Biogeophysical and Biogeochemical Processes at the Surface Determine Habitability in Two Extreme Climates

Greta Enid Marie Shum

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2024

Reading Committee:

Abigail L. S. Swann, Chair

Cecilia M. Bitz

Dargan M. W. Frierson

Program Authorized to Offer Degree:
Atmospheric Sciences

University of Washington

Abstract

Biogeophysical and Biogeochemical Processes at the Surface Determine Habitability in Two Extreme Climates

Greta Enid Marie Shum

Chair of the Supervisory Committee:
Abigail L. S. Swann
Departments of Atmospheric Sciences & Biology

In this dissertation we make two habitability assessments of Earth's climate. We examine first the habitability of Snowball Earth and second the habitability of Earth's future climate under different rates of decarbonization. In both cases, the concept of habitability can initially be seen as an interdisciplinary link between climate science and other fields; in using this metric, we gain insight into water-based life on Earth and on other worlds as well as mitigation policy. At a deeper level in both cases, habitability serves as a guiding light to illuminate other physical climate insights, in particular the power of biogeophysical and biogeochemical surface processes that couple the atmosphere, land, and ocean in regulating Earth's climate. Because these "surface-level" processes elicit responses across various components of the climate system, they have an strong influence on climate feedbacks and, as a result, the resilience or sensitivity of Earth's climate to extreme radiative changes.

In Chapter 2, we examine the influence of land surface albedo on habitability as it concerns eukaryotic, photosynthetic algae. We find that by regulating the flux of energy to the Earth system and, by extension, regulating the potential flux of water to the atmosphere, bare land surface albedo exerts a stronger control on habitability of likely refugia than the radiative influence of CO_2 concentration.

In Chapters 3 and 4, we investigate how the cycling of carbon among the atmosphere, land, and ocean, as determined explicitly by carbon cycle model structure and implicit model

assumptions about how carbon is fluxed across carbon pools, influences the habitability of future Earth for human populations.

In Chapter 3, we examine four simple climate models that represent the terrestrial and ocean carbon cycles differently. We test the models' responses to a variety of increasing and decreasing CO_2 emissions and find that the models behave similarly in an increasing-emissions regime and diverge in their response to decreasing emissions. We quantify how those differences in response to decarbonization change implied necessary mitigation policy through changes to the Remaining Carbon Budget (RCB).

In Chapter 4, we explore which aspects of modeled global biogeochemistry exert the strongest influence on the climate response to decarbonization. To do this, we compare two representations of the carbon cycle, selected from Chapter 3's intercomparison, and perform a comparison using a set of perturbed parameter ensembles. We find that differences in carbon cycle model structure lead to differences in the long-term carbon sinks but do not explain differences in climate response to decarbonization. This disconnect between carbon cycle modeling choices and climate response reflects a structural disconnect between carbon cycling and energy balance. To account for the role ocean circulation plays in both systems, we add a link between the two components and consequently change the coupling within the energy balance component and the carbon cycle component. This structural connection assumes different trade-offs in the observed carbon-climate system, leading to different behavior both in an increasing-emissions and decreasing-emissions regime. In particular, when we assume that exchange of heat to the deep ocean is linked to ocean circulation and thus the oceanic uptake of carbon, we find that in order to match present day observations of temperature and CO_2 the climate must warm more per unit of forcing, resulting in a smaller RCB.

These insights refine our understanding of how the atmosphere, biosphere, and hydrosphere are connected and have important consequences for how we represent their coupling in complex and simple climate models for other applications.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

What will follow is a piece of writing about what makes a place into a home for life as we know it. In many ways I felt out of place when I moved to Seattle with a literature degree and stepped into a graduate program alongside seasoned researchers and scientists. As I prepare to leave, I feel perhaps unreasonably proud of the metaphor in my conclusions: that habitability is made possible from strong but often hidden relationships. It feels happily true to the pursuit of this dissertation to highlight those relationships in my life, which are responsible for the very existence of this work and for the fact that it was able to survive to this stage.

First, I am grateful for the support and guidance I have received by being a part of various academic communities while here at UW. The research in this dissertation was generously funded by the National Science Foundation EAGER Program, the UW Program on Climate Change (PCC) Climate Research Accelerator Award, and the UW Astrobiology Program. The NCAR CESM community provided invaluable resources time and again. I lost count of the number of times David Bailey got me unstuck running CESM. The PCC became my sandbox for interdisciplinary projects, and I'm so grateful for their support of the ACORN program and for the many friends I made through the community. I am hugely grateful to the UW Atmospheric Sciences departmental staff, especially Erica Coleman, Jennifer Siembor, and Cathy Liao. I am also grateful to the ASE Union, UAW 4142, for supporting students tirelessly. The Ecoclimate Lab has been a terrific place to learn and share insights as a scientist, and I appreciate meeting so many friends and collaborators as part of it, especially Marysa Laguë, Claire Zarakas, Jennifer Hsiao, Amy Liu, Sam Pennypacker, Lily Zhang, Jin Kim, Lucas Vargas Zeppetello, Fiona Lo, Luke Parsons, Chad Small, Patrick Murphy, Haynes Stephens, and Ben Buchovechy.

I feel so lucky to be a part of the 2018 cohort of graduate students, whom I feel so

proud of and proud to be among. I'm also sincerely grateful to the Atmospherics, which has provided fellowship and entertainment as well as an outlet for some of my greatest contributions to the field of atmospheric sciences.

I am grateful for the friends who I have leaned on since before graduate school, especially the incredible community I had in Princeton. I'm grateful for the friendship of Jane Baldwin, Caroline Perkins, Mochi Liu, Mitchell Nahmias, Alex Kasdin, David Nicholls-Paul, Katherine Clifton, Silvia Mannes, Bert Jaeck, Chuck Witt, and Kelsea Witt. They encouraged me to apply to graduate school, listened to me, and steadied me from the other end of the phone. I've always been so humbled by the unwavering support of Michael Lemonick and John McPhee. They continue to inspire me with their ability to bring new dimensions to science with words. I was lucky enough to work with Sarah Jackson, Lynn Loo, Shari Bell, Jennifer Brady, Alyson Kenward, Bernadette Woods Placky, and Heidi Cullen – all incredibly inspirational women who made me feel like my career mattered. Tom Scherer, Jake Huang, Tamara Pico, Sierra Stites, Rima Abhyankar, and Ericka Howill have been there for me for many years, and I don't know who I would be without them.

I will always feel like I won the housing jackpot with Stephanie Rushley, Hannah Zanowski, Mika Malila, Katie Breen, and Logan Schumacher, with whom I lived during my early years as a student and who made those homes very happy ones – especially during lockdown.

I am grateful to the many friends who have been there for me here in Seattle and within UW, including Jessica Ouyang, Brian Nixon, Ursula Jongebloed, Vince Cooper, Adam Sokol, Carley Fredrickson, Sami Turbeville, Tyler Cox, Lily Hahn, Rudy Garcia, Zoe Snape, and Zho Ragen. There are not enough words to show my appreciation for Marysa and Claire, who have provided mentorship, guidance, laughter, friendship, collaboration, conspiracy, and their shoulder at so many points along this journey.

I am grateful to all of my mentors, teachers, collaborators, and advisors. Ed Waddington, Olivia Truax, James Yoon, and Charles Koven have been incredible collaborators and taught me by example how to be a scientist. My PhD committee is made of such inspirational

people, who have been more than committee members, giving this work and me so much of their time, intellect, feedback, and kindness: Stephen Riser provided encouragement and support from the beginning. Victoria Meadows has been an inspiration, tirelessly working to make the Astrobiology Program a paragon of interdisciplinary community, which it has become for me. Stephen Warren, who led the Snowball Earth habitability project, has essentially become a second advisor to me, graciously supplying his expertise, wisdom, time, and keen eye. I am especially grateful to my reading committee: Dargan Frierson has taught me so much as a science communicator, teacher, researcher, and mentor, and he has kept me grounded while encouraging me to pursue my scientific goals. Cecilia Bitz has been an exceptional mentor, helping me tackle technical challenges, decide on research directions, and articulate findings.

At the end of the day, as I descend to the basement level of ATG to collect my bike, more often than not, the question spinning in my head is how I am going to be able to thank my advisor, Abby Swann, for everything she has done for me. Abby has supported me across every research hurdle and inspired me in her leadership, mentorship, and friendship. I feel so grateful that I've been able to spend the last six years learning from her, and I look forward to many more of doing the same.

Finally, I am sincerely grateful to my family, which has grown significantly since six years ago. I'm inspired by my siblings' generosity, brilliance, and patience with me. Most of the breaks I took from this work were to play with my nieces. My partner's family has fed me and shown me so much love especially in the last few months. My parents have been my role models for their work ethic, humility, and love, and I cannot thank them enough for their seemingly unending support of my dreams (even as they have evolved). And finally, I am eternally thankful for my partner, Nick, and our dog, Fly, who are my home.

