5.2.: CMIP5 models used in this study, notations as in Table 5.1

Model	Institute	Atmosphere model	Land model	Comment	Main reference
BNU-ESM	College of Global Change and Earth System Science, China	CAM3.5	CoLM + BNU-DGVM	D	Ji et al. 2014
CanESM2	Canadian Center for Climate Modeling and Analysis, BC, Canada	CanAM4	CLASS2.7 + CTEM1		Arora et al. 2011
CESM1-BGC	National Center for Atmospheric Research, United States	CAM4	CLM4	N	Hurrell et al. 2013
GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory, USA	AM2	LM3.0	D	Dunne et al. 2012; Dunne et al. 2013
GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory, USA	AM2	LM3.0	D	Dunne et al. 2012; Dunne et al. 2013
HadGEM2-CC	Met Office Hadley Centre, United Kingdom	Unified Model v6.6	JULES + TRIFFID	D	Collins et al. 2011; Team et al. 2011
HadGEM2-ES	Met Office Hadley Centre, United Kingdom	Unified Model v6.6	JULES + TRIFFID	D	Collins et al. 2011; Team et al. 2011

Inmcm4	Institute for Numerical Mathematics, Russia	Inbuilt		Inbuilt			Volodin et al. 2010
FIO-ESM	The First Institute of Oceanography, SOA, China	CAM3.0		CLM3.5 CASA	+		Bao et al. 2012; Qiao et al. 2013
IPSL-CM5A-LR	L'Institut Pierre-Simon Laplace, France	LMDZ5		ORCHIDEE			Dufresne et al. 2013
IPSL-CM5B-LR	L'Institut Pierre-Simon Laplace, France	LMDZ5		ORCHIDEE			Dufresne et al. 2013
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Japan; Atmosphere and Ocean Research Institute, Japan	MIROC-AGCM SPRINTARS	+	MATSIRO SEIB-DGVM	+	D	Watanabe et al. 2011
MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, Japan; Atmosphere and Ocean Research Institute, Japan	MIROC-AGCM SPRINTARS	+	MATSIRO SEIB-DGVM	+	D	Watanabe et al. 2011
MPI-ESM-LR	Max Planck Institute for Meteorology, Germany	ECHAM6		JSBACH BETHY	+	D	Giorgetta et al. 2013
MPI-ESM-MR	Max Planck Institute for Meteorology, Germany	ECHAM6		JSBACH BETHY	+	D	Giorgetta et al. 2013
MRI-ESM1	Meteorological Research Institute, Japan	MRI-AGCM3.3		HAL			Yukimoto et al. 2011
NorESM1-ME	Norwegian Climate Center, Norway	CAM4-Oslo		CLM4		N	Tjiputra et al. 2013

While we split models into groups only dependent on the presence of an interactive nitrogen cycle in this study, vegetation dynamics is another important process for ESM comparison. Models interactively simulating vegetation cover may simulate trees or grasses in the wrong areas compared to models using observational land cover maps, impacting variables with a strong relation to land cover, such as LAI or GPP. While models with prescribed land cover may show better LAI in the present day, they cannot predict future changes in vegetation cover nor their impact on regional climate and carbon processes. For reference, Table 5.1 and Table 5.2 note models with dynamic vegetation with a D in the comment column.

5.1.2. Reference Data

A large range of observations and reanalysis data sets have been used to assess model performance. These data sets are listed in Table 5.3 along with their main reference(s), their source, the variables used and their temporal coverage. Both observational and reanalysis data sets will be referred to as "observations" from here on, to contrast the results from the CMIP model simulations. The longest observational records are derived from reanalyses, while satellite observations only provide data since the late 20th century. Since most reference data sets do not come with observational uncertainty, a common approach is to use several reference data sets per variable where available, as noted in Seiler et al. 2022. This approach is also taken in this study.

For the LAI, we use the LAI3g product (Zhu et al. 2013) that provides global monthly gridded data starting in the year 1981. It has been generated using an artificial neural network based on data from the Advanced Very High Resolution Radiometer (AVHRR) and the Moderate Resolution Imaging Spectroradiometer (MODIS). Furthermore, we also use the newly released GIMMS LAI4g data set (Cao et al. 2023) which is based on the same satellite data as LAI3g but employs a newer NDVI data set base which removes the effects of the satellite orbital drift and AVHRR sensor degradation, which plagued many other LAI data sets. Furthermore, LAI4g uses a large number of high-quality Landsat LAI samples to increase the spatiotemporal consistency of the data set. Lastly, the Global Land Surface Satellite (GLASS; Liang et al. 2021) is a product suite with 12 products, of which we employ both the LAI and GPP products. Similarly to LAI3g and LAI4g, GLASS is based on data obtained from AVHRR and MODIS. As newer GLASS data products only use MODIS and thus start from 2000, this paper uses a previous GLASS version (v4.0) which includes AVHRR data and thus starts in 1981 for LAI and 1982 for GPP respectively. GLASS LAI uses general regression neural networks trained on preprocessed reflectance data of an entire year to estimate the one-year LAI profile for each pixel. The LAI product is one of the variables used to estimate GLASS GPP with an Eddy Covariance-Light Use Efficiency model. Both GLASS products are available on a 0.05° grid with a frequency of 8 days.

Another GPP product, MTE (Jung et al. 2011), provides global monthly gridded data starting in 1982. It uses an upscaling of data from the FLUXNET eddy covariance tower network based

Table 5.3.: Reference data sets used in this study. data sets in bold are the main reference data set and those in italics the alternate reference for Figures 5.16–5.18.

Data set	Source	Variable	Start Year	Reference
JENA-CarboScope (sEXToc-NEET_v2020)	Inversion	NBP	1957	Rödenbeck 2005
CAMS (v20r2)	Inversion	NBP	1979	Chevallier et al. 2005; Chevallier et al. 2010; Chevallier 2013
GCP	Dynamic global vegetation and bookkeeping model averages	NBP	1959	Friedlingstein et al. 2022; Global Carbon Project 2021
FLUXCOM ANN-v1	Mix	GPP	1980	Jung et al. 2019
MTE	Upscaled in situ	GPP	1982	Jung et al. 2011
GLASS	Satellite	GPP, LAI	1982 (GPP), 1981 (LAI)	Liang et al. 2021; Yuan et al. 2007
LAI3g	Satellite	LAI	1981	Zhu et al. 2013
LAI4g	Satellite	LAI	1982	Cao et al. 2023
NDP-017b	Mix	cVeg	-	Gibbs 2006
HWSD+NCSCD	Empirical	cSoil	-	Hugelius et al. 2013; Wieder 2014

on the model tree ensembles (MTE) approach. Similarly, the FLUXCOM product (Jung et al. 2019) is also based on an upscaling of FLUXNET site level observations, but additionally incorporates a larger variety of machine learning methods, and also includes remote sensing (from MODIS) and meteorological data. Here, we use a global monthly gridded version of FLUXCOM (starting in 1980) from the RS+METEO setup. Due to the assumption of an unchanging average CO_2 level, both MTE and FLUXCOM data are known to have an unrealistic non-existent trend ($0.01\,\mathrm{PgC\ yr^{-2}}$ globally) (Anav et al. 2015). Thus trend analysis on GPP should exclude these data sets.

The main data set for the land-atmosphere carbon flux (NBP) is the JENA-CarboScope (version sEXTocNEET_v2020) product (Rödenbeck 2005), which provides global daily gridded data starting from the year 1957. This data set provides surface-atmosphere CO_2 fluxes based on atmospheric measurements calculated from an atmospheric transport inversion. The inversion used here (NEE-T inversion) involves a regression of interannual net ecosystem exchange (NEE) anomalies against air temperature anomalies (T). In total, JENA-CarboScope uses data from 156 atmospheric measurement sites distributed across the entire globe. The alternative data set for the land-atmosphere carbon flux is a further inversion product from the Copernicus Atmosphere Monitoring Service (CAMS; Chevallier et al. 2005; Chevallier et al. 2010; Chevallier 2013). CAMS provides global gridded data on a monthly resolution starting in 1979 (other temporal resolutions are also available). The inversion product we use here (v20r2) is based on surface measurements from more than 100 sites. A third data set used for comparing the global annual mean NBP is the Global Carbon Project (GCP; Friedlingstein et al. 2022), which estimates the global carbon budget using several observations and models. It provides estimates for emissions from fossil fuel combustion and industrial processes, emissions from land-use change, atmospheric CO_2 growth rate, ocean sink, land sink, cement carbonation sink and the budget imbalance from combining all these terms. The land-atmosphere carbon flux for GCP has to be calculated by subtracting the land-use change emissions from the residual land sink. The land sink was obtained from averaging the results from seventeen Dynamic Global Vegetation Models (DGVMs) which reproduce the observed mean total land uptake of the 1990s and is given with an uncertainty of ± 0.5 PgC yr^{-1} on average. The land use change emissions are estimated from the average of three bookkeeping models with an uncertainty of ± 0.7 PgC yr^{-1}, making it one of the only data sets with direct estimations for uncertainties.

For each of the remaining carbon cycle variables, only a single reference data set is taken into account. For carbon mass in vegetation (cVeg) and carbon mass in soil pool (cSoil), the NDP-017b (Gibbs 2006) and HWSD+NCSCD (Hugelius et al. 2013; Wieder 2014) products are used, respectively. Both data sets provide global gridded annual data for the single year 2000. NDP-017b uses an updated database that extends the methodology of Olson et al. 1985, who developed a global carbon stocks map of above and below ground biomass using 20 years of field investigations, consultations and literature analysis, to more up-to-data land cover conditions of the Global Land Cover Database (GLC2000). The Harmonized World Soil Database (HWSD) uses large volumes of regional and national soil information to create an empirical data set that provides soil parameter estimates for topsoil (0–30 cm) and subsoil (30–100 cm). Similar to Varney et al. 2022 we combine the HWSD data set with the Northern Circumpolar Soil Carbon Database (NCSCD; Hugelius et al. 2013) to complement the HWSD data in the polar region. It uses data on soil order coverage to calculate soil organic carbon content and mass with 1778 pedon (a three-dimensional body of soil) data. Wherever overlap between the two data sets occurs, the NCSCD data are chosen.

5.2. CMIP model performance

General climate variables, such as temperature and precipitation have a large influence on the carbon cycle. It is therefore important to assess how well these variables are simulated by the ESMs. If they are well reproduced but carbon cycle variables are not, it is likely due to a poor representation of processes specific to the carbon cycle, while a poor performance in the physical variables makes an attribution of the cause of poor performance in the carbon cycle variables more difficult. The CMIP6 and CMIP6 ensemble has been assessed compared to observations by Flato et al. 2013 and Eyring et al. 2021 respectively. A detailed analysis was also done by Bock et al. 2020 and references therein, who compare the surface temperature, pressure, precipitation, radiation, and clouds of CMIP3, CMIP5 and CMIP6 historical simulations for annual means. The CMIP6 models show better correlations for these variables than the CMIP5 models for all parameters, with smaller improvements for variables such as temperature which were already well represented in previous CMIP phases. However, the model spread is not significantly reduced but instead largely remains the same. Here, we expand the analysis to the carbon cycle. However it should be mentioned, that many carbon cycle processes are affected by physical variables on much smaller timescales, such as timing of precipitation throughout the day, or if surface temperatures fall below the freezing point at any time of the day may impact the growth of plants more than their monthly means suggest. This study uses monthly mean data which do not resolve many of these important events and thus does not investigate the impact of physical variables further, as any analysis would still be lacking many possible impacts of sub-frequency effects. Future studies using higher frequency data will put more emphasis also on the physical drivers of carbon processes.

5.2.1. Leaf Area Index

The LAI is the ratio of one-sided leaf area per unit ground area (Anav et al. 2013a) as a measure of the canopy structure. Models use LAI to calculate the photosynthetic uptake of the total canopy, also known as GPP. While LAI is an important building block for the carbon cycle, it was also one of the weaknesses of the carbon cycle in the CMIP5 ensemble and tended to be overestimated (Anav et al. 2013a; Anav et al. 2013b).

Carbon uptake by land follows a pronounced seasonal cycle, with CO_2 removed from the atmosphere through plant photosynthesis and released back through plant and soil respiration. With LAI describing the canopy structure and more plants thriving in summer, it is strongly linked to the seasonal cycle of atmospheric CO_2. The seasonal cycles for LAI for CMIP5 and CMIP6 MMMs for concentration and emission-driven simulations are shown in Figure 5.1, split into models with (Ncycle) and without interactive nitrogen cycle (non-Ncycle), as well as different regional means: Global, Northern Hemisphere (20°N - 90°N; NH), Southern Hemisphere (20°S - 90°S; SH) and Tropics (20°S - 20°N). From here on, concentration-driven simulations will be denoted by c, such as CMIP5c and CMIP6c, and emission-driven simulations by e (CMIP5e, CMIP6e). The BNU-ESM and MRI-ESM1 CMIP5 models were removed from the MMM due to featuring an unrealistically high mean LAI, almost doubling the LAI of the reference data and the other models in the SH and the tropics. A common mask is applied to all data sets which includes all missing values in any data set to allow for direct comparison between the models and reference data sets. This increases the LAI compared to unmasked regional means, as missing values are more common in desert and mountainous regions with low LAI (Figure 5.3). The chosen reference data sets LAI3g, LAI4g, and GLASS agree well across all regions, which was to be expected as they are all based on the same raw satellite

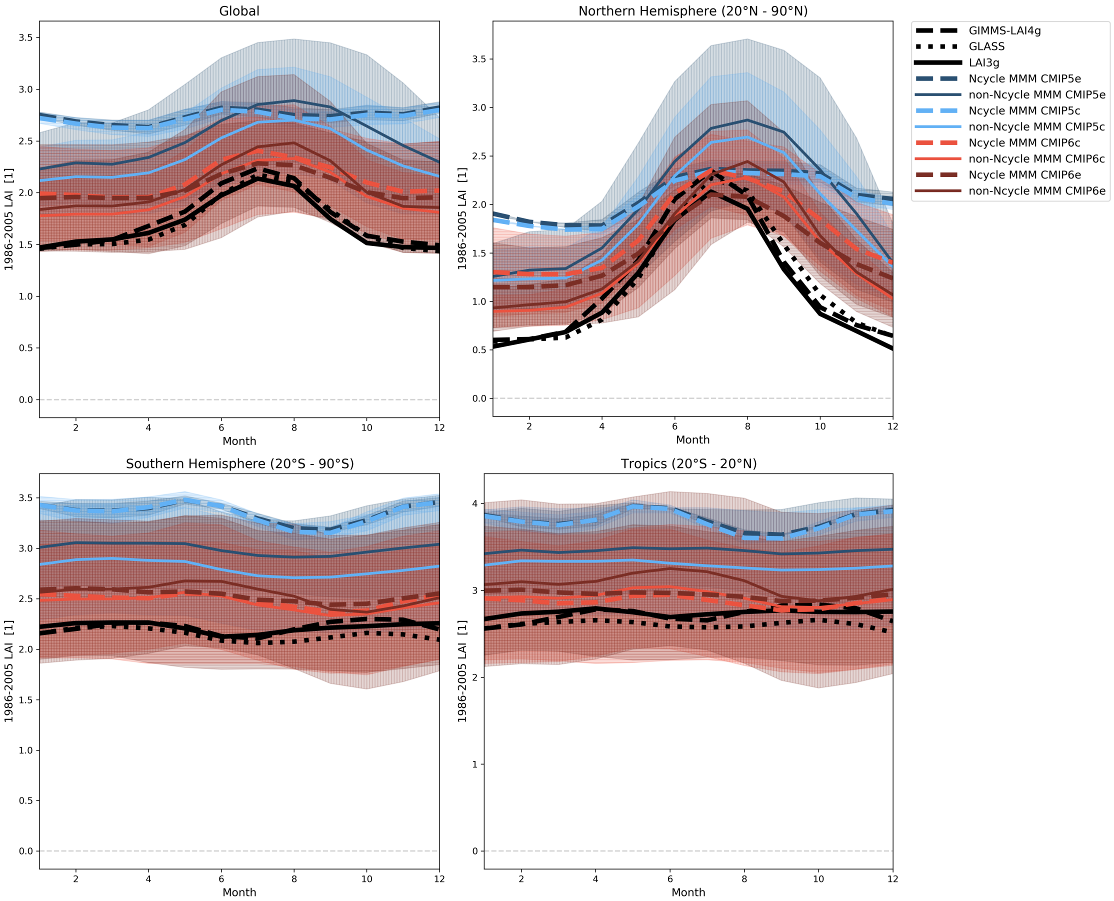


Figure 5.1.: Seasonal Cycle of LAI for a climatological mean of 1986-2005 for different regional averages: Global, Northern Hemisphere (20°N - 90°N), Southern Hemisphere (20°S - 90°S) and Tropics (20°S - 20°N). The reference data sets (LAI3g, solid line; LAI4g, dashed line; GLASS, dotted line) are shown in black, while the MMMs for the different project-experiment combinations are denoted by blue for CMIP5, red for CMIP6, with darker colors for the emission-driven simulations (dark blue CMIP5e, dark red CMIP6e) and lighter colors for concentration-driven simulations (light blue CMIP5c, light red CMIP6c). MMMs derived from models with coupled nitrogen cycle (Ncycle) are dashed, while solid lines represent MMMs of models without coupled nitrogen cycle (non-Ncycle). The shading represents the standard deviation of the MMMs, with vertical hatching for models without and horizontal hatching for models with coupled nitrogen cycle. For comparison with the reference data which contain many missing values, a common mask was applied to all data sets, removing values where any data set is missing a value. Adapted with permission from Gier et al. 2024.

data from AVHRR and MODIS. Xiao et al. 2017 found that the GLASS product outperformed other products, which included LAI3g, when compared to LAI from high-resolution reference maps. As such, GLASS, which is also the reference data set with the largest coverage, will be considered the main reference data set for our analysis of LAI.

There is a strong seasonal cycle in the NH, dominated by high-latitude vegetation in Eurasia and North America. The NH seasonal cycle is the dominant contribution to the global mean due to the higher relative land fraction in the higher latitudes of the global north compared to the global south. The tropics do not show a seasonal cycle due to the absence of strong seasons, while the vegetation in the SH is dominated by forests closer to the tropics, which also lack strong seasonality. All models overestimate LAI in all regions, but the CMIP6 models reproduce the reference data better than the CMIP5 models. According to Anav et al. 2013b, the overestimation in the mid-latitudes is likely partly due to a wet bias and its control on soil moisture, a saturation of satellite instrumentation, and missing parametrizations of disturbances. The annual mean precipitation wet bias is minimally reduced in CMIP6 (Bock et al. 2020) and new parametrizations such as nutrient limitations through interactive nitrogen cycle have been introduced in some models, leading to a reduced LAI in the CMIP6 MMMs compared to CMIP5. The seasonal cycle in the NH is reproduced, but while the CMIP5 non-Ncycle models reproduce the amplitude well with a positive offset of approximately $0.7\,\mathrm{m}^2/\mathrm{m}^2$, the CMIP6 non-Ncycle models are better at reproducing the peak value. Both CMIP5 Ncycle models (CESM1-BGC and NorESM1-ME) use the CLM4 land model with known issues regarding LAI, such as underestimating LAI in dry regions due to elevated CO_2 and overestimating LAI in moist regions (Lee et al. 2013), as well as an unrealistically strong nitrogen limitation (Wieder et al. 2019), hindering plant growth. This leads to these models showing a larger overestimation in LAI in both the Southern Hemisphere and Tropics dominated by moist rainforests. Additionally, the seasonal cycle amplitude is strongly reduced in the NH while the mean LAI is larger than for the reference data sets. Both CMIP projects show a weakness in simulating the end of the growing season, shown by the later decline of LAI in winter, which also leads to a smaller seasonal cycle amplitude, consistent with the findings of Park and Jeong 2021. The drawdown in autumn signifying the end of the growing season is smaller in Ncycle models compared to non-Ncyle models. The differences between the concentration and emission-driven simulations are small, with models participating in both simulations having very similar results (individual models not shown). Larger differences occurring here and in later analysis between the concentration and emission-driven simulations are likely due to the different subset of models in the historical simulations, and not due to the experiment design.

Figure 5.2 shows the mean and trend of LAI averaged over 1986-2005 and depicts project-experiment simulations with one type of marker each for a better overview. Filled symbols denote models with nitrogen cycle while the mean of the reference data is shown with black markers, with error bars showing the standard deviation of the reference data. Exact numbers for all data sets are found in the supplementary information (Tables B.1 and B.2). The errors given refer to the standard deviation of the mean as a measure for the interannual variability (IAV), while the standard error of the trend is the error of the linear regression calculating the trend. For the individual reference data, LAI3g and LAI4g agree well in mean and trend, while GLASS agrees with their mean but shows a significantly higher trend in all regions, leading to the large trend error bar. Xiao et al. 2017 analyzed the trend of several LAI products for different biome types for 1982-2011 and found GLASS to have significantly higher trends in savannahs and shrubs compared to LAI3g, but lower trends in deciduous broadleaf forests, evergreen needleleaf forests and even a negative trend for deciduous needleleaf forests, while grasses, cereal crops and evergreen broadleaf forests trends are similar for GLASS and LAI3g.

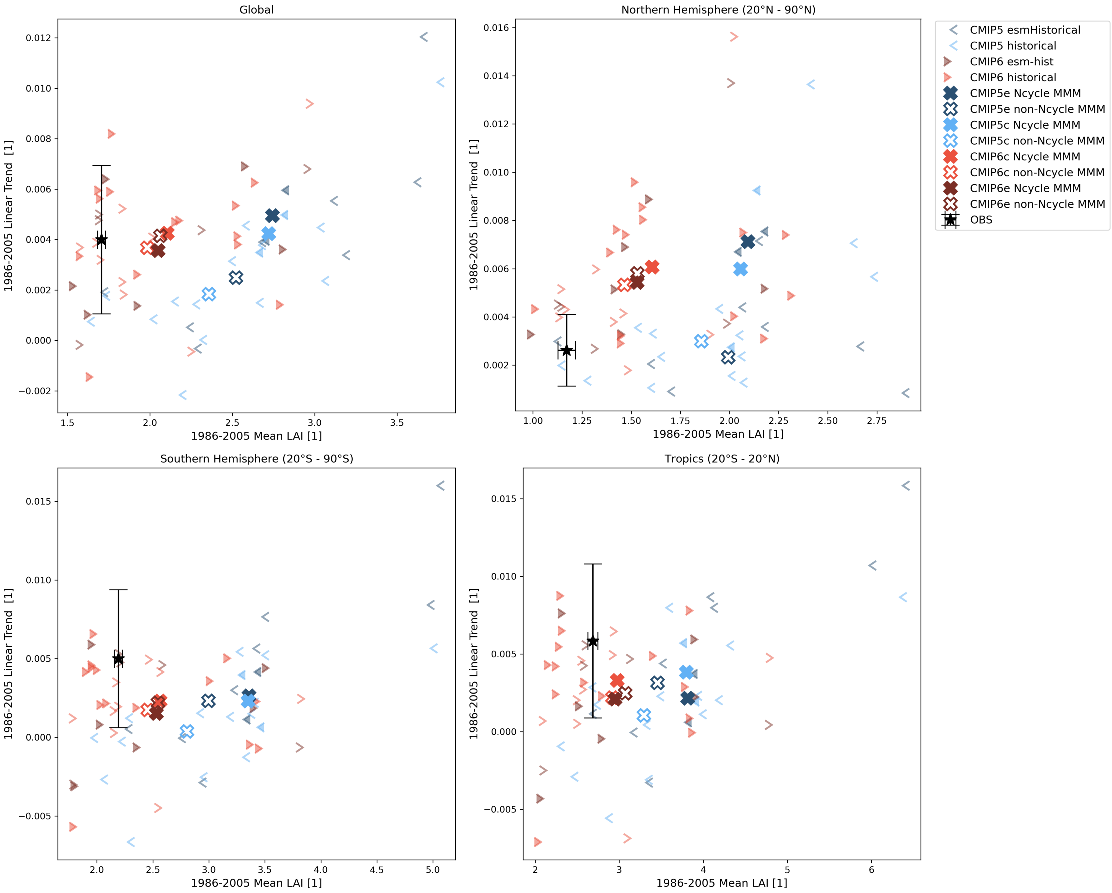


Figure 5.2.: Mean and Trend of LAI computed over 1986-2005 for different regions: Global, Northern Hemisphere (20°N - 90°N), Southern Hemisphere (20°S - 90°S) and Tropics (20°S - 20°N). The mean of the reference data sets (LAI3g, LAI4g and GLASS) is denoted by a black star, with errorbars for the standard deviation. Models for project-experiment combinations are shown with a single symbol each, blue smaller than (<) sign for CMIP5 and red greater than (>) sign for CMIP6, with darker colors for emission-driven and lighter colors for concentration-driven simulations, as well as Ncycle models being denoted with a filled symbol. MMMs are depicted with cross symbols, filled for Ncycle MMMs and using the color assigned to their project-experiment combination. Adapted with permission from **Gier** et al. 2024.

This results in larger trend differences in the SH (0.009 m^2/m^2 yr^{-1}) and Tropics (0.010 m^2/m^2 yr^{-1}) than the NH (0.003 m^2/m^2 yr^{-1}). MMMs span the range of 1.98 ± 0.02 m^2/m^2 (non-Ncycle CMIP6c) to 2.74 ± 0.05 m^2/m^2 (Ncycle CMIP5e), showing a significant overestimation compared to the reference data, with the overestimation of the CMIP5 MMMs of 0.7 m^2/m^2 reduced by half for CMIP6 MMMs, as seen in Figure 5.1. This large improvement for the CMIP6 models is not related to the difference between Ncycle and non-Ncycle models as their LAI MMMs means are comparable. Due to the large difference in trends in the reference data sets, the global mean trends of all CMIP MMMs lie within the range of the reference data. The model trends range between slightly negative -0.0022 ± 0.0010 m^2/m^2 yr^{-1} (IPSL-CM5A-LR CMIP5c) to strong positive 0.0120 ± 0.0013 m^2/m^2 yr^{-1} (BNU-ESM CMIP5e). As BNU-ESM was not considered in the MMM due to its large mean LAI compared to all the other models, the highest global LAI trend considered for the MMM is CMIP6c MRI-ESM2-0 at 0.0094 ± 0.0019 $\mathrm{m^2/m^2\ yr^{-1}}$, which is significantly larger than the trend in its CMIP6e simulation at 0.0068 ± 0.0019 $\mathrm{m^2/m^2\ yr^{-1}}$. The MMMs range between 0.0018 ± 0.0004 $\mathrm{m^2/m^2\ yr^{-1}}$ (non-Ncycle CMIP5c) and 0.0050 ± 0.0015 $\mathrm{m^2/m^2\ yr^{-1}}$ (Ncycle CMIP5e). Unlike the mean, the LAI trend does not show a strong difference between CMIP5 projects, nor any other grouping method we employed. The CMIP6 models only show a slightly smaller range in trend compared to the CMIP5 models, but more CMIP5 models have a lower trend compared to the reference data sets than CMIP6 models.

The other regions reflect these overall MMM comparisons as well. CMIP6 means are closer to the reference data than CMIP5 in all regions, albeit still overestimating LAI, and agree well with each other no matter the experiment or Ncycle status. The only exception is the NH CMIP6c Ncycle MMM, which shows a larger mean at 1.71 ± 0.04 $\mathrm{m^2/m^2}$ than the other three CMIP6 groupings with means between 1.47± 0.03 $\mathrm{m^2/m^2}$ and 1.53 ± 0.04 $\mathrm{m^2/m^2}$. This is due to the CMCC-ESM2 and CMCC-CM2-SR5 models, which show a much higher LAI in the NH compared to the reference data, but fit well in the other regions. These two models use the CLM4.5 land model, which Li et al. 2022 found to have a far longer peak growing season and to overestimate LAI in boreal forests compared to MODIS reference data, consistent with our results. The CLM5 models (CESM2, CESM2-WACCM, NorESM2-LM, NorESM2-MM) perform much better in mean LAI than CLM4.5 in the NH, but are still overestimating LAI compared to the reference data. The trend in the NH for the reference data set are 0.0014 ± 0.0003 $\mathrm{m^2/m^2\ yr^{-1}}$ (LAI4g), 0.0018 ± 0.0008 $\mathrm{m^2/m^2\ yr^{-1}}$ (LAI3g) and 0.0047 ± 0.0013 $\mathrm{m^2/m^2\ yr^{-1}}$ (GLASS), with the models showing a much larger range from 0.0008 ± 0.0008 $\mathrm{m^2/m^2\ yr^{-1}}$ (GFDL-ESM2M CMIP5e), with a significantly larger trend in CMIP5c at 0.0057 ± 0.0006 $\mathrm{m^2/m^2\ yr^{-1}}$, to 0.0156 ± 0.0023 $\mathrm{m^2/m^2\ yr^{-1}}$ (MRI-ESM2-0 CMIP6c), although most models have a trend below 0.01 $\mathrm{m^2/m^2\ yr^{-1}}$. For the MMMs, only the CMIP5 non-Ncycle MMMs fall between the reference data with 0.0023 ± 0.0003 $\mathrm{m^2/m^2\ yr^{-1}}$ for non-Ncycle CMIP6c and 0.0030 ± 0.0003 $\mathrm{m^2/m^2\ yr^{-1}}$ for non-Ncycle CMIP6e. The other MMMs show a larger trend than the reference data but comparable to each other, ranging from 0.0053 ± 0.0004 $\mathrm{m^2/m^2\ yr^{-1}}$ (non-Ncycle CMIP6c) to 0.0061 ± 0.0004 $\mathrm{m^2/m^2\ yr^{-1}}$ (Ncycle CMIP6c), with a slightly larger value of 0.0071 ± 0.0012 $\mathrm{m^2/m^2\ yr^{-1}}$ for Ncycle CMIP5e. CMIP6 MMM LAI mean ($\approx$ 2.5 $\mathrm{m^2/m^2}$ SH, 2.9 $\mathrm{m^2/m^2}$ tropics) and trend ($\approx$ 0.002 $\mathrm{m^2/m^2\ yr^{-1}}$ SH, 0.002 $\mathrm{m^2/m^2\ yr^{-1}}$ tropics) agree well with the LAI3g and LAI4g (mean 2.2 $\mathrm{m^2/m^2}$ SH, 2.7 $\mathrm{m^2/m^2}$ tropics, trend 0.002 $\mathrm{m^2/m^2\ yr^{-1}}$ SH, 0.002 $\mathrm{m^2/m^2\ yr^{-1}}$ tropics) reference data in the tropics and the SH, while the CMIP5 MMMs overestimate the mean by $\approx$ 0.7 $\mathrm{m^2/m^2}$ in the SH and $\approx$ 0.6 $\mathrm{m^2/m^2}$ in the tropics for non-Ncycle, as well as $\approx$ 1.2 $\mathrm{m^2/m^2}$ in the SH and $\approx$ 1.1 $\mathrm{m^2/m^2}$ in the tropics for Ncycle MMMs, but show a similar trend to LAI3g, LAI4g, and CMIP6 MMMs. The larger mean LAI in the CLM4 (CMIP5 Ncycle MMMs) can be traced back to the overestimation of LAI in moist regions mentioned before. Some models show a significant negative trend in LAI in the SH and the tropics resulting in a globally negative trend even with a positive trend in the NH.

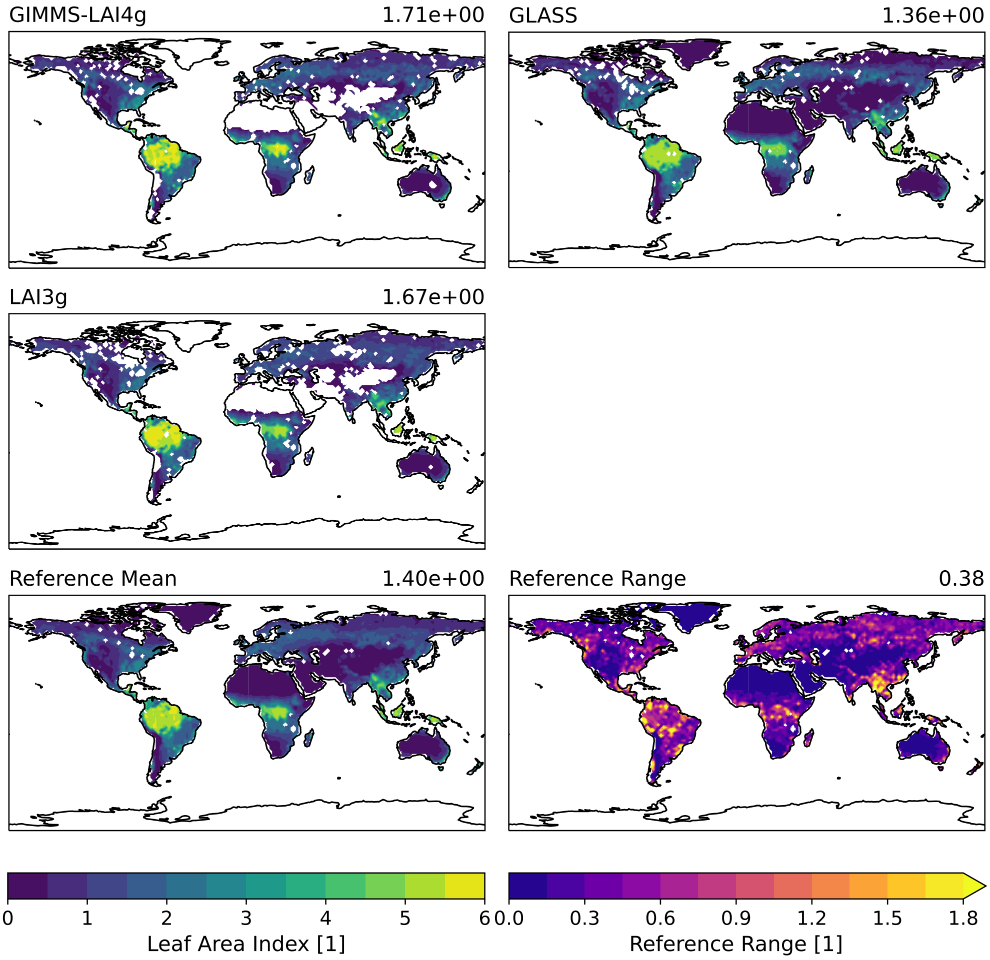


Figure 5.3.: Global maps of LAI averaged over 1986-2005 for all reference data sets, as well as the mean and range between lowest and highest values per grid cell of the reference data sets. Adapted with permission from **Gier** et al. 2024.

Maps of the LAI reference data are found in Figure 5.3, without the common mask to see the different coverages. Coverage of the different reference data sets varies a lot due to different quality control criteria and algorithms, with most missing values found in desert or mountainous regions such as the Sahara and the Himalayas. Additionally, a mean of the reference data and the range of the reference data per grid cell is shown, along with the global mean of the values in the upper right corner. GLASS has a larger coverage over desert and mountainous regions, which are regions with low plant coverage and thus low LAI, resulting in a lower global mean LAI of 1.36 m^2/m^2 compared to 1.67 m^2/m^2 and 1.71 m^2/m^2 from LAI3g and LAI4g respectively. This underlines the importance of the common mask used for Figure 5.1 and Figure 5.2 to obtain comparable results. The different data sets show the same pattern of LAI distribution, with the absolute values ranging between 1 and 6 m^2/m^2, while the differences are below 2 m^2/m^2 with the largest difference occurring in tropical rain forests and northern high latitudes, the regions with the largest absolute LAI values. For a gridcell bias comparison of the different model groupings (Figure 5.4) a combined reference data set was computed as the mean of the other reference data sets. Due to the different coverages, some areas are only calculated from the GLASS data, while others are an average of all three data sets. The range of values per gridcell going into the combined data set is plotted in the lower right of Figure 5.3. The largest difference occur in the areas with larger LAI as the tropical rainforests followed by boreal forests, with a global mean average range of 0.38 $\mathrm{m}^2/\mathrm{m}^2$. For the MMM bias maps shown in Figure 5.4 hatching is added where the MMMs agree with the reference mean within the MMM standard deviation. CMIP5 Ncycle MMMs show the issue of CLM4 in overestimating LAI in wet regions, with LAI in tropical rainforests almost doubling the reference value, while drier regions show a significant negative bias. While this results in a global mean bias of 0.66 to 0.67 $\mathrm{m}^2/\mathrm{m}^2$ smaller than 0.89 $\mathrm{m}^2/\mathrm{m}^2$ for CMIP5e non-Ncycle MMMs, it is still the worst performing model grouping when considering a gridcell basis due to its extreme biases in both directions. The hatching showing the agreement can be ignored in this case, as only two models contributed to the MMM std. The CMIP5e non-Ncycle MMM shows a strong overestimation across the northern latitudes besides Greenland. This can be attributed to the GFDL-ESM2G and GFDL-ESM2M models, which are known to have established coniferous trees in areas which should contain tundra or cold deciduous trees, as its vegetation spin-up only coniferous trees are allowed to grow in cold regions, but not grasses or deciduous trees which would have a lower LAI (Anav et al. 2013b). While the GFDL models show this problem in both CMIP5c and CMIP5e, due to the larger number of MMMs contributing to the concentration-driven simulations, their effect is reduced. In CMIP6, GFDL-ESM4 still has a large positive LAI bias throughout these areas, but it is significantly reduced compared to its CMIP5 predecessor. The second prominent overestimation is around the tropical rainforests, where models like BNU-ESM, MRI-ESM1, and in lesser extent also the GFDL models extend the LAI hotspot to larger areas around it compared to the reference mean. In CMIP5c HadGEM2-CC and HadGEM2-ES also show this overestimation. CMIP5c non-Ncycle MMM shows no special bias patterns, but instead a general overestimation in almost all areas with hatching present throughout the globe. CMIP6 MMMs show a reduced mean bias of less than half the CMIP5 overestimation, with almost no pronounced patterns and a bias reduction in all areas, with the largest improvements found in the northern high latitudes. The largest bias is in southeast asia for CMIP6c Ncycle MMMs which can be tracked to the CMCC-ESM2 and CMCC-CM2-SR5 models, and makes the mean bias of the CMIP6c Ncycle MMM higher than that of CMIP6c non-Ncycle MMM. Otherwise, the bias pattern looks similar for CMIP6 Ncycle and non-Ncycle models.