DEDICATION

To my parents

Chapter 1

INTRODUCTION

1.1 Climate models as tools and methodologies

Global climate models are a critical class of tools for exploring unobservable climate states, including paleoclimate and future climate, based on our knowledge of the Earth system. By modeling these unobservable states, we can also identify fundamental behaviors of the Earth system. This effort allows us to demarcate the limits of our knowledge by identifying which elements of the observable climate exert the strongest influence over the predictability of key climate metrics, such as habitability. We can gain insight not only from identifying those elements and describing how further measurements could increase our total understanding but also by building an awareness of how our characterization of the Earth system in models reflects how we interact with the climate system. This can be done by using different kinds of models or different configurations in our simulations of different climate settings to identify which aspects of the Earth system (and their connections) have the greatest impact.

In this dissertation, I examine habitability as the climate state of interest, and define two thresholds of habitability, one for past climate and one for future climate. The first has direct implications for understanding the history of life on Earth as well as understanding the possibility of Earth-like life elsewhere in the universe, and the second has direct immediate implications for human decisions in response to anthropogenic climate change. Both studies reveal limitations in our knowledge of the *observable* world, which, if pursued, would close the gap in what we can and cannot say about these unobservable worlds. In short, the dialectic of models and observations can be used to build knowledge about climates removed by time and space.

In both of our climate settings, Snowball Earth and future climate, the behavior of processes at the interface between Earth systems – atmosphere, ocean, and land – that is, processes at the surface, where water, energy, and carbon are exchanged and where biogeo-

chemical and biophysical feedbacks take place, exert a strong influence on their habitability. This "surface-level" weakness may reflect real-world obstacles that disciplinary boundaries impose as well as the potential leaps that intentionally cross-disciplinary research projects promise.

1.2 Habitability as a theme and a metric

In this dissertation, I have tried to approach a central question: how do the biogeochemical and biogeophysical processes that regulate land-atmosphere coupling define our understanding of habitability? While habitability might seem like an arbitrary metric with which to measure our and other climate systems, habitability has a rich history as the framing for many fundamental research questions in Earth Sciences, beginning with arguably the first work of popular science, *Entretiens sur la Pluralité des Mondes* or *Conversations on the Plurality of Worlds* by Bernard le Bovier de Fontenelle, written in 1686, which describes the leading theory of astrophysics at the time as well as the climates of various bodies in the Solar System (including Venus, Jupiter, Mars, and the Moon), primarily through the lens of what the inhabitants would experience and how they would appear, given their climates (Fontenelle, 1686).

Fontenelle engages the reader with the theme of habitability to illustrate a heliocentric view of the Solar System less than 50 years after the death of Galileo. It has since been argued that Fontenelle chose this narrative construction (and published in French, which was more accessible to non-academic audiences than Latin) not only to appeal to women readers but also to appeal more broadly to academic audiences and avoid any potential tension with the still scientifically radical contents of his work (Aït-Touati and Emanuel, 2011).

The format of *Entretiens* is a narration of six consecutive evening strolls in which a natural philosopher describes the workings of universe to a young and curious marquise, and through their didactic conversations, the climates of the various planets are described, most notably through the framing of what the inhabitants of those worlds would experience as well as what the characteristics of those beings would be, as a result of their climates. The work has a strong legacy in astrobiology, and it illustrates the power of the question

of "habitability" as a beacon that can be used to reveal and engage important research questions across the physical sciences from paleoclimate to climate change.

In this dissertation, I hope to continue Fontenelle's narrative legacy by exploring two stories of Earth's habitability that illuminate important aspects of our climate system.

Assessing Snowball Earth habitability

In the context of Snowball Earth, we ask whether early photosynthetic life could survive on the surface and what influences the answer the most. Our definition of habitability is based on known characteristics of life on Earth and partially motivated by the search for life on other planets, which requires liquid water to be present and on the surface (since any detection of life could only be made at the surface). Since evidence exists of surface-dwelling photosynthetic, eukaryotic algae prior to Snowball Earth, the mystery arises: how could life have survived a period when global mean surface temperature was significantly below zero, and what is the mechanism by which local temperature could have exceeded this threshold.

Assessing future climate habitability

While the concept of habitability has often been applied to exoplanets and past Earth climates, perhaps the most prescient application at the moment is its application to Earth's climate under anthropogenic emissions of CO_2 and other heat-trapping gases.

In this context, we can define habitability of the Earth system through the lens of the risks to global ecosystems and human health and well-being that anthropogenic climate change poses. In its Special Report on Global Warming of 1.5°C, the IPCC found that limiting global warming to below 1.5°C above pre-Industrial levels is critical because of the risks this level of warming poses to a host of natural and human systems including species disruption and loss in terrestrial, wetland, and ocean ecosystems, as well as loss of the environmental services those systems provide, extreme weather risk such as drought, extreme precipitation, and extreme heat, increased vulnerability to food insecurity, vector-borne diseases, poverty, and climate-driven displacement due to sea level rise and other climate change-driven impacts (SR1.5, 2018).

As a global collective, through the international Paris Agreement and other decarbonization pledges, nations and companies are now attempting to rapidly reduce to anthropogenic CO_2 emissions to net- or near-zero, but uncertainty still surrounds the amount of CO_2 that can be safely emitted and the related immediacy of the reductions needed (IPCC SR1.5, 2018; Rogelj et al., 2019). As public and private actors begin to scope decarbonization pathways, simplified models of the climate system are increasingly used to toe this line of decarbonization (IPCC Working Group I, 2021c; Sanderson et al., 2023). In Chapter 3, we ask whether these simplified representations of the climate system agree the reductions necessary to maintain habitability. We further explore how structural distinctions between models leads to differing climate outcomes. We seek to identify which aspects of these simplified models have the strongest influence on the relationship between remaining emissions and remaining warming and whether that influence reflects our complete understanding of the climate system.

While in each climate setting we define habitability differently, in both cases, "habitability" serves as a way to apply our understanding of the Earth system to wider disciplinary questions. In the case of Snowball Earth, understanding the habitability of the Snowball Earth climate not only expands our understanding of life on Earth, but it also allows us to expand our definition of habitability for exoplanets. In the case of future climate, understanding which processes determine total warming from anthropogenic climate change allow us to set informed mitigation targets.

1.3 Influence of biogeophysical coupling on habitability of Snowball Earth

The biogeophysical coupling between land and climate can cause changes in land properties such as surface albedo, evaporative resistance, surface roughness and other aspects of modern vegetation that influence water and energy fluxes, to be affected by and influence climate (Laguë et al., 2019; Bonan, 2015). The importance of interactivity between surface properties has been shown in simple conceptual models such as Daisyworld (Watson and Lovelock, 1983), which was used to demonstrate the plausibility of the Gaia Hypothesis (Lovelock, 1986; Wood et al., 2008). The importance of biogeophysical feedbacks has also been shown in global climate models: coupling a land surface model with an interactive

atmosphere yields larger temperature changes in response to changes in land surface properties, as well as different patterns in the climate response, than the same change in an uncoupled configuration, in which feedbacks are suppressed (Laguë et al., 2019). Boreal ecosystems in particular exhibit sensitivities to this coupling because of differences between the boreal forest and tundra in surface albedo, surface roughness (which alters the effect of snow), and the partitioning of energy between latent and sensible heat. For this reason, the spatial distribution of boreal forest is an important regulator of global climate (Bonan et al., 1995; Swann et al., 2010). In my Master's work, I showed that biogeophysical feedbacks in the boreal region can alter habitats and ecosystems over time by modeling the emergence of a new biome can influence neighboring ecosystems through their influence on regional climate (Shum et al., 2023). As a result, while they are a challenge to characterize given the uncertainty of past land surface types, biophysical feedbacks are key to include when assessing local climate in a paleoclimate context.

Earth system models, such as the Community Earth System Model (CESM) offer exciting opportunities to study worlds that would be greatly influenced by the inclusion of biogeophysical coupling between land and atmosphere, such as Snowball Earth. As in the modern boreal region, the contrast between snow-covered and snow-free land is stark.

In Chapter 2, we use an idealized land surface model called the Simple Land Interface Model (SLIM) (Laguë et al., 2019), which allows us to test the influence of land surface albedo on habitability. Using SLIM allows us to modify properties of the land surface in isolation, whereas modifying a land cover type in a more realistic land model might change two properties at the same time. By modifying land surface albedo in isolation, we can attribute any changes in climate directly to the radiative changes and accompanying biogeophysical feedbacks it kicks off.

Moreover, modeling Snowball Earth habitability not only adds context to early life on Earth, but also enhances our understanding of exoplanet habitability. In this context, habitability can be defined as hosting liquid water on the surface. The TRAPPIST-1 system, for example, is composed of a cool red dwarf star hosting seven approximately Earth-sized and Earth-density planets (Agol et al., 2021). Several planets lie within or just outside of the habitable zone; TRAPPIST-1e in particular has the strongest likelihood of hosting

liquid water despite having a equilibrium temperature several degrees below Earth's. Due to a lack of knowledge of their surface, modeling studies that have been conducted to assess the TRAPPIST planets' habitability use exclusively aquaplanet climates and are unable to include any representations of land or land-atmosphere feedbacks (Wolf, 2017; Sergeev et al., 2022). While the changes we make through atmospheric CO_2 and bare land albedo do not drastically shift global mean surface temperature, they do alter local climate in ways that could drastically shift habitability of the planets, suggesting that even the planets with very cold global mean surface temperatures could possibly host life.

1.4 Influence of biogeochemical coupling on future habitability with decarbonization

The cycling of chemical species, including greenhouse gases, between the biosphere, geosphere, and atmosphere is also regulated by processes at the surface.

On land, carbon is taken up from the atmosphere into the terrestrial biosphere through gross primary production (GPP) and returned to the atmosphere through plant respiration as well as decomposition, which together sum to the net uptake by the terrestrial biosphere, or net ecosystem production (NEP). In response to anthropogenic changes to the carbon cycle-climate system, the total net land–atmosphere flux balances land-use emissions from human management of ecosystems and changes to the land sink as a result of global environmental changes including higher atmospheric CO_2 burden, warming from climate change, and increased nitrogen deposition, all of which influence the biological processes that regulate the exchange of carbon (as well as water and energy). The total response of the terrestrial carbon-climate system to anthropogenic emissions-driven and ultimately the future land sink is largely in the hands of biological feedbacks to these changes, including CO_2 fertilization (increased GPP in response to increased CO_2) and climate-driven feedbacks (including increased respiration at higher temperatures).

By contrast, uptake of anthropogenic carbon from the atmosphere by the ocean reservoir is regulated by physical processes that modulate air-sea CO_2 exchange at the ocean surface as well as chemical and circulation feedbacks that regulate the transport of that carbon from surface to depth, though new evidence suggests that previously held assumptions about the

separation between anthropogenic and natural carbon cycling may no longer hold. Since biology plays a key part in natural ocean uptake of CO_2, until recently, the ocean's response to anthropogenic emissions was not considered sensitive to the operation of the biological carbon pump, which was assumed to be effectively in steady-state (Sarmiento and Gruber, 2006). However, new evidence suggests that the sensitivity of natural carbon ocean fluxes to forcing may be high (Gruber et al., 2023) and small changes to remineralization depth may lead have important consequences for future atmospheric CO_2 (Kwon et al., 2009).

In general, total uncertainty in carbon uptake by the land and ocean sinks is dominated by biological processes on land and physical process in the ocean. For this reason, in Chapter 4, we test the sensitivity of Simple Climate Models, or SCMs to parameters that regulate biological processes at the land surface and physical processes that regulate carbon exchange at the ocean surface.

1.5 Importance of identifying and quantifying structural assumptions within Simple Climate Models

In Chapters 3 and 4 we focus our attention on SCMs because of their critical role in setting and maintaining a global Remaining Carbon Budget (RCB). Biogeochemical processes in ESMs are represented with increasing complexity and remain poorly constrained (Seiler et al., 2022). Until this point, various significant challenges have prevented modeling centers from running coordinated mitigation scenarios in emissions-driven mode (e.g., not many models included biogeochemistry until recently; the computation burden of spinning up the carbon cycles can be significant; and past emissions are more uncertain than past CO_2 concentration).

As a consequence, mitigation decision-making has come to (perhaps unwittingly) rely heavily on simple models of the carbon cycle to determine decarbonization pathways (IPCC Working Group II, 2022). The added accessibility of SCMs in contrast ESMs elevates SCMs and intermediate-complexity models as tools for climate policy and as well as research direction scoping for ESMs and MIPs, as discussed in Chapter 4.

Since SCMs are charged with reproducing the general behavior of ESMs (in their evolution of temperature and atmospheric CO_2), it is important to characterize exactly how

changes to biogeochemical cycling within these models affects predictions for total carbon sink strength, and how that could shift our expectation for future habitability. Specifically, in tuning SCMs to ESM output, parameters combinations can be found that match observations of historical emissions and concentrations. What remains unclear is whether the processes these parameters represent reflect realistic biogeochemical functioning in ESMs.

In order to fully answer this question, it will be necessary to sample across multiple realizations of emissions-driven ESM simulations of decarbonization; however, in the absence of these simulations, we can explore potential trade-offs that occur in the process of constraining SCM ensembles to the historical record, setting the stage for future process-driven validation by ESMs.

1.6 Choosing simplicity amid the complexity of habitability

In both cases, we attempt to assess what controls global habitability. However, habitability is complex. In reality, the survival of life on Snowball Earth or other icy worlds is an outcome of much finer-scale processes than what we resolve. Even sea glacier dynamics play a critical role in the mechanics of the refugia we propose, yet we do not resolve them in our model experiments. Future climate habitability will largely be a function of how localities adapt to their unique climate impacts. Still, global-scale physical science-based assessments serve as critical points of engagement where local mitigation pathways intersect (Horton et al., 2021), and our Snowball habitability assessment can be easily used to refine future model simulations of Snowball Earth.

While simplicity will always come at the expense of realism, in these two cases, a simpler experimental design leads to conclusions that are more generalizable. Surface albedo may widen the margin of climate habitability enough to move an exoplanet from uninhabitable to a possible characterization target. And connecting carbon-climate processes may shift some mitigation pathways from optimistic to optimal.

Chapter 2

OCEAN BAYS SURROUNDED BY DESERT LAND COULD SUPPORT PHOTOSYNTHETIC LIFE ON SNOWBALL EARTH

(This chapter has been submitted for publication to AGU Advances as Shum, Greta Enid Marie, M. M. Laguë, A. L. S. Swann, C. M. Bitz, E. D. Waddington, S. G. Warren (2024). Ocean bays surrounded by desert land could support life on Snowball Earth.)

Abstract

Photosynthetic eukaryotic algae survived the Neoproterozoic Snowball Earth events, indicating that liquid-water refugia existed somewhere on the surface. We examine the potential for refugia at the coldest time of a snowball event, before CO_2 had risen and with high-albedo ice on the frozen ocean, before it became darkened by dust deposition. We use the Community Earth System Model to simulate a "modern" Snowball Earth (i.e., with continents in their current configuration), in which the ocean surface has frozen to the equator as "sea glaciers", hundreds of meters thick, flowing like ice shelves. Despite global mean surface temperatures below -60°C, some areas of the land surface reach above-freezing temperatures because they are darker than the ice-covered ocean. With low CO_2 (10 ppm) and land-surface albedo 0.4 (characteristic of bright sand-deserts), 0.1 percent of the land surface could host liquid water seasonally; this increases to 12 percent for darker land of albedo 0.2, characteristic of polar deserts. Narrow bays intruding from the ocean to these locations (such as the modern Red Sea) could provide a water source protected from sea-glacier invasion, where photosynthetic life could survive. The abundance of potential refugia increases more strongly in response to reducing the land albedo than to increasing the CO_2, for the same global radiative forcing.

2.1 Introduction

On Earth and Earth-like worlds, a large negative radiative forcing can initiate a positive ice-albedo feedback and ultimately lead to global glaciation (Budyko, 1969; Sellers, 1969). Geologic evidence indicates that the Earth has experienced several such events since the emergence of life (Harland, 1964; Kirschvink, 1992; Hoffman et al., 2017; Evans, 2000). These events were likely caused by a reduction of the atmospheric greenhouse effect, resulting from disturbance of the global carbon cycle (Hoffman et al., 2017). During the Neoproterozoic era (600-800 Ma), two "Snowball Earth" events occurred: the Sturtian, with a duration of 58 million years, and the Marinoan, with a duration of 10 million years (Macdonald et al., 2010). The oceans would likely have been covered by ice hundreds of meters thick, but photosynthetic eukaryotic algae were able to survive (Porter, 2004; Knoll, 2011, 2014), indicating that some liquid water was maintained at or near the surface where light was available for photosynthesis.

In this paper, we focus on the "hard" Snowball Earth, in which the equatorial ocean would be covered by thick ice. That ice differed in several ways from sea ice on the polar oceans of modern Earth. Modern sea-ice thickness is limited to a few meters by summertime melting and by a heat flux F_0 of several watts per square meter from the ocean water below, which originally gained its heat by absorption of solar energy at lower latitudes.

But at the onset of a snowball event, when sea ice reached the equator it would shut off solar heating of the ocean water below. After a few thousand years, the ocean would have lost its reservoir of heat, leaving only geothermal heat, $F_0 \approx 0.08$ W m^{-2} (about one-hundredth that of the modern oceanic flux F_0 to the ice bottom), increasing the equilibrium ice thickness from a few meters to a few hundred meters (Warren et al., 2002).