While CMIP6 LAI has improved compared to CMIP5, especially a significantly reduced mean bias, a general overestimation of LAI remains, along with issues of correctly reproducing the length of the growing season in the northern hemisphere, and a large model spread in mean and trend LAI. Neither the introduction of an interactive nitrogen cycle nor the comparison

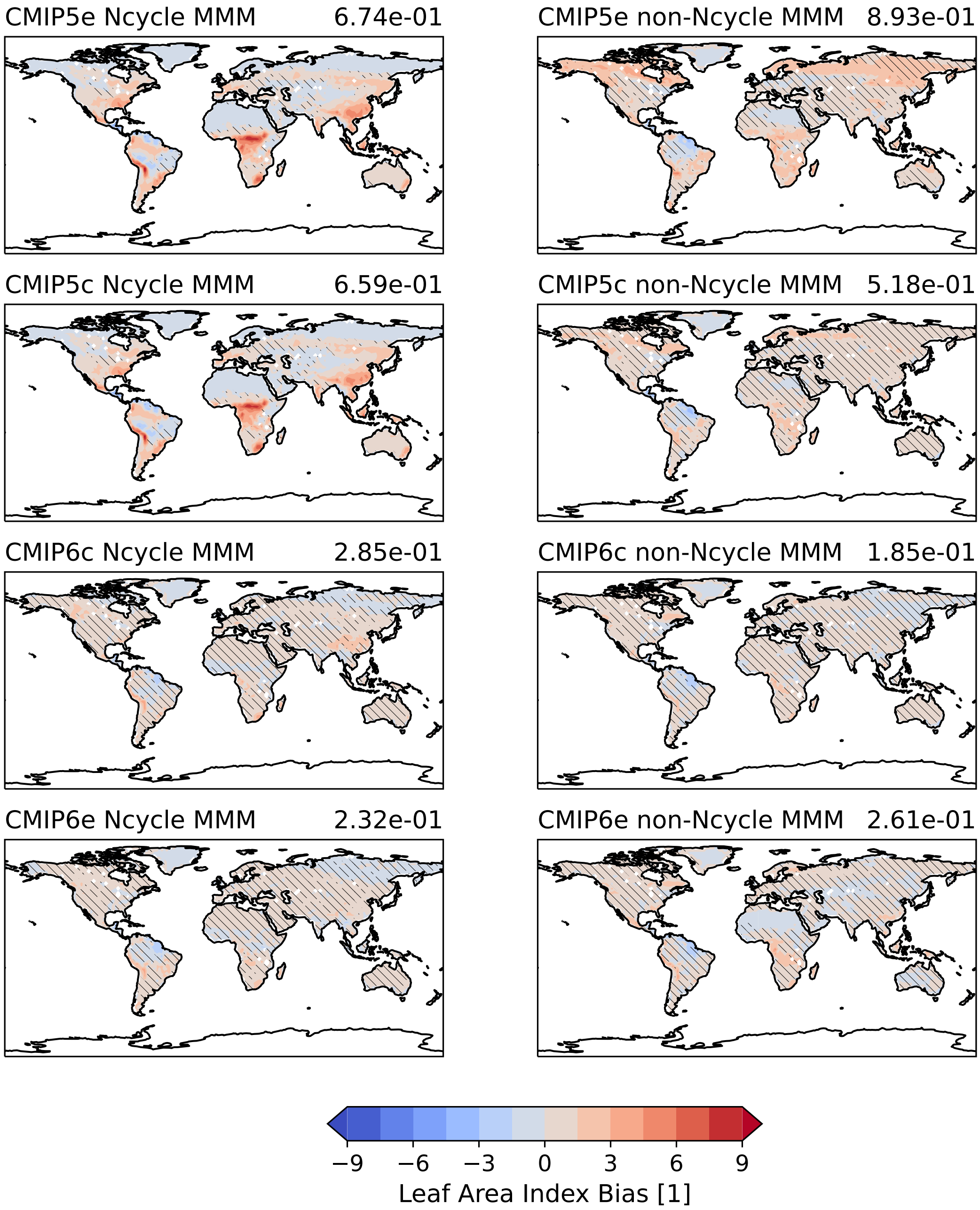


Figure 5.4.: Global maps of LAI bias for 1986-2005 with respect to the reference data set mean shown in Figure 5.3. The panels show the MMMs of the models with (left) and without (right) coupled nitrogen cycle for the different project-experiment combinations. The hatching represents the areas where the MMM of the models and reference mean agree within the MMM std. Adapted with permission from **Gier** et al. 2024.

between emission and concentration-driven simulations show large differences in the overall quality of the CMIP6 simulations for LAI.

5.2.2. Gross Primary Productivity

GPP represents the CO_2 uptake on land due to photosynthesis. This was one of the biggest weaknesses of the CMIP5 ensemble, with most models overestimating photosynthesis as well as leaf area index (Anav et al. 2013a). The seasonal cycle of GPP (Figure 5.5) shows good agreement between the GLASS and MTE reference data, while the FLUXCOM data show a lower GPP in all regions, as well as a shorter growing season in the Northern Hemisphere. All models reproduce a similar shape of the NH seasonal cycle to the GLASS and MTE data in both model generations and experiments. As found in Anav et al. 2013a the CMIP5 non-Ncycle models overestimate GPP in all regions, while the CMIP5 Ncycle models strongly underestimate the peak of the seasonal cycle in the NH. CMIP6 models perform better than CMIP5, but while the CMIP6 non-Ncycle models still overestimate the GPP peak in summer similarly to CMIP5 for the NH, the Ncycle models show a very good agreement for both CMIP6c and CMIP6e. Nitrogen limitation is stronger in northern latitudes through boreal forests and tundra (Du et al. 2020), as compared to tropical and subtropical forests, which are more limited by phosphorus. However, from the CMIP models used in this study, only ACCESS-ESM1-5 includes an interactive phosphorus cycle. It therefore makes sense that Ncycle models show a decreased GPP compared to non-Ncycle models in the NH, closer to reference data. The model spread remains large in CMIP6, albeit smaller for Ncycle models, which is denoted by horizontal hatching. Following LAI, there is no strong discernible seasonal cycle in neither the SH nor the Tropics, but the CMIP6 models are closer to the mean GPP than the CMIP5 models, with lower values for Ncycle models.

The temporal mean and linear trend of the spatial sums for GPP during the time period 1986-2005 is shown in Figure 5.6. MTE and FLUXCOM data are known to have an unrealistic non-existent trend (0.01 PgC yr^{-2} globally) due to the assumption of an unchanging average CO_2 level (Anav et al. 2015) in these data sets. As such, the model trend should not be compared to the trend of these two reference data sets, and we have omitted these data sets from the calculation of the reference trend. The mean GPP of all regions of these data sets along with all other numerical values from the plot can be found in Tables B.3 and B.4. The trend of GLASS (0.45 ± 0.09 PgC yr^{-2} globally) is closely linked to the high trend of LAI GLASS, one of the main influences on GPP. The reference data sets agree well in mean GPP with the largest difference being a lower mean for FLUXCOM in the NH. Globally the reference values range from 93.0 ± 0.4 PgC yr^{-1} for FLUXCOM and 102.6 ± 1.2 PgC yr^{-1} for MTE to 108.3 ± 3.4 PgC yr^{-1} for GLASS. The models show a large range from 83.5 ± 2.5 PgC yr^{-1} (BNU-ESM CMIP5e) to 152.0 ± 4.6 PgC yr^{-1} with an even larger mean for MRI-ESM1 CMIP5e as an outlier, which shows a large GPP in all regions. The CMIP5 models are on the higher side of this range, as seen by the MMMs of 120.6 ± 2.2 PgC yr^{-1} (non-Ncycle CMIP5c)

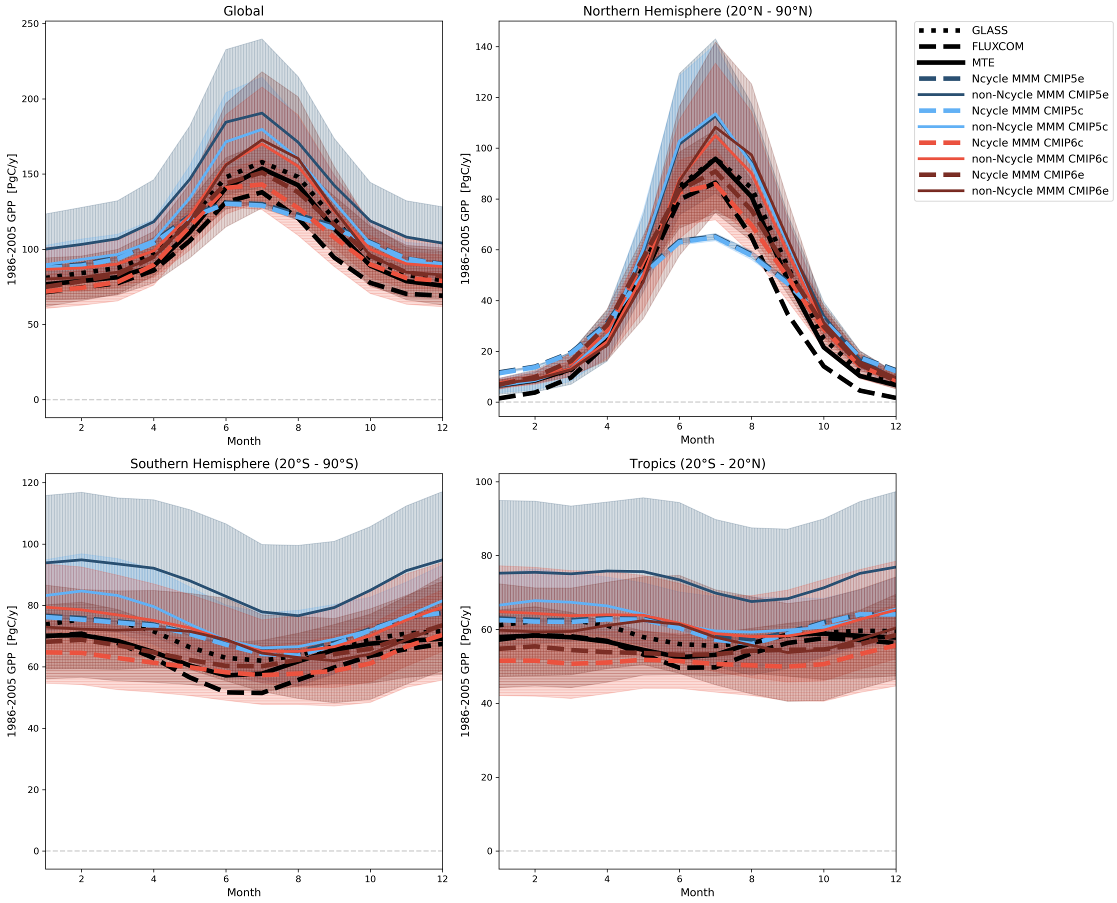


Figure 5.5.: As Figure 5.1 but for gross primary production using GLASS, FLUXCOM and MTE reference data. Additionally, the regional GPP is calculated as the area weighted sum instead of the mean of the gridcells used for LAI. Adapted with permission from **Gier** et al. 2024.

and 132.9 ± 2.1 PgC yr^{-1} (non-Ncycle CMIP5e) with the Ncycle CMIP5 models much lower at 106.5 ± 1.6 PgC yr^{-1} (Ncycle CMIP5c) and 107.3 ± 1.9 PgC yr^{-1} (Ncycle CMIP5e) due to their underestimation in the NH as seen in Figure 5.5. The CMIP6 Ncycle MMMs agree very well with the reference data, while the CMIP6 non-Ncycle MMMs show a larger mean GPP. The global trend of GLASS (0.45 ± 0.09 PgC yr^{-2}) is positive, which is well matched by the non-Ncycle CMIP6c MMM with the other MMMs showing a smaller trend. In the NH, more CMIP5 models overestimate the mean GPP than CMIP6 models. The MMMs span a large spread, with the Ncycle MMMs showing a lower mean GPP than the non-Ncycle MMMs which are overestimating GPP compared to the reference data. The models are clustered around the GLASS trend, with outliers for MRI-ESM2-0 in both its CMIP6c and CMIP6e runs. The CMIP6 non-Ncycle MMM shows a higher trend than the other MMMs due to the MRI-ESM2-0 outliers. In the SH mean GPP, the CMIP6 MMMs match the reference data well with lower values for Ncycle than for non-Ncycle MMMs. The CMIP5 MMMs are slightly above these, with the non-Ncycle CMIP5e having a much larger value due to the outlier of MRI-ESM1 and FIO-ESM both with values well above 100 PgC yr^{-1}. Compared to the GLASS trend of

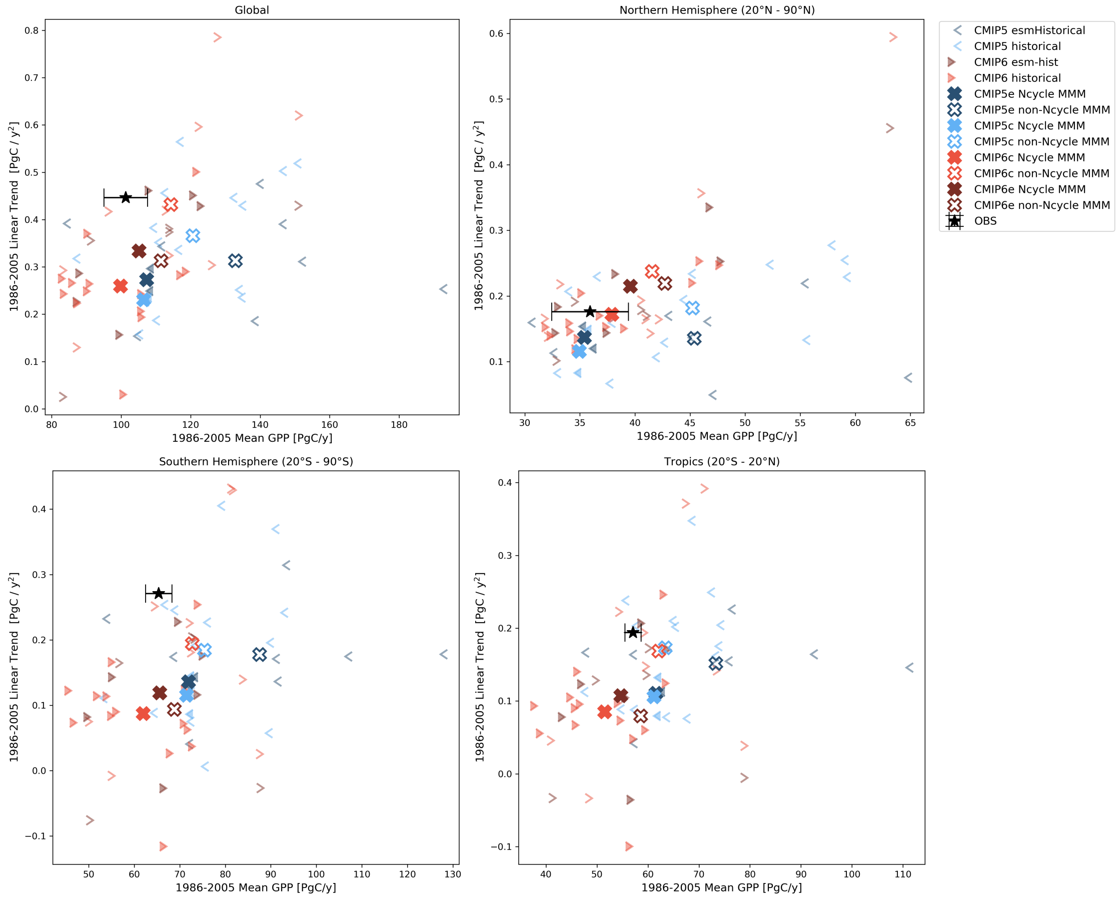


Figure 5.6.: As Figure 5.2 but for gross primary production using GLASS, FLUXCOM and MTE reference data. Adapted with permission from **Gier** et al. 2024.

0.27 ± 0.07 PgC yr^{-2}, the MMMs underestimate the trend. The distribution in the tropics is very similar as the SH. The GLASS trend of 0.19 ± 0.05 PgC yr^{-2} is underestimated by MMMs. In summary, the Ncycle MMM shows a better performance than the non-Ncycle MMM in the NH, while it shows a slight underestimation in the tropics and a similar performance in the SH. The model spread over the trend in GPP stays similar throughout the model generations, with the mean trend being largely consistent with the GLASS reference data set in all regions but underestimated everywhere but in the NH.

As for LAI, the coverage of the GLASS data is larger than for the other reference data sets, with missing values for FLUXCOM and MTE mainly found over the Sahara and the Himalayas (Figure 5.7), GLASS shows a larger GPP in the tropical rainforests and boreal forests, explaining the larger mean GPP seen before. The reference mean data set has hardly any missing values left and the largest difference in the data sets are at places with the highest GPP, with discrepancies in the areas bordering the hotspots of the rainforests and boreal forests. Even though the GPP bias maps (Figure 5.8) for Ncycle CMIP5 MMMs have a global mean bias almost two magnitudes smaller than the CMIP5 non-Ncycle MMMs, they show the same pattern of overestimation in wet regions and underestimation in dry regions

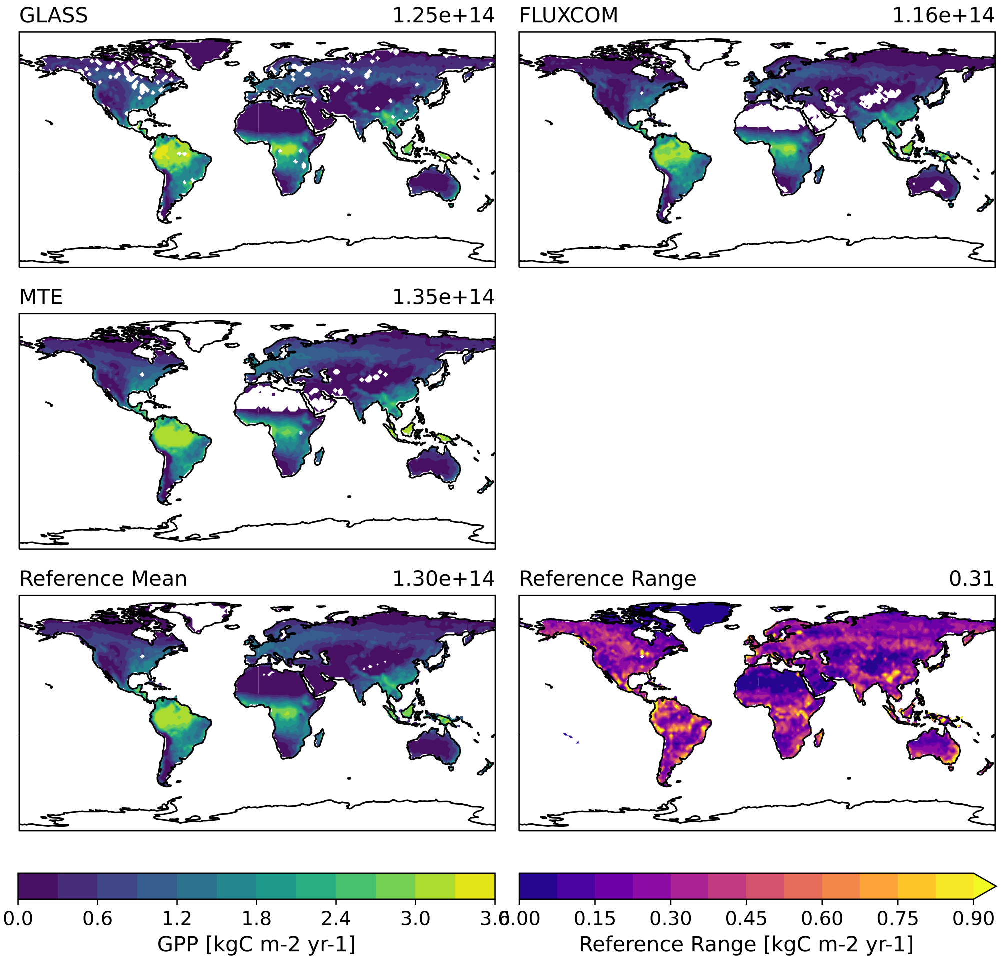


Figure 5.7.: Global maps of GPP averaged over 1986-2005 for all reference data sets (GLASS, FLUXCOM and MTE), as well as the mean and range between lowest and highest values per grid cell of the reference data sets. The number in the top right denotes the global GPP flux. Adapted with permission from **Gier** et al. 2024.

found for LAI (Figure 5.4), underlining the strong influence of LAI on GPP. The CMIP5 non-Ncycle MMMs also shows similar patterns to the LAI bias maps, but the overestimation in the areas around the tropical rainforests is strongly reinforced, while the previous strong overestimation of LAI in the northern high latitudes for the non-Ncycle CMIP5e MMM is reduced. The CMIP5c non-Ncycle MMM additionally shows a strong underestimation at the northeastern coast of South America. The global mean bias for CMIP5 non-Ncycle MMMs lies at $3.6{\cdot}10^{-13}$ PgC m^{-2} yr^{-1} for CMIP5e and is reduced by half for CMIP5c. This bias is further reduced to approximately $1.0{\cdot}10^{-13}$ PgC m^{-2} yr^{-1} for CMIP6 non-Ncycle MMMs, which show similar bias patterns to CMIP5 non-Ncycle MMMs but overall reduction to the bias patterns, with a larger reduction in the savannah regions of Africa. The global mean bias for the CMIP6 Ncycle MMMs is further reduced to $0.01{\cdot}10^{-13}$ PgC m^{-2} yr^{-1} for CMIP6e and a negative bias of

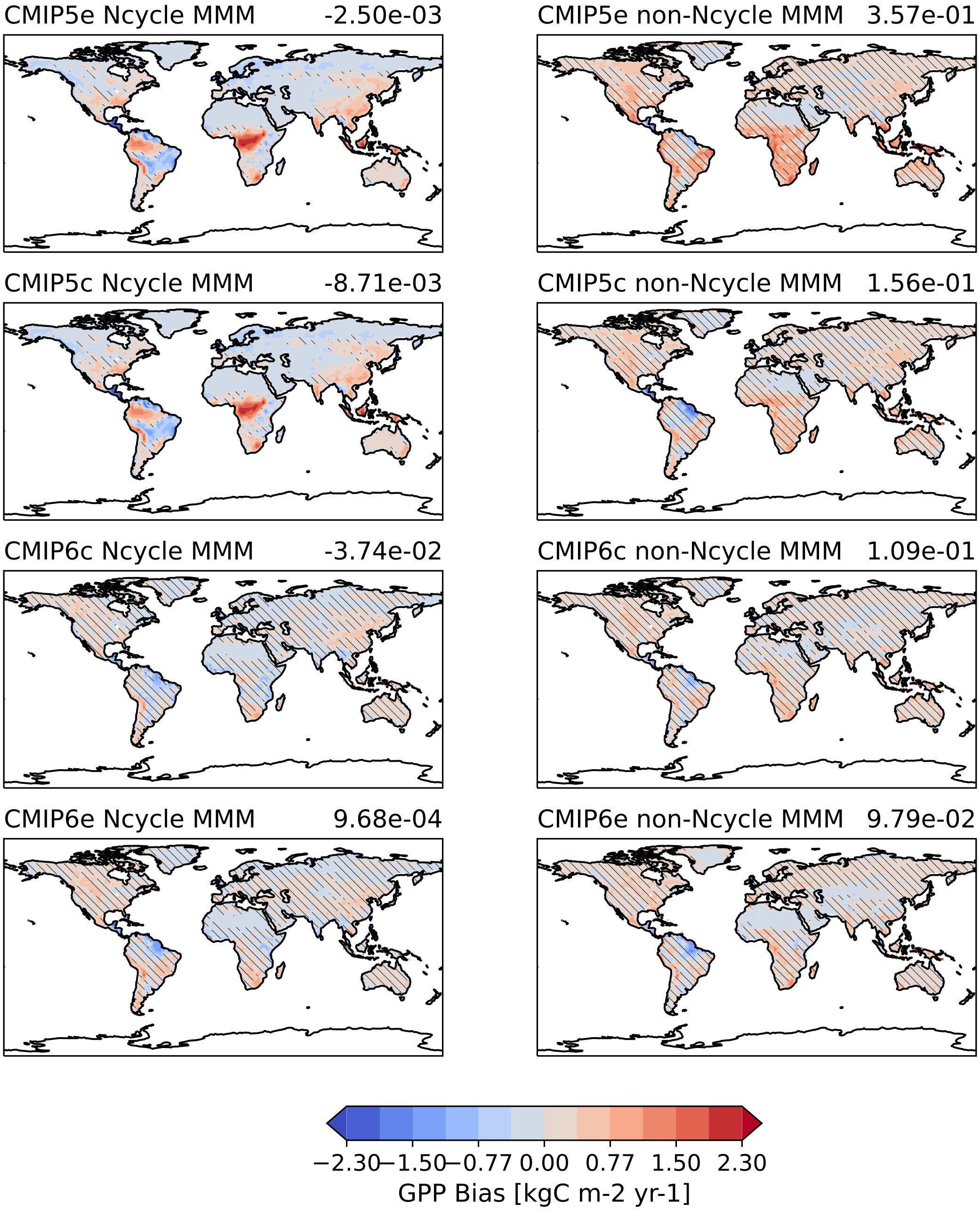


Figure 5.8.: Global maps of GPP bias for 1986-2005 with respect to the reference data set mean shown in Figure 5.7. The panels show the MMMs of the models with (left) and without (right) coupled nitrogen cycle for the different project-experiment combinations. The hatching represents the areas where the MMM of the models and reference mean agree within the MMM std, while the number in the top right denotes the global mean bias. Adapted with permission from **Gier** et al. 2024.

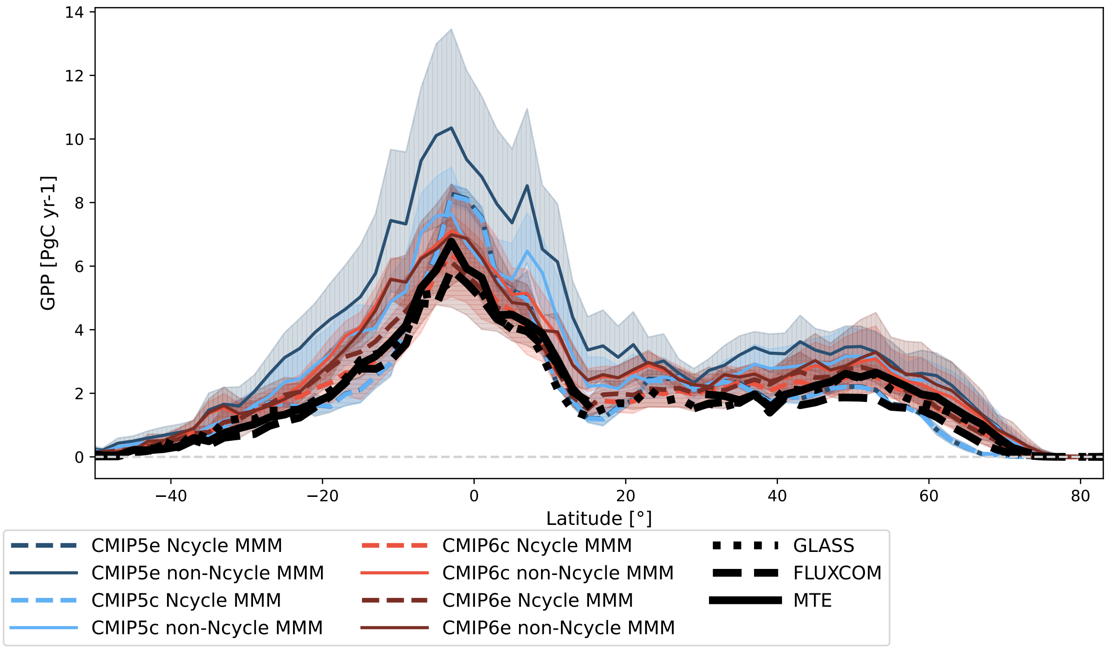


Figure 5.9.: Area weighted zonal sums of GPP and the reference data sets GLASS, FLUXCOM and MTE. No common masking is applied, but latitudes were set to missing if more than 15% of the land grid cells contained missing data. The hatching depicts the MMM standard deviation, with a horizontal hatching for models with and vertical hatching for models without interactive nitrogen cycle. Adapted with permission from **Gier** et al. 2024.

$-0.4\cdot10^{-13}$ PgC m^{-2} yr^{-1} for CMIP6c. Both show a reduction in the northern hemisphere bias of the open shrublands, turning some into a negative bias, as well as southwest Africa, while the slight overestimation in North America remains, as well as the underestimation at the north eastern part of South America. This is summarized in Figure 5.9, which shows the zonal sums of the reference data and the MMMs. Unlike the seasonal cycle and scatterplots shown before, a common mask is not applied here, but instead values are masked out if a data set has more than 15% of a latitudes land points set to missing values. The large overestimation of non-Ncycle CMIP5e can be seen which is reduced in non-Ncycle CMIP5c, with both showing a peak slightly north of the equator which is not seen in the reference data and which is due to the overestimation of the shrublands south of the Sahara. The CMIP6 non-Ncycle MMMs show a much better approximation across all latitudes, with a slight reduction of the bias in the NH, but still show a significant overestimation in the tropics. This is remedied in the CMIP6 Ncycle models, which show a very good agreement with the reference data across all latitudes, now with slight underestimations at high latitudes.

5.2.3. Land-Atmosphere Flux

The net carbon flux from the atmosphere into the land (net biome productivity, NBP) characterizes the balance between carbon uptake due to photosynthesis and carbon release by

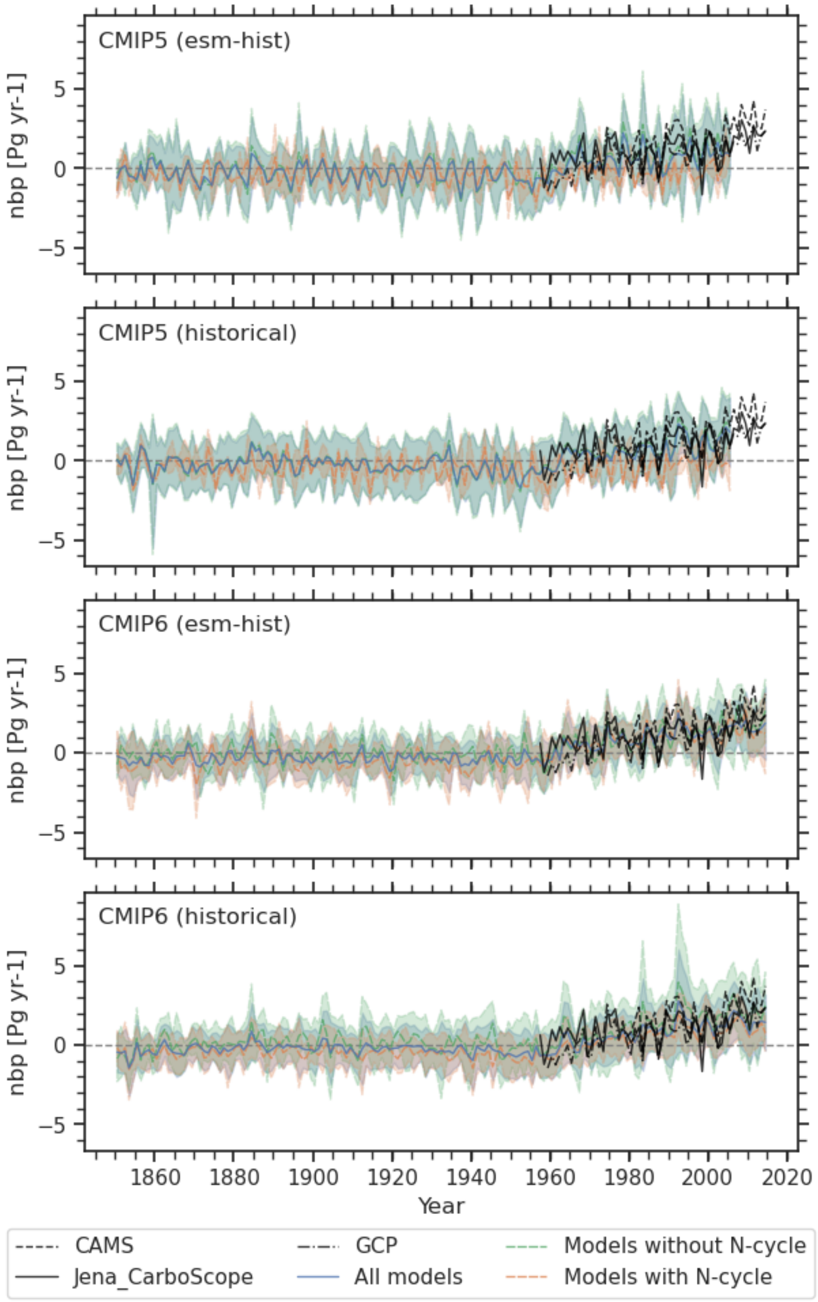


Figure 5.10.: Global Land-Atmosphere carbon flux time series for CMIP5 (top two panels) and CMIP6 (bottom two panels) concentration (panels 2 & 4) and emission-driven (panels 1 & 3) historical simulations. Model results are separated into MMMs of models with (orange dashed line) and without interactive nitrogen cycle (green dashed line), with the MMM of all models shown in blue and the reference data sets CarboScope (solid), CAMS (dashed), and GCP (dash-dotted) shown in black. The standard deviation of the MMM is given by the shaded areas. Adapted with permission from **Gier** et al. 2024.

respiration, as well as other processes like fires and de- and afforestation. Positive values of NBP denote carbon uptake by land. The CMIP6 EC-Earth models (EC-Earth3-CC, EC-Earth-Veg) are excluded from the MMM for NBP because they show a very strong land source in December in seemingly random grid cells all over the globe. MIROC-ESM and MIROC-ESM-CHEM are also removed for a similar reason: they contain grid cells which seemingly randomly show large sources and sinks popping up in random months. Due to their appearance at random months instead of only in December like in EC-Earth, it does not influence regional means or climatologies as much, but can be seen in mean map plots very well.

Figure 5.10 shows the global evolution of the land-atmosphere flux for CMIP5e, CMIP5c, CMIP6e and CMIP6c simulations in order from top to bottom. Each panel shows the MMM of all models with simulations in the respective project and experiment combinations (blue line), as well as the two MMMs of models with (dashed orange) and without (dashed green) nitrogen cycle. The colored shading represents the standard deviation of their respective MMMs. For comparison with observations, data from CAMS (dashed black), Jena CarboScope (solid black), and GCP (dash-dotted black) are added to each panel. There is a large year-to-year variability, which can also be seen in the models. All project-experiment combinations agree with all reference data sets, with the CMIP6 simulations showing a better agreement in the 1990s, during which the CMIP5 models generally underestimate the land carbon sink. From 1850 to 1970 the models do not show a significant carbon source or sink, only the CMIP6e MMM shows a small carbon source in this time period. However, the model variance as indicated by the shaded areas is large enough to be in agreement with a neutral state. Since the 1980s the land has been acting as a carbon sink which is increasing over time. This increase has previously been attributed mainly to the fertilization effect from rising atmospheric CO_2 concentrations (Canadell et al. 2021). In the CMIP5 simulations, Ncycle and non-Ncycle models do not show any significant differences before 1980, after which the Ncycle models show a slightly lower carbon flux. Ncycle CMIP6 models show a slightly lower land-atmosphere carbon flux over the full time period compared to non-Ncycle models.

The global seasonal cycle for NBP (Figure 5.11) is dominated by the northern hemisphere, with almost no discernible cycle in the southern hemisphere or the tropics. There is generally a good agreement between the two inversions, but the CAMS inversions shows a larger seasonal cycle in the SH and tropics of approximately 6 PgC yr^{-1}, where CarboScope shows no clear seasonal cycle. In the NH, and due to its large contribution to the total also globally, CAMS has a higher NBP at the start of the year and to a lesser degree at the end of the year, where CarboScope shows a larger negative NBP and thus carbon sink. The models agree with the carbon sink of CarboScope in these months, but have a weaker carbon sink (higher NBP) in NH autumn. The CMIP5 Ncycle MMMs have a smaller seasonal cycle amplitude compared to any of the other MMMs and the reference data, carried over from GPP. The other MMMs reproduce the seasonal cycle well, while the non-Ncycle CMIP6e MMM is shifted late by a month, showing possible issues with the start and end of the growing season. In the SH and tropics, where CarboScope found no significant cycle and CAMS had a slightly larger one

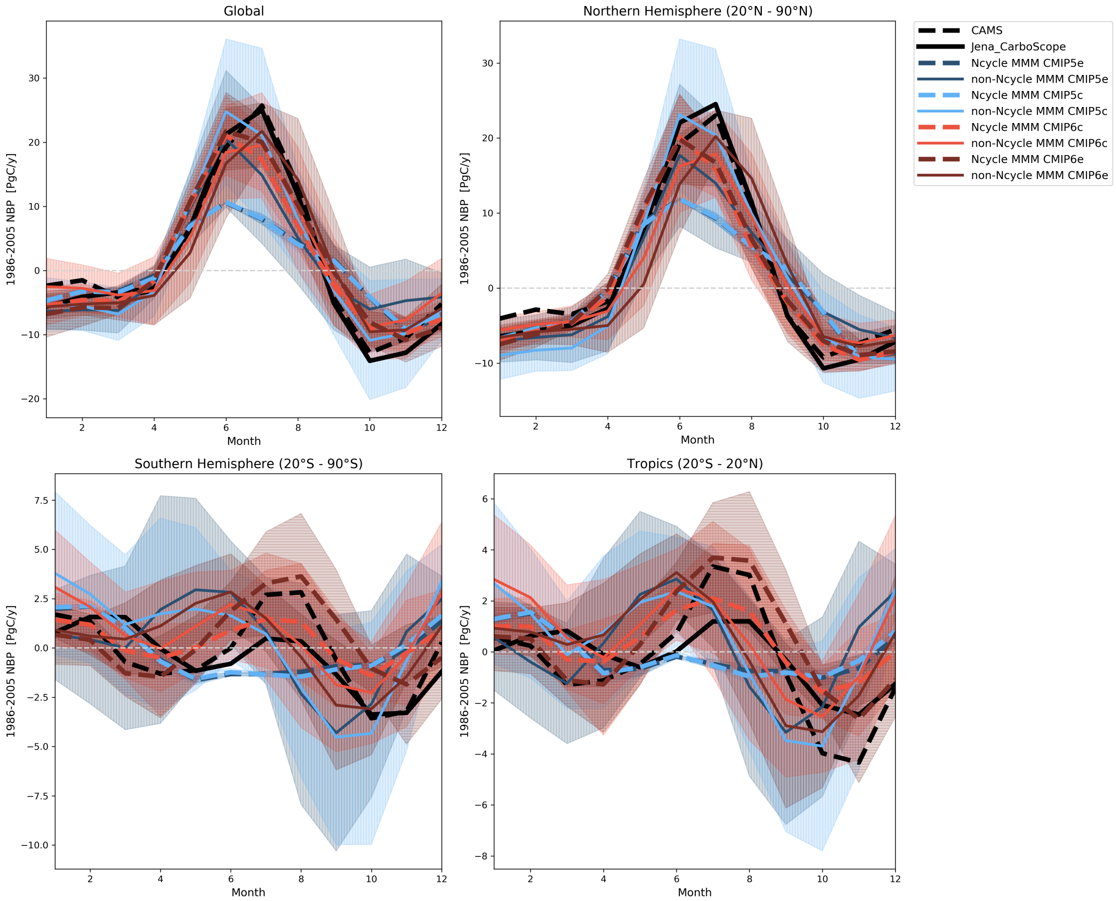


Figure 5.11.: As Figure 5.1 but for land-atmosphere carbon flux and the reference data sets CAMS and CarboScope. Adapted with permission from **Gier** et al. 2024.

in the tropics, the Ncycle models follow the shape and timing of the CAMS data, while the non-Ncycle models have a seasonal cycle shifted to two months earlier. There is no significant difference between CMIP5 and CMIP6 nor between c and e experiments in the MMM.

The temporal mean and trend of spatially summed NBP is shown in Figure 5.12 with numbers given in Tables B.5 and B.6. CAMS shows a larger mean NBP compared to CarboScope in all regions but the tropics, which is consistent with the NBP averages from Seiler et al. 2022 using this data set and who found its NBP to be larger than comparable data sets and model results. GCP as a global average is only available as reference data set for the global panel. Globally, the reference data sets have a mean NBP of 0.71 ± 0.94 PgC yr^{-1} for CarboScope, 1.00 ± 0.84 PgC yr^{-1} for GCP and 1.72 ± 0.94 PgC yr^{-1} for CAMS. The models show a far larger range with outliers for the INM-CM CMIP6c models. The Ncycle CMIP5 means show negative trends, while the other MMMs range between 0.84 ± 0.34 PgC yr^{-1} (Ncycle CMIP6c) and 1.56 ± 0.62 PgC yr^{-1} (non-Ncycle CMIP6c), with Ncycle MMMs showing significantly smaller mean NBP similar to the GCP reference, while the non-Ncycle MMMs have a larger NBP but still smaller than the CAMS data. The relatively good overall agreement of the models' mean NBP with the reference data does not hold for the different regions. Most models and all MMMs

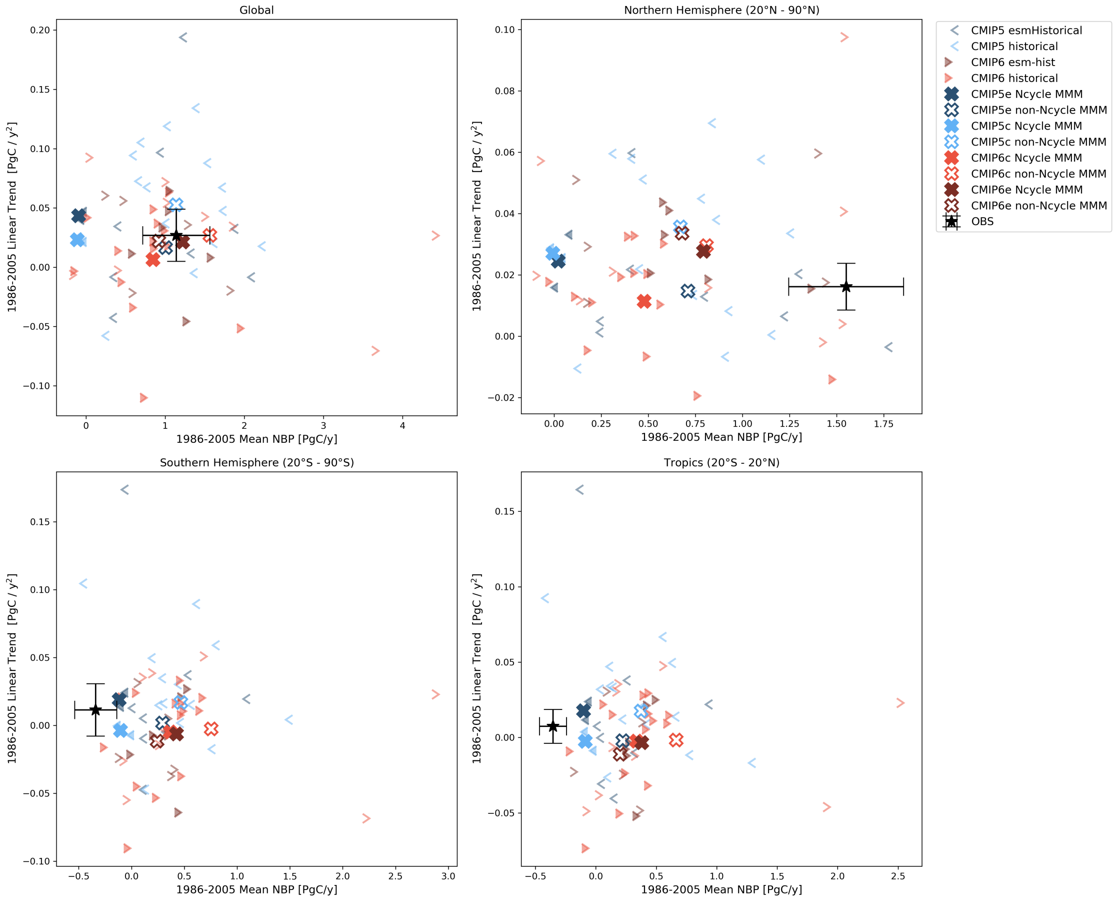


Figure 5.12.: As Figure 5.2 but for land-atmosphere carbon flux and the reference data sets CAMS and CarboScope. The GCP data are a globally averaged timeseries and thus only appears in the global plot. Adapted with permission from **Gier** et al. 2024.

simulate a lower carbon sink in the northern hemisphere when compared to the inversions, with Ncycle models generally showing a smaller mean NBP, but no large discernible differences between the different groupings. Conversely, while the inversions estimate both the southern hemisphere as well as the tropics to be a slight carbon source due to deforestation, the MMMs with the exception of the CMIP5 Ncycle show a carbon uptake by land in these regions. The large values for the non-Ncycle CMIP6c MMM are again due to the overestimation of the INM-CM4-8 and INM-CM5-0 models, but as their mean NBP in the NH is not a large outlier, we did not remove these from the MMM. The underestimation in the NH combined with the overestimation in the SH and tropics leads to the good global agreement of the total carbon sink. This is in agreement with the findings from IPCC AR6 (Canadell et al. 2021; Eyring et al. 2021). The inclusion of a nitrogen cycle and therefore the inclusion of nitrogen limitations on CO_2 fertilization was expected to address this discrepancy of the distribution of the carbon sinks (Canadell et al. 2021), but the data do not support this, as the Ncycle MMMs do not show a different performance to the non-Ncycle MMMs in CMIP6. While the models show a large range of trends, the MMMs agree well with the reference data, and this continues

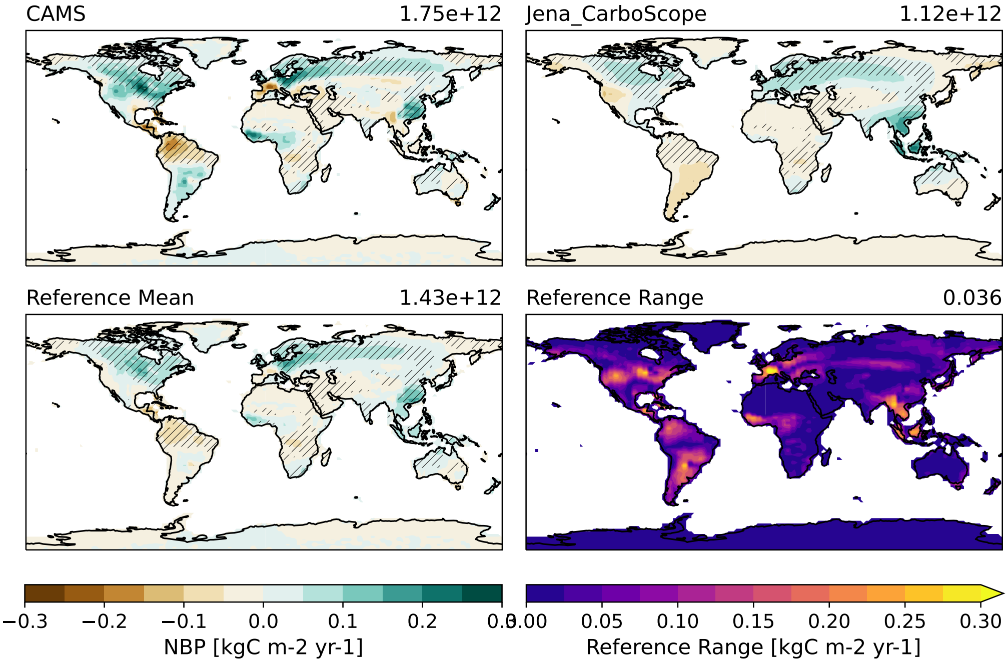


Figure 5.13.: Similar to Figure 5.7 but for land-atmosphere carbon flux. Additionally, the forward slash hatching symbolizes areas where the reference data sets agree on the sign or where the difference is smaller than the size of one bin of the contour plot. Adapted with permission from **Gier** et al. 2024.

in the other regions as well. While N-limitation is not expected to be substantial at present day, it represents a major limitation on future land-carbon uptake (Zaehle et al. 2015), and thus its inclusion a major advance in being able to robustly simulate future carbon balance of the terrestrial carbon cycle.

As seen before, the reference data sets show different means, trends and slightly different seasonal cycles in the different regions. For a more detailed look, Figure 5.13 shows maps of the reference data. The hatched area is the area where the data sets agree on the sign or within a margin of half the bin size of the contour plot. While in large parts of the globe the data sets agree in sign, there are significant differences. In North America CAMS shows a much larger carbon sink than CarboScope which instead shows some carbon sources along the west coast and throughout South America. In CAMS South America is split into a much stronger carbon source in the amazonian rainforests and a carbon sink in the southern part. The data sets also disagree in Europe, which is a carbon source according to CAMS but a sink in CarboScope. Literature found Europe to be a carbon sink for the first part of the 21st century (Ciais et al. 2013; Reuter et al. 2014) using different reference data sets. This would support the CarboScope data set but due to the different time frames considered it is not definitive. CAMS also sees a carbon sink in tropical Africa where CarboScope has a slight carbon source. In South East Asia, CarboScope has a large carbon sink, where CAMS shows a neutral carbon flux. The area of most agreement is in the NH as a large carbon sink where there is no deforestation and a bit of aforestation. The amazon region was a large carbon sink which is becoming a source

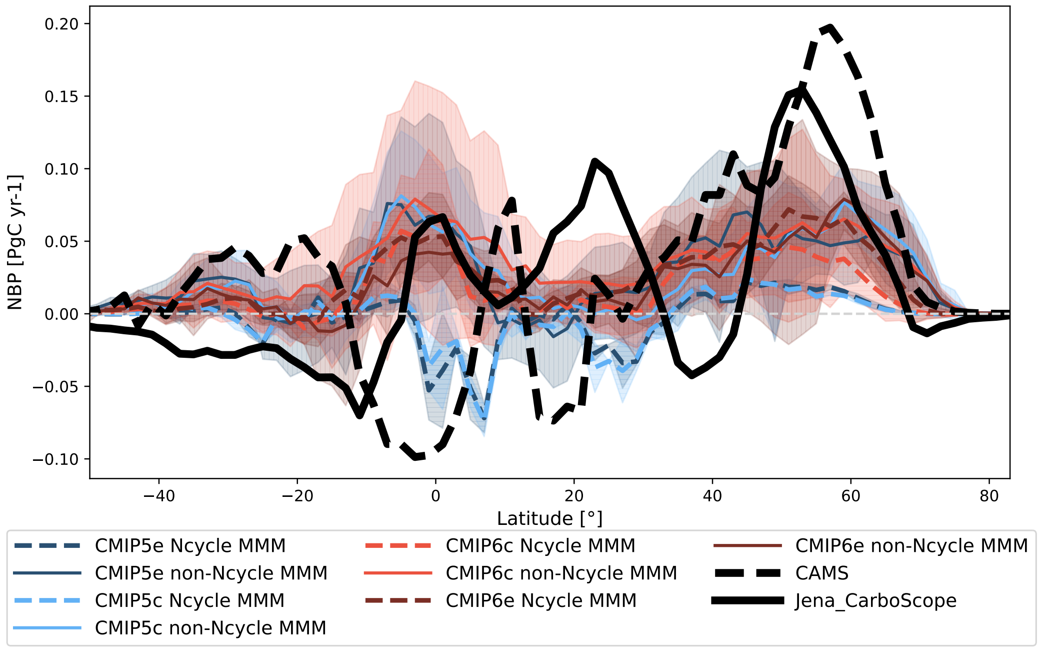


Figure 5.14.: As Figure 5.9 but for land-atmosphere carbon flux with CAMS and CarboScope reference data. Adapted with permission from **Gier** et al. 2024.

due to deforestation (Gatti et al. 2021). CAMS sees the amazon as a strong carbon source, while CarboScope shows a smaller source, with a sink in the northwestern region. Keenan and Williams 2018 found inverse models to show south America as both a carbon source, as well as a carbon sink and is thus a hotly debated area. Kou-Giesbrecht et al. 2023 attribute the weak agreement between CarboScope and CAMS to differences in the inversion models and atmospheric CO_2 measurements used, with larger differences at latitudes with smaller land areas. Due to the difference between the observational data sets, the bias maps to a reference mean shown for the other considered variables so far have been omitted. Instead, the area weighted zonal sums are plotted in Figure 5.14 for comparison of the MMMs with both reference data sets. The reference data sets disagree for almost all latitudes, thus making the model comparison to the reference data in these regions not very meaningful. The area where the reference data are in most agreement is in the northern high latitudes (50–80 °N), where both of them show a strong carbon sink, about double of that shown in both CMIP5 and CMIP6 models. While the issues in CLM4 and thus the CMIP5 Ncycle models are clearly visible (similarly to LAI and GPP), the other MMMs show similar NBP in all latitudes.

5.2.4. Carbon Stocks

Another large uncertainty in CMIP5 was the amount of carbon stored in soil and vegetation. This leads to large uncertainties in land-use change emissions which are important for quantifying cumulative emissions as well as climate mitigation strategies (Friedlingstein et al. 2023).

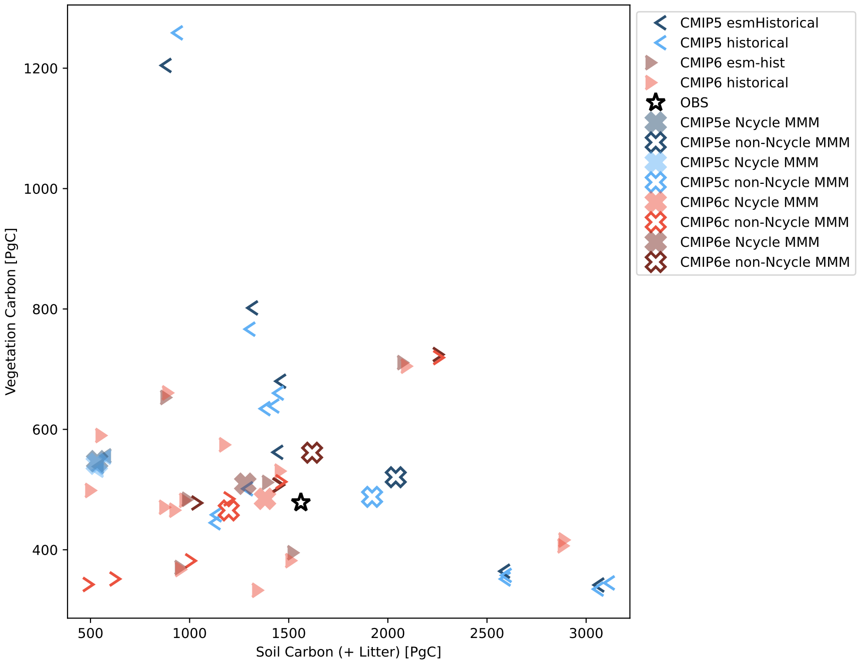


Figure 5.15.: Scatter plot of global mean vegetation and soil carbon over 1986-2005, with observations from NDP (vegetation carbon) and HWSD+NCSCD (soil carbon). As in Figure 5.2, filled symbols denote models with nitrogen cycle. Adapted with permission from **Gier** et al. 2024.

Varney et al. 2023b investigated the carbon-climate feedbacks of soil and vegetation carbon and found soil carbon to be the dominant response of the land surface, highlighting the need to reduce the uncertainty in carbon storage to better quantify future changes of the climate system. Figure 5.15 shows a scatterplot of the global sums vegetation against the combined soil and litter carbon. The observational soil (HWSD+NCSCD) and vegetation carbon (NDP) data sets are derived from in situ measurements taken over a long period of time and are thus given without a time coordinate, while the models were averaged over 1986-2005. Note that some models (CanESM5-CanOE CMIP6c, GFDL-ESM4 CMIP6e, INM-CM4-8 CMIP6c, INM-CM5-0 CMIP6c, FIO-ESM CMIP5e, CanESM2 CMIP5c and inmcm4 CMIP5c) did not have data on the ESGF nodes for soil or vegetation carbon and are thus missing from the carbon stocks analysis. BNU-ESM CMIP5c and CMIP5e shows a far larger vegetation carbon than the other models and is thus removed in the calculation of the mean. Additionally, CLM5 and thus CESM2, CESM2-WACCM, NorESM2-LM, and NorESM2-MM include a full vertical soil profile. For these models, the *cSoilAbove1m* variable is used for better comparison with the other models, as done in Varney et al. 2022. The large spread in the global carbon stocks still remains in CMIP6 as shown in Figure 5.15 with values for each data set listed in Table B.7. In CMIP5 vegetation carbon was spread between 335 PgC (MPI-ESM-LR CMIP5c) and 802 PgC (GFDL-ESM2M CMIP5e) with the outlier of BNU-ESM even reaching values above 1200 PgC. The spread has only marginally been reduced in CMIP6 to a range of 333 PgC (EC-Earth3-Veg CMIP6c) to 724 PgC (CNRM-ESM2-1 CMIP6e). The reference data are at a value of 478 PgC, in the lower range of the models, with the MMMs ranging between 465 PgC (non-Ncycle CMIP6c) and 547 (Ncycle CMIP5e). The spread in soil carbon is even larger with a CMIP5 range of 513 PgC (CESM1-BGC CMIP5c) up to 3092 PgC (MPI-ESM-MR CMIP5c). The overestimation by MPI-ESM is due to its decomposition parameterization depending on soil moisture and showing maxima in continental dry lands. In CMIP6 MPI-ESM1-2 the soil carbon model was changed to YASSO which simulates more plausible soil carbon content (Mauritsen et al. 2019). The spread in soil carbon was not significantly reduced in CMIP6 with a range of 514 PgC (GFDL-ESM4 CMIP6c) to 2913 PgC (CMCC-ESM2 CMIP6c), with the reference value for HWSD+NCSCD at 1561 PgC. The CMIP5 Ncycle models have a soil carbon on the lower end of the range, consistent with the CMIP6 models TaiESM1 and SAM0-UNICON which also use CLM4. The CMIP5 Ncycle MMMs are on the very low end of the range with 532 PgC and 534 PgC for CMIP5 and CMIP5e respectively, while the other MMMs are closer to the reference data and range between 1197 PgC for non-Ncycle CMIP6c and 2040 PgC for non-Ncycle CMIP5e. While the CMIP6 Ncycle MMMs are closer to the reference data, no significant improvement due to the inclusion of the interactive nitrogen cycle can be seen when considering the whole spread of the models. This is consistent with Wang et al. 2022 who found that changing models from C to CN coupling often result in lowered ecosystem storage, but due to different parametrizations simulate similar carbon pools. Varney et al. 2022 suggest that much of the uncertainty in carbon stocks is due to the simulation of belowground processes - this is backed up by the differences in soil carbon being much greater than in GPP, and thus implicating differences in simulated residence times (Carvalhais et al. 2014; Todd-Brown et al. 2014). For a more in-depth discussion on we would like to refer to dedicated studies, such as Varney et al. 2022 and Wei et al. 2022. Furthermore, Varney et al. 2023a found that while the CMIP6 future soil carbon projections have a lower model spread compared to CMIP5, the structure of soil carbon models within CMIP6 ESMs has likely contributed towards this reduction.

5.2.5. Overall Model Performance

In this Section we assess the overall performance of CMIP5 and CMIP6 models with respect to carbon cycle variables. Figure 5.16 shows a performance metrics (portrait) plot similar to Gleckler et al. 2008. It is produced by calculating the normalized relative space-time RMSD of the climatological seasonal cycle of a model variable with respect to a reference observation. The normalization is done relative to the ensemble median of both CMIP5 and CMIP6 models, with positive values (red) denoting a higher RMSD and thus worse performance while negative values (blue) denote a lower RMSD than the ensemble median and thus a better performance. As the carbon stocks from the observations do not vary in time, the calculation of the RMSD as done here is not meaningful and thus only NBP, GPP and LAI are shown in the plot for all four considered regions (global, northern hemisphere, southern hemisphere, tropics). MMMs for both Ncycle and non-Ncycle models were added, with the models which were excluded

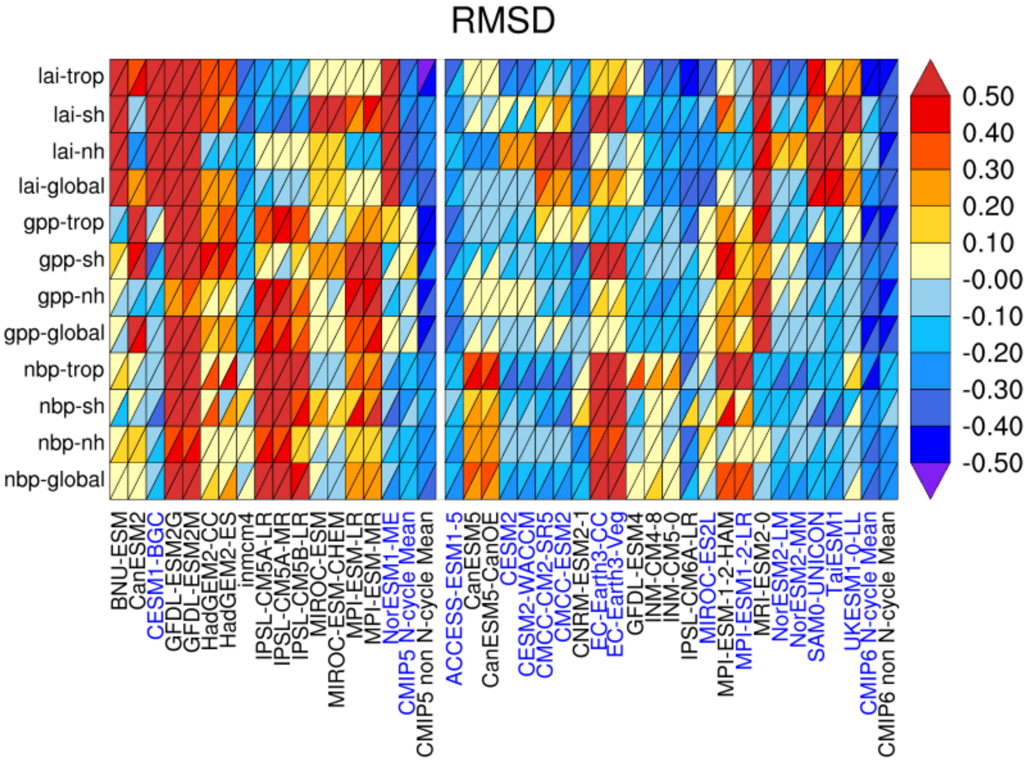


Figure 5.16.: Relative space-time root-mean-square deviation (RMSD) performance metrics for CMIP5 (left) and CMIP6 (right) concentration-driven simulations for variables relevant to the carbon cycle compared to reference data sets. Blue shading indicates a performance better than the median RMSD of all models in the plot, while the redder the color, the worse the performance. The RMSD is normalized relative to the median of all models. The considered time periods depend on the start of the observational data (see Table 5.3) and end in 2005 to accommodate the end of the CMIP5 data. When using two observational references, a diagonal split is introduced, with the default reference data set being shown on the lower right, while the alternate data set are used for the top left triangle. The default and alternate reference data sets are marked in Table 5.3 and are as follows: LAI: LAI4g (main), GLASS (alt); GPP: FLUXCOM (main), GLASS(alt); NBP: CarboScope (main), CAMS (alt). Adapted with permission from **Gier** et al. 2024.

from MMMs due to various issues as stated in the previous sections also removed from the MMM here. Note that the MMMs were calculated on the climatologies prior to calculation of the RMSDs, so over- and underestimations can cancel each other out. This is the standard for the performance metrics plot implemented in ESMValTool and kept for consistency. Variables with two reference data sets show the main reference in the lower right triangle, while the alternate reference is shown for the upper left triangle. data sets marked in bold in Table 5.3 are the main references. CMIP5 models are shown on the left and CMIP6 models on the right, with figures for both concentration-driven (Figure 5.16) and emission-driven models (Figure 5.17). Models with a nitrogen cycle are marked with blue labels.

Most models have similar scores when compared to the different observations for GPP and LAI, showing that the inter-model spread is CMIP6 is larger than the observational uncertainty

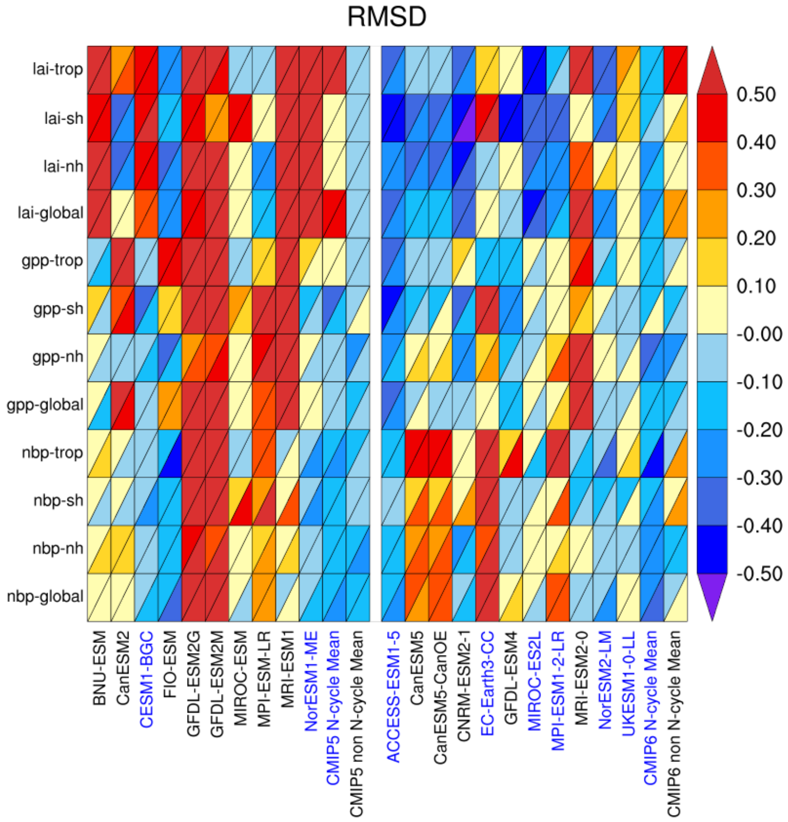


Figure 5.17.: As Figure 5.16 but for emission-driven simulations. Adapted with permission from **Gier** et al. 2024.

in these variables. For NBP however, models can have different scores to the considered reference data, which is due to the difference in the reference data found in the previous sections. Models on average perform much better than CMIP5 models, with models that had a predecessor in CMIP5 improving on their CMIP5 performance in almost all variables, such as GFDL and IPSL in all variables and CESM and NorESM in LAI, with the exception of CanESM which shows a reduced performance for NBP in CMIP6. Large improvements can be found in all variables going from CMIP5 to CMIP6, especially in LAI with the exception of the models using the older CLM4 land component (SAM0-UNICON and TaiESM1) in CMIP6 and GPP, which were previously identified as weaknesses in CMIP5. Only the MRI-ESM2-0 model shows a bad performance in both these variables. As mentioned before, dynamic vegetation in models plays a large role in their ability to simulate variables directly related to it, with models interactively simulating vegetation cover (marked with a D in Table 5.1 and Table 5.2), showing a below average score for LAI and GPP RMSD.

Models which were remarked upon in the previous sections as having good or bad agreement with the observations in specific areas, such as the NBP problems in December for EC-Earth3 have RMSDs that reflect these statements, making this metric a well-suited measure for overall performance. Most models have similar RMSDs in the different regions, with the global value reflecting a mean of the different regions. There does not seem to be a qualitative difference between Ncycle and non-Ncycle models as a whole, but the MMMs perform better than any individual model. The better performance of the MMMs is mathematically expected as long as the assumption that both observations and model simulations draw from the same distribution holds true (Christiansen 2018). The global NBP, LAI, and GPP are also found in Figure 42 of chapter 3 of the IPCC AR6 (Eyring et al. 2021), which showed not only carbon cycle variables but also other land, ocean and atmosphere variables averaged over 1980-1999 for comparison across models from CMIP3 to CMIP6. Their results are compared to reference data from JMA-TRANSCOM for NBP, LAI3g for LAI, and MTE and FLUXCOM for GPP. As other than NBP these reference sets are the same as the ones considered in this paper, the results are also the same. For NBP despite the different data set, the performance of the models is very similar to the one found for CAMS, our alternative data set. The ILAMB benchmark used in chapter 5 of the IPCC AR6 (Canadell et al. 2021) also comes to the conclusion of model improvement from CMIP5 to CMIP6. No qualitative difference can be found between models that have both emission-driven and concentration-driven simulations compared to models with only concentration-driven simulations, and models with both simulations have similar RMSDs in both. This indicates that carbon exchanges are well simulated in these models as the freely evolving fluxes are comparable to results with prescribed atmospheric concentrations.

Centered pattern correlations for these variables and regions are shown in Figure 5.18, with a score of 1 meaning perfect similarity of a model to the reference data, while a value of 0 signifies no relationship. The longer lines denote the MMM, while the grey circle shows the similarity of the alternate data set to the main reference data set. For GPP and LAI the reference data sets show very good similarity of above 0.9, while for NBP the differences of the references highlighted in Figure 5.13 is highlighted through a small correlation of up to 0.3, with a high anti-correlation in the southern hemisphere. Due to this, the precise value of the correlation coefficient between models and reference data set is not a good measure, but it can be seen that the models show a large spread. For GPP, the CMIP6 performance in the tropics is vastly improved, with even higher correlation values for Ncycle models. Other than in the NH, the CMIP5 models show a large spread in correlation values, which has reduced for CMIP6. The correlation distribution for LAI is similar as GPP, with the highest correlation values found in the tropics and globally, but the difference between Ncycle and non-Ncycle models is not as prominent. These overall performance plots underline the specific conclusions from the separate sections above.

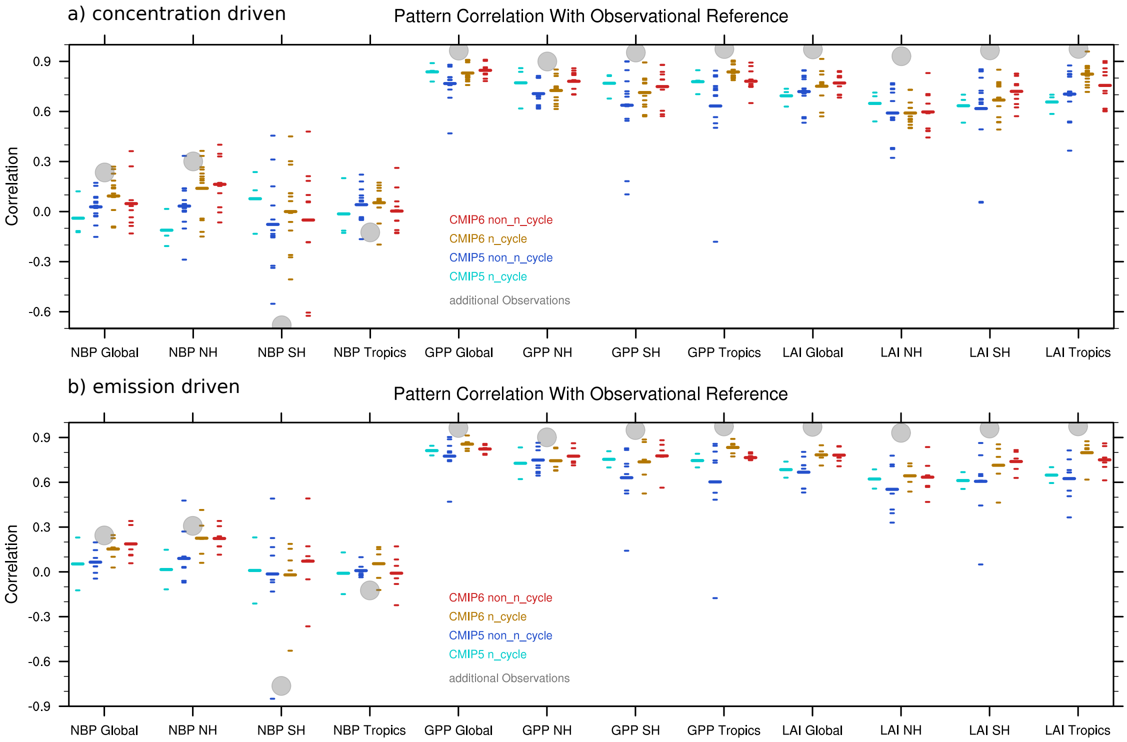


Figure 5.18.: Centered pattern correlations between models and reference data sets for annual mean climatology for concentration-driven (a) and emission-driven (b) CMIP5 and CMIP6 models, split into Ncycle and non-Ncycle models. Main and alternate observations are the same as in Figure 5.16. Adapted with permission from **Gier** et al. 2024.