The geothermal flux is essentially independent of latitude, but the ice surface on the snowball ocean would be colder at high latitude than at low latitude, resulting in thicker ice at higher latitude. The latitudinal thickness gradient would cause the ice to flow (Goodman and Pierrehumbert, 2003). In this state, the thick ice on the frozen ocean would be growing from above by snowfall (the original sea ice having melted off the bottom), and therefore can be classified as glacier ice rather than sea ice. This ice, flowing like the modern Antarctic ice

shelves but not dependent on continental glaciation, is called a sea glacier (Warren et al., 2002). Sea glaciers are computed to flow as much as 7-50 meters per year even when they cover the entire ocean (Goodman, 2006; Li and Pierrehumbert, 2011). If a small area of the ocean were to open up, it would be quickly filled by inflow of the sea glacier. How, then, could liquid water be maintained at the surface? Several hypotheses for refugia have been proposed, which we now list.

2.1.1 Types of proposed refugia

Five ideas have been proposed for liquid-water refugia at the ocean surface.

(a) **Hotspots.** Hoffman and Schrag (2000, 2002) noted that geological hotspots at the ocean floor under shallow water, as occur near the coasts of Hawaii and Iceland, would melt inflowing ice fast enough to maintain pools of liquid seawater. These pools would be small in area, and would not be stable for millions of years, so any life would have to survive many long and deep migrations.

(b) **Thin ice.** McKay (2000), using a broadband model for solar radiation, proposed that absorption of sunlight within the ice might be able to limit the tropical ice thickness to 10 m. Warren et al. (2002) pointed out that the visible and ultraviolet wavelengths, which penetrate deeply, are not absorbed but eventually are scattered back out to space, whereas the near-infrared wavelengths, which are indeed absorbed, are absorbed in the top few millimeters, so their heat is easily conducted up to the atmosphere. By modifying McKay's model to compute the radiation spectrally, Warren et al. found that the equilibrium ice thickness in a typical example grew from 1 m to 800 m. McKay joined as a coauthor on that paper, agreeing that the thin-ice solution was not viable. Pollard and Kasting (2005) tried to find a thin-ice solution that would even hold off sea glaciers, and succeeded only when three parameters were pushed beyond their acceptable limits (Warren and Brandt, 2006). An improved model (Pollard et al., 2017) convincingly rejected the thin-ice solution.

(c) **Waterbelt.** Some models of snowball initiation have found that sea ice could reach

the outer tropics but still leave a wide belt of open water centered on the equator, spanning tens of degrees of latitude and circling the globe, if the sea-ice (or sea-glacier) albedo is low enough. Abbot, Voigt, and Koll (2011) found that this "waterbelt" state could exist with sea-glacier albedo of 0.45 but is inaccessible for sea-glacier albedo > 0.55. Dadic et al.'s (2013) measurements of modern surrogates for sea-glacier surfaces found albedos $0.57 - 0.80$ under clear sky, and even higher under cloudy sky, arguing against the waterbelt idea. A follow-on investigation by Voigt's group (Braun et al., 2022) found the waterbelt to be unviable, even with sea-glacier albedo as low as 0.45. Most recently, Hörner and Voigt (2023) showed that the waterbelt in earlier models resulted from inadequate vertical resolution in the sea ice.

(d) **Ice surface.** Vincent and Howard-Williams (2000) and Vincent et al. (2000) suggested that microbial life could survive on the ice surface of Snowball Earth, pointing to the widespread microbial communities that thrive both in surface meltwater pools and in brine pockets, on modern Arctic and Antarctic sea ice and ice shelves. These communities can persist even if only a few days per year have temperatures above freezing. Such communities could indeed have been active during the rapid advance of sea ice at the onset of Snowball Earth. But after the ice reaches the equator, the strong positive albedo-temperature feedback causes dramatic cooling. An early general circulation model (GCM) of the hard snowball by Pollard and Kasting (2004) obtained a global average surface temperature of -49°C. The warmest temperature on the ocean surface was found on a summer afternoon in the subtropics, ~ -30°C, which seemed to rule out any surface life. However, we will see below that Vincent's proposal can be resurrected if it is considered in combination with the next proposed refugium *e*.

(e) **Narrow bay.** One place where ocean water could be safe from sea-glacier inflow is at the innermost end of a narrow bay resulting from continental rifting, like the modern Red Sea. When flowing into a narrow bay, nearly enclosed by dry land, ice flow can be slowed by resistive shear stresses from the side-walls, and by obstacles such as islands,

shoals, or narrows in the bay. If the bay is long enough and the sublimation rate is high enough, the ice thickness can taper down to zero before the end of the bay is reached (Campbell et al., 2011, 2014). If ocean-sourced water flowing under the ice can find its way to the end of the narrow bay, it could provide a refugium safe from sea glaciers. If the surrounding land is net-evaporative (i.e., potential sublimation outpaces precipitation, as in deserts), this place would be safe from land glaciers as well.

Refugia in narrow bays would be larger than the isolated geothermal hotspots around volcanic islands, and would have long lifetimes, similar to timescales of continental drift. But even if the end of the bay is safe from sea-glacier inflow, there is the risk that the climate might still be so cold that thick sea ice would grow locally. Such a refugium would be feasible only if local temperatures reach above freezing, which could occur because the albedo of nearby bare land surfaces would be lower than that of the ice-covered ocean. Land surfaces during the Snowball Earth events were not vegetated: possible snow-free surfaces would be bare rock, bare soil, and sand. Most of them are brighter than vegetated land, but with albedos 0.1-0.4 they are much darker than ice or snow. To evaluate the feasibility of this refugium, climate modeling is needed, and that is the subject of this paper.

2.2 Investigation of the narrow bay refugium

Given the high uncertainty of Neoproteorozoic paleocontinental reconstructions, for this investigation we apply an earth-system model to the modern continental configuration. This approach has been used in prior investigations, called "modern Snowball Earth" (Voigt and Marotzke, 2010; Liu et al., 2018). It has the advantage of familiar geography, allowing comparisons of atmospheric circulation and climate with the familiar regional climates of the present. We take the modern continental configuration as a representative arrangement of continents of various size and shape, scattered across a range of latitudes.

As potential refugia, in addition to nearly-enclosed seas such as the Red Sea and Mediterranean, we also seek locations on land where the local temperature exceeds the freezing point

of seawater at least once during the year; i.e. $T_{max} > -2°C$. With the right coastline geometry in a paleocontinental configuration (to allow for ocean-water access), these locations would be potential oases for photosynthetic eukaryotes.

Continental positioning itself is uncertain, and coastline geometry of the Neoproterozoic is even more uncertain. The period was tectonically active, and thus we make an explicit assumption that narrow bays are likely to have occurred and therefore seasonally above-freezing temperatures on land in our "modern Snowball Earth" allow for potential refugia even if these areas are not currently near a modern narrow bay. Additionally, many narrow bays would be smaller than the resolution of a typical global Earth System Model; thus, we focus on land surface temperatures as an indicator of possible refugia. To investigate whether life could survive the harshest conditions of the Snowball climate, in this paper we test the hypothesis that above-freezing land temperatures can exist in an Earth System Model in the "hard" Snowball Earth limit, in which even the tropical oceans are covered by ice hundreds of meters thick (no waterbelt).

The name "snowball" is somewhat misleading, in that the ocean was not entirely snow-covered. On the modern Earth, evaporation (E) exceeds precipitation (P) over nearly half the ocean, mostly in the subtropics. A large region of negative P E would also have existed on the Snowball Earth, according to general circulation models, although the hydrological cycle was probably weakened by a factor of ~ 30 (Pollard and Kasting, 2004). At high and middle latitudes the sea glaciers would have been covered by thick snow. But as sea glaciers flowed equatorward into the tropical region of net sublimation, their surface snow (albedo ~ 0.8) would sublimate away, exposing old snow ("firn", albedo ~ 0.7). Then the firn would likewise sublimate away, exposing bare glacier ice (albedo ~ 0.6) to the atmosphere and to solar radiation. These albedos were measured on modern surrogates in the Allan Hills of East Antarctica: firn and glacier ice exposed by sublimation, which have never experienced melting (Figure 2.1 and Table 2.1) (Dadic et al., 2013).