5.3. Conclusion

To be able to have confidence in model projections of climate change, Earth System Models first need to show the ability to simulate observed climatologies and trends of the carbon cycle in the present day climate. In the Coupled Model Intercomparison Project Phase 5 (CMIP5, Taylor et al. 2012), several weaknesses of the simulated carbon cycle were found, such as a general overestimation of photosynthesis and a wide range of values for carbon stocks, which became one of the main areas of focus for improvement for some model groups (Delire et al. 2020). In this study, we have analysed the land carbon cycle of models participating in the Coupled Model Intercomparison Project Phase 6 (CMIP6, Eyring et al. 2016c) to investigate whether these weaknesses were improved in the newer model generation, with a special focus on differences arising due to inclusion of an interactive terrestrial nitrogen cycle in some of the CMIP6 models. Concentration and emission-driven simulations from CMIP5 and CMIP6 models were compared to reference data sets, with 2 out of 18 CMIP5 models and 15 out of 23 CMIP6 models including carbon-nitrogen interactions. We assessed means, trends and seasonal cycles of LAI, GPP, and NBP. We furthermore looked at land carbon stocks to see if the large range of values simulated in CMIP5 was reduced in CMIP6.

In general, CMIP6 models show a better performance across all assessed land carbon cycle variables to differing degrees, and no significant differences between the concentration-driven and emission-driven simulation were found in the considered variables, that cannot be explained by the number of different models. While there is a bias towards the CLM land component in the CMIP6 models, the different versions (4, 4.5, 5) do not perform the same and thus these versions can be seen as independent components for the multi-model mean.

The leaf area index was a weakness of the CMIP5 simulation, as its seasonal cycle was not well captured and its absolute value was generally overestimated. While the peak of the climatological seasonal cycle of LAI is much better reproduced in CMIP6, the amplitude of the seasonal cycle is weaker in CMIP6 compared to observations due to a weaker drawdown in winter. Thus LAI should remain an area of focus for future model development. Mean LAI is much better reproduced in CMIP6, while the range of trends in the observations is large enough to cover most models for both CMIP5 and CMIP6. It should be noted that due to correlations between parameters, there are often tradeoffs for better reproducing one variable. In CLM5 such a tradeoff had to be weighed between biases for GPP and LAI against high PFT survivability rates (Lawrence et al. 2019). Therefore, looking at one variable separately instead of the whole model performance can lead to wrong conclusions about the model's ability of reproducing the carbon cycle, depending on which choices were made in the tuning. Similarly, models interactively simulating vegetation cover perform worse in the evaluation of present-day LAI compared to models using observationally derived landcover maps due to simulating trees and grasses in the wrong areas. However, only these models with dynamic vegetation can account for future changes in vegetation and the impact of these changes on climate and carbon processes in future projections.

One of the largest improvements due to the inclusion of an interactive nitrogen cycle was seen in GPP, where the CMIP6 nitrogen cycle models were able to capture the seasonal cycle in the northern hemisphere well, which was previous overestimated. Beside the improvements in the NH, bias patterns in the tropics showing larger GPP overestimations bordering tropical rainforests are reduced in CMIP6 models, with some of these biases wholly removed in the multi-model mean of the CMIP6 models with interactive nitrogen cycle.

The land carbon sink is underestimated in the northern hemisphere regardless of CMIP phase or inclusion of nitrogen cycle. The models compensate for this by simulating a larger carbon sink in the tropics and the southern hemisphere for a global average close to the observed value. An improvement is seen in CMIP6 in capturing the amplitude of the seasonal cycle, which is controlled by carbon uptake through photosynthesis in the growth season and carbon release by respiration. This improvement can largely be attributed to the improved seasonal cycle of GPP.

The large range of soil and vegetation carbon was another large weakness of CMIP5, with inter-model differences of 900 PgC for vegetation carbon and 2500 PgC for soil carbon. This range has not significantly decreased in CMIP6, and it remains an area for improvement.

While we find a significant improvement of the inclusion of the nitrogen cycle for photosynthesis, the effects are reduced for the leaf area index and the land-atmosphere carbon flux. Despite similar NBP for models with and without interactive nitrogen cycle, models without interactive nitrogen overestimate carbon fertilization, leading to large differences of atmospheric carbon content for future scenario simulations (Kou-Giesbrecht and Arora 2023). Therefore, the inclusion of further limiting nutrients like phosphorus is important, as they will likely have substantial impacts on future carbon uptake (Yang et al. 2023). Model performance overall has improved from CMIP5 to CMIP6 even with the added complexity introducing more degrees of freedom into the models, as also found in the latest IPCC report (Canadell et al. 2021; Eyring et al. 2021). This is a positive outlook for the future, as many aspects have to be considered when increasing model complexity, such as a need to adjust existing parametrisations after model structural changes from carbon-only to carbon-nitrogen coupling. Without such adjustments, lowered ecosystem carbon storage simulated by models with N processes would lead to an underestimation of carbon pools (Wang et al. 2022). The increased overall model performance confirms results from the individual model groups who found improved performance in carbon cycle variables compared to previous model configurations, with the biggest improvements seen in LAI and GPP (Danabasoglu et al. 2020; Ziehn et al. 2017). Many areas requiring improvement remain, such as simulated carbon stocks which saw no significant reduction in the simulated range between CMIP5 and CMIP6, or the inclusion of more nutrient limitations like an interactive phosphorus cycle. The improvement of the carbon cycle in the models since CMIP5 is a step in the right direction for a better understanding and a more accurate simulation of future trends. Based on our analysis, due to the small differences between historical concentration and emission-driven simulations despite the increased process-realism, we recommend ESMs in future CMIP phases to be based on emission-driven simulations to fully account for climate-carbon feedbacks in future projections, supporting the message from Sanderson et al. 2023. Similarly, due the significant improvements in GPP with the inclusion of an interactive nitrogen cycle and no detrimental change in the present day evaluation of any carbon cycle variable, we suggest that the nitrogen cycle should be seen as a necessary part of carbon cycle models in the future.

6. Conclusion

6.1. Overall Summary

The analysis of the terrestrial carbon cycle and its representation in climate models is vital to improve the models' ability to correctly simulate future projections and thus assess future climate change. This thesis evaluated the terrestrial carbon cycle in CMIP6 models compared to its predecessor phase CMIP5 with satellite observations to assess the performance of the latest generation of climate models. For the evaluation, ESMValTool, an open-source community developed diagnostics and performance metrics tool for routine evaluation of CMIP models, has been expanded with all diagnostics that are presented in this thesis. The author's involvement with ESMValTool as a member of the scientific lead development team also led to co-authorship in several peer reviewed studies (Eyring et al. 2020; Lauer et al. 2020; Weigel et al. 2021). The code contributed to ESMValTool is publicly available and beneficial to the entire scientific community, which has become evident by its use in the latest IPCC AR6. The author has also had several further contributions to the IPCC AR6 as acknowledged graphic developer in Chapter 3 (Eyring et al. 2021) and contributing author in Chapter 5 (Eyring et al. 2021), contributing to both chapters with figures and data related to the carbon cycle of ESMs.

In this thesis, several key science questions have been addressed. The first question ("How can recent satellite observations be used to evaluate climate model simulations, and how does sparse data coverage affect results?") is addressed in Chapter 4 based on **Gier** et al. 2020. Satellite observations shine in spatially resolved comparisons and perform well in general analysis of short- and mid-term features, such as the seasonal cycle amplitude (SCA) of the carbon cycle. However, for comparable analysis, model simulations should be sampled in the same way as the corresponding observations. Sparse data coverage can affect results in different ways. In the case of SCA, peaks or troughs might occur at times of missing data resulting in a lowered SCA. Furthermore, in the case of this study, the use of a composite timeseries of different satellites with different spatial coverage can result in artificial trends over the whole timeseries. This study helped to resolve a previous discrepancy of rising SCA with CO_2 in the midlatitudes found in models while satellite data showed a negative trend for the SCA with increasing CO_2. In reality, this negative trend was artificially induced by the different data coverage of models and observations. The short timeseries available for satellite data can also limit the analysis of other phenomena, such as the trend of growing season temperature and interannual variability (IAV) of CO_2 growth rate (GR), which is strongly affected by ENSO.

The second key science question ("How do CMIP6 models compare to CMIP5 models in simulating atmospheric CO_2 and terrestrial carbon cycle variables, on the basis of observational and reanalysis data?") has been present throughout all studies presented in this thesis, starting from the contributions to the IPCC AR6, as well as the two studies presented in Chapters 4 and 5 based on **Gier** et al. 2020 and **Gier** et al. 2024. CMIP6 models perform similar or better than CMIP5 models for both column-average CO_2 mole fraction (XCO_2) as well as terrestrial carbon variables with different degrees of improvement. Previously identified weaknesses of CMIP5 models were addressed in CMIP6, and due to some of the newly included processes some of the biases were reduced. However, large intermodel ranges and biases remain in the simulation of carbon stocks. Models remain unable to correctly model the strength and timing of the seasonal cycle of leaf area index (LAI), and the distribution of carbon sinks in CMIP6 models still shows an underestimation in the NH which is compensated by an overestimation of the sinks in the SH and the tropics.

Chapter 5 based on **Gier** et al. 2024 answers the third key science question of this thesis ("What effect does increased process complexity, through the inclusion of an interactively coupled nitrogen cycle and the forcing of carbon emissions instead of carbon concentrations, have on model performance?"). In this study, models with and without nitrogen cycle were compared. The inclusion of an interactive nitrogen cycle and the corresponding increased process complexity does not decrease model performance across all variables. Instead, a large improvement is found in simulating the seasonal cycle in the NH only for models with interactive nitrogen cycle, while an overestimation of the photosynthesis remains for models without a nitrogen cycle. Similarly, the comparison of concentration and emission-driven simulations shows only small differences. As only the emission-driven simulations can fully account for climate-carbon feedback in future projections, these should become the standard simulations to perform in CMIP7 and beyond, with the nitrogen cycle as a necessary part of the carbon cycle in these models.

Overall, the recommendations of this thesis provide important guidance of focus areas for future model development and for model experiment designs.

6.2. Outlook

Since the publication of **Gier** et al. 2020 (see Chapter 4), the observational products from that study have been extended to 2022 and now include two additional satellites: OCO-2 since 2014 and GOSAT-2 since 2019. An analysis of their spatial coverage should be done to see if similar effects occur in this time frame, like the artificially induced negative trend of the SCA with CO_2 which was found for the combination of the other satellites. This could be compared with further observational data sets using methods for fusing timeseries of different satellites, such as that of Wang et al. 2023. Many observational datasets are now obtained with the help of machine learning methods. The difference between these datasets to more traditional datasets should be analysed, as these may open new applications for the analysis of climate model data. Promoting the exchange of ideas and needs of modeling groups and people working on observational data will help in the exchange of expertise and streamline future developments according to the needs of both groups.

The ultimate aim of the development of ESMs is to accurately and realistically simulate the future climate under different forcing scenarios. To this end, the performance of models is usually measured by assessing their capability of simulating the present-day climate. However, Earth system models differ in their complexity in terms of representing important processes or components. Models that prescribe important drivers such as atmospheric CO_2 concentrations instead of prescribing CO_2 emissions and interactively simulating important related Earth system processes, responses and feedbacks might perform better in present-day comparisons than the more complex models. Therefore, if performance remains similar, it can already called a success, as including important Earth system feedbacks is relevant under changing climate conditions, but might not be fully apparent in present-day simulations.

Examples in this thesis include the comparison between models with and without an interactive nitrogen cycle. Without enough nitrogen, plant photosynthesis can become nutrient-limited and the CO_2 uptake is reduced despite rising CO_2 availability. Including an interactive nitrogen cycle coupled to the carbon cycle can take into account this nitrogen limitation, which will become even more relevant with rising CO_2 concentration and temperatures. Therefore, it is crucial to include an interactive nitrogen cycle for accurate simulations of future photosynthesis. This is even more underlined by the fact that models with a coupled nitrogen cycle already show a better estimate of the seasonal cycle of photosynthesis in the present day.

Another example is the comparison between concentration-driven and emission-driven simulations. Only the emission-driven simulations allow climate-carbon cycle feedbacks to influence the atmospheric CO_2 concentration and thus the warming pathways. In concentration-driven models this has to be done in the models deriving the CO_2 concentrations, which may model feedbacks differently than the models using these prescribed concentrations. Taking these feedbacks into account in a fully consistent way is vital for an accurate prediction of climate warming. Thus, based on the results of this thesis (i.e., models show similar performances in emission and concentration-driven simulations), it is recommended to standardize running the models in emission-driven mode in future CMIP generations, in agreement with Sanderson et al. 2023.

Running emission-driven simulations and including an interactive nitrogen cycle are not the only novelties introduced to CMIP models in recent years. The study of this thesis presented in Chapter 5 found clear weaknesses in models with dynamical vegetation to represent present-day LAI in the performance metrics plots. Similarly to the considerations about the nitrogen cycle and the emission-driven simulations, models without dynamic vegetation prescribe vegetation based on observations and require external vegetation projections for future projections, which may not be consistent with the feedbacks included in the model itself. Thus, a detailed analysis of dynamical vegetation in models, as well as further processes like land-use change is vital to further inform model developers on where to focus their improvements. Studies on the carbon cycle also benefit from a detailed comparison of the physical variables influencing the carbon cycle, such as temperature and precipitation. However, many of the relevant processes occur on smaller timescales, calling for the need to analyze simulations with hourly time steps.

Through adding diagnostics for the carbon cycle to ESMValTool, it becomes easier and more time efficient to compare the general performance of models in this area, allowing for more detailed analysis of specific processes, as well as easier to create information for modeling groups to benchmark their models during development. The evaluation and developed ESMValTool diagnostics of this thesis can also support the annual Global carbon budget project (Friedlingstein et al. 2023) that supports the Global Stocktake and Annual Conferences of the Parties (COP) convened under the United Nations Framework Convention on Climate Change (UNFCCC).

Appendix

A. Supplementary Materials for Chapter 4

The supplementary material for Chapter 4 is taken from the appendix of **Gier** et al. 2020, with the plots made by the author of this thesis.

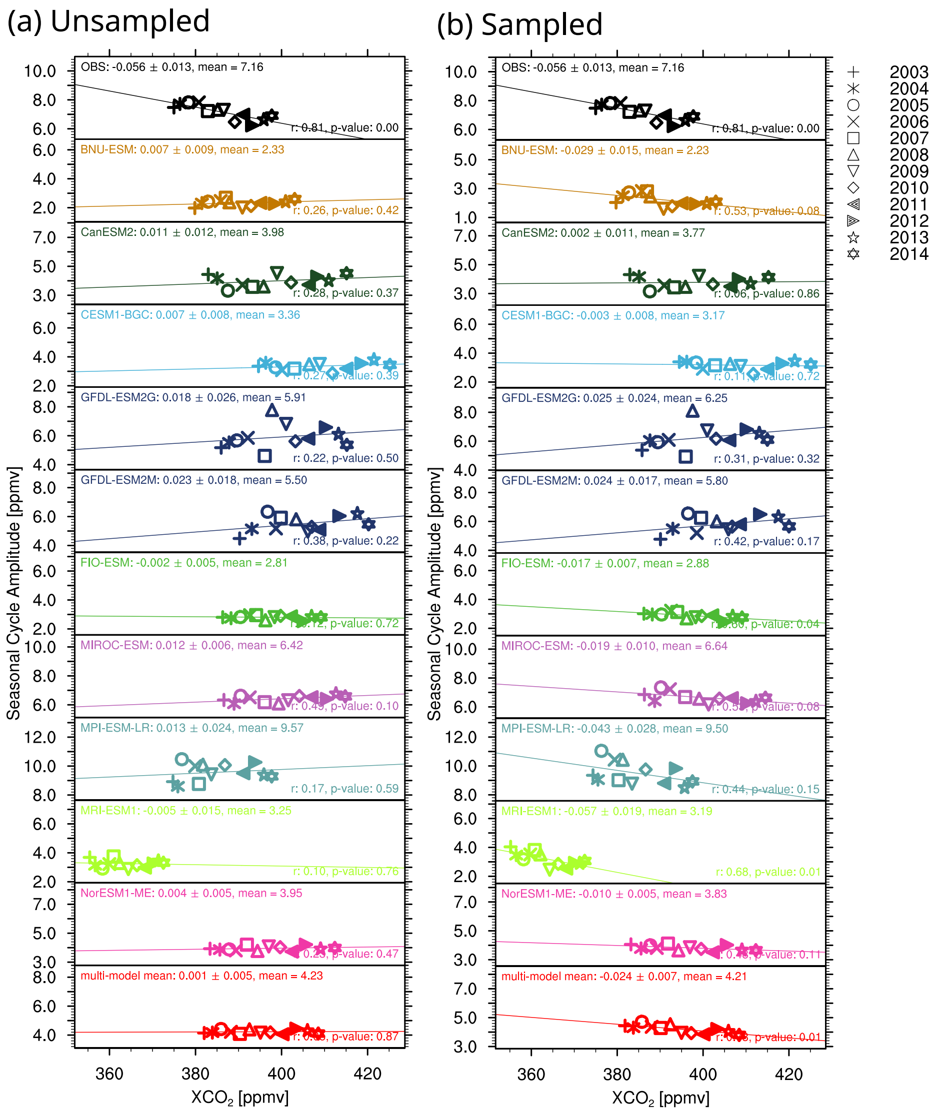


Figure A.1.: Same as Fig. 4.8 but for CMIP5 models. Adapted with permission from **Gier** et al. 2020. Adapted with permission from **Gier** et al. 2020.

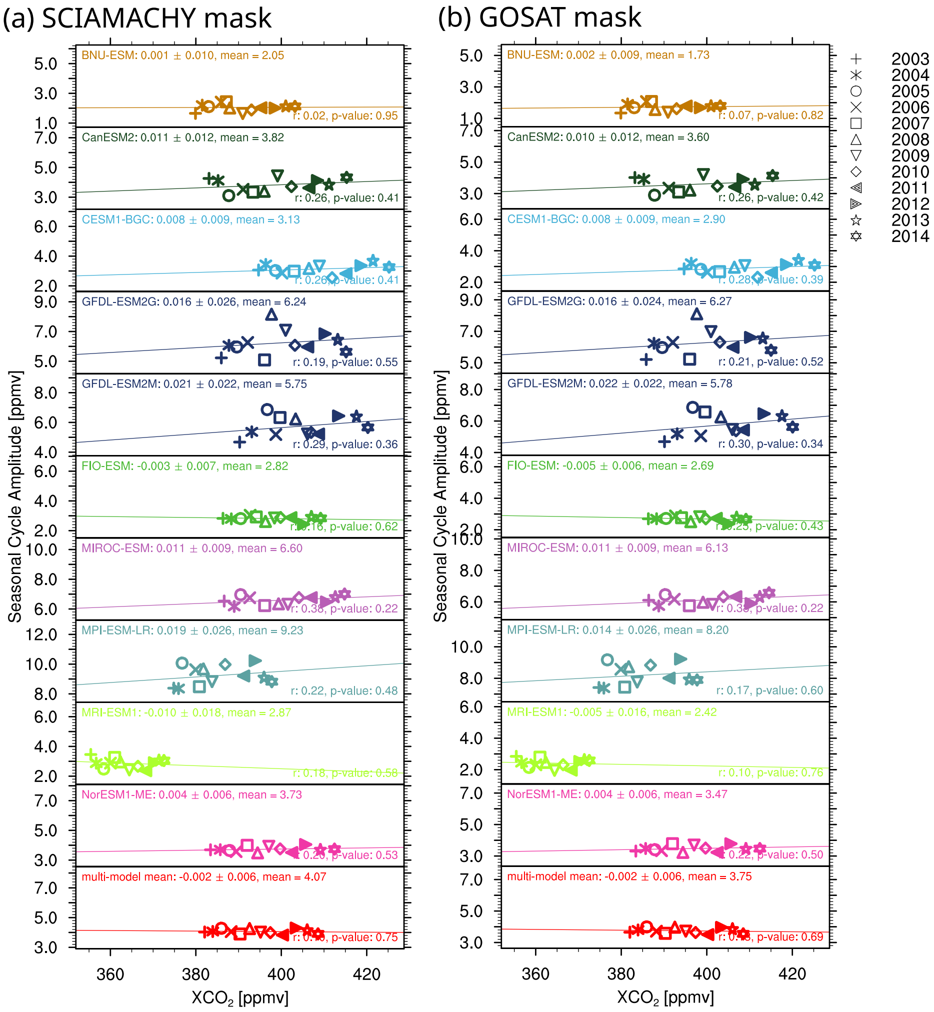


Figure A.2.: Same as Fig. 4.10 but for CMIP5 models. Adapted with permission from **Gier** et al. 2020.

B. Supplementary Materials for Chapter 5

The supplementary material for Chapter 5 is taken from the appendix of **Gier** et al. 2024, which are the tables including the values of each individual model and reference dataset for the scatterplots. These are automatically generated by the diagnostic script written by the author.

B.1. Means and Trends

Data values for scatterplots Figures 5.2, 5.6 and 5.12.

Table B.1.: LAI Mean for 1986-2005 [1]

Model	Global	NH	SH	Tropics
BNU-ESM CMIP5e	3.64 ± 0.08	2.13 ± 0.06	5.04 ± 0.1	6.37 ± 0.1
BNU-ESM CMIP5c	3.75 ± 0.06	2.4 ± 0.09	4.98 ± 0.04	6.34 ± 0.06
CESM1-BGC CMIP5e	2.68 ± 0.06	2.03 ± 0.06	3.31 ± 0.09	3.78 ± 0.1
CESM1-BGC CMIP5c	2.64 ± 0.04	1.99 ± 0.05	3.27 ± 0.05	3.72 ± 0.05
CanESM2 CMIP5e	1.71 ± 0.06	1.11 ± 0.02	2.26 ± 0.11	2.65 ± 0.13
CanESM2 CMIP5c	1.72 ± 0.04	1.13 ± 0.02	2.26 ± 0.08	2.65 ± 0.1
FIO-ESM CMIP5e	2.23 ± 0.01	1.69 ± 0.01	2.73 ± 0.0	3.14 ± 0.0
GFDL-ESM2G CMIP5e	3.1 ± 0.06	2.65 ± 0.02	3.48 ± 0.1	4.09 ± 0.1
GFDL-ESM2G CMIP5c	3.02 ± 0.04	2.62 ± 0.05	3.35 ± 0.06	3.96 ± 0.05
GFDL-ESM2M CMIP5e	3.17 ± 0.05	2.88 ± 0.02	3.4 ± 0.1	4.05 ± 0.11
GFDL-ESM2M CMIP5c	3.05 ± 0.07	2.72 ± 0.04	3.31 ± 0.12	3.9 ± 0.12
HadGEM2-CC CMIP5c	2.48 ± 0.02	1.52 ± 0.03	3.38 ± 0.02	4.15 ± 0.02
HadGEM2-ES CMIP5c	2.57 ± 0.03	1.6 ± 0.03	3.48 ± 0.04	4.28 ± 0.04
IPSL-CM5A-LR CMIP5c	2.18 ± 0.03	2.05 ± 0.02	2.28 ± 0.06	2.84 ± 0.07
IPSL-CM5A-MR CMIP5c	2.13 ± 0.03	2.04 ± 0.03	2.2 ± 0.04	2.7 ± 0.05
IPSL-CM5B-LR CMIP5c	2.0 ± 0.02	1.93 ± 0.03	2.04 ± 0.05	2.43 ± 0.07
MIROC-ESM-CHEM CMIP5c	2.66 ± 0.03	2.0 ± 0.02	3.25 ± 0.06	3.56 ± 0.07
MIROC-ESM CMIP5e	2.66 ± 0.05	2.05 ± 0.03	3.2 ± 0.08	3.48 ± 0.09
MIROC-ESM CMIP5c	2.65 ± 0.03	2.06 ± 0.02	3.16 ± 0.05	3.45 ± 0.05
MPI-ESM-LR CMIP5e	2.27 ± 0.02	1.59 ± 0.02	2.92 ± 0.04	3.32 ± 0.04
MPI-ESM-LR CMIP5c	2.26 ± 0.03	1.59 ± 0.01	2.89 ± 0.07	3.29 ± 0.07
MPI-ESM-MR CMIP5c	2.3 ± 0.02	1.64 ± 0.02	2.93 ± 0.04	3.32 ± 0.05
MRI-ESM1 CMIP5e	3.61 ± 0.07	2.16 ± 0.06	4.95 ± 0.1	5.98 ± 0.13
NorESM1-ME CMIP5e	2.8 ± 0.05	2.16 ± 0.07	3.41 ± 0.07	3.85 ± 0.08
NorESM1-ME CMIP5c	2.8 ± 0.05	2.12 ± 0.07	3.43 ± 0.06	3.87 ± 0.06
inmcm4 CMIP5c	1.62 ± 0.02	1.26 ± 0.02	1.95 ± 0.03	2.27 ± 0.03
ACCESS-ESM1-5 CMIP6e	1.64 ± 0.02	1.43 ± 0.03	1.82 ± 0.03	2.09 ± 0.04
ACCESS-ESM1-5 CMIP6c	1.65 ± 0.02	1.46 ± 0.02	1.82 ± 0.05	2.07 ± 0.06
CESM2-WACCM CMIP6c	1.78 ± 0.04	1.49 ± 0.05	2.02 ± 0.03	2.35 ± 0.04
CESM2 CMIP6c	1.79 ± 0.05	1.53 ± 0.06	2.0 ± 0.04	2.33 ± 0.06
CMCC-CM2-SR5 CMIP6c	2.17 ± 0.06	2.33 ± 0.07	1.98 ± 0.08	2.18 ± 0.07
CMCC-ESM2 CMIP6c	2.2 ± 0.05	2.3 ± 0.06	2.06 ± 0.07	2.28 ± 0.06
CNRM-ESM2-1 CMIP6e	1.59 ± 0.03	1.33 ± 0.03	1.82 ± 0.05	2.13 ± 0.04
CNRM-ESM2-1 CMIP6c	1.59 ± 0.03	1.34 ± 0.05	1.81 ± 0.04	2.12 ± 0.03

Continued on next page

Table B.1.: LAI Mean for 1986-2005 [1]

Model	Global	NH	SH	Tropics
CanESM5-CanOE CMIP6e	1.71 ± 0.04	1.14 ± 0.03	2.24 ± 0.06	2.64 ± 0.07
CanESM5-CanOE CMIP6c	1.69 ± 0.04	1.15 ± 0.03	2.2 ± 0.07	2.58 ± 0.08
CanESM5 CMIP6e	1.71 ± 0.04	1.16 ± 0.03	2.22 ± 0.06	2.61 ± 0.06
CanESM5 CMIP6c	1.72 ± 0.02	1.18 ± 0.03	2.22 ± 0.03	2.6 ± 0.03
EC-Earth3-CC CMIP6e	2.59 ± 0.06	1.6 ± 0.06	3.53 ± 0.08	3.93 ± 0.08
EC-Earth3-CC CMIP6c	2.54 ± 0.05	1.58 ± 0.05	3.45 ± 0.09	3.82 ± 0.07
EC-Earth3-Veg CMIP6c	2.55 ± 0.04	1.57 ± 0.06	3.47 ± 0.05	3.87 ± 0.04
GFDL-ESM4 CMIP6e	2.33 ± 0.04	2.0 ± 0.03	2.61 ± 0.06	3.17 ± 0.08
GFDL-ESM4 CMIP6c	2.27 ± 0.04	1.92 ± 0.03	2.58 ± 0.09	3.13 ± 0.09
INM-CM4-8 CMIP6c	1.86 ± 0.02	1.5 ± 0.02	2.18 ± 0.02	2.53 ± 0.02
INM-CM5-0 CMIP6c	1.85 ± 0.02	1.48 ± 0.03	2.18 ± 0.02	2.53 ± 0.01
IPSL-CM6A-LR CMIP6c	1.85 ± 0.04	1.16 ± 0.04	2.49 ± 0.05	2.97 ± 0.06
MIROC-ES2L CMIP6e	1.55 ± 0.01	1.01 ± 0.02	2.05 ± 0.01	2.56 ± 0.01
MIROC-ES2L CMIP6c	1.59 ± 0.02	1.03 ± 0.03	2.11 ± 0.02	2.62 ± 0.02
MPI-ESM-1-2-HAM CMIP6c	2.03 ± 0.04	1.43 ± 0.03	2.59 ± 0.07	2.96 ± 0.08
MPI-ESM1-2-LR CMIP6e	1.94 ± 0.04	1.46 ± 0.02	2.38 ± 0.08	2.82 ± 0.09
MPI-ESM1-2-LR CMIP6c	1.94 ± 0.03	1.46 ± 0.02	2.38 ± 0.05	2.83 ± 0.06
MRI-ESM2-0 CMIP6e	2.97 ± 0.09	2.02 ± 0.1	3.84 ± 0.15	4.81 ± 0.2
MRI-ESM2-0 CMIP6c	2.99 ± 0.07	2.04 ± 0.11	3.85 ± 0.09	4.83 ± 0.12
NorESM2-LM CMIP6e	1.75 ± 0.04	1.48 ± 0.04	1.98 ± 0.04	2.35 ± 0.05
NorESM2-LM CMIP6c	1.7 ± 0.04	1.44 ± 0.05	1.93 ± 0.03	2.28 ± 0.04
NorESM2-MM CMIP6c	1.71 ± 0.04	1.41 ± 0.04	1.98 ± 0.05	2.32 ± 0.06
SAM0-UNICON CMIP6c	2.65 ± 0.06	2.09 ± 0.07	3.19 ± 0.1	3.87 ± 0.12
TaiESM1 CMIP6c	2.55 ± 0.04	2.04 ± 0.04	3.03 ± 0.07	3.44 ± 0.07
UKESM1-0-LL CMIP6e	2.82 ± 0.02	2.19 ± 0.04	3.42 ± 0.03	3.94 ± 0.02
UKESM1-0-LL CMIP6c	2.81 ± 0.02	2.19 ± 0.03	3.39 ± 0.02	3.9 ± 0.02
GIMMS-LAI4g	1.74 ± 0.01	1.22 ± 0.01	2.22 ± 0.02	2.72 ± 0.02
GLASS	1.69 ± 0.06	1.19 ± 0.04	2.14 ± 0.08	2.6 ± 0.1
LAI3g	1.69 ± 0.03	1.11 ± 0.02	2.22 ± 0.04	2.74 ± 0.05
Ncycle MMM CMIP5e	2.74 ± 0.05	2.09 ± 0.05	3.36 ± 0.07	3.81 ± 0.07
non-Ncycle MMM CMIP5e	2.52 ± 0.02	1.99 ± 0.02	3.0 ± 0.04	3.46 ± 0.04
Ncycle MMM CMIP5c	2.72 ± 0.04	2.06 ± 0.05	3.35 ± 0.04	3.8 ± 0.04
non-Ncycle MMM CMIP5c	2.36 ± 0.02	1.86 ± 0.02	2.81 ± 0.02	3.29 ± 0.02
Ncycle MMM CMIP6c	2.12 ± 0.03	1.71 ± 0.04	2.49 ± 0.02	2.87 ± 0.02
non-Ncycle MMM CMIP6c	1.98 ± 0.02	1.46 ± 0.03	2.45 ± 0.02	2.92 ± 0.02

Continued on next page

Table B.1.: LAI Mean for 1986-2005 [1]

Model	Global	NH	SH	Tropics
Ncycle MMM CMIP6e	2.05 ± 0.02	1.53 ± 0.03	2.53 ± 0.02	2.95 ± 0.02
non-Ncycle MMM CMIP6e	2.06 ± 0.03	1.53 ± 0.04	2.54 ± 0.04	3.07 ± 0.05

Table B.2.: LAI Trend for 1986-2005 [yr^{-1}]

Model	Global	NH	SH	Tropics
BNU-ESM CMIP5e	0.012 ± 0.0013	0.0072 ± 0.002	0.016 ± 0.0013	0.0158 ± 0.0015
BNU-ESM CMIP5c	0.0102 ± 0.0008	0.0136 ± 0.0016	0.0057 ± 0.0012	0.0087 ± 0.0012
CESM1-BGC CMIP5e	0.004 ± 0.0022	0.0067 ± 0.002	0.0011 ± 0.0034	0.0006 ± 0.0039
CESM1-BGC CMIP5c	0.0035 ± 0.0016	0.0027 ± 0.0017	0.004 ± 0.0017	0.0057 ± 0.0016
CanESM2 CMIP5e	0.0019 ± 0.0022	0.003 ± 0.0005	0.0005 ± 0.0043	0.0012 ± 0.005
CanESM2 CMIP5c	0.0018 ± 0.0017	0.002 ± 0.0006	0.0012 ± 0.0033	0.0028 ± 0.0038
FIO-ESM CMIP5e	0.0005 ± 0.0002	0.0009 ± 0.0003	-0.0001 ± 0.0001	-0.0001 ± 0.0001
GFDL-ESM2G CMIP5e	0.0055 ± 0.0019	0.0028 ± 0.0008	0.0077 ± 0.0035	0.008 ± 0.0035
GFDL-ESM2G CMIP5c	0.0045 ± 0.0012	0.0071 ± 0.001	0.0015 ± 0.0022	0.0011 ± 0.002
GFDL-ESM2M CMIP5e	0.0034 ± 0.0019	0.0008 ± 0.0008	0.0056 ± 0.0038	0.0087 ± 0.004
GFDL-ESM2M CMIP5c	0.0024 ± 0.0026	0.0057 ± 0.0006	-0.0013 ± 0.0049	0.0023 ± 0.0049
HadGEM2-CC CMIP5c	0.0031 ± 0.0005	0.0036 ± 0.0007	0.0022 ± 0.0006	0.002 ± 0.0008
HadGEM2-ES CMIP5c	0.0046 ± 0.0006	0.0033 ± 0.0008	0.0052 ± 0.001	0.0055 ± 0.0009
IPSL-CM5A-LR CMIP5c	-0.0022 ± 0.001	0.0024 ± 0.0008	-0.0067 ± 0.0018	-0.0056 ± 0.0022

Continued on next page

Table B.2.: LAI Trend for 1986-2005 [yr^{-1}]

Model	Global	NH	SH	Tropics
IPSL-CM5A-MR CMIP5c	0.0015 ± 0.001	0.0032 ± 0.0011	-0.0003 ± 0.0017	0.0017 ± 0.0021
IPSL-CM5B-LR CMIP5c	0.0008 ± 0.0009	0.0043 ± 0.0007	-0.0027 ± 0.0019	-0.0029 ± 0.0025
MIROC-ESM-CHEM CMIP5c	0.0038 ± 0.0011	0.0016 ± 0.0008	0.0054 ± 0.0019	0.008 ± 0.0022
MIROC-ESM CMIP5e	0.0039 ± 0.0017	0.0044 ± 0.0009	0.003 ± 0.0031	0.0044 ± 0.0033
MIROC-ESM CMIP5c	0.0015 ± 0.0011	0.0013 ± 0.0008	0.0013 ± 0.002	0.0023 ± 0.002
MPI-ESM-LR CMIP5e	-0.0003 ± 0.0008	0.0021 ± 0.0004	-0.0029 ± 0.0016	-0.0033 ± 0.0015
MPI-ESM-LR CMIP5c	0.0014 ± 0.0013	0.0011 ± 0.0004	0.0015 ± 0.0028	0.0004 ± 0.0028
MPI-ESM-MR CMIP5c	0.0 ± 0.0008	0.0023 ± 0.0005	-0.0025 ± 0.0015	-0.0031 ± 0.0017
MRI-ESM1 CMIP5e	0.0063 ± 0.0023	0.0036 ± 0.0023	0.0084 ± 0.0034	0.0107 ± 0.0047
NorESM1-ME CMIP5e	0.006 ± 0.0015	0.0075 ± 0.0024	0.0041 ± 0.0024	0.0037 ± 0.0029
NorESM1-ME CMIP5c	0.005 ± 0.0014	0.0092 ± 0.0019	0.0006 ± 0.0024	0.0019 ± 0.0022
inmcm4 CMIP5c	0.0007 ± 0.0007	0.0014 ± 0.0009	-0.0 ± 0.0012	-0.0009 ± 0.0013
ACCESS-ESM1-5 CMIP6e	0.001 ± 0.0006	0.0051 ± 0.0006	-0.0031 ± 0.0011	-0.0043 ± 0.0014
ACCESS-ESM1-5 CMIP6c	-0.0014 ± 0.0008	0.0029 ± 0.0004	-0.0057 ± 0.0015	-0.0071 ± 0.0018
CESM2-WACCM CMIP6c	0.0059 ± 0.0005	0.0074 ± 0.0006	0.0043 ± 0.0007	0.0065 ± 0.0007
CESM2 CMIP6c	0.0082 ± 0.0005	0.0096 ± 0.0008	0.0066 ± 0.0008	0.0087 ± 0.0011
CMCC-CM2-SR5 CMIP6c	0.0047 ± 0.0021	0.0049 ± 0.0025	0.0044 ± 0.003	0.0043 ± 0.0026
CMCC-ESM2 CMIP6c	0.0047 ± 0.0016	0.0074 ± 0.0017	0.0021 ± 0.0028	0.0024 ± 0.0023