In our modeling we do not attempt to simulate the regional evolution of ocean surfaces from snow, through firn, to glacier ice, as the sea glaciers flow equatorward. Instead, everywhere that the ocean is not snow-covered, we assign its albedo to that of firn, thus biasing the global climate to a cold extreme and exaggerating the difficulty of maintaining

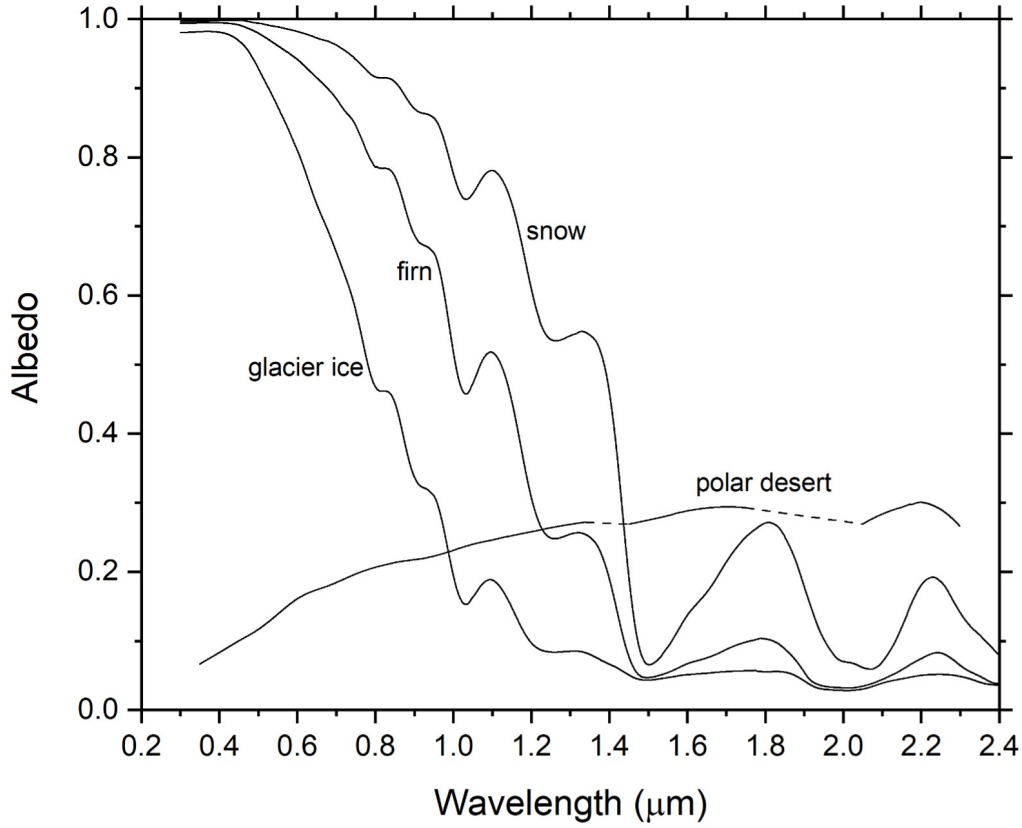


Figure 2.1: Spectral albedos of representative surface types. Cold fine-grained snow was measured at Dome C on the East Antarctic Plateau (Figure 6 of Hudson et al., 2006). Firn was measured just upstream of the Allan Hills blue-ice field in East Antarctica (Site R9 of Dadic et al., 2013). Glacier ice is from the Allan Hills blue-ice field (Site R1 of Dadic et al., 2013). These firn and ice sites can represent sea-glacier surfaces on Snowball Earth because they were originally formed by snow accumulation and exposed by sublimation, never having experienced melting. The "polar desert" site is an unvegetated surface of soil and stones in northeast Greenland (photograph shown in Figure 6 of Bøggild et al., 2010). For the "polar desert" surface, albedo measurements were not possible from 1.35 to 1.45 μm, and from 1.75 to 2.05 μm, because the incident solar radiation flux was near zero at these wavelengths due to atmospheric water-vapor absorption; the dashed lines interpolate across these regions.

Surface type	Albedo			Reference
	0.3-0.7 μm (UV, Visible)	0.7-3.0 μm (near-IR)	0.3-3.0 μm (total solar; broadband	
Snow	0.98	0.68	0.83	Hudson et al. (2006), Grenfell et al. (1994)
Firn	0.94	0.44	0.69	Dadic et al. (2013)
Glacier ice	0.89	0.26	0.58	Dadic et al. (2013)
Polar desert (gravel and soil of northeast Greenland)	0.13	0.21	0.17	Bøggild et al. (2010)
Sand desert (Arabia)			0.40	Smith (1986)
Stony desert (Gobi)			0.10-0.15	Abell et al. (2020a,b)

Table 2.1: Band-albedos of representative snow, ice, and land surfaces. The spectral albedos for four of these surface types are shown in Figure 2.1.

refugia. We then investigate how the albedo of bare land surfaces, and the atmospheric CO_2 level, influence the potential habitability of a snowball climate. We test a range of CO_2 levels from 10 to 200 ppm and a range of uniform surface albedos for bare land from 0.2 to 0.4, using a climate model with the modern continental configuration.

During a snowball event, volcanic CO_2 accumulates in the atmosphere because its removal mechanisms (dissolving in rainwater and reacting with surface rocks) are suppressed. As the climate warmed with rising atmospheric CO_2 during the progression of a snowball event, refugia would have become more widespread. In addition, wind-erosion of bare land would have lifted dust that could accumulate on the ice, lowering its albedo, leading to additional potential mechanisms for refugia. Here we instead focus on the most extreme bottleneck of the cold early phase of a snowball event, the most critical time for survival of surface life.

2.3 Methods

2.3.1 Experimental Design

We use a modified version of the Community Earth System Model, version 2.1.0 (Danabasoglu et al., 2020), with the Simple Land Interface Model (SLIM) (Laguë et al., 2019) coupled to the Community Atmosphere Model, version 4 (CAM4) (Neale et al., 2010), the Los Alamos sea ice model CICE5 (Hunke et al., 2015) in its thermodynamic-only mode (Bitz and Lipscomb, 1999), and a slab ocean model (SOM) (Bitz et al., 2012). Ocean heat flux convergence is set to zero everywhere, and sea surface temperatures are allowed to evolve. Simulations are run at a nominal 2° resolution.

SLIM is an idealized land model, designed for assessing the interactive roles of discrete land properties (e.g. bare-ground albedo, evaporative resistance, heat capacity of the soil, etc.); using it allows us to directly assess the climate response of specified land-surface albedos. Heat diffusion in the model is represented on a vertical soil grid that is separate from the water budget. Hydrology is represented using a simple bucket model that combines a user-specified "lid" resistance with a resistance related to the fill level of the water bucket. Snow can accumulate on the surface, and can be removed by sublimation to the atmosphere

or melting into the land. The model solves a linearized surface energy budget to calculate surface temperature, surface fluxes of radiation, turbulent heat fluxes, and ground heat storage.

When snow falls on the surface of land or sea ice, it masks the albedo of the underlying surface. On land, snow masks the albedo of bare ground when it exceeds a mass of 10 kg/m2 liquid-equivalent (about 3 to 10 cm of snow, for typical snow densities 0.1-0.3 g cm-3). Land-surface models for the modern Earth normally use a larger snow-masking depth because of the presence of grass, bushes, and trees, but these plants did not appear until long after the Neoproterozoic snowball events.

In the midlatitudes and polar regions of a snowball climate, kilometer-thick ice sheets covering the oceans would accumulate, thicken, and flow like modern ice shelves towards thinner regions of net-sublimation as "sea glaciers" (Goodman and Pierrehumbert, 2003). It would therefore be glacier ice, not sea ice, that would cover the ocean surface in a snowball climate after the initial global freezing had taken place, and would have albedos ranging from that of snow in areas of snow accumulation to exposed firn and finally bare glacier ice in regions of net sublimation (Figure 2.1 and Table 2.1). Rather than predicting the detailed state of ice surface conditions, which is typical in CICE5, for simplicity we revert back to the CCSM3 shortwave radiative transfer formulation. This option allows us to prescribe "sea-ice" surface albedos. For bare (snow-free) ice we set visible and near-infrared band albedos to values appropriate for firn, so as to bias the global climate to its cold extreme. As snow accumulates, snow masks the bare ice, and the band albedos transition to values appropriate for snow (Table 2.1).

In all simulations, sea ice is initialized with 100% concentration and 20 m thick in all ocean gridcells. Sea ice rapidly grows thicker, but would take thousands of years to reach an equilibrium. We do not expect sea ice ever to reach an equilibrium thickness in our simulations even if we extended them, since geothermal heating is not represented; that heat source would be necessary to limit the freezing of seawater to the base of thick ice (McKay, 2000; Warren et al., 2002; Goodman and Pierrehumbert, 2003). Our surface temperatures could be seen as too warm. For example, if the model ice thickness is only 50 m but in equilibrium would be $>$ 500 m, there is an excessive conductive heat flux of 1.4

Wm^{-2} upward through the ice, causing the surface temperature to be too high by ∼0.5 K.

Over Earth's history, the Sun has brightened by about 1% every 100 million years, so at 600 Ma the solar constant was ∼94% of its present value (Crowley and Baum, 1993). That value, 94%, has been used to initiate the snowball state in models with Neoproterozoic continents clustered at low latitude. The low-latitude land facilitates snowball initiation in those models, because the albedo of bare land (0.2-0.4) exceeds that of open ocean (0.07). For a "modern" Snowball Earth, with most of the continental area at middle or high latitude, a lower solar constant, about 91%, is needed to initiate the snowball (Voigt and Marotzke, 2010), and that is the value we use for this work.

In models, the critical CO_2 level required for snowball onset depends on several modeling choices: sea-ice dynamics (Lewis and Eby, 2006; Voigt and Abbot, 2012), land topography (Liu et al., 2018), continental configuration (Liu et al., 2013), mountains (Walsh et al., 2019), atmospheric dust (Liu et al., 2020, 2021), and cloud radiative forcing (Voigt and Marotzke, 2010). In prior work, modeled snowball climates have been initiated at CO_2 mixing ratios as low as 2 ppm (Voigt and Abbot, 2012) and as high as 600 ppm (Liu et al., 2017) but generally fall between 50 and 300 ppm, with exact values dependent on the ice coverage on sea and land, the solar constant, and the land area (Schrag et al., 2002; Yang et al., 2012). For the coldest early stage of a snowball event, we therefore specify a variety of CO_2 levels, along with several choices for the albedo of snow-free land, and examine the resulting climatic patterns.

For initiation of the snowball state, we set the CO_2 mixing ratio to 100 ppm, and the albedo of snow-free land to a broadband value of 0.4. The albedo of bare land (rocks or soil) increases with wavelength across the solar spectrum from the ultraviolet (UV) to the infrared (IR) (Figure 2.1). We specify the albedo in two bands: 0.3 in the UV and visible (wavelengths 0.3-0.7 μm), and 0.5 in the near-IR (wavelengths 0.7-5.0 μm). Under these conditions, our initial simulation establishes the conditions necessary to maintain a frozen ocean in approximately 20 years. We run the initiation simulation for a total of 100 years to ensure that the frozen-ocean state is not transient, and that the atmosphere is in steady state. Based on these conditions sufficient to generate a snowball climate, further runs are initialized with a fully ice-covered ocean to test the sensitivity of surface temperatures to

variations in bare-land surface albedo and atmospheric CO_2 concentration (details below).

We run the model at the global scale and thus do not resolve the fine-scale dynamics of a sea glacier invading a bay. As described above, we assume that if a land gridcell experiences a monthly average surface temperature that exceeds the melting point, that gridcell could potentially support liquid water if a narrow arm of the sea were to reach it in a paleocontinental configuration, thus rendering it a potential refugium for photosynthetic life.

2.3.2 Sampling across a range of atmospheric CO_2 concentrations and land albedos

In the absence of land plants, which break up stones into sand or silt with higher albedo, the Neoproterozoic land surface was probably darker than modern deserts (the fossil record suggests that land plants did not evolve until 461–472 Ma (Kenrick et al., 2012; Morris et al., 2018)). The highest broadband albedo for modern deserts is 0.40 for the fine sand of the Arabian Empty Quarter (Smith, 1986); the lowest albedo is 0.10-0.15 for the stony desert of the western Gobi (Abell et al., 2020a; Figure S1 of Abell et al., 2020b). Within this range, we test five different sets of snow-free land albedo values, listed here from brightest to darkest as [UV-visible albedo/near-IR albedo]: 0.3/0.5, 0.25/0.45, 0.2/0.4, 0.15/0.35, and 0.1/0.3. Approximately half the solar energy is in the near-IR, so the total solar (broadband) albedo for each case is the average of the two values given; i.e., 0.40, 0.35, 0.30, 0.25, 0.20. As a shorthand to identify the cases, we simply give the broadband values.

We sample CO_2 levels of 10, 25, 50, 100, and 200 ppm. We do not test every combination of albedo and CO_2 values, but we do test the edge cases, as well as the full range of albedos for a 50 ppm CO_2 atmosphere, and the full range of CO_2 values for land albedo 0.3 (Table 2.2).

Albedo of bare (snow-free) land		CO_2 (ppm)	Global average planetary albedo	Global mean surface temperature (°C)	Radiative forcing (Wm^{-2}) relative to 10 ppm, albedo 0.4	Land area with $T_{max} > -2°C$ in warmest month (% of total)	Land area with PE>P in annual mean (% of total)	Land area with PE>P and $T_{max} > -2°C$ (% of total)	Ocean area with $T_{max} > -2°C$ in warmest month (% of total)
Broadband solar	Visible/ near-IR								
0.4	0.3/0.5	10	0.683	-69.2	0	0.1	50.8	0.1	0
0.4	0.3/0.5	50	0.681	-67.9	2.75	0.23	54.16	0.23	0.01
0.4	0.3/0.5	200	0.68	-66.5	5.16	0.54	59.83	0.54	0.04
0.35	0.25/0.45	50	0.675	-66.6	3.78	1.41	57.67	1.41	0.11
0.3	0.2/0.4	10	0.67	-66.7	2.05	3.1	56.77	3.1	0.24
0.3	0.2/0.4	25	0.669	-66.0	3.6	3.48	59.9	3.45	0.31
0.3	0.2/0.4	50	0.668	-65.3	4.8	4.04	63.14	4.04	0.36
0.3	0.2/0.4	100	0.667	-64.6	5.99	5.27	66.32	5.27	0.49
0.3	0.2/0.4	200	0.667	-63.9	7.21	5.72	70.4	5.7	0.40
0.25	0.15/0.35	50	0.662	-64.0	5.81	9.15	68.16	9.02	0.87
0.2	0.1/0.3	10	0.657	-64.2	4.06	12.15	67.87	12.02	1.17
0.2	0.1/0.3	50	0.655	-62.7	6.81	14.99	74.57	14.82	1.53
0.2	0.1/0.3	200	0.653	-61.2	9.22	17.22	85.03	17.05	1.83

Table 2.2: Characteristics of the model runs. The consequences of radiative forcing shown for combinations of changes to the bare-land albedo and the CO_2 level are relative to the coldest case of CO_2 = 10 ppm and bare-land albedo = 0.4 broadband (0.3 visible / 0.5 near-IR). Land is "net-evaporative" if potential evaporation (PE) exceeds precipitation (P).

We use the Parallel Offline Radiation Tool (PORT) (Conley et al., 2013) to calculate total global radiative forcing associated with each experiment relative to a base case in the middle range at CO_2 = 50 ppm and bare (snow-free) land albedo 0.3. [In Table 2.2 the radiative forcings are shown instead relative to the coldest case (CO_2 = 10 ppm, albedo = 0.4) for ease of comparison.] Table 2.2 shows that dropping the albedo of bare land from the brightest case (0.4) to the darkest case (0.2) at 50 ppm causes a radiative forcing (RF) of 4.06 Wm^{-2}, resulting in a 5.2 K increase in global mean surface temperature, implying a climate sensitivity of 1.28 $K/(Wm^{-2})$. Increasing CO_2 from 10 ppm to 200 ppm (with land albedo 0.3) causes RF = 5.16 Wm^{-2}, and results in a temperature increase of 2.8 K, implying a lower climate sensitivity of 0.54 $K/(Wm^{-2})$. These climate sensitivities may be compared to a median of 0.5 $K/(Wm^{-2})$ in 19 GCMs for the modern Earth (Cess et al., 1990).

These climate sensitivities indicate that albedo-driven forcing kicks off stronger feedbacks than CO_2-driven forcing. The snowball climate has been shown to be relatively insensitive to CO_2-driven forcing; at such low temperatures, the positive feedback from the water-vapor greenhouse effect is weak (Pierrehumbert, 2005). Outside of the tropics, the wintertime greenhouse effect is negative, resembling the modern Antarctic plateau (Sejas et al., 2018). The greenhouse effect can be calculated as $G = \sigma {T_S}^4 - OLR$, where σ is the Stefan-Boltzmann constant, and OLR is outgoing longwave radiation. Surface temperature change due greenhouse warming is $\Delta T_g = [T_S - \frac{OLR}{\sigma}]^{1/4}$. At 50 ppm, G ranges from 0.7 Wm^{-2} to 2.6 Wm^{-2}, and ΔT_g, is between 0.9 and 2.3 K (higher warming at lower bare-land albedo). At 200 ppm and 0.2 albedo, G is 4.4 W m^{-2}, and ΔT_g reaches 3.5 K. Pierrehumbert (2005) obtained a similarly small value for the global average on a hard snowball; his Figure 4 shows a clear-sky greenhouse effect of $\sim$ 8 Wm^{-2}. These snowball values are much smaller than those for the modern Earth, where $G \approx 150$ W m^{-2} and $\Delta T_g = 33$ K.

2.4 Results

2.4.1 No land surface has above-freezing mean-annual temperatures, but some areas of the land reach above-freezing temperatures seasonally.

In our runs with lower bare-land albedo, we find no areas of land that are above-freezing on annual average. Places with annual mean temperatures above freezing would have the potential of "open water" refugia. But refugia do not require open water. If the end of an oceanic bay is below freezing on annual average, but the warmest month is above freezing, sea ice would form in the bay, and it would partially melt in summer. Neoproterozoic algae could have survived in the temporary meltwater pools on the ice, as has been observed by mat-forming eukaryotic algae on the McMurdo Ice Shelf and on the Ward Hunt Ice Shelf in the Canadian Arctic (Vincent et al., 2000; Vincent and Howard-Williams, 2000). In these modern analogs, organisms can survive perennially in ice that is deeply frozen for all but a few weeks or days per year. This would be the "ice surface" refugium described above in Section 1.1(d).