Continued on next page

Table B.2.: LAI Trend for 1986-2005 [yr^{-1}]

Model	Global	NH	SH	Tropics
CNRM-ESM2-1 CMIP6e	-0.0002 ± 0.0011	0.0027 ± 0.001	-0.003 ± 0.0019	-0.0025 ± 0.0016
CNRM-ESM2-1 CMIP6c	0.0037 ± 0.0009	0.006 ± 0.0013	0.0012 ± 0.0015	0.0007 ± 0.0013
CanESM5-CanOE CMIP6e	0.0048 ± 0.0011	0.0045 ± 0.0007	0.0048 ± 0.0021	0.0056 ± 0.0025
CanESM5-CanOE CMIP6c	0.0039 ± 0.0013	0.004 ± 0.0006	0.0035 ± 0.0025	0.0046 ± 0.003
CanESM5 CMIP6e	0.005 ± 0.0009	0.0044 ± 0.0006	0.0053 ± 0.0019	0.0042 ± 0.0022
CanESM5 CMIP6c	0.0032 ± 0.0006	0.0043 ± 0.0006	0.0019 ± 0.001	0.0027 ± 0.0011
EC-Earth3-CC CMIP6e	0.0069 ± 0.0016	0.0089 ± 0.0011	0.0044 ± 0.003	0.0059 ± 0.0029
EC-Earth3-CC CMIP6c	0.0053 ± 0.0017	0.008 ± 0.001	0.0023 ± 0.0034	0.0029 ± 0.0029
EC-Earth3-Veg CMIP6c	0.0041 ± 0.0013	0.0085 ± 0.0009	-0.0007 ± 0.0021	0.0009 ± 0.0016
GFDL-ESM4 CMIP6e	0.0044 ± 0.0012	0.0037 ± 0.0008	0.0046 ± 0.0023	0.0047 ± 0.0029
GFDL-ESM4 CMIP6c	-0.0004 ± 0.0017	0.0033 ± 0.0007	-0.0045 ± 0.0032	-0.0069 ± 0.0034
INM-CM4-8 CMIP6c	0.0018 ± 0.0005	0.0018 ± 0.0008	0.0017 ± 0.0006	0.002 ± 0.0008
INM-CM5-0 CMIP6c	0.0023 ± 0.0005	0.0041 ± 0.0007	0.0003 ± 0.0006	0.0005 ± 0.0006
IPSL-CM6A-LR CMIP6c	0.0052 ± 0.0008	0.0052 ± 0.0007	0.0049 ± 0.0017	0.0064 ± 0.0018
MIROC-ES2L CMIP6e	0.0022 ± 0.0003	0.0033 ± 0.0005	0.0008 ± 0.0004	0.0016 ± 0.0005
MIROC-ES2L CMIP6c	0.0033 ± 0.0003	0.0043 ± 0.0005	0.0022 ± 0.0005	0.0032 ± 0.0006
MPI-ESM-1-2-HAM CMIP6c	0.0041 ± 0.0014	0.0038 ± 0.0005	0.0042 ± 0.0025	0.0049 ± 0.0031

Continued on next page

Table B.2.: LAI Trend for 1986-2005 [yr^{-1}]

Model	Global	NH	SH	Tropics
MPI-ESM1-2-LR CMIP6e	0.0014 ± 0.0015	0.0033 ± 0.0004	-0.0006 ± 0.003	-0.0005 ± 0.0037
MPI-ESM1-2-LR CMIP6c	0.0026 ± 0.0011	0.0032 ± 0.0004	0.0019 ± 0.0021	0.0023 ± 0.0024
MRI-ESM2-0 CMIP6e	0.0068 ± 0.003	0.0137 ± 0.0022	-0.0006 ± 0.0058	0.0004 ± 0.0079
MRI-ESM2-0 CMIP6c	0.0094 ± 0.0019	0.0156 ± 0.0023	0.0024 ± 0.0034	0.0047 ± 0.0045
NorESM2-LM CMIP6e	0.0064 ± 0.0007	0.0069 ± 0.0007	0.0059 ± 0.001	0.0076 ± 0.0012
NorESM2-LM CMIP6c	0.0059 ± 0.0007	0.0076 ± 0.001	0.0042 ± 0.001	0.0042 ± 0.0013
NorESM2-MM CMIP6c	0.0056 ± 0.0008	0.0067 ± 0.0005	0.0046 ± 0.0015	0.0055 ± 0.0018
SAM0-UNICON CMIP6c	0.0063 ± 0.0021	0.0075 ± 0.0022	0.005 ± 0.0039	0.0078 ± 0.0046
TaiESM1 CMIP6c	0.0038 ± 0.0012	0.004 ± 0.0013	0.0036 ± 0.0025	0.0049 ± 0.0026
UKESM1-0-LL CMIP6e	0.0036 ± 0.0005	0.0052 ± 0.0008	0.0018 ± 0.0009	0.0022 ± 0.0006
UKESM1-0-LL CMIP6c	0.0014 ± 0.0006	0.0031 ± 0.0009	-0.0005 ± 0.0009	-0.0001 ± 0.0006
GIMMS-LAI4g	0.002 ± 0.0003	0.0014 ± 0.0003	0.0021 ± 0.0006	0.0027 ± 0.0006
GLASS	0.0082 ± 0.0015	0.0047 ± 0.0013	0.0112 ± 0.002	0.0128 ± 0.0023
LAI3g	0.0019 ± 0.0011	0.0018 ± 0.0008	0.0017 ± 0.0017	0.0019 ± 0.0021
Ncycle MMM CMIP5e	0.005 ± 0.0015	0.0071 ± 0.0012	0.0026 ± 0.0026	0.0022 ± 0.0028
non-Ncycle MMM CMIP5e	0.0025 ± 0.0007	0.0023 ± 0.0003	0.0023 ± 0.0014	0.0031 ± 0.0015
Ncycle MMM CMIP5c	0.0042 ± 0.0011	0.006 ± 0.0012	0.0023 ± 0.0017	0.0038 ± 0.0015
non-Ncycle MMM CMIP5c	0.0018 ± 0.0004	0.003 ± 0.0003	0.0004 ± 0.0008	0.001 ± 0.0007
Ncycle MMM CMIP6c	0.0043 ± 0.0003	0.0061 ± 0.0004	0.0024 ± 0.0005	0.0033 ± 0.0005

Continued on next page

Table B.2.: LAI Trend for 1986-2005 [yr^{-1}]

Model	Global	NH	SH	Tropics
non-Ncycle MMM CMIP6c	0.0037 ± 0.0004	0.0053 ± 0.0004	0.0017 ± 0.0006	0.0022 ± 0.0008
Ncycle MMM CMIP6e	0.0036 ± 0.0004	0.0054 ± 0.0004	0.0015 ± 0.0007	0.0021 ± 0.0008
non-Ncycle MMM CMIP6e	0.0041 ± 0.0008	0.0058 ± 0.0005	0.0022 ± 0.0015	0.0025 ± 0.0019

Table B.3.: GPP Mean for 1986-2005 [PgC yr^{-1}]

Model	Global	NH	SH	Tropics
BNU-ESM CMIP5e	83.54 ± 2.48	30.35 ± 1.3	53.19 ± 1.54	47.13 ± 1.18
BNU-ESM CMIP5c	86.26 ± 2.0	33.66 ± 1.35	52.6 ± 0.92	46.93 ± 0.92
CESM1-BGC CMIP5e	107.42 ± 2.23	34.94 ± 1.1	72.49 ± 1.3	61.98 ± 1.12
CESM1-BGC CMIP5c	106.12 ± 1.7	34.5 ± 0.78	71.62 ± 1.06	61.24 ± 0.93
CanESM2 CMIP5e	103.74 ± 2.72	32.32 ± 0.86	71.42 ± 2.76	56.61 ± 2.07
CanESM2 CMIP5c	104.37 ± 2.41	32.66 ± 0.74	71.71 ± 2.09	56.66 ± 1.78
FIO-ESM CMIP5e	151.2 ± 2.28	44.86 ± 0.99	106.34 ± 1.69	91.94 ± 1.59
GFDL-ESM2G CMIP5e	139.01 ± 4.89	46.29 ± 1.23	92.72 ± 4.42	75.82 ± 2.78
GFDL-ESM2G CMIP5c	134.03 ± 3.53	44.88 ± 1.67	89.15 ± 2.86	73.24 ± 1.98
GFDL-ESM2M CMIP5e	137.59 ± 5.15	46.76 ± 1.42	90.83 ± 5.5	75.22 ± 3.86
GFDL-ESM2M CMIP5c	132.97 ± 5.94	44.11 ± 1.4	88.86 ± 6.12	72.58 ± 4.27
HadGEM2-CC CMIP5c	111.57 ± 3.19	36.26 ± 1.82	75.31 ± 2.08	64.71 ± 2.07
HadGEM2-ES CMIP5c	116.03 ± 3.77	37.6 ± 1.35	78.44 ± 3.04	67.96 ± 2.77
IPSL-CM5A-LR CMIP5c	133.84 ± 2.22	58.96 ± 1.67	74.87 ± 1.59	67.0 ± 1.47
IPSL-CM5A-MR CMIP5c	131.41 ± 3.23	58.77 ± 1.99	72.64 ± 2.0	64.27 ± 1.9
IPSL-CM5B-LR CMIP5c	115.58 ± 2.65	51.95 ± 1.69	63.63 ± 1.87	54.03 ± 1.64
MIROC-ESM-CHEM CMIP5c	109.79 ± 2.5	41.62 ± 0.96	68.17 ± 1.95	57.0 ± 1.64
MIROC-ESM CMIP5e	110.68 ± 2.51	42.76 ± 1.43	67.93 ± 1.85	56.45 ± 1.52
MIROC-ESM CMIP5c	108.34 ± 2.73	42.38 ± 1.05	65.96 ± 2.05	54.98 ± 1.75
MPI-ESM-LR CMIP5e	145.6 ± 3.12	55.14 ± 1.81	90.46 ± 2.43	72.51 ± 1.3
MPI-ESM-LR CMIP5c	145.66 ± 3.99	55.28 ± 1.19	90.38 ± 3.64	71.76 ± 1.91
MPI-ESM-MR CMIP5c	149.88 ± 3.69	57.58 ± 2.1	92.29 ± 2.32	73.62 ± 1.58
MRI-ESM1 CMIP5e	191.82 ± 3.1	64.52 ± 1.47	127.31 ± 2.0	110.68 ± 1.73
NorESM1-ME CMIP5e	107.22 ± 1.74	35.87 ± 0.89	71.34 ± 1.16	61.16 ± 1.09

Continued on next page

Table B.3.: GPP Mean for 1986-2005 [PgC yr^{-1}]

Model	Global	NH	SH	Tropics
NorESM1-ME CMIP5c	106.83 ± 1.66	35.42 ± 1.0	71.4 ± 0.95	61.12 ± 0.82
inmcm4 CMIP5c	109.19 ± 1.97	37.4 ± 0.8	71.78 ± 1.69	62.96 ± 1.45
ACCESS-ESM1-5 CMIP6e	100.33 ± 1.3	33.27 ± 1.15	67.06 ± 0.93	57.17 ± 0.88
ACCESS-ESM1-5 CMIP6c	101.48 ± 1.28	34.39 ± 0.91	67.09 ± 1.45	56.94 ± 1.3
CESM2-WACCM CMIP6c	90.96 ± 1.73	34.31 ± 1.12	56.66 ± 0.88	47.18 ± 0.77
CESM2 CMIP6c	91.09 ± 2.44	35.39 ± 1.39	55.7 ± 1.38	46.66 ± 1.13
CMCC-CM2-SR5 CMIP6c	83.79 ± 2.16	37.64 ± 1.26	46.15 ± 1.4	38.25 ± 0.89
CMCC-ESM2 CMIP6c	84.36 ± 1.65	37.06 ± 1.23	47.3 ± 0.93	39.37 ± 0.63
CNRM-ESM2-1 CMIP6e	84.11 ± 2.15	33.15 ± 1.12	50.96 ± 2.05	41.88 ± 1.15
CNRM-ESM2-1 CMIP6c	84.18 ± 2.36	33.49 ± 1.67	50.7 ± 1.62	41.55 ± 0.93
CanESM5-CanOE CMIP6e	114.7 ± 2.8	40.82 ± 1.31	73.88 ± 1.76	60.75 ± 1.44
CanESM5-CanOE CMIP6c	113.66 ± 3.02	40.84 ± 1.3	72.83 ± 2.08	59.87 ± 1.72
CanESM5 CMIP6e	114.73 ± 2.6	41.34 ± 1.21	73.39 ± 1.74	60.32 ± 1.21
CanESM5 CMIP6c	114.82 ± 2.16	41.68 ± 1.09	73.15 ± 1.25	60.19 ± 1.02
EC-Earth3-CC CMIP6e	121.52 ± 3.19	47.03 ± 2.28	74.49 ± 1.47	56.11 ± 0.96
EC-Earth3-CC CMIP6c	119.45 ± 2.5	46.11 ± 1.75	73.34 ± 1.47	55.17 ± 1.11
EC-Earth3-Veg CMIP6c	117.81 ± 2.27	45.46 ± 1.53	72.36 ± 1.28	54.63 ± 1.0
GFDL-ESM4 CMIP6e	92.19 ± 2.39	34.82 ± 1.26	57.38 ± 1.56	50.38 ± 1.34
GFDL-ESM4 CMIP6c	88.07 ± 1.89	32.36 ± 1.01	55.71 ± 1.61	49.03 ± 1.2
INM-CM4-8 CMIP6c	123.16 ± 4.08	41.22 ± 1.16	81.94 ± 3.0	71.68 ± 2.69
INM-CM5-0 CMIP6c	126.98 ± 2.59	42.43 ± 1.25	84.54 ± 1.51	74.08 ± 1.36
IPSL-CM6A-LR CMIP6c	97.25 ± 2.63	32.08 ± 1.14	65.17 ± 1.86	54.93 ± 1.54
MIROC-ES2L CMIP6e	88.02 ± 1.45	37.68 ± 1.03	50.33 ± 0.59	43.68 ± 0.53
MIROC-ES2L CMIP6c	91.73 ± 1.73	39.27 ± 1.16	52.46 ± 0.84	45.33 ± 0.76
MPI-ESM-1-2-HAM CMIP6c	128.61 ± 5.77	46.29 ± 2.63	82.32 ± 3.65	67.99 ± 2.89
MPI-ESM1-2-LR CMIP6e	123.7 ± 3.79	48.04 ± 1.64	75.66 ± 3.0	64.42 ± 2.21
MPI-ESM1-2-LR CMIP6c	122.46 ± 3.66	47.92 ± 1.83	74.55 ± 2.43	63.63 ± 1.91
MRI-ESM2-0 CMIP6e	151.86 ± 4.17	63.44 ± 3.0	88.42 ± 2.98	79.52 ± 2.71
MRI-ESM2-0 CMIP6c	151.98 ± 4.62	63.74 ± 3.92	88.25 ± 2.09	79.5 ± 1.83
NorESM2-LM CMIP6e	88.78 ± 2.01	33.03 ± 1.0	55.75 ± 1.19	47.44 ± 1.03
NorESM2-LM CMIP6c	86.71 ± 1.82	32.15 ± 1.12	54.56 ± 1.04	46.21 ± 0.91
NorESM2-MM CMIP6c	88.13 ± 1.8	32.6 ± 0.91	55.53 ± 1.36	46.37 ± 1.09
SAM0-UNICON CMIP6c	106.57 ± 2.12	34.85 ± 0.94	71.72 ± 1.77	64.01 ± 1.77
TaiESM1 CMIP6c	106.56 ± 1.48	35.05 ± 1.03	71.51 ± 0.8	60.03 ± 0.65
UKESM1-0-LL CMIP6e	108.72 ± 3.39	38.47 ± 1.76	70.25 ± 2.79	59.24 ± 2.45

Continued on next page

Table B.3.: GPP Mean for 1986-2005 [PgC yr^{-1}]

Model	Global	NH	SH	Tropics
UKESM1-0-LL CMIP6c	106.73 ± 2.97	38.3 ± 1.47	68.43 ± 2.69	57.66 ± 2.18
GLASS	108.26 ± 3.43	39.19 ± 1.34	69.08 ± 2.28	59.14 ± 1.7
FLUXCOM	93.04 ± 0.41	31.09 ± 0.2	61.95 ± 0.49	55.34 ± 0.31
MTE	102.63 ± 1.17	37.5 ± 0.75	65.13 ± 0.68	56.72 ± 0.5
Ncycle MMM CMIP5e	107.32 ± 1.88	35.4 ± 0.9	71.92 ± 1.12	61.57 ± 0.98
non-Ncycle MMM CMIP5e	132.9 ± 2.07	45.37 ± 0.91	87.52 ± 1.53	73.29 ± 1.18
Ncycle MMM CMIP5c	106.47 ± 1.6	34.96 ± 0.82	71.51 ± 0.87	61.18 ± 0.78
non-Ncycle MMM CMIP5c	120.64 ± 2.23	45.22 ± 1.1	75.41 ± 1.26	63.41 ± 1.1
Ncycle MMM CMIP6c	99.85 ± 1.59	37.89 ± 1.05	61.95 ± 0.64	51.53 ± 0.57
non-Ncycle MMM CMIP6c	114.3 ± 2.6	41.57 ± 1.45	72.73 ± 1.22	62.09 ± 1.04
Ncycle MMM CMIP6e	105.18 ± 2.15	39.59 ± 1.32	65.59 ± 1.09	54.68 ± 0.91
non-Ncycle MMM CMIP6e	111.52 ± 2.12	42.71 ± 1.38	68.81 ± 1.05	58.57 ± 0.87

Table B.4.: GPP Trend for 1986-2005 [PgC yr^{-2}]

Model	Global	NH	SH	Tropics
BNU-ESM CMIP5e	0.3915 ± 0.0353	0.1594 ± 0.0357	0.2321 ± 0.0283	0.1665 ± 0.026
BNU-ESM CMIP5c	0.3177 ± 0.0272	0.2067 ± 0.023	0.111 ± 0.0258	0.1123 ± 0.0257
CESM1-BGC CMIP5e	0.2964 ± 0.0552	0.1536 ± 0.0246	0.1428 ± 0.0395	0.1127 ± 0.0359
CESM1-BGC CMIP5c	0.2269 ± 0.0419	0.0828 ± 0.0242	0.1441 ± 0.0254	0.1322 ± 0.02
CanESM2 CMIP5e	0.1535 ± 0.1023	0.113 ± 0.0215	0.0405 ± 0.1098	0.0422 ± 0.0818
CanESM2 CMIP5c	0.1575 ± 0.0886	0.0826 ± 0.0222	0.075 ± 0.0812	0.0879 ± 0.0681
FIO-ESM CMIP5e	0.3111 ± 0.0536	0.1364 ± 0.0228	0.1746 ± 0.0536	0.1641 ± 0.0502
GFDL-ESM2G CMIP5e	0.4753 ± 0.1595	0.161 ± 0.0312	0.3143 ± 0.1598	0.2258 ± 0.0971
GFDL-ESM2G CMIP5c	0.4292 ± 0.098	0.2338 ± 0.0376	0.1955 ± 0.1043	0.1751 ± 0.0672

Continued on next page

Table B.4.: GPP Trend for 1986-2005 [PgC yr^{-2}]

Model	Global	NH	SH	Tropics
GFDL-ESM2M CMIP5e	0.1855 ± 0.2006	0.0493 ± 0.0553	0.1362 ± 0.2171	0.1544 ± 0.1496
GFDL-ESM2M CMIP5c	0.2511 ± 0.2294	0.1939 ± 0.0321	0.0571 ± 0.2438	0.1613 ± 0.166
HadGEM2-CC CMIP5c	0.4561 ± 0.0713	0.2295 ± 0.0498	0.2266 ± 0.0649	0.2017 ± 0.069
HadGEM2-ES CMIP5c	0.5642 ± 0.0741	0.159 ± 0.0395	0.4051 ± 0.0774	0.3473 ± 0.0763
IPSL-CM5A-LR CMIP5c	0.2351 ± 0.0689	0.2287 ± 0.0394	0.0064 ± 0.0635	0.0757 ± 0.0559
IPSL-CM5A-MR CMIP5c	0.4459 ± 0.0747	0.2547 ± 0.0519	0.1913 ± 0.0657	0.2099 ± 0.0574
IPSL-CM5B-LR CMIP5c	0.3361 ± 0.0697	0.2477 ± 0.0336	0.0884 ± 0.0716	0.0885 ± 0.0619
MIROC-ESM-CHEM CMIP5c	0.3517 ± 0.0552	0.1065 ± 0.0287	0.2452 ± 0.0517	0.2028 ± 0.0443
MIROC-ESM CMIP5e	0.3434 ± 0.0589	0.1697 ± 0.0404	0.1738 ± 0.0613	0.1635 ± 0.0469
MIROC-ESM CMIP5c	0.3824 ± 0.061	0.129 ± 0.0285	0.2534 ± 0.0556	0.2379 ± 0.0416
MPI-ESM-LR CMIP5e	0.3901 ± 0.0835	0.2193 ± 0.0505	0.1709 ± 0.0881	0.1509 ± 0.0378
MPI-ESM-LR CMIP5c	0.5023 ± 0.1059	0.1328 ± 0.0358	0.3695 ± 0.1157	0.2491 ± 0.0484
MPI-ESM-MR CMIP5c	0.5187 ± 0.0816	0.2771 ± 0.0523	0.2415 ± 0.0727	0.2042 ± 0.0405
MRI-ESM1 CMIP5e	0.2532 ± 0.1083	0.0754 ± 0.0558	0.1777 ± 0.0678	0.1459 ± 0.0596
NorESM1-ME CMIP5e	0.249 ± 0.037	0.12 ± 0.0214	0.129 ± 0.0349	0.1087 ± 0.0349
NorESM1-ME CMIP5c	0.2346 ± 0.0367	0.1483 ± 0.0192	0.0862 ± 0.0318	0.0794 ± 0.0268
inmcm4 CMIP5c	0.1868 ± 0.0651	0.0665 ± 0.0276	0.1203 ± 0.0611	0.0776 ± 0.0546
ACCESS-ESM1-5 CMIP6e	0.1563 ± 0.0361	0.1833 ± 0.0153	-0.0269 ± 0.0366	-0.0358 ± 0.034

Continued on next page

Table B.4.: GPP Trend for 1986-2005 [PgC yr^{-2}]

Model	Global	NH	SH	Tropics
ACCESS-ESM1-5 CMIP6c	0.03 ± 0.0503	0.146 ± 0.011	-0.116 ± 0.0509	-0.0997 ± 0.046
CESM2-WACCM CMIP6c	0.2483 ± 0.0368	0.1584 ± 0.0245	0.0899 ± 0.028	0.0957 ± 0.0206
CESM2 CMIP6c	0.3699 ± 0.0433	0.2041 ± 0.0273	0.1658 ± 0.0385	0.1402 ± 0.0306
CMCC-CM2-SR5 CMIP6c	0.275 ± 0.0566	0.1526 ± 0.0351	0.1224 ± 0.0477	0.0931 ± 0.0278
CMCC-ESM2 CMIP6c	0.2432 ± 0.032	0.1699 ± 0.0283	0.0732 ± 0.0328	0.0556 ± 0.0212
CNRM-ESM2-1 CMIP6e	0.0254 ± 0.0855	0.1013 ± 0.0376	-0.0759 ± 0.0795	-0.0334 ± 0.045
CNRM-ESM2-1 CMIP6c	0.2927 ± 0.0639	0.2179 ± 0.0424	0.0749 ± 0.062	0.0456 ± 0.0354
CanESM5-CanOE CMIP6e	0.3806 ± 0.0666	0.1786 ± 0.0312	0.2019 ± 0.0517	0.1724 ± 0.0406
CanESM5-CanOE CMIP6c	0.4188 ± 0.0691	0.1931 ± 0.0252	0.2256 ± 0.0636	0.1936 ± 0.051
CanESM5 CMIP6e	0.3736 ± 0.0545	0.1698 ± 0.0268	0.2038 ± 0.0499	0.1356 ± 0.0364
CanESM5 CMIP6c	0.3238 ± 0.0402	0.143 ± 0.0272	0.1808 ± 0.026	0.1474 ± 0.0211
EC-Earth3-CC CMIP6e	0.4508 ± 0.0697	0.335 ± 0.0451	0.1159 ± 0.052	0.1045 ± 0.029
EC-Earth3-CC CMIP6c	0.2899 ± 0.0724	0.2531 ± 0.0357	0.0369 ± 0.0578	0.0732 ± 0.0407
EC-Earth3-Veg CMIP6c	0.2826 ± 0.0613	0.2198 ± 0.0319	0.0628 ± 0.0489	0.097 ± 0.0325
GFDL-ESM4 CMIP6e	0.3558 ± 0.0449	0.1912 ± 0.0223	0.1645 ± 0.0485	0.1283 ± 0.0438
GFDL-ESM4 CMIP6c	0.1299 ± 0.069	0.1378 ± 0.0236	-0.0079 ± 0.0643	-0.0335 ± 0.0472
INM-CM4-8 CMIP6c	0.5962 ± 0.0823	0.1647 ± 0.0249	0.4314 ± 0.0631	0.3916 ± 0.0546
INM-CM5-0 CMIP6c	0.3037 ± 0.0746	0.1643 ± 0.0314	0.1394 ± 0.0504	0.1419 ± 0.0427

Continued on next page

Table B.4.: GPP Trend for 1986-2005 [PgC yr^{-2}]

Model	Global	NH	SH	Tropics
IPSL-CM6A-LR CMIP6c	0.4168 ± 0.0369	0.1655 ± 0.0233	0.2512 ± 0.0446	0.2223 ± 0.0316
MIROC-ES2L CMIP6e	0.2258 ± 0.0223	0.1439 ± 0.0232	0.0819 ± 0.0134	0.078 ± 0.0103
MIROC-ES2L CMIP6c	0.2643 ± 0.0295	0.1505 ± 0.0295	0.1138 ± 0.0201	0.105 ± 0.0173
MPI-ESM-1-2-HAM CMIP6c	0.7854 ± 0.136	0.3564 ± 0.0625	0.429 ± 0.1046	0.371 ± 0.075
MPI-ESM1-2-LR CMIP6e	0.4284 ± 0.1123	0.2526 ± 0.0273	0.1758 ± 0.1121	0.1707 ± 0.0784
MPI-ESM1-2-LR CMIP6c	0.5008 ± 0.0854	0.2469 ± 0.044	0.2539 ± 0.0763	0.2459 ± 0.0491
MRI-ESM2-0 CMIP6e	0.4292 ± 0.1318	0.4558 ± 0.0525	-0.0266 ± 0.1186	-0.0055 ± 0.108
MRI-ESM2-0 CMIP6c	0.6198 ± 0.1118	0.5943 ± 0.069	0.0254 ± 0.083	0.0384 ± 0.0723
NorESM2-LM CMIP6e	0.2863 ± 0.0433	0.1435 ± 0.0213	0.1429 ± 0.0332	0.1232 ± 0.029
NorESM2-LM CMIP6c	0.2658 ± 0.0363	0.1524 ± 0.0266	0.1135 ± 0.0317	0.0902 ± 0.0296
NorESM2-MM CMIP6c	0.2231 ± 0.0491	0.1395 ± 0.0152	0.0836 ± 0.0505	0.0667 ± 0.0406
SAM0-UNICON CMIP6c	0.2424 ± 0.0624	0.1189 ± 0.0247	0.1234 ± 0.0642	0.1242 ± 0.064
TaiESM1 CMIP6c	0.2062 ± 0.0334	0.1346 ± 0.0259	0.0716 ± 0.0269	0.0599 ± 0.0216
UKESM1-0-LL CMIP6e	0.461 ± 0.0834	0.2334 ± 0.0454	0.2277 ± 0.0993	0.2065 ± 0.0864
UKESM1-0-LL CMIP6c	0.1936 ± 0.1111	0.1672 ± 0.0442	0.0264 ± 0.1086	0.0476 ± 0.0874
GLASS	0.4466 ± 0.0871	0.1759 ± 0.0337	0.2707 ± 0.0646	0.1939 ± 0.0502
FLUXCOM	-	-	-	-
MTE	-	-	-	-
Ncycle MMM CMIP5e	0.2727 ± 0.0388	0.1368 ± 0.0158	0.1359 ± 0.031	0.1107 ± 0.0293

Continued on next page

Table B.4.: GPP Trend for 1986-2005 [PgC yr^{-2}]

Model	Global	NH	SH	Tropics
non-Ncycle MMM CMIP5e	0.3127 ± 0.0367	0.1354 ± 0.0176	0.1774 ± 0.0444	0.1516 ± 0.0306
Ncycle MMM CMIP5c	0.2307 ± 0.0332	0.1156 ± 0.0179	0.1152 ± 0.0217	0.1058 ± 0.0187
non-Ncycle MMM CMIP5c	0.3657 ± 0.0224	0.1816 ± 0.0092	0.1841 ± 0.0251	0.1731 ± 0.0163
Ncycle MMM CMIP6c	0.2594 ± 0.0162	0.1722 ± 0.0098	0.0871 ± 0.0148	0.0853 ± 0.0106
non-Ncycle MMM CMIP6c	0.4317 ± 0.0191	0.2374 ± 0.014	0.1943 ± 0.0164	0.1686 ± 0.012
Ncycle MMM CMIP6e	0.3336 ± 0.0335	0.2147 ± 0.0146	0.119 ± 0.0334	0.1073 ± 0.0258
non-Ncycle MMM CMIP6e	0.3128 ± 0.0413	0.2193 ± 0.0187	0.0935 ± 0.0357	0.0794 ± 0.0293

Table B.5.: NBP Mean for 1986-2005 [PgC yr^{-1}]

Model	Global	NH	SH	Tropics
BNU-ESM CMIP5e	0.36 ± 0.92	0.39 ± 0.69	-0.03 ± 0.94	0.01 ± 0.96
BNU-ESM CMIP5c	1.32 ± 1.15	0.89 ± 0.59	0.43 ± 1.04	0.39 ± 0.99
CESM1-BGC CMIP5e	-0.08 ± 0.97	0.06 ± 0.52	-0.14 ± 0.65	-0.11 ± 0.55
CESM1-BGC CMIP5c	-0.14 ± 0.67	0.02 ± 0.36	-0.17 ± 0.6	-0.12 ± 0.41
CanESM2 CMIP5e	0.3 ± 2.47	0.23 ± 0.54	0.08 ± 2.59	0.12 ± 1.94
CanESM2 CMIP5c	0.21 ± 2.07	0.11 ± 0.34	0.1 ± 1.89	0.07 ± 1.4
FIO-ESM CMIP5e	1.84 ± 0.64	0.78 ± 0.17	1.06 ± 0.61	0.91 ± 0.62
GFDL-ESM2G CMIP5e	0.9 ± 2.95	0.39 ± 0.78	0.5 ± 2.83	0.23 ± 2.01
GFDL-ESM2G CMIP5c	0.62 ± 2.16	0.39 ± 0.9	0.23 ± 1.86	0.01 ± 1.3
GFDL-ESM2M CMIP5e	0.31 ± 3.0	0.22 ± 1.09	0.08 ± 3.54	0.02 ± 2.74
GFDL-ESM2M CMIP5c	0.55 ± 3.91	0.29 ± 0.94	0.26 ± 4.04	0.09 ± 3.22
HadGEM2-CC CMIP5c	0.98 ± 1.29	0.82 ± 0.67	0.16 ± 1.06	0.09 ± 1.02
HadGEM2-ES CMIP5c	1.35 ± 1.21	0.76 ± 0.39	0.58 ± 1.17	0.53 ± 1.09
IPSL-CM5A-LR CMIP5c	0.73 ± 1.16	0.45 ± 0.6	0.27 ± 1.07	0.36 ± 0.99
IPSL-CM5A-MR CMIP5c	0.97 ± 1.46	0.44 ± 0.77	0.53 ± 1.26	0.63 ± 1.12
IPSL-CM5B-LR CMIP5c	1.57 ± 1.31	0.84 ± 0.48	0.73 ± 1.33	0.75 ± 1.22

Continued on next page

Table B.5.: NBP Mean for 1986-2005 [PgC yr^{-1}]

Model	Global	NH	SH	Tropics
MIROC-ESM-CHEM CMIP5c	1.68 ± 1.41	0.91 ± 0.51	0.77 ± 1.27	0.61 ± 1.1
MIROC-ESM CMIP5e	1.19 ± 1.66	1.28 ± 0.59	-0.1 ± 1.53	-0.16 ± 1.31
MIROC-ESM CMIP5c	0.65 ± 1.52	1.14 ± 0.61	-0.49 ± 1.37	-0.44 ± 1.06
MPI-ESM-LR CMIP5e	1.28 ± 1.16	1.2 ± 0.69	0.08 ± 0.94	-0.02 ± 0.45
MPI-ESM-LR CMIP5c	1.49 ± 1.72	1.08 ± 0.64	0.41 ± 1.39	0.11 ± 0.77
MPI-ESM-MR CMIP5c	1.69 ± 1.25	1.24 ± 0.63	0.45 ± 1.01	0.19 ± 0.64
MRI-ESM1 CMIP5e	2.05 ± 0.98	1.76 ± 0.54	0.29 ± 0.8	0.29 ± 0.74
NorESM1-ME CMIP5e	-0.11 ± 0.55	-0.02 ± 0.42	-0.1 ± 0.55	-0.09 ± 0.45
NorESM1-ME CMIP5c	-0.08 ± 0.84	-0.04 ± 0.37	-0.04 ± 0.71	-0.05 ± 0.58
inmcm4 CMIP5c	2.18 ± 1.15	0.72 ± 0.5	1.46 ± 1.07	1.27 ± 0.97
ACCESS-ESM1-5 CMIP6e	1.31 ± 0.72	0.84 ± 0.34	0.47 ± 0.75	0.36 ± 0.64
ACCESS-ESM1-5 CMIP6c	0.77 ± 1.13	0.78 ± 0.31	-0.01 ± 0.97	-0.06 ± 0.84
CESM2-WACCM CMIP6c	1.03 ± 0.54	0.51 ± 0.34	0.52 ± 0.35	0.5 ± 0.23
CESM2 CMIP6c	1.08 ± 0.58	0.6 ± 0.32	0.49 ± 0.4	0.46 ± 0.35
CMCC-CM2-SR5 CMIP6c	0.89 ± 1.12	0.19 ± 0.58	0.7 ± 0.89	0.63 ± 0.54
CMCC-ESM2 CMIP6c	0.89 ± 0.63	0.22 ± 0.52	0.67 ± 0.57	0.62 ± 0.35
CNRM-ESM2-1 CMIP6e	1.86 ± 1.02	1.46 ± 0.45	0.4 ± 0.82	0.25 ± 0.52
CNRM-ESM2-1 CMIP6c	1.89 ± 0.69	1.55 ± 0.53	0.34 ± 0.65	0.16 ± 0.44
CanESM5-CanOE CMIP6e	0.51 ± 1.46	0.13 ± 0.51	0.38 ± 1.15	0.41 ± 1.0
CanESM5-CanOE CMIP6c	0.08 ± 1.87	-0.05 ± 0.52	0.13 ± 1.54	0.21 ± 1.36
CanESM5 CMIP6e	0.62 ± 1.29	0.19 ± 0.32	0.43 ± 1.12	0.39 ± 0.95
CanESM5 CMIP6c	0.44 ± 0.69	0.15 ± 0.32	0.29 ± 0.51	0.35 ± 0.47
EC-Earth3-CC CMIP6e	0.62 ± 1.77	0.6 ± 0.86	0.01 ± 1.48	0.28 ± 1.02
EC-Earth3-CC CMIP6c	0.44 ± 1.8	0.51 ± 1.0	-0.07 ± 1.73	0.16 ± 1.36
EC-Earth3-Veg CMIP6c	0.49 ± 1.69	0.41 ± 0.79	0.08 ± 1.36	0.27 ± 1.12
GFDL-ESM4 CMIP6e	0.28 ± 1.28	0.19 ± 0.35	0.09 ± 1.12	0.12 ± 1.04
GFDL-ESM4 CMIP6c	-0.13 ± 1.63	-0.08 ± 0.38	-0.05 ± 1.44	0.05 ± 1.15
INM-CM4-8 CMIP6c	4.46 ± 2.21	1.55 ± 0.75	2.91 ± 1.54	2.55 ± 1.23
INM-CM5-0 CMIP6c	3.7 ± 2.05	1.44 ± 0.85	2.25 ± 1.36	1.94 ± 1.19
IPSL-CM6A-LR CMIP6c	1.06 ± 0.72	0.83 ± 0.21	0.22 ± 0.71	0.19 ± 0.57
MIROC-ES2L CMIP6e	1.61 ± 0.64	1.38 ± 0.34	0.23 ± 0.52	0.27 ± 0.42
MIROC-ES2L CMIP6c	1.99 ± 1.24	1.49 ± 0.4	0.5 ± 1.03	0.46 ± 0.95
MPI-ESM-1-2-HAM CMIP6c	1.04 ± 1.31	0.33 ± 0.4	0.71 ± 1.19	0.58 ± 0.91
MPI-ESM1-2-LR CMIP6e	1.09 ± 1.47	0.59 ± 0.42	0.5 ± 1.46	0.45 ± 1.15
MPI-ESM1-2-LR CMIP6c	0.89 ± 1.06	0.44 ± 0.52	0.45 ± 1.03	0.42 ± 0.82