We find that in our coldest case (10 ppm CO_2, bare-land albedo 0.4), 0.1% of the land surface area reaches temperatures above the freezing temperature of seawater (-2°C) in at least one month of the year, despite a global mean surface temperature of -69°C (Figure 2.2a). In our warmest case (200 ppm CO_2, bare-land albedo 0.2) in which global mean surface temperature is -61°C, 17% of land surface area reaches temperatures warm enough to host liquid water in the warmest month. (Our specification of the freezing temperature as -2°C is a conservative choice. With perhaps 20% of the ocean water converted to land glaciers and sea glaciers, the salinity of the remaining seawater would increase, lowering its freezing temperature closer to -3°C.)

To investigate how refugia might form, we calculate the potential evaporation (PE, mm/day), a measure of the rate at which the atmosphere could evaporate or sublimate water from the surface, if that surface had unlimited water availability. Over land, we calculate PE using a modified version of the Penman-Monteith equation (Penman, 1948; Monteith, 1981; Scheff and Frierson, 2014). Over the ocean, it is equal to the latent heat flux from the ocean to the atmosphere converted to units of water flux (mm/day). The dry

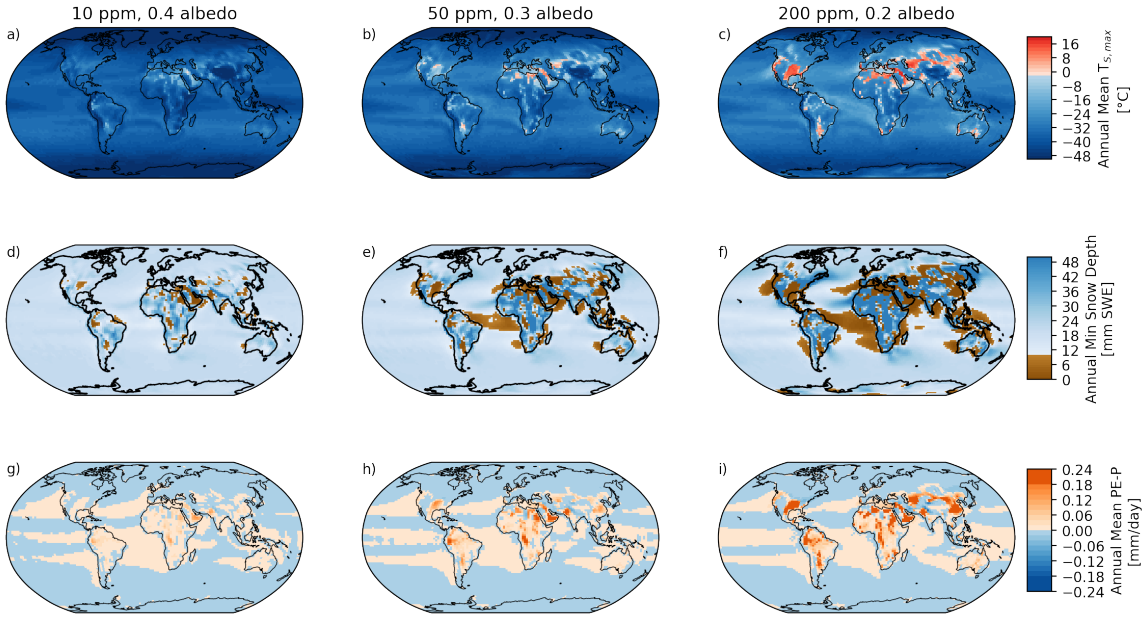


Figure 2.2: Habitability of the simulated snowball climate. **(a, b, c)** Surface temperature of the warmest month in the coldest simulated case (10 ppm CO_2, bare-land albedo 0.4), the midpoint simulation (50 ppm CO_2, bare-land albedo 0.3) and the warmest case (200 ppm CO2, bare-land albedo 0.1), respectively. Red areas indicate locations with seasonal melting, suggesting the possibility of ice-surface refugia. **(d, e, f)** Annual minimum snow depth (monthly mean) from the same cases, given as mm snow water equivalent (SWE), which is equivalent to kg/m^2. Brown areas indicate places where snow depth drops below the threshold for surface-albedo change at some point during the year. Note that ocean areas without snow accumulation would nonetheless be covered by sea glaciers. **(g, h, i)** Annual mean potential evaporation (PE) minus precipitation (P) for the same cases.

snowball atmosphere over land creates demand for water resulting in large areas of the land surface that are snow-free for part of the year (Figure 2.2d-f). Above-freezing land surfaces are concentrated in places where PE outpaces precipitation – in particular, parts of the Arabian Peninsula, the modern Sahara Desert and eastern Asia (compare Figures 2.2a-c and 2.2g-i). Without snow cover, low-albedo land absorbs more solar radiation, allowing for above-freezing local land temperatures and the potential for unfrozen water and refugia if an ocean bay were to intrude to those locations.

A net-evaporative location experiencing mean-annual temperatures above freezing may thus serve as a refugium if it is connected via a narrow bay to the ocean, allowing seawater from below the sublimating sea glacier to flow into the bay and replace the water lost by evaporation.

Our model does not represent the growth of ice sheets on land. However, we can infer that snow-covered regions in Figure 2.2d-f are where ice sheets would grow; where they would flow would depend on the land's topography. Ice sheets have been directly simulated on Neoproterozoic continents (Donnadieu et al., 2003; Mitchell et al., 2015); they cover only parts of the continents, allowing large regions to be ice-free land.

2.4.2 Land surface albedo exerts control on the habitability of nearly-enclosed bays.

Land surface temperatures increase more strongly in response to decreases in bare-land surface albedo than to increases in CO_2, for the same magnitude of global radiative forcing. This response occurs across all latitudes and is stronger in regions with more total land area (Figure 3.3a). Starting from a baseline of albedo 0.4 and 10 ppm CO_2, increasing CO_2 from 10 to 200 ppm constitutes a global radiative forcing of 5.16 Wm^{-2}, while decreasing snow-free land albedo from 0.4 to 0.2 constitutes a smaller forcing of 4.06 Wm^{-2} yet causes a greater increase of warm land area. The change of land area per unit of radiative forcing is shown in Figure 3.3b, for four cases of similar RF, two caused by increasing CO_2 and the others by decreasing albedo. Despite smaller globally mean radiative forcing from albedo changes, land surface temperatures are more responsive to albedo since the radiative forcing is concentrated over snow-free land.

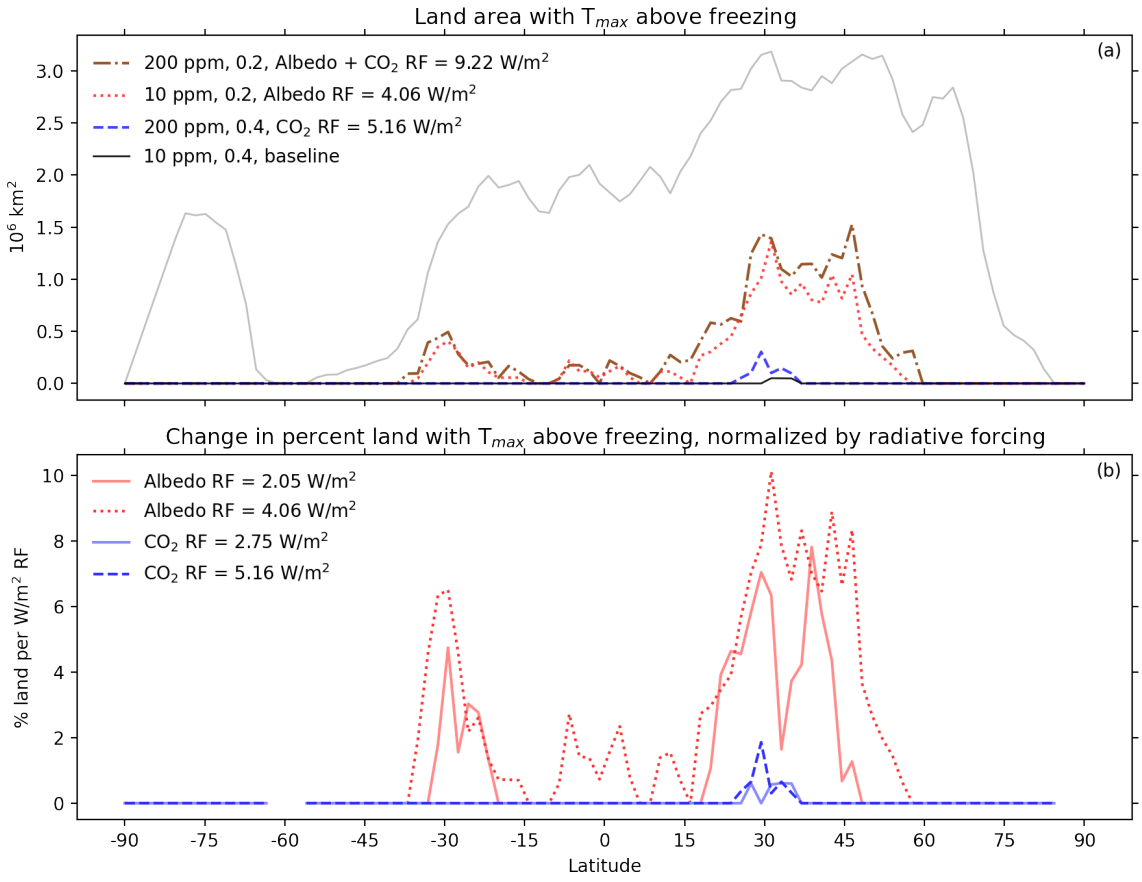


Figure 2.3: **(a)** Area of land (per 1.9 degrees of latitude increment) with temperature above freezing in the warmest month. Total land area is shown in grey. **(b)** Change in percent land area above freezing in the warmest month, per unit radiative forcing, relative to the coldest case (10 ppm CO_2, albedo 0.4).