Continued on next page

Table B.5.: NBP Mean for 1986-2005 [PgC yr^{-1}]

Model	Global	NH	SH	Tropics
MRI-ESM2-0 CMIP6e	1.33 ± 1.42	1.42 ± 0.65	-0.09 ± 1.3	-0.16 ± 1.25
MRI-ESM2-0 CMIP6c	1.54 ± 1.21	1.56 ± 0.72	-0.02 ± 1.16	-0.05 ± 1.05
NorESM2-LM CMIP6e	1.08 ± 0.66	0.53 ± 0.37	0.55 ± 0.39	0.52 ± 0.36
NorESM2-LM CMIP6c	0.94 ± 0.53	0.44 ± 0.3	0.5 ± 0.39	0.46 ± 0.39
NorESM2-MM CMIP6c	1.07 ± 0.69	0.58 ± 0.25	0.49 ± 0.5	0.44 ± 0.44
SAM0-UNICON CMIP6c	0.06 ± 1.02	-0.01 ± 0.58	0.07 ± 1.14	0.09 ± 0.97
TaiESM1 CMIP6c	-0.11 ± 0.8	0.13 ± 0.45	-0.24 ± 0.83	-0.19 ± 0.4
UKESM1-0-LL CMIP6e	1.0 ± 1.06	0.63 ± 0.38	0.38 ± 1.18	0.3 ± 0.96
UKESM1-0-LL CMIP6c	0.63 ± 1.01	0.37 ± 0.29	0.26 ± 1.1	0.22 ± 0.93
CAMS	1.72 ± 0.98	1.86 ± 0.47	-0.14 ± 0.74	-0.47 ± 0.59
Jena_CarboScope	0.71 ± 0.94	1.25 ± 0.36	-0.54 ± 0.72	-0.24 ± 0.62
GCP	1.0 ± 0.84	NaN	NaN	NaN
Ncycle MMM CMIP5e	-0.09 ± 0.59	0.02 ± 0.36	-0.12 ± 0.45	-0.1 ± 0.41
non-Ncycle MMM CMIP5e	1.0 ± 0.68	0.71 ± 0.3	0.29 ± 0.76	0.22 ± 0.59
Ncycle MMM CMIP5c	-0.11 ± 0.48	-0.01 ± 0.29	-0.1 ± 0.43	-0.09 ± 0.34
non-Ncycle MMM CMIP5c	1.14 ± 0.66	0.67 ± 0.29	0.47 ± 0.52	0.37 ± 0.43
Ncycle MMM CMIP6c	0.84 ± 0.34	0.48 ± 0.16	0.37 ± 0.28	0.34 ± 0.22
non-Ncycle MMM CMIP6c	1.56 ± 0.62	0.81 ± 0.25	0.76 ± 0.47	0.66 ± 0.39
Ncycle MMM CMIP6e	1.22 ± 0.57	0.79 ± 0.24	0.43 ± 0.54	0.38 ± 0.42
non-Ncycle MMM CMIP6e	0.92 ± 0.48	0.68 ± 0.27	0.24 ± 0.39	0.2 ± 0.36

Table B.6.: NBP Trend for 1986-2005 [PgC yr^{-2}]

Model	Global	NH	SH	Tropics
BNU-ESM CMIP5e	0.0343 ± 0.0358	0.0217 ± 0.0269	0.0127 ± 0.0373	-0.0003 ± 0.0385
BNU-ESM CMIP5c	-0.0049 ± 0.0457	-0.0067 ± 0.0235	0.0018 ± 0.0415	0.0087 ± 0.0396
CESM1-BGC CMIP5e	0.0468 ± 0.037	0.0331 ± 0.0192	0.0137 ± 0.0256	0.0117 ± 0.0219
CESM1-BGC CMIP5c	0.0251 ± 0.0259	0.0255 ± 0.0131	-0.0003 ± 0.024	0.0037 ± 0.0164
CanESM2 CMIP5e	-0.0427 ± 0.0979	0.0048 ± 0.0216	-0.0475 ± 0.1028	-0.0404 ± 0.0767

Continued on next page

Table B.6.: NBP Trend for 1986-2005 [PgC yr^{-2}]

Model	Global	NH	SH	Tropics
CanESM2 CMIP5c	-0.0579 ± 0.0815	-0.0105 ± 0.0135	-0.0473 ± 0.0746	-0.0264 ± 0.0555
FIO-ESM CMIP5e	0.0323 ± 0.0243	0.0128 ± 0.006	0.0195 ± 0.024	0.0218 ± 0.0242
GFDL-ESM2G CMIP5e	0.0966 ± 0.1154	0.0597 ± 0.0276	0.037 ± 0.1126	0.0378 ± 0.0796
GFDL-ESM2G CMIP5c	0.0726 ± 0.0845	0.0578 ± 0.0332	0.0148 ± 0.074	0.0319 ± 0.0514
GFDL-ESM2M CMIP5e	-0.0084 ± 0.1195	0.0012 ± 0.0434	-0.0096 ± 0.1412	-0.0308 ± 0.1091
GFDL-ESM2M CMIP5c	0.0941 ± 0.1543	0.0595 ± 0.0349	0.0347 ± 0.1607	0.047 ± 0.1277
HadGEM2-CC CMIP5c	0.1189 ± 0.0439	0.0694 ± 0.0214	0.0495 ± 0.0411	0.034 ± 0.0405
HadGEM2-ES CMIP5c	0.1342 ± 0.0374	0.0448 ± 0.0118	0.0893 ± 0.0422	0.0667 ± 0.0411
IPSL-CM5A-LR CMIP5c	0.0674 ± 0.0435	0.0511 ± 0.0209	0.0162 ± 0.0424	0.0205 ± 0.0391
IPSL-CM5A-MR CMIP5c	0.0369 ± 0.0575	0.0219 ± 0.0302	0.0151 ± 0.0503	0.0138 ± 0.0446
IPSL-CM5B-LR CMIP5c	0.0204 ± 0.0518	0.0379 ± 0.0169	-0.0175 ± 0.0527	-0.0116 ± 0.0484
MIROC-ESM-CHEM CMIP5c	0.0672 ± 0.054	0.0082 ± 0.0201	0.0591 ± 0.0484	0.0493 ± 0.0424
MIROC-ESM CMIP5e	0.1939 ± 0.0477	0.0203 ± 0.0229	0.1736 ± 0.0453	0.1644 ± 0.035
MIROC-ESM CMIP5c	0.105 ± 0.0553	0.0004 ± 0.0242	0.1046 ± 0.0487	0.0925 ± 0.0364
MPI-ESM-LR CMIP5e	0.0116 ± 0.046	0.0065 ± 0.0275	0.0051 ± 0.0372	0.0074 ± 0.0179
MPI-ESM-LR CMIP5c	0.0878 ± 0.0655	0.0575 ± 0.0218	0.0303 ± 0.0551	0.033 ± 0.0296
MPI-ESM-MR CMIP5c	0.0474 ± 0.0487	0.0335 ± 0.0236	0.0139 ± 0.0402	0.0119 ± 0.0254

Continued on next page

Table B.6.: NBP Trend for 1986-2005 [PgC yr^{-2}]

Model	Global	NH	SH	Tropics
MRI-ESM1 CMIP5e	-0.0085 ± 0.0391	-0.0036 ± 0.0215	-0.0049 ± 0.0318	-0.0103 ± 0.0294
NorESM1-ME CMIP5e	0.0399 ± 0.0198	0.0158 ± 0.0162	0.024 ± 0.0211	0.0238 ± 0.0172
NorESM1-ME CMIP5c	0.0213 ± 0.033	0.0286 ± 0.0132	-0.0073 ± 0.0282	-0.0087 ± 0.0232
inmcm4 CMIP5c	0.0176 ± 0.0457	0.0135 ± 0.0198	0.0042 ± 0.0427	-0.0169 ± 0.0384
ACCESS-ESM1-5 CMIP6e	-0.0457 ± 0.0266	0.0185 ± 0.0129	-0.0642 ± 0.0257	-0.0521 ± 0.0222
ACCESS-ESM1-5 CMIP6c	-0.11 ± 0.0366	-0.0195 ± 0.0115	-0.0906 ± 0.0323	-0.0735 ± 0.0288
CESM2-WACCM CMIP6c	0.0307 ± 0.0202	0.0203 ± 0.0127	0.0105 ± 0.0137	0.011 ± 0.0088
CESM2 CMIP6c	0.0631 ± 0.0179	0.0302 ± 0.0104	0.0329 ± 0.0141	0.0292 ± 0.0122
CMCC-CM2-SR5 CMIP6c	0.0156 ± 0.0447	-0.0046 ± 0.0233	0.0202 ± 0.0351	0.0145 ± 0.0213
CMCC-ESM2 CMIP6c	0.0217 ± 0.0246	0.011 ± 0.0206	0.0107 ± 0.0226	0.0091 ± 0.0137
CNRM-ESM2-1 CMIP6e	-0.0198 ± 0.0405	0.0175 ± 0.0174	-0.0373 ± 0.0315	-0.0247 ± 0.0201
CNRM-ESM2-1 CMIP6c	0.0343 ± 0.0263	0.0406 ± 0.0189	-0.0064 ± 0.0259	-0.0063 ± 0.0175
CanESM5-CanOE CMIP6e	0.0559 ± 0.0568	0.0509 ± 0.0162	0.005 ± 0.046	0.01 ± 0.0399
CanESM5-CanOE CMIP6c	0.0923 ± 0.0712	0.0572 ± 0.0159	0.0352 ± 0.0608	0.0353 ± 0.0535
CanESM5 CMIP6e	-0.0218 ± 0.0512	0.0109 ± 0.0124	-0.0327 ± 0.0439	-0.0484 ± 0.0362
CanESM5 CMIP6c	-0.0028 ± 0.0273	0.0117 ± 0.0127	-0.0145 ± 0.0198	-0.0122 ± 0.0184
EC-Earth3-CC CMIP6e	0.0115 ± 0.0703	0.033 ± 0.0332	-0.0216 ± 0.0587	-0.0121 ± 0.0406
EC-Earth3-CC CMIP6c	0.0138 ± 0.0715	-0.0066 ± 0.0397	0.0205 ± 0.0686	0.0151 ± 0.0539

Continued on next page

Table B.6.: NBP Trend for 1986-2005 [PgC yr^{-2}]

Model	Global	NH	SH	Tropics
EC-Earth3-Veg CMIP6c	-0.0126 ± 0.0672	0.0324 ± 0.0306	-0.045 ± 0.0533	-0.0237 ± 0.0444
GFDL-ESM4 CMIP6e	0.0603 ± 0.0492	0.0292 ± 0.0119	0.0311 ± 0.0442	0.0304 ± 0.0409
GFDL-ESM4 CMIP6c	-0.0066 ± 0.0649	0.0197 ± 0.0143	-0.0263 ± 0.057	-0.0383 ± 0.045
INM-CM4-8 CMIP6c	0.0267 ± 0.088	0.004 ± 0.0298	0.0228 ± 0.0611	0.0229 ± 0.0487
INM-CM5-0 CMIP6c	-0.0705 ± 0.0801	-0.002 ± 0.034	-0.0685 ± 0.0519	-0.0462 ± 0.0462
IPSL-CM6A-LR CMIP6c	0.0542 ± 0.0255	0.0158 ± 0.0077	0.0384 ± 0.0269	0.0305 ± 0.0217
MIROC-ES2L CMIP6e	0.0081 ± 0.0253	0.0155 ± 0.0131	-0.0074 ± 0.0208	-0.0078 ± 0.0165
MIROC-ES2L CMIP6c	-0.0516 ± 0.0477	-0.0141 ± 0.0156	-0.0375 ± 0.0402	-0.0319 ± 0.037
MPI-ESM-1-2-HAM CMIP6c	0.0718 ± 0.0496	0.021 ± 0.0152	0.0508 ± 0.0459	0.0473 ± 0.0344
MPI-ESM1-2-LR CMIP6e	0.0644 ± 0.0565	0.0436 ± 0.0134	0.0208 ± 0.0578	0.0212 ± 0.0456
MPI-ESM1-2-LR CMIP6c	0.0487 ± 0.0407	0.0327 ± 0.0191	0.016 ± 0.041	0.0279 ± 0.0319
MRI-ESM2-0 CMIP6e	0.0356 ± 0.0561	0.0595 ± 0.0215	-0.0239 ± 0.0514	-0.0229 ± 0.0494
MRI-ESM2-0 CMIP6c	0.0425 ± 0.0473	0.0975 ± 0.0169	-0.055 ± 0.0444	-0.0489 ± 0.0404
NorESM2-LM CMIP6e	0.0472 ± 0.024	0.0205 ± 0.0141	0.0266 ± 0.0141	0.0249 ± 0.0129
NorESM2-LM CMIP6c	0.0367 ± 0.0191	0.0206 ± 0.0111	0.0161 ± 0.0151	0.0157 ± 0.0152
NorESM2-MM CMIP6c	0.0179 ± 0.0273	0.0103 ± 0.0098	0.0076 ± 0.0198	0.0052 ± 0.0174
SAM0-UNICON CMIP6c	0.0417 ± 0.0394	0.0177 ± 0.0226	0.024 ± 0.0452	0.0219 ± 0.0383
TaiESM1 CMIP6c	-0.0034 ± 0.032	0.0129 ± 0.0176	-0.0162 ± 0.0327	-0.0092 ± 0.0159

Continued on next page

Table B.6.: NBP Trend for 1986-2005 [PgC yr^{-2}]

Model	Global	NH	SH	Tropics
UKESM1-0-LL CMIP6e	0.0335 ± 0.042	0.0409 ± 0.0122	-0.0074 ± 0.0475	-0.0034 ± 0.039
UKESM1-0-LL CMIP6c	-0.0341 ± 0.0399	0.0193 ± 0.011	-0.0534 ± 0.0425	-0.0504 ± 0.0356
CAMS	0.0545 ± 0.037	0.0238 ± 0.0178	0.0307 ± 0.0285	0.0186 ± 0.0232
Jena_CarboScope	0.0006 ± 0.0374	0.0085 ± 0.0143	-0.0079 ± 0.0287	-0.0038 ± 0.0246
GCP	0.0255 ± 0.033	NaN	NaN	NaN
Ncycle MMM CMIP5e	0.0433 ± 0.0213	0.0245 ± 0.0133	0.0189 ± 0.0175	0.0177 ± 0.0158
non-Ncycle MMM CMIP5e	0.0164 ± 0.027	0.0147 ± 0.0113	0.0017 ± 0.0304	-0.0021 ± 0.0236
Ncycle MMM CMIP5c	0.0232 ± 0.0185	0.0271 ± 0.0094	-0.0038 ± 0.0171	-0.0025 ± 0.0136
non-Ncycle MMM CMIP5c	0.0526 ± 0.0232	0.0357 ± 0.0082	0.0169 ± 0.0204	0.0176 ± 0.0167
Ncycle MMM CMIP6c	0.0064 ± 0.0133	0.0114 ± 0.0057	-0.0049 ± 0.0109	-0.0025 ± 0.0086
non-Ncycle MMM CMIP6c	0.0269 ± 0.024	0.0295 ± 0.007	-0.0026 ± 0.0187	-0.0018 ± 0.0155
Ncycle MMM CMIP6e	0.0214 ± 0.022	0.0277 ± 0.0068	-0.0063 ± 0.0216	-0.0034 ± 0.0168
non-Ncycle MMM CMIP6e	0.022 ± 0.0184	0.0336 ± 0.0076	-0.0116 ± 0.0153	-0.0111 ± 0.014

B.2. Mean Carbon stocks

Data values for Figure 5.15.

Table B.7.: Carbon Storage Totals [PgC]

	cVeg	cSoil + cLitter
BNU-ESM CMIP5e	1205	856
BNU-ESM CMIP5c	1259	913
CESM1-BGC CMIP5e	538	515
CESM1-BGC CMIP5c	532	513

Continued on next page

Table B.7.: Carbon Storage Totals [PgC]

	cVeg	cSoil + cLitter
GFDL-ESM2G CMIP5e	680	1434
GFDL-ESM2G CMIP5c	660	1422
GFDL-ESM2M CMIP5e	802	1292
GFDL-ESM2M CMIP5c	766	1279
HadGEM2-CC CMIP5c	445	1103
HadGEM2-ES CMIP5c	458	1109
IPSL-CM5A-LR CMIP5c	634	1355
IPSL-CM5A-MR CMIP5c	639	1400
IPSL-CM5B-LR CMIP5c	501	1267
MIROC-ESM-CHEM CMIP5c	351	2567
MIROC-ESM CMIP5e	364	2564
MIROC-ESM CMIP5c	357	2572
MPI-ESM-LR CMIP5e	341	3039
MPI-ESM-LR CMIP5c	335	3036
MPI-ESM-MR CMIP5c	345	3092
MRI-ESM1 CMIP5e	562	1419
NorESM1-ME CMIP5e	556	552
NorESM1-ME CMIP5c	554	550
ACCESS-ESM1-5 CMIP6e	653	904
ACCESS-ESM1-5 CMIP6c	661	914
CESM2-WACCM CMIP6c	574	1201
CESM2 CMIP6c	482	1000
CMCC-CM2-SR5 CMIP6c	406	2908
CMCC-ESM2 CMIP6c	416	2913
CNRM-ESM2-1 CMIP6e	724	2279
CNRM-ESM2-1 CMIP6c	719	2281
CanESM5-CanOE CMIP6e	508	1475
CanESM5 CMIP6e	507	1470
CanESM5 CMIP6c	513	1486
EC-Earth3-CC CMIP6e	395	1546
EC-Earth3-CC CMIP6c	382	1535
EC-Earth3-Veg CMIP6c	333	1367
GFDL-ESM4 CMIP6c	342	514
IPSL-CM6A-LR CMIP6c	352	649
MIROC-ES2L CMIP6e	512	1417

Continued on next page

Table B.7.: Carbon Storage Totals [PgC]

	cVeg	cSoil + cLitter
MIROC-ES2L CMIP6c	531	1481
MPI-ESM-1-2-HAM CMIP6c	382	1030
MPI-ESM1-2-LR CMIP6e	370	977
MPI-ESM1-2-LR CMIP6c	367	980
MRI-ESM2-0 CMIP6e	478	1062
MRI-ESM2-0 CMIP6c	485	1224
NorESM2-LM CMIP6e	484	1014
NorESM2-LM CMIP6c	466	950
NorESM2-MM CMIP6c	471	898
SAM0-UNICON CMIP6c	590	577
TaiESM1 CMIP6c	498	525
UKESM1-0-LL CMIP6e	711	2100
UKESM1-0-LL CMIP6c	705	2118
OBS OBS historical	478	1561
Ncycle_mmm CMIP5e	547	534
non-Ncycle_mmm CMIP5e	521	2040
Ncycle_mmm CMIP5c	543	532
non-Ncycle_mmm CMIP5c	488	1920
Ncycle_mmm CMIP6c	485	1380
non-Ncycle_mmm CMIP6c	465	1197
Ncycle_mmm CMIP6e	510	1281
non-Ncycle_mmm CMIP6e	561	1618

List of Abbreviations

AOGCM	Atmosphere-Ocean General Circulation Model	9
AR4	Fourth Assessment Report	9
AR5	Fifth Assessment Report	5
AR6	Sixth Assessment Report	9
C3S	Copernicus Climate Change Service	38
C4MIP	Coupled Climate-Carbon Cycle Model Intercomparison Project	11
CMIP	Coupled Model Intercomparison Project	10
CMOR	Climate Model Output Rewriter	27
CLM	Community Land Model	15
CO_2	carbon dioxide	5
COP	Conferences of the Parties	104
cSoil	carbon mass in soil pool	71
CV	controlled vocabulary	11
cVeg	carbon mass in vegetation	71
EC	Emergent constraint	49
ESGF	Earth System Grid Federation	11
ESM	Earth system model	5
EMMA	ensemble median algorithm	38
ESMValCore	core functionalities of the ESMValTool	27
ESMValTool	Earth System Model Evaluation Tool	27
ESRL	Earth System Research Laboratories	32
GHG	greenhouse gas	5
GISS	Goddard Institute for Space Studies	44
GISTEMP	GISS Surface Temperature Analysis	44
GOSAT	Greenhouse Gases Observing Satellite	38
GPP	gross primary production	7

GR	growth rate	37
IAV	interannual variability	49
IPCC	Intergovernmental Panel on Climate Change	5
LAI	leaf area index	15
LSM	land surface model	13
LT	Local Time	60
MIP	Model Intercomparison Project	10
MLO	Mauna Loa Observatory	33
MMM	multi-model mean	45
NBP	net biome productivity	15
NetCDF	Network Common Data Form	28
NH	Northern Hemisphere	8
NIR	near infrared	25
NOAA	National Oceanic and Atmospheric Administration	32
NPP	net primary production	15
OCO-2	Orbiting Carbon Observatory-2	61
obs4MIPs	Observations for Model Intercomparisons Project	38
PFT	plant functional type	13
ppm	parts per million	1
RCP	Representative Concentration Pathway	11
RMSD	root-mean-square deviation	94
SCA	seasonal cycle amplitude	31
ScenarioMIP	Scenario Model Intercomparison Project	11
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric CHartographY	38
SH	Southern Hemisphere	44
SMPI	Single Model Performance Index	29
SSP	Shared Socioeconomic Pathway	11
SWIR	short-wave infrared	25
TANSO-FTS	Thermal And Near infrared Sensor for carbon Observation Fourier transform spectrometer	38
TCCON	Total Carbon Column Observing Network	38

UNFCCC	United Nations Framework Convention on Climate Change	104
WCRP	World Climate Research Programme	10
WGCM	Working Group on Coupled Modelling	10
XCO_2	column-average CO_2 mole fraction	25

List of Figures

List of Tables

References

Adachi, Y., Yukimoto, S., Deushi, M., Obata, A., Nakano, H., Tanaka, T. Y., Hosaka, M., Sakami, T., Yoshimura, H., Hirabara, M., Shindo, E., Tsujino, H., Mizuta, R., Yabu, S., Koshiro, T., Ose, T., & Kitoh, A. (2013). Basic performance of a new earth system model of the Meteorological Research Institute (MRI-ESM1). *Papers in Meteorology and Geophysics*, *64*(0), 1–19. https://doi.org/10.2467/mripapers.64.1

Alexander, K., & Easterbrook, S. M. (2015). The software architecture of climate models: a graphical comparison of CMIP5 and EMICAR5 configurations. *Geoscientific Model Development*, *8*(4), 1221–1232. https://doi.org/10.5194/gmd-8-1221-2015

Ali, A. A., Xu, C., Rogers, A., Fisher, R. A., Wullschleger, S. D., Massoud, E. C., Vrugt, J. A., Muss, J. D., McDowell, N. G., Fisher, J. B., Reich, P. B., & Wilson, C. J. (2016). A global scale mechanistic model of photosynthetic capacity (LUNA V1.0). *Geoscientific Model Development*, *9*(2), 587–606. https://doi.org/10.5194/gmd-9-587-2016

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., Jones, C., Jung, M., Myneni, R., & Zhu, Z. (2013a). Evaluating the Land and Ocean Components of the Global Carbon Cycle in the CMIP5 Earth System Models. *Journal of Climate*, *26*(18), 6801–6843. https://doi.org/10.1175/jcli-d-12-00417.1

Anav, A., Murray-Tortarolo, G., Friedlingstein, P., Sitch, S., Piao, S., & Zhu, Z. (2013b). Evaluation of Land Surface Models in Reproducing Satellite Derived Leaf Area Index over the High-Latitude Northern Hemisphere. Part II: Earth System Models. *Remote Sensing*, *5*(8), 3637–3661. https://doi.org/10.3390/rs5083637

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., & Zhao, M. (2015). Spatiotemporal patterns of terrestrial gross primary production: A review. *Reviews of Geophysics*, *53*(3), 785–818. https://doi.org/10.1002/2015RG000483

Andela, B., Broetz, B., de Mora, L., Drost, N., Eyring, V., Koldunov, N., Lauer, A., Mueller, B., Predoi, V., Righi, M., Schlund, M., Vegas-Regidor, J., Zimmermann, K., Adeniyi, K., Arnone, E., Bellprat, O., Berg, P., Bock, L., Caron, L.-P., Carvalhais, N., Cionni, I., Cortesi, N., Corti, S., Crezee, B., Davin, E. L., Davini, P., Deser, C., Diblen, F., Docquier, D., Dreyer, L., Ehbrecht, C., Earnshaw, P., Gier, B., Gonzalez-Reviriego, N., Goodman, P., Hagemann, S., von Hardenberg, J., Hassler, B., Hunter, A., Kadow, C., Kindermann, S., Koirala, S., Lledó, L., Lejeune, Q., Lembo, V., Little, B., Loosveldt-Tomas, S., Lorenz, R., Lovato, T., Lucarini, V., Massonnet, F., Mohr, C. W., Moreno-Chamarro, E., Amarjiit, P., Pérez-Zanón, N., Phillips, A., Russell, J., Sandstad, M., Sellar, A., Senftleben, D., Serva, F., Sillmann, J., Stacke, T., Swaminathan, R., Torralba, V., Weigel, K., Roberts, C., Kalverla, P., Alidoost, S., Verhoeven, S., Vreede, B., Smeets, S., Soares Siqueira, A., & Kazeroni, R. (2023a). *ESMValTool* (Version v2.9.0). Zenodo. https://doi.org/10.5281/zenodo.8120970

Andela, B., Broetz, B., de Mora, L., Drost, N., Eyring, V., Koldunov, N., Lauer, A., Predoi, V., Righi, M., Schlund, M., Vegas-Regidor, J., Zimmermann, K., Bock, L., Diblen, F., Dreyer, L., Earnshaw, P., Hassler, B., Little, B., Loosveldt-Tomas, S., Smeets, S., Camphuijsen, J., Gier, B. K., Weigel, K., Hauser, M., Kalverla, P., Galytska, E., Cos-Espuña, P., Pelupessy, I., Koirala, S., Stacke, T., Alidoost, S., Jury, M., Sénési, S., Crocker, T., Vreede, B., Soares Siqueira, A., Kazeroni, R., & Bauer, J. (2023b). *ESMValCore* (Version v2.9.0). Zenodo. https://doi.org/10.5281/zenodo.8112684

Anderson, J., Balaji, V., Broccoli, A., Cooke, W., Delworth, T., Dixon, K., Donner, L., Dunne, K., Freidenreich, S., Garner, S., Gudgel, R., Gordon, C., Held, I., Hemler, R., Horowitz, L., Klein, S., Knutson, T., Kushner, P., Langenhost, A., Lau, N., Liang, Z., Malyshev, S., Milly, P., Nath, M., Ploshay, J., Ramaswamy, V., Schwarzkopf, M., Shevliakova, E., Sirutis, J., Soden, B., Stern, W., Thompson, L., Wilson, R., Wittenberg, A., & Wyman, B. (2004). The new GFDL global atmosphere and land model AM2-LM2: Evaluation with prescribed SST simulations. *Journal of Climate*, *17*(24), 4641–4673. https://doi.org/10.1175/JCLI-3223.1

Arora, V. K., & Boer, G. J. (2003). A Representation of Variable Root Distribution in Dynamic Vegetation Models. *Earth Interactions*, *7*(6), 1–19. https://doi.org/10.1175/1087-3562(2003)007<0001:AROVRD>2.0.CO;2

Arora, V. K., & Boer, G. J. (2005). A parameterization of leaf phenology for the terrestrial ecosystem component of climate models. *Global Change Biology*, *11*(1), 39–59. https://doi.org/10.1111/j.1365-2486.2004.00890.x

Arora, V. K., Scinocca, J. F., Boer, G. J., Christian, J. R., Denman, K. L., Flato, G. M., Kharin, V. V., Lee, W. G., & Merryfield, W. J. (2011). Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophysical Research Letters*, *38*(5), n/a–n/a. https://doi.org/10.1029/2010gl046270

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., Schwinger, J., Bopp, L., Boucher, O., Cadule, P., Chamberlain, M. A., Christian, J. R., Delire, C., Fisher, R. A., Hajima, T., Ilyina, T., Joetzjer, E., Kawamiya, M., Koven, C. D., Krasting, J. P., Law, R. M., Lawrence, D. M., Lenton, A., Lindsay, K., Pongratz, J., Raddatz, T., Séférian, R., Tachiiri, K., Tjiputra, J. F., Wiltshire, A., Wu, T., & Ziehn, T. (2020). Carbon–concentration and carbon–climate feedbacks in CMIP6 models and their comparison to CMIP5 models. *Biogeosciences*, *17*(16), 4173–4222. https://doi.org/10.5194/bg-17-4173-2020

Arora, V. K. (2003). Simulating energy and carbon fluxes over winter wheat using coupled land surface and terrestrial ecosystem models. *Agricultural and Forest Meteorology*, *118*(1), 21–47. https://doi.org/10.1016/S0168-1923(03)00073-X

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., & Betts, A. K. (2009). A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System. *Journal of Hydrometeorology*, *10*(3), 623–643. https://doi.org/10.1175/2008JHM1068.1

Bao, Y., Qiao, F.-L., & Song, Z. (2012). The historical global carbon cycle simulation of FIO-ESM, 6834.

Barnes, E. A., Parazoo, N., Orbe, C., & Denning, A. S. (2016). Isentropic transport and the seasonal cycle amplitude of CO2. *Journal of Geophysical Research-Atmospheres*, *121*(13), 8106–8124. https://doi.org/10.1002/2016jd025109

Bastos, A., Ciais, P., Chevallier, F., Rodenbeck, C., Ballantyne, A. P., Maignan, F., Yin, Y., Fernandez-Martinez, M., Friedlingstein, P., Penuelas, J., Piao, S. L., Sitch, S., Smith, W. K., Wang, X. H., Zhu, Z. C., Haverd, V., Kato, E., Jain, A. K., Lienert, S., Lombardozzi, D., Nabel, J. E. M. S., Peylin, P., Poulter, B., & Zhu, D. (2019). Contrasting effects of CO2 fertilization, land-use change and warming on seasonal amplitude of Northern Hemisphere CO2 exchange. *Atmospheric Chemistry and Physics*, *19*(19), 12361–12375. https://doi.org/10.5194/acp-19-12361-2019

Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M., Biraud, S., Stephens, B., Andrews, A., & Worthy, D. (2013). Global CO2 fluxes estimated from GOSAT retrievals of total column CO2. *Atmospheric Chemistry and Physics*, *13*(17), 8695–8717. https://doi.org/10.5194/acp-13-8695-2013

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., & Harding, R. J. (2011). The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes. *Geoscientific Model Development*, *4*(3), 677–699. https://doi.org/10.5194/gmd-4-677-2011

Bock, L., Lauer, A., Schlund, M., Barreiro, M., Bellouin, N., Jones, C., Meehl, G. A., Predoi, V., Roberts, M. J., & Eyring, V. (2020). Quantifying Progress Across Different CMIP Phases With the ESMValTool [e2019JD032321 2019JD032321]. *Journal of Geophysical Research: Atmospheres*, *125*(21), e2019JD032321. https://doi.org/10.1029/2019JD032321

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, E., Lionel, Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., & Vuichard, N. (2020). Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. *Journal of Advances in Modeling Earth Systems*, *12*(7), e2019MS002010. https://doi.org/10.1029/2019MS002010

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noel, S., Rozanov, V. V., Chance, K. V., & Goede, A. P. H. (1999). SCIAMACHY: Mission objectives and measurement modes. *Journal of the Atmospheric Sciences*, *56*(2), 127–150. https://doi.org/Doi10.1175/1520-0469(1999)056<0127:Smoamm>2.0.Co;2

Bronselaer, B., Winton, M., Russell, J., Sabine, C. L., & Khatiwala, S. (2017). Agreement of CMIP5 Simulated and Observed Ocean Anthropogenic CO2 Uptake. *Geophysical Research Letters*, *44*(24), 12, 298–12, 305. https://doi.org/10.1002/2017GL074435

Buchwitz, M., de Beek, R., Burrows, J. P., Bovensmann, H., Warneke, T., Notholt, J., Meirink, J. F., Goede, A. P. H., Bergamaschi, P., Körner, S., Heimann, M., & Schulz, A. (2005). Atmospheric methane and carbon dioxide from SCIAMACHY satellite data: initial comparison with chemistry and transport models. *Atmospheric Chemistry and Physics*, *5*(4), 941–962. https://doi.org/10.5194/acp-5-941-2005

Buchwitz, M., & Reuter, M. (2016). Merged SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT atmospheric column-average dry-air mole fraction of CO2 (XCO2) (XCO2_CRDP3_001) - Technical Document. http://www.esa-ghg-cci.org/?q=webfm_send/330

Buchwitz, M., Reuter, M., Schneising-Weigel, O., Aben, I., Detmers, R. G., Hasekamp, O. P., Boesch, H., Anand, J., Crevoisier, C., & Armante, R. (2017a). Product Quality Assessment Report (PQAR) – Main document. *Technical Report Copernicus Climate Change Service (C3S)*, 1–103. http://www.iup.uni-bremen.de/carbon_ghg/docs/C3S/CDR1_2003-2016/PQAR/C3S_D312a_Lot6.3.1.7-v1_PQAR_MAIN_v1.1.pdf

Buchwitz, M., Reuter, M., Schneising-Weigel, O., Aben, I., Detmers, R. G., Hasekamp, O. P., Boesch, H., Anand, J., Crevoisier, C., & Armante, R. (2017b). Product User Guide and Specification (PUGS) – Main document. *Technical Report Copernicus Climate Change Service (C3S)*, 1–91. http://www.iup.uni-bremen.de/carbon_ghg/docs/C3S/CDR1_2003-2016/PUGS/C3S_D312a_Lot6.3.1.5-v1_PUGS_MAIN_v1.3.pdf

Buchwitz, M., Reuter, M., Schneising, O., Noël, S., **Gier**, **B.**, Bovensmann, H., Burrows, J. P., Boesch, H., Anand, J., Parker, R. J., Somkuti, P., Detmers, R. G., Hasekamp, O. P., Aben, I., Butz, A., Kuze, A., Suto, H., Yoshida, Y., Crisp, D., & O'Dell, C. (2018). Computation and analysis of atmospheric carbon dioxide annual mean growth rates from satellite observations during 2003–2016. *Atmospheric Chemistry and Physics*, *18*(23), 17355–17370. https://doi.org/10.5194/acp-18-17355-2018

Burrows, J. P., Holzle, E., Goede, A. P. H., Visser, H., & Fricke, W. (1995). Sciamachy - Scanning Imaging Absorption Spectrometer for Atmospheric Chartography. *Acta Astronautica*, *35*(7), 445–451. https://doi.org/Doi10.1016/0094-5765(94)00278-T

Canadell, J., Monteiro, P., Costa, M., Cotrim da Cunha, L., Cox, P., Eliseev, A., Henson, S., Ishii, M., Jaccard, S., Koven, C., Lohila, A., Patra, P., Piao, S., Rogelj, J., Syampungani, S., Zaehle, S., & Zickfeld, K. (2021). Global Carbon and other Biogeochemical Cycles and

Feedbacks. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 673–816). Cambridge University Press. https://doi.org/10.1017/9781009157896.007

Cao, S., Li, M., Zhu, Z., Wang, Z., Zha, J., Zhao, W., Duanmu, Z., Chen, J., Zheng, Y., Chen, Y., Myneni, R. B., & Piao, S. (2023). Spatiotemporally consistent global dataset of the GIMMS leaf area index (GIMMS LAI4g) from 1982 to 2020. *Earth System Science Data*, *15*(11), 4877–4899. https://doi.org/10.5194/essd-15-4877-2023

Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., Mu, M., Saatchi, S., Santoro, M., Thurner, M., Weber, U., Ahrens, B., Beer, C., Cescatti, A., Randerson, J. T., & Reichstein, M. (2014). Global covariation of carbon turnover times with climate in terrestrial ecosystems. *Nature*, *514*(7521), 213–217. https://doi.org/10.1038/nature13731

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., Masina, S., Scoccimarro, E., Materia, S., Bellucci, A., & Navarra, A. (2019). Global Mean Climate and Main Patterns of Variability in the CMCC-CM2 Coupled Model. *Journal of Advances in Modeling Earth Systems*, *11*(1), 185–209. https://doi.org/10.1029/2018MS001369

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., & Ciais, P. (2005). Inferring CO2sources and sinks from satellite observations: Method and application to TOVS data. *J. Geophys. Res.*, *110*(D24).

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morguí, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., & Worthy, D. (2010). CO2surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements. *J. Geophys. Res.*, *115*(D21).

Chevallier, F., Palmer, P. I., Feng, L., Boesch, H., O'Dell, C. W., & Bousquet, P. (2014). Toward robust and consistent regional CO2 flux estimates from in situ and spaceborne measurements of atmospheric CO2. *Geophysical Research Letters*, *41*(3), 1065–1070. https://doi.org/10.1002/2013gl058772

Chevallier, F. (2013). On the parallelization of atmospheric inversions of CO_2 surface fluxes within a variational framework. *Geosci. Model Dev.*, *6*(3), 783–790.

Christiansen, B. (2018). Ensemble Averaging and the Curse of Dimensionality. *Journal of Climate*, *31*(4), 1587–1596. https://doi.org/10.1175/JCLI-D-17-0197.1

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., Piao, S., & Thornton, P. (2013). Carbon and Other Biogeochemical Cycles. In T. Stocker, D. Qin, G.-K. Plattner,

M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. Midgley (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 465–570). Cambridge University Press. https://doi.org/10.1017/CBO9781107415324.015

Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., Huntingford, C., & Cox, P. M. (2011). The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, *4*(3), 701–722. https://doi.org/10.5194/gmd-4-701-2011

Collatz, G., Ribas-Carbo, M., & Berry, J. (1992). Coupled Photosynthesis-Stomatal Conductance Model for Leaves of C4 Plants. *Functional Plant Biology*, *19*(5), 519. https://doi.org/10.1071/pp9920519

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., & Woodward, S. (2011). Development and evaluation of an Earth-System model – HadGEM2. *Geosci. Model Dev.*, *4*(4), 1051–1075. https://doi.org/10.5194/gmd-4-1051-2011

Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A., & Totterdell, I. J. (2000). Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, *408*(6809), 184–187. https://doi.org/10.1038/35041539

Cox, P. M., Pearson, D., Booth, B. B., Friedlingstein, P., Huntingford, C., Jones, C. D., & Luke, C. M. (2013). Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature*, *494*(7437), 341–4. https://doi.org/10.1038/nature11882

Cox, P. M. (2001). Description of the TRIFFID Dynamic Global Vegetation Model. *Hadley Centre, Met Office, UK, Technical Note 24*.

Crisp, D., Dolman, H., Tanhua, T., McKinley, G. A., Hauck, J., Bastos, A., Sitch, S., Eggleston, S., & Aich, V. (2022). How Well Do We Understand the Land-Ocean-Atmosphere Carbon Cycle? [e2021RG000736 2021RG000736]. *Reviews of Geophysics*, *60*(2), e2021RG000736. https://doi.org/10.1029/2021RG000736

Dai, Y., Zeng, X., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., Denning, A. S., Dirmeyer, P. A., Houser, P. R., Niu, G., Oleson, K. W., Schlosser, C. A., & Yang, Z.-L. (2003). The Common Land Model. *Bulletin of the American Meteorological Society*, *84*(8), 1013–1024. https://doi.org/10.1175/bams-84-8-1013

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., &

Strand, W. G. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. https://doi.org/10.1029/2019ms001916

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J.-P., Alias, A., Saint-Martin, D., Séférian, R., Sénési, S., & Voldoire, A. (2019). Recent Changes in the ISBA-CTRIP Land Surface System for Use in the CNRM-CM6 Climate Model and in Global Off-Line Hydrological Applications. *Journal of Advances in Modeling Earth Systems*, *11*(5), 1207–1252. https://doi.org/10.1029/2018MS001545

Delire, C., Séférian, R., Decharme, B., Alkama, R., Calvet, J.-C., Carrer, D., Gibelin, A.-L., Joetzjer, E., Morel, X., Rocher, M., & Tzanos, D. (2020). The global land carbon cycle simulated with ISBA-CTRIP: improvements over the last decade. *Journal of Advances in Modeling Earth Systems*, *n/a*(n/a), e2019MS001886. https://doi.org/10.1029/2019ms001886

Dlugokencky, E. J., Lang, P. M., Mund, J. W., M., C. A., Crotwell, M. J., & Thoning, K. W. (2018). *Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network*. Retrieved 2018, from ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flask/surface/

Dlugokencky, E. J., & Tans, P. P. (2020). *Trends in Atmospheric Carbon Dioxide*. Retrieved 2022, from www.esrl.noaa.gov/gmd/ccgg/trends/gl%5C_data.html

Dlugokencky, E., Mund, J., Crotwell, A., Crotwell, M., & Thoning, K. (2020). Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA GML Carbon Cycle Cooperative Global Air Sampling Network, 1968–2019. https://doi.org/10.15138/wkgj-f215

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L.-P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., Fuentes-Franco, R., Gröger, M., v. Hardenberg, J., Hieronymus, J., Karami, M. P., Keskinen, J.-P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., van Noije, T., Nolan, P., O'Donnell, D., Ollinaho, P., van den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F., Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., & Zhang, Q. (2022). The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6. *Geoscientific Model Development*, *15*(7), 2973–3020. https://doi.org/10.5194/gmd-15-2973-2022

Du, E., Terrer, C., Pellegrini, A. F., Ahlström, A., van Lissa, C. J., Zhao, X., Xia, N., Wu, X., & Jackson, R. B. (2020). Global patterns of terrestrial nitrogen and phosphorus limitation. *Nature Geoscience*, *13*(3), 221–226.

Dufresne, J. L., Foujols, M. A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J. P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J. Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M. P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N., & Vuichard, N. (2013). Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate Dynamics*, *40*(9), 2123–2165. https://doi.org/10.1007/s00382-012-1636-1

Dunne, J. P., John, J. G., Adcroft, A. J., Griffies, S. M., Hallberg, R. W., Shevliakova, E., Stouffer, R. J., Cooke, W., Dunne, K. A., Harrison, M. J., Krasting, J. P., Malyshev, S. L., Milly, P. C. D., Phillipps, P. J., Sentman, L. T., Samuels, B. L., Spelman, M. J., Winton, M., Wittenberg, A. T., & Zadeh, N. (2012). GFDL's ESM2 Global Coupled Climate-Carbon Earth System Models. Part I: Physical Formulation and Baseline Simulation Characteristics. *Journal of Climate*, *25*(19), 6646–6665. https://doi.org/10.1175/Jcli-D-11-00560.1

Dunne, J. P., John, J. G., Shevliakova, E., Stouffer, R. J., Krasting, J. P., Malyshev, S. L., Milly, P. C. D., Sentman, L. T., Adcroft, A. J., Cooke, W., Dunne, K. A., Griffies, S. M., Hallberg, R. W., Harrison, M. J., Levy, H., Wittenberg, A. T., Phillips, P. J., & Zadeh, N. (2013). GFDL's ESM2 Global Coupled Climate-Carbon Earth System Models. Part II: Carbon System Formulation and Baseline Simulation Characteristics. *Journal of Climate*, *26*(7), 2247–2267. https://doi.org/10.1175/Jcli-D-12-00150.1

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P. G., Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng, Y., & Zhao, M. (2020). The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2019MS002015. https://doi.org/10.1029/2019ms002015

Eyring, V., Gleckler, P. J., Heinze, C., Stouffer, R. J., Taylor, K. E., Balaji, V., Guilyardi, E., Joussaume, S., Kindermann, S., Lawrence, B. N., Meehl, G. A., Righi, M., & Williams, D. N. (2016a). Towards improved and more routine Earth system model evaluation in CMIP. *Earth System Dynamics*, *7*(4), 813–830. https://doi.org/10.5194/esd-7-813-2016

Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., Anav, A., Andrews, O., Cionni, I., Davin, E. L., Deser, C., Ehbrecht, C., Friedlingstein, P., Gleckler, P., Gottschaldt, K. D., Hagemann, S., Juckes, M., Kindermann, S., Krasting, J., Kunert, D., Levine, R., Loew, A., Makela, J., Martin, G., Mason, E., Phillips, A. S., Read, S., Rio, C., Roehrig, R., Senftleben, D., Sterl, A., van Ulft, L. H., Walton, J., Wang, S. Y., & Williams, K. D. (2016b). ESMValTool (v1.0) - a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP. *Geoscientific Model Development*, *9*(5), 1747–1802. https://doi.org/10.5194/gmd-9-1747-2016

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016c). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., Collins, W. D., **Gier**, **B. K.**, Hall, A. D., Hoffman, F. M., Hurtt, G. C., Jahn, A., Jones, C. D., Klein, S. A., Krasting, J. P., Kwiatkowski, L., Lorenz, R., Maloney, E., Meehl, G. A., Pendergrass, A. G., Pincus, R., Ruane, A. C., Russell, J. L., Sanderson, B. M., Santer, B. D., Sherwood, S. C., Simpson, I. R., Stouffer, R. J., & Williamson, M. S. (2019). Taking climate model evaluation to the next level. *Nature Climate Change*, *9*(2), 102–110. https://doi.org/10.1038/s41558-018-0355-y

Eyring, V., Bock, L., Lauer, A., Righi, M., Schlund, M., Andela, B., Arnone, E., Bellprat, O., Brötz, B., Caron, L.-P., Carvalhais, N., Cionni, I., Cortesi, N., Crezee, B., Davin, E. L., Davini, P., Debeire, K., de Mora, L., Deser, C., Docquier, D., Earnshaw, P., Ehbrecht, C., **Gier**, **B. K.**, Gonzalez-Reviriego, N., Goodman, P., Hagemann, S., Hardiman, S., Hassler, B., Hunter, A., Kadow, C., Kindermann, S., Koirala, S., Koldunov, N., Lejeune, Q., Lembo, V., Lovato, T., Lucarini, V., Massonnet, F., Müller, B., Pandde, A., Pérez-Zanón, N., Phillips, A., Predoi, V., Russell, J., Sellar, A., Serva, F., Stacke, T., Swaminathan, R., Torralba, V., Vegas-Regidor, J., von Hardenberg, J., Weigel, K., & Zimmermann, K. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 – an extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of Earth system models in CMIP. *Geoscientific Model Development*, *13*(7), 3383–3438. https://doi.org/10.5194/gmd-13-3383-2020

Eyring, V., Gillett, N., Achuta Rao, K., Barimalala, R., Barreiro Parrillo, M., Bellouin, N., Cassou, C., Durack, P., Kosaka, Y., McGregor, S., Min, S., Morgenstern, O., & Sun, Y. (2021). Human Influence on the Climate System. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 423–552). Cambridge University Press. https://doi.org/10.1017/9781009157896.005

Farquhar, G. D., von Caemmerer, S., & Berry, J. A. (1980). A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. *Planta*, *149*(1), 78–90. https://doi.org/10.1007/bf00386231

Fernández-Martínez, M., Sardans, J., Chevallier, F., Ciais, P., Obersteiner, M., Vicca, S., Canadell, J. G., Bastos, A., Friedlingstein, P., Sitch, S., Piao, S. L., Janssens, I. A., & Peñuelas, J. (2019). Global trends in carbon sinks and their relationships with CO2 and temperature. *Nature Climate Change*, *9*(1), 73–79. https://doi.org/10.1038/s41558-018-0367-7

Ferraro, R., Waliser, D. E., Gleckler, P., Taylor, K. E., & Eyring, V. (2015). Evolving Obs4MIPs to Support Phase 6 of the Coupled Model Intercomparison Project (CMIP6). *Bulletin of the American Meteorological Society*, *96*(8), ES131–ES133. https://doi.org/10.1175/BAMS-D-14-00216.1

Fisher, J. B., Huntzinger, D. N., Schwalm, C. R., & Sitch, S. (2014). Modeling the Terrestrial Biosphere. *Annual Review of Environment and Resources*, *39*(1), 91–123. https://doi.org/10.1146/annurev-environ-012913-093456

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems [e2018MS001453 2018MS001453]. *Journal of Advances in Modeling Earth Systems*, *12*(4), e2018MS001453. https://doi.org/10.1029/2018MS001453

Flato, G. M., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W. D., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., & Rummukainen, M. (2013). Evaluation of Climate Models. Cambridge University Press. https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter09_FINAL.pdf

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. https://doi.org/10.1016/j.rse.2009.08.016

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 C4MIP Model Intercomparison. *Journal of Climate*, *19*(14), 3337–3353. https://doi.org/10.1175/jcli3800.1

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*, *27*(2), 511–526. https://doi.org/10.1175/Jcli-D-12-00579.1

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L. P., Currie, K. I., Feely, R. A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D. S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Joetzjer, E., Kaplan, J. O., Kato, E., Goldewijk, K. K., Korsbakken, J. I., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P. C., Melton, J. R., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Omar, A. M., Ono, T., Peregon, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Werf, G. R., Wiltshire, A. J., & Zaehle, S. (2019). Global Carbon Budget 2019. *Earth System Science Data*, *11*(4), 1783–1838. https://doi.org/10.5194/essd-11-1783-2019

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C. E., Hauck, J., Le Quéré, C., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Bates, N. R., Becker, M., Bellouin, N., Bopp, L., Chau, T. T. T., Chevallier, F., Chini, L. P., Cronin, M., Currie, K. I., Decharme, B., Djeutchouang, L. M., Dou, X., Evans, W., Feely, R. A., Feng, L., Gasser, T., Gilfillan, D., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Luijkx, I. T., Jain, A., Jones, S. D., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lienert, S., Liu, J., Marland, G., McGuire, P. C., Melton, J. R., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., Ono, T., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T. M., Schwinger, J., Schwingshackl, C., Séférian, R., Sutton, A. J., Sweeney, C., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F., van der Werf, G. R., Vuichard, N., Wada, C., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, C., Yue, X., Zaehle, S., & Zeng, J. (2022). Global Carbon Budget 2021. *Earth System Science Data*, *14*(4), 1917–2005. https://doi.org/10.5194/essd-14-1917-2022

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Barbero, L., Bates, N. R., Becker, M., Bellouin, N., Decharme, B., Bopp, L., Brasika, I. B. M., Cadule, P., Chamberlain, M. A., Chandra, N., Chau, T.-T.-T., Chevallier, F., Chini, L. P., Cronin, M., Dou, X., Enyo, K., Evans, W., Falk, S., Feely, R. A., Feng, L., Ford, D. J., Gasser, T., Ghattas, J., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Heinke, J., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jacobson, A. R., Jain, A., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Joos, F., Kato, E., Keeling, R. F., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Körtzinger, A., Lan, X., Lefèvre, N., Li, H., Liu, J., Liu, Z., Ma, L., Marland, G., Mayot, N., McGuire, P. C., McKinley, G. A., Meyer, G., Morgan, E. J., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O'Brien, K. M., Olsen, A., Omar, A. M., Ono, T., Paulsen, M., Pierrot, D., Pocock, K., Poulter, B., Powis, C. M., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T. M., Schwinger, J., Séférian, R., . . . Zheng, B. (2023). Global Carbon Budget 2023. *Earth System Science Data*, *15*(12), 5301–5369. https://doi.org/10.5194/essd-15-5301-2023

Gatti, L. V., Basso, L. S., Miller, J. B., Gloor, M., Gatti Domingues, L., Cassol, H. L. G., Tejada, G., Aragão, L. E. O. C., Nobre, C., Peters, W., Marani, L., Arai, E., Sanches, A. H., Corrêa, S. M., Anderson, L., Von Randow, C., Correia, C. S. C., Crispim, S. P., & Neves, R. A. L. (2021). Amazonia as a carbon source linked to deforestation and climate change. *Nature*, *595*(7867), 388–393. https://doi.org/10.1038/s41586-021-03629-6

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z.-L., & Zhang, M. (2011). The Community Climate System Model Version 4. *Journal of Climate*, *24*(19), 4973–4991. https://doi.org/10.1175/2011jcli4083.1

Gettelman, A., & Rood, R. B. (2016). Essence of a Climate Model. In *Demystifying Climate Models: A Users Guide to Earth System Models* (pp. 37–58). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-48959-8_4

Gibbs, H. K. (2006). Olson's major world ecosystem complexes ranked by carbon in live vegetation: An updated database using the GLC2000 land cover product [Title of the publication associated with this dataset: Carbon Dioxide Information Analysis Center (CDIAC) Datasets].

Gibelin, A.-L., Calvet, J.-C., & Viovy, N. (2008). Modelling energy and CO2 fluxes with an interactive vegetation land surface model-Evaluation at high and middle latitudes. *Agricultural and Forest Meteorology*, *148*(10), 1611–1628. https://doi.org/10.1016/j.agrformet.2008.05.013

Gier, B. K., Buchwitz, M., Reuter, M., Cox, P. M., Friedlingstein, P., & Eyring, V. (2020). Spatially resolved evaluation of Earth system models with satellite column-averaged CO_2. *Biogeosciences*, *17*(23), 6115–6144. https://doi.org/10.5194/bg-17-6115-2020

Gier, B. K., Schlund, M., Friedlingstein, P., Jones, C. D., Jones, C., Zaehle, S., & Eyring, V. (2024). Representation of the Terrestrial Carbon Cycle in CMIP6. *EGUsphere*, *2024*, 1–63. https://doi.org/10.5194/egusphere-2024-277

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Bottinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H. D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K. D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K. H., Claussen, M., Marotzke, J., & Stevens, B. (2013). Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of Advances in Modeling Earth Systems*, *5*(3), 572–597. https://doi.org/10.1002/jame.20038

Gleckler, P. J., Taylor, K. E., & Doutriaux, C. (2008). Performance metrics for climate models. *Journal of Geophysical Research: Atmospheres*, *113*(D6). https://doi.org/10.1029/2007jd008972

Global Carbon Project. (2021). Supplemental data of Global Carbon Project 2021. https://doi.org/10.18160/GCP-2021

Goll, D. S., Winkler, A. J., Raddatz, T., Dong, N., Prentice, I. C., Ciais, P., & Brovkin, V. (2017). Carbon–nitrogen interactions in idealized simulations with JSBACH (version 3.10). *Geoscientific Model Development*, *10*(5), 2009–2030. https://doi.org/10.5194/gmd-10-2009-2017

Graven, H. D., Keeling, R. F., Piper, S. C., Patra, P. K., Stephens, B. B., Wofsy, S. C., Welp, L. R., Sweeney, C., Tans, P. P., Kelley, J. J., Daube, B. C., Kort, E. A., Santoni, G. W., & Bent, J. D. (2013). Enhanced seasonal exchange of CO2 by northern ecosystems since 1960. *Science*, *341*(6150), 1085–9. https://doi.org/10.1126/science.1239207

Hajima, T., Tachiiri, K., Ito, A., & Kawamiya, M. (2014). Uncertainty of Concentration-Terrestrial Carbon Feedback in Earth System Models. *Journal of Climate*, *27*(9), 3425–3445. https://doi.org/10.1175/Jcli-D-13-00177.1

Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., Ohgaito, R., Ito, A., Yamazaki, D., Okajima, H., Ito, A., Takata, K., Ogochi, K., Watanabe, S., & Kawamiya, M. (2020a). Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks. *Geosci. Model Dev.*, *13*(5), 2197–2244. https://doi.org/10.5194/gmd-13-2197-2020

Hajima, T., Abe, M., Arakawa, O., Suzuki, T., Komuro, Y., Ogochi, K., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Ohgaito, R., Ito, A., Yamazaki, D., Ito, A., Takata, K., Watanabe, S., Kawamiya, M., & Tachiiri, K. (2020b). MIROC MIROC-ES2L model output prepared for CMIP6 CMIP esm-hist. https://doi.org/10.22033/ESGF/CMIP6.5496

Hansen, J., Ruedy, R., Sato, M., & Lo, K. (2010). Global Surface Temperature Change. *Reviews of Geophysics*, *48*. https://doi.org/ArtnRg400410.1029/2010rg000345

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*(2), 141–162. https://doi.org/10.1002/2013jg002381

Hosaka, M. (2011). A new MRI land surface model HAL. *AGU Fall Meeting Abstracts*, *2011*, Article GC43B-0901, GC43B–0901.

Houweling, S., Baker, D., Basu, S., Boesch, H., Butz, A., Chevallier, F., Deng, F., Dlugokencky, E. J., Feng, L., Ganshin, A., Hasekamp, O., Jones, D., Maksyutov, S., Marshall, J., Oda, T., O'Dell, C. W., Oshchepkov, S., Palmer, P. I., Peylin, P., Poussi, Z., Reum, F., Takagi, H., Yoshida, Y., & Zhuravlev, R. (2015). An intercomparison of inverse models for estimating sources and sinks of CO2using GOSAT measurements. *Journal of Geophysical Research: Atmospheres*, *120*(10), 5253–5266. https://doi.org/10.1002/2014jd022962

Huang, Y., Lu, X., Shi, Z., Lawrence, D., Koven, C. D., Xia, J., Du, Z., Kluzek, E., & Luo, Y. (2018a). Matrix approach to land carbon cycle modeling: A case study with the Community Land Model. *Global Change Biology*, *24*(3), 1394–1404. https://doi.org/10.1111/gcb.13948

Huang, Y., Zhu, D., Ciais, P., Guenet, B., Huang, Y., Goll, D. S., Guimberteau, M., Jornet-Puig, A., Lu, X., & Luo, Y. (2018b). Matrix-Based Sensitivity Assessment of Soil Organic Carbon Storage: A Case Study from the ORCHIDEE-MICT Model. *Journal of Advances in Modeling Earth Systems*, *10*(8), 1790–1808. https://doi.org/10.1029/2017MS001237

Hugelius, G., Tarnocai, C., Broll, G., Canadell, J. G., Kuhry, P., & Swanson, D. K. (2013). The Northern Circumpolar Soil Carbon Database: spatially distributed datasets of soil coverage and soil carbon storage in the northern permafrost regions. *Earth System Science Data*, *5*(1), 3–13. https://doi.org/10.5194/essd-5-3-2013

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J.-F., Large, W. G., Lawrence, D., Lindsay, K., Lipscomb, W. H., Long, M. C., Mahowald, N., Marsh, D. R., Neale, R. B., Rasch, P., Vavrus, S., Vertenstein, M., Bader, D., Collins, W. D., Hack, J. J., Kiehl, J., & Marshall, S. (2013). The Community Earth System Model: A Framework for Collaborative Research. *Bulletin of the American Meteorological Society*, *94*(9), 1339–1360. https://doi.org/10.1175/bams-d-12-00121.1

IPCC. (2021a). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Vol. In Press). Cambridge University Press. https://doi.org/10.1017/9781009157896

IPCC. (2021b). Summary for Policymakers. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 3–32). Cambridge University Press. https://doi.org/10.1017/9781009157896.001

Ito, A., & Inatomi, M. (2012). Water-Use Efficiency of the Terrestrial Biosphere: A Model Analysis Focusing on Interactions between the Global Carbon and Water Cycles. *Journal of Hydrometeorology*, *13*(2), 681–694. https://doi.org/10.1175/JHM-D-10-05034.1

Ji, D., Wang, L., Feng, J., Wu, Q., Cheng, H., Zhang, Q., Yang, J., Dong, W., Dai, Y., Gong, D., Zhang, R. H., Wang, X., Liu, J., Moore, J. C., Chen, D., & Zhou, M. (2014). Description and basic evaluation of Beijing Normal University Earth System Model (BNU-ESM) version 1. *Geoscientific Model Development*, *7*(5), 2039–2064. https://doi.org/10.5194/gmd-7-2039-2014

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*(8), 2853–2880. https://doi.org/10.5194/gmd-9-2853-2016

Jones, C. D. (2020). So What Is in an Earth System Model? *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001967. https://doi.org/10.1029/2019ms001967

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., & Williams, C. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research: Biogeosciences*, *116*(G3). https://doi.org/10.1029/2010jg001566

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., & Reichstein, M. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific Data*, *6*(1), 74. https://doi.org/10.1038/s41597-019-0076-8

Kaminski, T., Knorr, W., Schurmann, G., Scholze, M., Rayner, P. J., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J. P., Heimann, M., Hooker-Stroud, A., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E. N., Kostler, C., Mathieu, P. P., Pinty, B., Reick, C. H., Rodenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., van Scheltinga, A. T., Vossbeck, M., Widmann, H., & Ziehn, T. (2013). The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges. *Journal of Geophysical Research-Biogeosciences*, *118*(4), 1414–1426. https://doi.org/10.1002/jgrg.20118

Keeling, C. D., Bacastow, R. B., Bainbridge, A. E., Ekdahl, C. A., Guenther, P. R., Waterman, L. S., & Chin, J. F. S. (1976). Atmospheric Carbon-Dioxide Variations at Mauna-Loa Observatory, Hawaii. *Tellus*, *28*(6), 538–551. https://doi.org/10.1111/j.2153-3490.1976.tb00701.x

Keeling, C. D., Bacastow, R. B., Carter, A. F., Piper, S. C., Whorf, T. P., Heimann, M., Mook, W. G., & Roeloffzen, H. (1989). A three-dimensional model of atmospheric CO2transport based on observed winds: 1. Analysis of observational data. In *Aspects of Climate Variability in the Pacific and the Western Americas* (pp. 165–236). https://doi.org/10.1029/GM055p0165

Keeling, C. D., Whorf, T. P., Wahlen, M., & Vanderplicht, J. (1995). Interannual Extremes in the Rate of Rise of Atmospheric Carbon-Dioxide since 1980. *Nature*, *375*(6533), 666–670. https://doi.org/DOI10.1038/375666a0

Keeling, C. D., Chin, J. F. S., & Whorf, T. P. (1996). Increased activity of northern vegetation inferred from atmospheric CO2 measurements. *Nature*, *382*(6587), 146–149. https://doi.org/10.1038/382146a0

Keenan, T., & Williams, C. (2018). The Terrestrial Carbon Sink. *Annual Review of Environment and Resources*, *43*(1), 219–243. https://doi.org/10.1146/annurev-environ-102017-030204

Kou-Giesbrecht, S., & Arora, V. K. (2023). Compensatory Effects Between CO2, Nitrogen Deposition, and Nitrogen Fertilization in Terrestrial Biosphere Models Without Nitrogen Compromise Projections of the Future Terrestrial Carbon Sink [e2022GL102618 2022GL102618]. *Geophysical Research Letters*, *50*(8), e2022GL102618. https://doi.org/10.1029/2022GL102618

Kou-Giesbrecht, S., Arora, V. K., Seiler, C., Arneth, A., Falk, S., Jain, A. K., Joos, F., Kennedy, D., Knauer, J., Sitch, S., O'Sullivan, M., Pan, N., Sun, Q., Tian, H., Vuichard, N., & Zaehle, S. (2023). Evaluating nitrogen cycling in terrestrial biosphere models: a disconnect between the carbon and nitrogen cycles. *Earth System Dynamics*, *14*(4), 767–795. https://doi.org/10.5194/esd-14-767-2023

Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S., Collins, W. D., Bonan, G. B., Lawrence, D. M., & Swenson, S. C. (2013). The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4. *Biogeosciences*, *10*(11), 7109–7131. https://doi.org/10.5194/bg-10-7109-2013

Kowalczyk, E., Stevens, L., Law, R., Dix, M., Wang, Y., Harman, I., Haynes, K., Srbinovsky, J., Pak, B., & Ziehn, T. (2013). The land surface model component of ACCESS: Description and impact on the simulated surface climatology. *Australian Meteorological and Oceanographic Journal*, *63*, 65–82.

Krasting, J. P., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., Rand, K., Zadeh, N. T., Balaji, V., Durachta, J., Dupuis, C., Menzel, R., Robinson, T., Underwood, S., Vahlenkamp, H., Dunne, K. A., Gauthier, P. P., Ginoux, P., Griffies, S. M., Hallberg, R., Harrison, M., Hurlin, W., Malyshev, S., Naik, V., Paulot, F., Paynter, D. J., Ploshay, J., Reichl, B. G., Schwarzkopf, D. M., Seman, C. J., Silvers, L., Wyman, B., Zeng, Y., Adcroft, A., Dunne, J. P., Dussin, R., Guo, H., He, J., Held, I. M., Horowitz, L. W., Lin, P., Milly, P., Shevliakova, E., Stock, C., Winton, M., Wittenberg, A. T., Xie, Y., & Zhao, M. (2018). NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP esm-hist. https://doi.org/10.22033/ESGF/CMIP6.8522

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., & Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, *19*(1). https://doi.org/10.1029/2003GB002199

Kuma, P., Bender, F. A.-M., & Jönsson, A. R. (2023). Climate Model Code Genealogy and Its Relation to Climate Feedbacks and Sensitivity [e2022MS003588 2022MS003588]. *Journal of Advances in Modeling Earth Systems*, *15*(7), e2022MS003588. https://doi.org/10.1029/2022MS003588

Kuze, A., Suto, H., Nakajima, M., & Hamazaki, T. (2009). Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. *Appl Opt*, *48*(35), 6716–33. https://doi.org/10.1364/AO.48.006716

Lauer, A., Eyring, V., Bellprat, O., Bock, L., **Gier, B. K.**, Hunter, A., Lorenz, R., Pérez-Zanón, N., Righi, M., Schlund, M., Senftleben, D., Weigel, K., & Zechlau, S. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 – diagnostics for emergent constraints and future projections from Earth system models in CMIP. *Geoscientific Model Development*, *13*(9), 4205–4228. https://doi.org/10.5194/gmd-13-4205-2020

Law, R. M., Ziehn, T., Matear, R. J., Lenton, A., Chamberlain, M. A., Stevens, L. E., Wang, Y. P., Srbinovsky, J., Bi, D., Yan, H., & Vohralik, P. F. (2017). The carbon cycle in the Australian Community Climate and Earth System Simulator (ACCESS-ESM1) – Part 1: Model description and pre-industrial simulation. *Geosci. Model Dev.*, *10*(7), 2567–2590. https://doi.org/10.5194/gmd-10-2567-2017

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., & Slater, A. G. (2011). Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *Journal of Advances in Modeling Earth Systems*, *3*(3). https://doi.org/10.1029/2011ms000045

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., van den Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., & Zeng, X. (2019). The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. https://doi.org/10.1029/2018ms001583

Le Quéré, C., Raupach, M. R., Canadell, J. G., Marland, G., Bopp, L., Ciais, P., Conway, T. J., Doney, S. C., Feely, R. A., Foster, P., Friedlingstein, P., Gurney, K., Houghton, R. A., House, J. I., Huntingford, C., Levy, P. E., Lomas, M. R., Majkut, J., Metzl, N., Ometto, J. P., Peters, G. P., Prentice, I. C., Randerson, J. T., Running, S. W., Sarmiento, J. L., Schuster, U., Sitch, S., Takahashi, T., Viovy, N., van der Werf, G. R., & Woodward, F. I. (2009). Trends in the sources and sinks of carbon dioxide. *Nature Geoscience*, *2*(12), 831–836. https://doi.org/10.1038/ngeo689

Le Treut, H., Somerville, R., Cubasch, U., Ding, Y., Mauritzen, C., Mokssit, A., Peterson, T., & Prather, M. (2007). Historical Overview of Climate Change Science. In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, & H. L. Miller (Eds.), *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. https://www.ipcc.ch/site/assets/uploads/2018/03/ar4-wg1-chapter1.pdf

Lee, E., Felzer, B. S., & Kothavala, Z. (2013). Effects of nitrogen limitation on hydrological processes in CLM4-CN. *Journal of Advances in Modeling Earth Systems*, *5*(4), 741–754. https://doi.org/10.1002/jame.20046

Lee, W. L., Wang, Y. C., Shiu, C. J., Tsai, I., Tu, C. Y., Lan, Y. Y., Chen, J. P., Pan, H. L., & Hsu, H. H. (2020). Taiwan Earth System Model Version 1: description and evaluation of mean state. *Geosci. Model Dev.*, *13*(9), 3887–3904. https://doi.org/10.5194/gmd-13-3887-2020

Lenssen, N. J. L., Schmidt, G. A., Hansen, J. E., Menne, M. J., Persin, A., Ruedy, R., & Zyss, D. (2019). Improvements in the GISTEMP Uncertainty Model. *Journal of Geophysical Research: Atmospheres*, *124*(12), 6307–6326. https://doi.org/10.1029/2018jd029522

Li, X., Melaas, E., Carrillo, C. M., Ault, T., Richardson, A. D., Lawrence, P., Friedl, M. A., Seyednasrollah, B., Lawrence, D. M., & Young, A. M. (2022). A Comparison of Land Surface Phenology in the Northern Hemisphere Derived from Satellite Remote Sensing and the Community Land Model. *Journal of Hydrometeorology*, *23*(6), 859–873. https://doi.org/10.1175/JHM-D-21-0169.1

Liang, S., Cheng, J., Jia, K., Jiang, B., Liu, Q., Xiao, Z., Yao, Y., Yuan, W., Zhang, X., Zhao, X., & Zhou, J. (2021). The Global Land Surface Satellite (GLASS) Product Suite. *Bulletin of the American Meteorological Society*, *102*(2), E323–E337. https://doi.org/10.1175/BAMS-D-18-0341.1

Lindeskog, M., Arneth, A., Bondeau, A., Waha, K., Seaquist, J., Olin, S., & Smith, B. (2013). Implications of accounting for land use in simulations of ecosystem carbon cycling in Africa. *Earth System Dynamics*, *4*(2), 385–407. https://doi.org/10.5194/esd-4-385-2013

Lindsay, K., Bonan, G. B., Doney, S. C., Hoffman, F. M., Lawrence, D. M., Long, M. C., Mahowald, N. M., Moore, J. K., Randerson, J. T., & Thornton, P. E. (2014). Preindustrial-Control and Twentieth-Century Carbon Cycle Experiments with the Earth System Model CESM1(BGC). *Journal of Climate*, *27*(24), 8981–9005. https://doi.org/10.1175/Jcli-D-12-00565.1

Lovato, T., Peano, D., Butenschon, M., Materia, S., Iovino, D., Scoccimarro, E., Fogli, P. G., Cherchi, A., Bellucci, A., Gualdi, S., Masina, S., & Navarra, A. (2022). CMIP6 Simulations With the CMCC Earth System Model (CMCC-ESM2). *Journal of Advances in Modeling Earth Systems*, *14*(3), e2021MS002814. https://doi.org/ARTNe2021MS00281410.1029/2021MS002814

Lu, X., Du, Z., Huang, Y., Lawrence, D., Kluzek, E., Collier, N., Lombardozzi, D., Sobhani, N., Schuur, E. A. G., & Luo, Y. (2020). Full Implementation of Matrix Approach to Biogeochemistry Module of CLM5 [e2020MS002105 2020MS002105]. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2020MS002105. https://doi.org/10.1029/2020MS002105

Luo, Y., & Smith, B. (2022). *Land Carbon Cycle Modeling: Matrix Approach, Data Assimilation, & Ecological Forecasting (1st ed.)* CRC Press. https://doi.org/10.1201/9780429155659

Maki, T., Ikegami, M., Fujita, T., Hirahara, T., Yamada, K., Mori, K., Takeuchi, A., Tsutsumi, Y., Suda, K., & Conway, T. J. (2010). New technique to analyse global distributions of CO_2 concentrations and fluxes from non-processed observational data. *Tellus B: Chemical and Physical Meteorology*, *62*(5), 797–809. https://doi.org/10.1111/j.1600-0889.2010.00488.x

Manabe, S., Bryan, K., & Spelman, M. J. (1975). A Global Ocean-Atmosphere Climate Model. Part I. The Atmospheric Circulation. *Journal of Physical Oceanography*, *5*(1), 3–29. https://doi.org/10.1175/1520-0485(1975)005<0003:AGOACM>2.0.CO;2

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la-Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., & Roeckner, E. (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2. *Journal of Advances in Modeling Earth Systems*, *11*(4), 998–1038. https://doi.org/10.1029/2018ms001400

Mauritsen, T., & Roeckner, E. (2020). Tuning the MPI-ESM1.2 Global Climate Model to Improve the Match With Instrumental Record Warming by Lowering Its Climate Sensitivity. *Journal of Advances in Modeling Earth Systems*, *12*(5), e2019MS002037. https://doi.org/10.1029/2019ms002037

Meehl, G. A., Boer, G. J., Covey, C., Latif, M., & Stouffer, R. J. (2000). The Coupled Model Intercomparison Project (CMIP). *Bulletin of the American Meteorological Society*, *81*(2), 313–318. www.jstor.org/stable/26215108

Morice, C. P., Kennedy, J. J., Rayner, N. A., & Jones, P. D. (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *Journal of Geophysical Research: Atmospheres*, *117*(D8), n/a–n/a. https://doi.org/10.1029/2011jd017187

Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G., & Nemani, R. R. (1997). Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature*, *386*(6626), 698–702. https://doi.org/10.1038/386698a0

Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Stier, P., Partridge, D. G., Tegen, I., Bey, I., Stanelle, T., Kokkola, H., & Lohmann, U. (2019). The global aerosol–climate model ECHAM6.3–HAM2.3 – Part 2: Cloud evaluation, aerosol radiative forcing, and climate sensitivity. *Geosci. Model Dev.*, *12*(8), 3609–3639. https://doi.org/10.5194/gmd-12-3609-2019

NOAA. (2005). *Global Monitoring Laboratory - Carbon Cycle Greenhouse Gases*. https://gml.noaa.gov/ccgg/trends/mlo.html

Oleson, K. W., Niu, G.-Y., Yang, Z.-L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., Stöckli, R., Dickinson, R. E., Bonan, G. B., Levis, S., Dai, A., & Qian, T. (2008). Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal* *of Geophysical Research: Biogeosciences*, *113*(G1), n/a–n/a. https://doi.org/10.1029/2007jg000563

Olson, J. S., Watts, J. A., & Allison, L. J. (1985). Major world ecosystem complexes ranked by carbon in live vegetation: a database. https://www.osti.gov/biblio/6944260

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

Park, S., Shin, J., Kim, S., Oh, E., & Kim, Y. (2019). Global Climate Simulated by the Seoul National University Atmosphere Model Version 0 with a Unified Convection Scheme (SAM0-UNICON). *Journal of Climate*, *32*(10), 2917–2949. https://doi.org/10.1175/jcli-d-18-0796.1

Park, H., & Jeong, S. (2021). Leaf area index in Earth system models: how the key variable of vegetation seasonality works in climate projections. *Environmental Research Letters*, *16*(3), 034027. https://doi.org/10.1088/1748-9326/abe2cf

PCMDI. (2024). *CMIP6 - Coupled Model Intercomparison Project Phase 6*. https://pcmdi.llnl.gov/CMIP6/

Petrie, R., Denvil, S., Ames, S., Levavasseur, G., Fiore, S., Allen, C., Antonio, F., Berger, K., Bretonnière, P.-A., Cinquini, L., Dart, E., Dwarakanath, P., Druken, K., Evans, B., Franchistéguy, L., Gardoll, S., Gerbier, E., Greenslade, M., Hassell, D., Iwi, A., Juckes, M., Kindermann, S., Lacinski, L., Mirto, M., Nasser, A. B., Nassisi, P., Nienhouse, E., Nikonov, S., Nuzzo, A., Richards, C., Ridzwan, S., Rixen, M., Serradell, K., Snow, K., Stephens, A., Stockhause, M., Vahlenkamp, H., & Wagner, R. (2021). Coordinating an operational data distribution network for CMIP6 data. *Geoscientific Model Development*, *14*(1), 629–644. https://doi.org/10.5194/gmd-14-629-2021

Piao, S., Liu, Z., Wang, Y., Ciais, P., Yao, Y., Peng, S., Chevallier, F., Friedlingstein, P., Janssens, I. A., Peñuelas, J., Sitch, S., & Wang, T. (2018). On the causes of trends in the seasonal amplitude of atmospheric CO2. *Global Change Biology*, *24*(2), 608–616. https://doi.org/10.1111/gcb.13909

Qiao, F. L., Song, Z. Y., Bao, Y., Song, Y. J., Shu, Q., Huang, C. J., & Zhao, W. (2013). Development and evaluation of an Earth System Model with surface gravity waves. *Journal of Geophysical Research-Oceans*, *118*(9), 4514–4524. https://doi.org/10.1002/jgrc.20327

Rabin, S. S., Ward, D. S., Malyshev, S. L., Magi, B. I., Shevliakova, E., & Pacala, S. W. (2018). A fire model with distinct crop, pasture, and non-agricultural burning: use of new data and a model-fitting algorithm for FINAL.1. *Geoscientific Model Development*, *11*(2), 815–842. https://doi.org/10.5194/gmd-11-815-2018

Reich, P. B., Hobbie, S. E., Lee, T., Ellsworth, D. S., West, J. B., Tilman, D., Knops, J. M. H., Naeem, S., & Trost, J. (2006). Nitrogen limitation constrains sustainability of ecosystem response to CO2. *Nature*, *440*(7086), 922–925. https://doi.org/10.1038/nature04486

Reichler, T., & Kim, J. (2008). How Well Do Coupled Models Simulate Today's Climate? *Bulletin of the American Meteorological Society*, *89*(3), 303–312. https://doi.org/10.1175/bams-89-3-303

Reuter, M., Bösch, H., Bovensmann, H., Bril, A., Buchwitz, M., Butz, A., Burrows, J. P., amp, apos, Dell, C. W., Guerlet, S., Hasekamp, O., Heymann, J., Kikuchi, N., Oshchepkov, S., Parker, R., Pfeifer, S., Schneising, O., Yokota, T., & Yoshida, Y. (2013). A joint effort to deliver satellite retrieved atmospheric CO2 concentrations for surface flux inversions: the ensemble median algorithm EMMA. *Atmospheric Chemistry and Physics*, *13*(4), 1771–1780. https://doi.org/10.5194/acp-13-1771-2013

Reuter, M., Buchwitz, M., Hilker, M., Heymann, J., Schneising, O., Pillai, D., Bovensmann, H., Burrows, J. P., Bösch, H., Parker, R., Butz, A., Hasekamp, O., O'Dell, C. W., Yoshida, Y., Gerbig, C., Nehrkorn, T., Deutscher, N. M., Warneke, T., Notholt, J., Hase, F., Kivi, R., Sussmann, R., Machida, T., Matsueda, H., & Sawa, Y. (2014). Satellite-inferred European carbon sink larger than expected. *Atmospheric Chemistry and Physics*, *14*(24), 13739–13753. https://doi.org/10.5194/acp-14-13739-2014

Reuter, M., Buchwitz, M., & Schneising-Weigel, O. (2017). Product User Guide and Specification (PUGS) – ANNEX D for products XCO2_EMMA and XCH4_EMMA, 1–19. http://www.iup.uni-bremen.de/carbon_ghg/docs/C3S/CDR1_2003-2016/PUGS/C3S_D312a_Lot6.3.1.5-v1_PUGS_ANNEX-D_v1.3.pdf

Reuter, M., Buchwitz, M., Schneising, O., Noël, S., Bovensmann, H., Burrows, J. P., Boesch, H., Di Noia, A., Anand, J., Parker, R. J., Somkuti, P., Wu, L., Hasekamp, O. P., Aben, I., Kuze, A., Suto, H., Shiomi, K., Yoshida, Y., Morino, I., Crisp, D., O'Dell, C. W., Notholt, J., Petri, C., Warneke, T., Velazco, V. A., Deutscher, N. M., Griffith, D. W. T., Kivi, R., Pollard, D. F., Hase, F., Sussmann, R., Té, Y. V., Strong, K., Roche, S., Sha, M. K., De Mazière, M., Feist, D. G., Iraci, L. T., Roehl, C. M., Retscher, C., & Schepers, D. (2020). Ensemble-based satellite-derived carbon dioxide and methane column-averaged dry-air mole fraction data sets (2003–2018) for carbon and climate applications. *Atmos. Meas. Tech.*, *13*(2), 789–819. https://doi.org/10.5194/amt-13-789-2020

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Da Silva, L. A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J. C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., & Tavoni, M. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, *42*, 153–168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Righi, M., Andela, B., Eyring, V., Lauer, A., Predoi, V., Schlund, M., Vegas-Regidor, J., Bock, L., Brötz, B., de Mora, L., Diblen, F., Dreyer, L., Drost, N., Earnshaw, P., Hassler, B., Koldunov, N., Little, B., Tomas, S. L., & Zimmermann, K. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 – technical overview. *Geoscientific Model Development*, *13*(3), 1179–1199. https://doi.org/10.5194/gmd-13-1179-2020

Rödenbeck, C. (2005). Estimating CO2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport Technical Report 6. *Max Planck Institute for Biogeochemistry, Jena*. http://www.bgc-jena.mpg.de/CarboScope/s/tech%5C_report6.pdf

Rosie, F., & Koven, C. (2020). Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems. *Journal of Advances in Modeling Earth Systems*, *12*. https://doi.org/10.1029/2018MS001453

Sanderson, B. M., Booth, B. B. B., Dunne, J., Eyring, V., Fisher, R. A., Friedlingstein, P., Gidden, M. J., Hajima, T., Jones, C. D., Jones, C., King, A., Koven, C. D., Lawrence, D. M., Lowe, J., Mengis, N., Peters, G. P., Rogelj, J., Smith, C., Snyder, A. C., Simpson, I. R., Swann, A. L. S., Tebaldi, C., Ilyina, T., Schleussner, C.-F., Seferian, R., Samset, B. H., van Vuuren, D., & Zaehle, S. (2023). The need for carbon emissions-driven climate projections in CMIP7. *EGUsphere*, *2023*, 1–51. https://doi.org/10.5194/egusphere-2023-2127

Sato, H., Itoh, A., & Kohyama, T. (2007). SEIB–DGVM: A new Dynamic Global Vegetation Model using a spatially explicit individual-based approach. *Ecological Modelling*, *200*(3), 279–307. https://doi.org/10.1016/j.ecolmodel.2006.09.006

Schneising, O., Reuter, M., Buchwitz, M., Heymann, J., Bovensmann, H., & Burrows, J. P. (2014). Terrestrial carbon sink observed from space: variation of growth rates and seasonal cycle amplitudes in response to interannual surface temperature variability. *Atmospheric Chemistry and Physics*, *14*(1), 133–141. https://doi.org/10.5194/acp-14-133-2014

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., Decharme, B., Delire, C., Berthet, S., Chevallier, M., Sénési, S., Franchisteguy, L., Vial, J., Mallet, M., Joetzjer, E., Geoffroy, O., Guérémy, J.-F., Moine, M.-P., Msadek, R., Ribes, A., Rocher, M., Roehrig, R., Salas-y-Mélia, D., Sanchez, E., Terray, L., Valcke, S., Waldman, R., Aumont, O., Bopp, L., Deshayes, J., Éthé, C., & Madec, G. (2019). Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4182–4227. https://doi.org/10.1029/2019MS001791

Seferian, R. (2019). CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 CMIP esm-hist. https://doi.org/10.22033/ESGF/CMIP6.4003

Seiler, C., Melton, J. R., Arora, V. K., Sitch, S., Friedlingstein, P., Anthoni, P., Goll, D., Jain, A. K., Joetzjer, E., Lienert, S., Lombardozzi, D., Luyssaert, S., Nabel, J. E. M. S., Tian, H., Vuichard, N., Walker, A. P., Yuan, W., & Zaehle, S. (2022). Are Terrestrial Biosphere Models Fit for Simulating the Global Land Carbon Sink? [e2021MS002946 2021MS002946]. *Journal of Advances in Modeling Earth Systems*, *14*(5), e2021MS002946. https://doi.org/10.1029/2021MS002946

Seland, ?, Bentsen, M., Olivià, D. J. L., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O. A., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., & Schulz, M. (2019). NCC NorESM2-LM model output prepared for CMIP6 CMIP esm-hist. https://doi.org/10.22033/ESGF/CMIP6.7924

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., & Schulz, M. (2020). Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geoscientific Model Development*, *13*(12), 6165–6200. https://doi.org/10.5194/gmd-13-6165-2020

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., de Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., & Zerroukat, M. (2019). UKESM1: Description and Evaluation of the U.K. Earth System Model. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4513–4558. https://doi.org/10.1029/2019MS001739

Sellers, P. J., Schimel, D. S., Moore, B., Liu, J., & Eldering, A. (2018). Observing carbon cycle–climate feedbacks from space. *Proceedings of the National Academy of Sciences*, *115*(31), 7860–7868. https://doi.org/10.1073/pnas.1716613115

Shi, M., Fisher, J. B., Brzostek, E. R., & Phillips, R. P. (2016). Carbon cost of plant nitrogen acquisition: global carbon cycle impact from an improved plant nitrogen cycle in the Community Land Model. *Global change biology*, *22*(3), 1299–1314. https://doi.org/10.1111/gcb.13131

Simpkins, G. (2017). Progress in climate modelling. *Nature Climate Change*, *7*(10), 684–685. https://doi.org/10.1038/nclimate3398

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., & Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences*, *11*(7), 2027–2054. https://doi.org/10.5194/bg-11-2027-2014

Still, C. J., Riley, W. J., Biraud, S. C., Noone, D. C., Buenning, N. H., Randerson, J. T., Torn, M. S., Welker, J., White, J. W. C., Vachon, R., Farquhar, G. D., & Berry, J. A. (2009). Influence of clouds and diffuse radiation on ecosystem-atmosphere CO2 and CO18O exchanges. *Journal of Geophysical Research: Biogeosciences*, *114*(G1). https://doi.org/10.1029/2007jg000675

Sulman, B. N., Phillips, R. P., Oishi, A. C., Shevliakova, E., & Pacala, S. W. (2014). Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO2. *Nature Climate Change*, *4*(12), 1099–1102. https://doi.org/10.1038/nclimate2436

Sulman, B. N., Shevliakova, E., Brzostek, E. R., Kivlin, S. N., Malyshev, S., Menge, D. N., & Zhang, X. (2019). Diverse Mycorrhizal Associations Enhance Terrestrial C Storage in a Global Model. *Global Biogeochemical Cycles*, *33*(4), 501–523. https://doi.org/10.1029/2018GB005973

Sun, Y., Goll, D. S., Chang, J., Ciais, P., Guenet, B., Helfenstein, J., Huang, Y., Lauerwald, R., Maignan, F., Naipal, V., Wang, Y., Yang, H., & Zhang, H. (2021). Global evaluation of the nutrient-enabled version of the land surface model ORCHIDEE-CNP v1.2 (r5986). *Geoscientific Model Development*, *14*(4), 1987–2010. https://doi.org/10.5194/gmd-14-1987-2021

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., & Winter, B. (2019a). The Canadian Earth System Model version 5 (CanESM5.0.3). *Geosci. Model Dev.*, *12*(11), 4823–4873. https://doi.org/10.5194/gmd-12-4823-2019

Swart, N. C., Cole, J. N., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Solheim, L., von Salzen, K., Yang, D., Winter, B., & Sigmond, M. (2019b). CCCma CanESM5 model output prepared for CMIP6 CMIP esm-hist. https://doi.org/10.22033/ESGF/CMIP6.3583

Swart, N. C., Cole, J. N., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Solheim, L., von Salzen, K., Yang, D., Winter, B., & Sigmond, M. (2019c). CCCma CanESM5-CanOE model output prepared for CMIP6 CMIP esm-hist. https://doi.org/10.22033/ESGF/CMIP6.10232

Takata, K., Emori, S., & Watanabe, T. (2003). Development of the minimal advanced treatments of surface interaction and runoff [Project for Intercomparison of Land-surface Parameterization Schemes, Phase 2(e)]. *Global and Planetary Change*, *38*(1), 209–222. https://doi.org/10.1016/S0921-8181(03)00030-4

Tang, Y., Rumbold, S., Ellis, R., Kelley, D., Mulcahy, J., Sellar, A., Walton, J., & Jones, C. (2019). MOHC UKESM1.0-LL model output prepared for CMIP6 CMIP esm-hist. https://doi.org/10.22033/ESGF/CMIP6.5929

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/bams-d-11-00094.1

Team, T. H. D., Martin, G., Bellouin, N., Collins, W. J., Culverwell, I. D., Halloran, P. R., Hardiman, S. C., Hinton, T. J., Jones, C. D., McDonald, R. E., McLaren, A. J., O'Connor, F. M., Roberts, M. J., Rodriguez, J. M., Woodward, S., Best, M. J., Brooks, M. E., Brown, A. R., Butchart, N., Dearden, C., Derbyshire, S. H., Dharssi, I., Doutriaux-Boucher, M., Edwards, J. M., Falloon, P. D., Gedney, N., Gray, L. J., Hewitt, H. T., Hobson, M., Huddleston, M. R., Hughes, J., Ineson, S., Ingram, W. J., James, P. M., Johns, T. C., Johnson, C. E., Jones, A., Jones, C. P., Joshi, M. M., Keen, A. B., Liddicoat, S., Lock, A. P., Maidens, A. V., Manners, J. C., Milton, S. F., Rae, J. G. L., Ridley, J. K., Sellar, A., Senior, C. A., Totterdell, I. J., Verhoef, A., Vidale, P. L., & Wiltshire, A. (2011). The HadGEM2 family of Met Office Unified Model climate configurations. *Geosci. Model Dev.*, *4*(3), 723–757. https://doi.org/10.5194/gmd-4-723-2011

Tegen, I., Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Bey, I., Schutgens, N., Stier, P., Watson-Parris, D., Stanelle, T., Schmidt, H., Rast, S., Kokkola, H., Schultz, M., Schroeder, S., Daskalakis, N., Barthel, S., Heinold, B., & Lohmann, U. (2019). The global aerosol–climate model ECHAM6.3–HAM2.3 – Part 1: Aerosol evaluation. *Geosci. Model Dev.*, *12*(4), 1643–1677. https://doi.org/10.5194/gmd-12-1643-2019

Teixeira, J., Waliser, D., Ferraro, R., Gleckler, P., Lee, T., & Potter, G. (2014). Satellite Observations for CMIP5: The Genesis of Obs4MIPs. *Bulletin of the American Meteorological Society*, *95*(9), 1329–1334. https://doi.org/10.1175/BAMS-D-12-00204.1

Tjiputra, J. F., Roelandt, C., Bentsen, M., Lawrence, D. M., Lorentzen, T., Schwinger, J., Seland, O., & Heinze, C. (2013). Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM). *Geoscientific Model Development*, *6*(2), 301–325. https://doi.org/10.5194/gmd-6-301-2013

Todd-Brown, K. E. O., 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*(8), 2341–2356. https://doi.org/10.5194/bg-11-2341-2014

Tuomi, M., Laiho, R., Repo, A., & Liski, J. (2011). Wood decomposition model for boreal forests. *Ecological Modelling*, *222*(3), 709–718. https://doi.org/10.1016/j.ecolmodel.2010.10.025

UCAR. (2020). CLM technical note [last access: 2022-09]. https://escomp.github.io/ctsm-docs/versions/release-clm5.0/html/tech_note/

Varney, R. M., Chadburn, S. E., Burke, E. J., & Cox, P. M. (2022). Evaluation of soil carbon simulation in CMIP6 Earth system models. *Biogeosciences*, *19*(19), 4671–4704. https://doi.org/10.5194/bg-19-4671-2022

Varney, R. M., Chadburn, S. E., Burke, E. J., Jones, S., Wiltshire, A. J., & Cox, P. M. (2023a). Simulated responses of soil carbon to climate change in CMIP6 Earth system models: the role of false priming. *Biogeosciences*, *20*(18), 3767–3790. https://doi.org/10.5194/bg-20-3767-2023

Varney, R. M., Friedlingstein, P., Chadburn, S. E., Burke, E. J., & Cox, P. M. (2023b). Soil carbon-concentration and carbon-climate feedbacks in CMIP6 Earth system models. *EGUsphere*, *2023*, 1–21. https://doi.org/10.5194/egusphere-2023-2666

Verseghy, D. L., McFarlane, N. A., & Lazare, M. (1993). Class—A Canadian land surface scheme for GCMS, II. Vegetation model and coupled runs. *International Journal of Climatology*, *13*(4), 347–370. https://doi.org/10.1002/joc.3370130402

Verseghy, D. L. (1991). Class—A Canadian land surface scheme for GCMS. I. Soil model. *International Journal of Climatology*, *11*(2), 111–133. https://doi.org/10.1002/joc.3370110202

Verseghy, D. L. (2000). The Canadian land surface scheme (CLASS): Its history and future. *Atmosphere-Ocean*, *38*(1), 1–13. https://doi.org/10.1080/07055900.2000.9649637

Volodin, E. M., Dianskii, N. A., & Gusev, A. V. (2010). Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations. *Izvestiya, Atmospheric and Oceanic Physics*, *46*(4), 414–431. https://doi.org/10.1134/S000143381004002X

Volodin, E. M., Mortikov, E. V., Kostrykin, S. V., Galin, V. Y., Lykosov, V. N., Gritsun, A. S., Diansky, N. A., Gusev, A. V., & Yakovlev, N. G. (2017a). Simulation of modern climate with the new version of the INM RAS climate model. *Izvestiya, Atmospheric and Oceanic Physics*, *53*(2), 142–155. https://doi.org/10.1134/S0001433817020128

Volodin, E. M., Mortikov, E. V., Kostrykin, S. V., Galin, V. Y., Lykossov, V. N., Gritsun, A. S., Diansky, N. A., Gusev, A. V., & Iakovlev, N. G. (2017b). Simulation of the present-day climate with the climate model INMCM5. *Climate Dynamics*, *49*(11), 3715–3734. https://doi.org/10.1007/s00382-017-3539-7

Volodin, E. M., Mortikov, E. V., Kostrykin, S. V., Galin, V. Y., Lykossov, V. N., Gritsun, A. S., Diansky, N. A., Gusev, A. V., Iakovlev, N. G., Shestakova, A. A., & Emelina, S. V. (2018). Simulation of the modern climate using the INM-CM48 climate model. *Russian Journal of Numerical Analysis and Mathematical Modelling*, *33*(6), 367–374. https://doi.org/10.1515/rnam-2018-0032

Volodin, E. M. (2007). Atmosphere-ocean general circulation model with the carbon cycle. *Izvestiya, Atmospheric and Oceanic Physics*, *43*(3), 266–280. https://doi.org/10.1134/s0001433807030024

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., & Rose, S. K. (2011). The representative concentration pathways: an overview. *Climatic Change*, *109*(1–2), 5–31. https://doi.org/10.1007/s10584-011-0148-z

Waliser, D., Gleckler, P. J., Ferraro, R., Taylor, K. E., Ames, S., Biard, J., Bosilovich, M. G., Brown, O., Chepfer, H., Cinquini, L., Durack, P. J., Eyring, V., Mathieu, P. P., Lee, T., Pinnock, S., Potter, G. L., Rixen, M., Saunders, R., Schulz, J., Thépaut, J. N., & Tuma, M. (2020). Observations for Model Intercomparison Project (Obs4MIPs): status for CMIP6. *Geosci. Model Dev.*, *13*(7), 2945–2958. https://doi.org/10.5194/gmd-13-2945-2020

Wang, Y. P., Law, R. M., & Pak, B. (2010). A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. *Biogeosciences*, *7*(7), 2261–2282. https://doi.org/10.5194/bg-7-2261-2010

Wang, S., Luo, Y., & Niu, S. (2022). Reparameterization Required After Model Structure Changes From Carbon Only to Carbon-Nitrogen Coupling [e2021MS002798 2021MS002798]. *Journal of Advances in Modeling Earth Systems*, *14*(4), e2021MS002798. https://doi.org/10.1029/2021MS002798

Wang, Y., Yuan, Q., Li, T., Yang, Y., Zhou, S., & Zhang, L. (2023). Seamless mapping of long-term (2010–2020) daily global XCO_2 and XCH_4 from the Greenhouse Gases Observing Satellite (GOSAT), Orbiting Carbon Observatory 2 (OCO-2), and CAMS global greenhouse gas reanalysis (CAMS-EGG4) with a spatiotemporally self-supervised fusion method. *Earth System Science Data*, *15*(8), 3597–3622. https://doi.org/10.5194/essd-15-3597-2023

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T., Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori, S., & Kawamiya, M. (2011). MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments. *Geoscientific Model Development*, *4*(4), 845–872. https://doi.org/10.5194/gmd-4-845-2011

WCRP. (2024). *WCRP Coupled Model Intercomparison Project (CMIP)*. https://www.wcrp-climate.org/wgcm-cmip

Wei, N., Xia, J., Zhou, J., Jiang, L., Cui, E., Ping, J., & Luo, Y. (2022). Evolution of Uncertainty in Terrestrial Carbon Storage in Earth System Models from CMIP5 to CMIP6. *Journal of Climate*, *35*(17), 5483–5499. https://doi.org/10.1175/JCLI-D-21-0763.1

Weigel, K., Bock, L., **Gier**, **B. K.**, Lauer, A., Righi, M., Schlund, M., Adeniyi, K., Andela, B., Arnone, E., Berg, P., Caron, L.-P., Cionni, I., Corti, S., Drost, N., Hunter, A., Lledó, L., Mohr, C. W., Paçal, A., Pérez-Zanón, N., Predoi, V., Sandstad, M., Sillmann, J., Sterl, A., Vegas-Regidor, J., von Hardenberg, J., & Eyring, V. (2021). Earth System Model Evaluation Tool (ESMValTool) v2.0 – diagnostics for extreme events, regional and impact evaluation, and analysis of Earth system models in CMIP. *Geoscientific Model Development*, *14*(6), 3159–3184. https://doi.org/10.5194/gmd-14-3159-2021

Weng, E. S., Malyshev, S., Lichstein, J. W., Farrior, C. E., Dybzinski, R., Zhang, T., Shevliakova, E., & Pacala, S. W. (2015). Scaling from individual trees to forests in an Earth system modeling framework using a mathematically tractable model of height-structured competition. *Biogeosciences*, *12*(9), 2655–2694. https://doi.org/10.5194/bg-12-2655-2015

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*(5), 794–807. https://doi.org/10.1002/2013jg002591

Wenzel, S., Cox, P. M., Eyring, V., & Friedlingstein, P. (2016). Projected land photosynthesis constrained by changes in the seasonal cycle of atmospheric CO2. *Nature*, *538*(7626), 499–501. https://doi.org/10.1038/nature19772

Wieder, W. R., Lawrence, D. M., Fisher, R. A., Bonan, G. B., Cheng, S. J., Goodale, C. L., Grandy, A. S., Koven, C. D., Lombardozzi, D. L., Oleson, K. W., & Thomas, R. Q. (2019). Beyond Static Benchmarking: Using Experimental Manipulations to Evaluate Land Model Assumptions. *Global Biogeochemical Cycles*, *33*(10), 1289–1309. https://doi.org/10.1029/2018gb006141

Wieder, W. (2014). Regridded Harmonized World Soil Database v1.2.

Wieners, K.-H., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M., Bittner, M., Legutke, S., Schupfner, M., Wachsmann, F., Gayler, V., Haak, H., de Vrese, P., Raddatz, T., Mauritsen, T., von Storch, J.-S., Behrens, J., Brovkin, V., Claussen, M., Crueger, T., Fast, I., Fiedler, S., Hagemann, S., Hohenegger, C., Jahns, T., Kloster, S., Kinne, S., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Müller, W., Nabel, J., Notz, D., Peters-von Gehlen, K., Pincus, R., Pohlmann, H., Pongratz, J., Rast, S., Schmidt, H., Schnur, R., Schulzweida, U., Six, K., Stevens, B., Voigt, A., & Roeckner, E. (2019). MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP esm-hist. https://doi.org/10.22033/ESGF/CMIP6.6545

Wiltshire, A. J., Burke, E. J., Chadburn, S. E., Jones, C. D., Cox, P. M., Davies-Barnard, T., Friedlingstein, P., Harper, A. B., Liddicoat, S., Sitch, S., & Zaehle, S. (2021). JULES-CN: a coupled terrestrial carbon–nitrogen scheme (JULES vn5.1). *Geoscientific Model Development*, *14*(4), 2161–2186. https://doi.org/10.5194/gmd-14-2161-2021

Wunch, D., Toon, G. C., Blavier, J. F., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W., Sherlock, V., & Wennberg, P. O. (2011). The total carbon column observing network. *Philos Trans A Math Phys Eng Sci*, *369*(1943), 2087–112. https://doi.org/10.1098/rsta.2010.0240

Xiao, Z., Liang, S., & Jiang, B. (2017). Evaluation of four long time-series global leaf area index products. *Agricultural and Forest Meteorology*, *246*, 218–230. https://doi.org/10.1016/j.agrformet.2017.06.016

Yang, X., Thornton, P., Ricciuto, D., Wang, Y., & Hoffman, F. (2023). Global evaluation of terrestrial biogeochemistry in the Energy Exascale Earth System Model (E3SM) and the role of the phosphorus cycle in the historical terrestrial carbon balance. *Biogeosciences*, *20*(14), 2813–2836. https://doi.org/10.5194/bg-20-2813-2023

Yin, Y., Ciais, P., Chevallier, F., Li, W., Bastos, A., Piao, S. L., Wang, T., & Liu, H. Y. (2018). Changes in the Response of the Northern Hemisphere Carbon Uptake to Temperature Over the Last Three Decades. *Geophysical Research Letters*, *45*(9), 4371–4380. https://doi.org/10.1029/2018gl077316

Yuan, W., Liu, S., Zhou, G., Zhou, G., Tieszen, L. L., Baldocchi, D., Bernhofer, C., Gholz, H., Goldstein, A. H., Goulden, M. L., Hollinger, D. Y., Hu, Y., Law, B. E., Stoy, P. C., Vesala, T., & Wofsy, S. C. (2007). Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. *Agricultural and Forest Meteorology*, *143*(3), 189–207. https://doi.org/10.1016/j.agrformet.2006.12.001

Yukimoto, S., Yoshimura, H., Hosaka, M., Sakami, T., Tsujino, H., Hirabara, M., Tanaka, T. Y., Deushi, M., Obata, A., Nakano, H., Adachi, Y., Shindo, E., Yabu, S., Ose, T., & Kitoh, A. (2011). Meteorological Research Institute-Earth System Model Version 1 (MRI-ESM1). *Technical Reports*, *64*, 88. https://doi.org/10.11483/mritechrepo.64

Yukimoto, S., Adachi, Y., Hosaka, M., Sakami, T., Yoshimura, H., Hirabara, M., Tanaka, T. Y., Shindo, E., Tsujino, H., Deushi, M., Mizuta, R., Yabu, S., Obata, A., Nakano, H., Koshiro, T., Ose, T., & Kitoh, A. (2012). A New Global Climate Model of the Meteorological Research Institute: MRI-CGCM3 - Model Description and Basic Performance -. *Journal of the Meteorological Society of Japan*, *90A*(0), 23–64. https://doi.org/10.2151/jmsj.2012-A02

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yabu, S., Yoshimura, H., Shindo, E., Mizuta, R., Obata, A., Adachi, Y., & Ishii, M. (2019a). The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component. *Journal of the Meteorological Society of Japan. Ser. II*, *97*(5), 931–965. https://doi.org/10.2151/jmsj.2019-051

Yukimoto, S., Koshiro, T., Kawai, H., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yoshimura, H., Shindo, E., Mizuta, R., Ishii, M., Obata, A., & Adachi, Y. (2019b). MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP esm-hist. https://doi.org/10.22033/ESGF/CMIP6.6807

Zaehle, S., Jones, C. D., Houlton, B., Lamarque, J.-F., & Robertson, E. (2015). Nitrogen Availability Reduces CMIP5 Projections of Twenty-First-Century Land Carbon Uptake. *Journal of Climate*, *28*(6), 2494–2511. https://doi.org/10.1175/JCLI-D-13-00776.1

Zhang, Q., Pitman, A. J., Wang, Y. P., Dai, Y. J., & Lawrence, P. J. (2013). The impact of nitrogen and phosphorous limitation on the estimated terrestrial carbon balance and warming of land use change over the last 156 yr. *Earth System Dynamics*, *4*(2), 333–345. https://doi.org/10.5194/esd-4-333-2013

Zhao, F., Zeng, N., Asrar, G., Friedlingstein, P., Ito, A., Jain, A., Kalnay, E., Kato, E., Koven, C. D., Poulter, B., Rafique, R., Sitch, S., Shu, S. J., Stocker, B., Viovy, N., Wiltshire, A., & Zaehle, S. (2016). Role of CO2, climate and land use in regulating the seasonal amplitude increase of carbon fluxes in terrestrial ecosystems: a multimodel analysis. *Biogeosciences*, *13*(17), 5121–5137. https://doi.org/10.5194/bg-13-5121-2016

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R., & Myneni, R. (2013). Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of Photosynthetically Active Radiation (FPAR)3g Derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the Period 1981 to 2011. *Remote Sensing*, *5*(2), 927–948. https://doi.org/10.3390/rs5020927

Ziehn, T., Lenton, A., Law, R. M., Matear, R. J., & Chamberlain, M. A. (2017). The carbon cycle in the Australian Community Climate and Earth System Simulator (ACCESS-ESM1) – Part 2: Historical simulations. *Geosci. Model Dev.*, *10*(7), 2591–2614. https://doi.org/10.5194/gmd-10-2591-2017

Ziehn, T., Chamberlain, M., Lenton, A., Law, R., Bodman, R., Dix, M., Wang, Y., Dobrohotoff, P., Srbinovsky, J., Stevens, L., Vohralik, P., Mackallah, C., Sullivan, A., O'Farrell, S., & Druken, K. (2019). CSIRO ACCESS-ESM1.5 model output prepared for CMIP6 CMIP esm-hist. https://doi.org/10.22033/ESGF/CMIP6.4242

Ziehn, T., Chamberlain, M. A., Law, R. M., Lenton, A., Bodman, R. W., Dix, M., Stevens, L., Wang, Y.-P., & Srbinovsky, J. (2020). The Australian Earth System Model: ACCESS-ESM1.5. *Journal of Southern Hemisphere Earth Systems Science*, *70*(1), 193. https://doi.org/10.1071/es19035

Acknowledgments

This work has been supported by the European Union's Horizon 2020 project *Climate-Carbon Interactions in the Coming Century (4C)* under Grant Agreement Number 821003, and the *Advanced Earth System Model Evaluation for CMIP* (EVal4CMIP) project funded by the Helmholtz Society. I acknowledge the World Climate Research Programme (WCRP), which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP5 and CMIP6. I thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP and ESGF. The computational resources of the Deutsches Klimarechenzentrum (DKRZ, Germany) granted by its Scientific Steering Committee (WLA) under project ID 854 (Erdsystemmodellevaluierung) were used to compute these results and are kindly acknowledged.

Firstly, I would like to give a huge thank you to my supervisor, Prof. Dr. Veronika Eyring. You gave me the opportunity to spearhead an amazing group in a new location with you. Even when things got tougher, you gave me courage with your unending patience, support, and enthusiasm. Thanks to you, I was able to attend many interesting conferences and workshop, as well as meet and befriend many leading figures and good colleagues.

Among these are my second examiner, Prof. Dr. Gerrit Lohmann from the AWI, who I still owe an introductory ESMValTool workshop at AWI - we'll get to it.. sometime!

Thank you to all the co-authors on my publications: Dr. Michael Buchwitz, Dr. Max Reuter, Prof. Dr. Peter M. Cox, Prof. Dr. Pierre Friedlingstein, Dr. Chris Jones, Prof. Dr. Colin Jones, Prof. Dr. Sönke Zaehle. Your input and deep insights into the topics helped me elevate my research to the next level, and you never lost sight of the big picture! Michael, Max, you adopted me into your group until ours grew big enough to warrant our own group meetings, but even afterwards you still took time for me to discuss the satellite data whenever I had questions! I hope I could in turn give you some insights into climate models. Peter, Pierre, thank you for hosting me for a fun week in Exeter, where we finally had the spark of inspiration to solve the SCA trend conundrum.

Many thanks to the DLR-EVA group for helping me start my journey into climate science, and imparting all the knowledge I needed to work in the new Bremen group. Birgit Hassler and Manuel Schlund, you were my shining lights when I needed to vent, when I got stuck and didn't know where to go, and most of all, you were just amazing company! I hope we get to work together for many more years to come. Thank you for your work in proofreading this

thesis, even on tight schedules and short notices. Thank you to the climate modelling group here in Bremen, we had some fun times! We got thrown out of our offices during COVID, had what felt like 1465753 different dates for moving back, and in the end I get to hand in my thesis on the day we move back. As if that isn't destiny. I got to see the group grow from just Veronika and me to what we are today! My first office mate, Katja Weigel, is still my office mate all these years later! Thank you Katja for always lending an ear and your help, for making me laugh so hard we entertained the whole floor sometimes. If I wanted to list all my colleagues with whom I enjoyed my discussions, who elevated my office life to the next level, I'd be basically be listing them all! So just have a collective thank you, you know how much you mean to me!

I also want to give a huge thank you to the ESMValTool community. We've had the most amazing workshops, fun virtual meetings where we ended up giving house tours or where Bouwe makes me show my newly made plush on camera. Never change!

Thank you to all my friends, both online and offline, for their support during these years, for keeping me sane during insane times, and for making my life a better place.

Last, but certainly not least, I want to thank my family. Without your support, your love, your encouragement, but also your teasing, your pushing and kicking, I would never have gotten this far! I know you will always have my back. And a final word to my dear brother: Thomas, I beat you to submission after all!