Decreasing bare-land albedo facilitates potential refugia in two ways: (1) warming of bare land and (2) exposing new bare land that then becomes warm. In our simulations, both mechanisms occur to effect a change in habitability, with approximately 45% of the newly exposed land above freezing (in the warmest month) having become warm enough to host refugia through the first mechanism, and 55% through the second mechanism at constant CO_2. (Partitioning is similar at 10, 50, and 200 ppm CO_2). We do not see above-freezing temperatures in locations that have snow cover. If a warm, net-precipitating location did exist, it would become a warm ice sheet, like a modern temperate glacier or the wet-snow zone of modern Greenland.

Either decreasing bare-land albedo or raising CO_2 expands the area of potentially habitable land in coastal gridcells; continental interiors already meet the criteria for potential habitability at albedo 0.4. A world with low bare-land albedo would therefore be more likely to host life in narrow bays that intrude into the dark land. As mentioned above, the bare land in the Neoproterozoic probably resembled modern stony deserts rather than sand or soil, so its albedo may have been even lower than the lowest case we modeled (broadband albedo 0.2).

Figure 2.4 and Table 2.2 show, for the various combinations of CO_2 and albedo, the percent of land area capable of hosting refugia, were an arm of the sea to reach it, demonstrating again the relative importance of global radiative forcing by land albedo and CO_2. Starting from the coldest case (bare-land albedo 0.4, CO_2=10 ppm), positive radiative forcing results from either increasing CO_2 or reducing land albedo. For the same radiative forcing, a change of land albedo is more effective than a change of CO_2. Figure 2.4a shows this for the annual maximum temperature (T_{max}); the area of above-freezing land ranges from 0.1% to 12%, by decreasing bare-land albedo alone. [Even in the warmest case (bare-land albedo 0.2, CO_2=200 ppm), we do not see places with annual mean temperature $\overline{T} > -2$°C. That would allow for "open-water" refugia, because during at least part of the year the bay would be ice-free.]

A temperature criterion is not sufficient. We also need PE > P so that land glaciers will not form at these locations. Figure 2.4b shows that the percent of land area with PE > P increases slightly with darkening of the land or increasing CO_2. Combining these criteria,

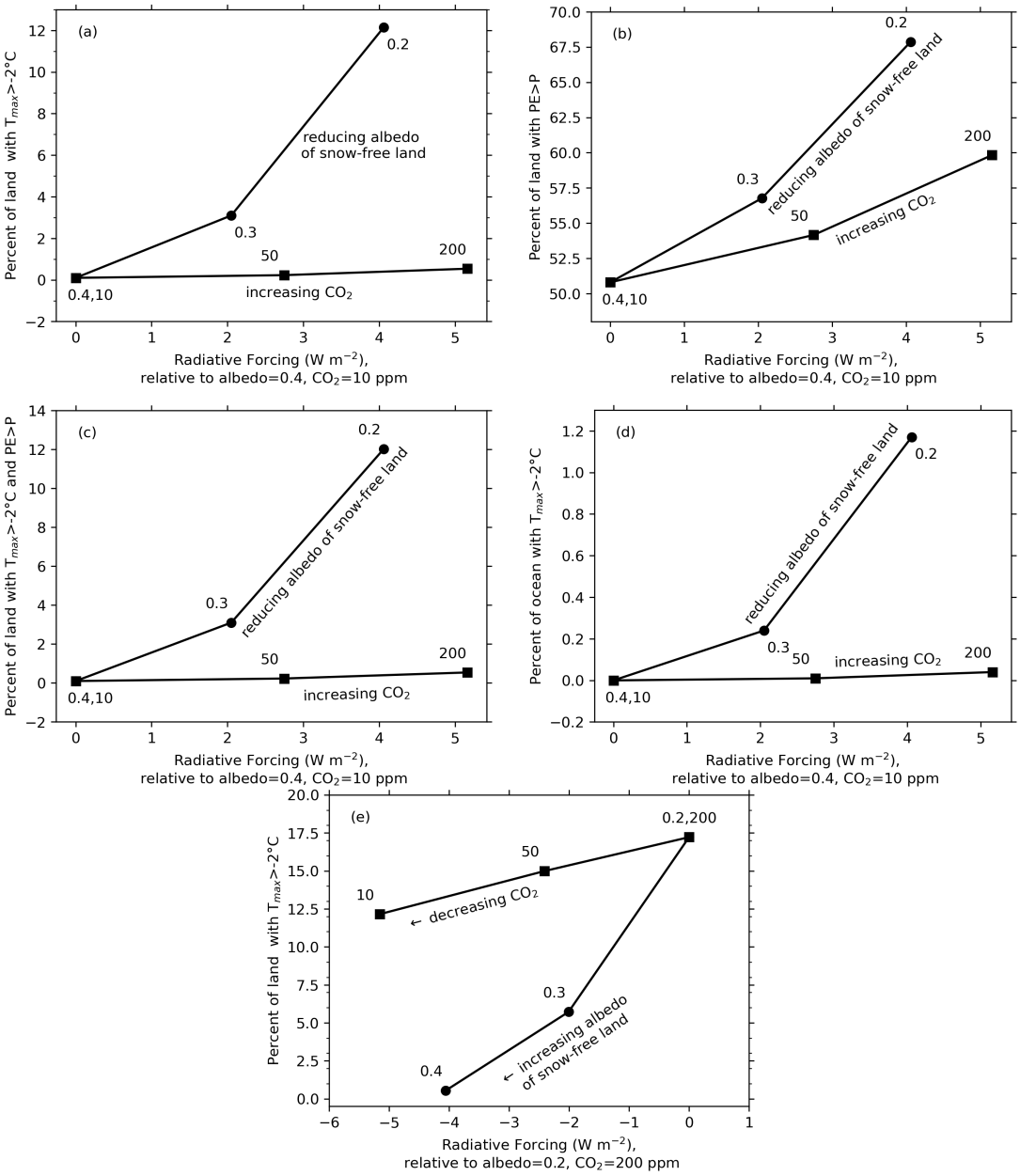


Figure 2.4: Percent of land area capable of hosting refugia, were an arm of the sea to reach it, for several combinations of CO_2 mixing ratio and bare-land albedo. Relative to the coldest case (albedo=0.4 and CO_2=10 ppm), the change in suitable land area is shown as a function of the radiative forcing, caused either by darkening the surface or by increasing CO_2. **(a)** Percent of land area with temperature of the warmest month $T_{max} > -2$°C. **(b)** Percent of land area with mean annual PE¿P. **(c)** Percent of land area with $T_{max} > -2$°C and PE>P. **(d)** Percent of *ocean* area with temperature of the warmest month $T_{max} > -2$°C. **(e)** Percent of land area with temperature of the warmest month $T_{max} > -2$°C relative to the warmest case (albedo=0.2 and CO_2=200 ppm).

Figure 2.4c shows the percent of land area with $T_{max} > -2$°C and PE > P.

There are also some small areas of the ocean, all on coastlines, where the surface temperature exceeds −2°C in the warmest month, but in all cases these areas represent less than 2% of the ocean area (Table 2.2, Figure 2.4d).

With high-albedo land, the climate is so cold that very little of the land reaches above freezing even with 200 ppm CO_2, as shown in Figure 2.4a. But with the warmer climate for darker bare land (albedo 0.2), the above-freezing land area does become sensitive to the CO_2 level (Figure 2.4e).

2.4.3 Refugia on desert land?

Besides inhabiting the ends of narrow oceanic bays, photosynthetic eukaryotes may also have been active on unglaciated land surfaces. Reviewing "early life on land", Lenton and Daines (2017) emphasized microbial mats powered by oxygenic photosynthesis. In their words: "Initially, such mats would have been dominated by [prokaryotic] cyanobacteria. Sometime during the Proterozoic Eon (2.5-0.54 Ga) they probably gained eukaryotic algae and fungi... Today a mixture of cyanobacteria, algae, fungi, lichens and nonvascular plants are found in terrestrial mats, often termed 'biological soil crusts' or 'cryptogamic cover'." Lenton and Daines cited evidence that "by the start of the Neoproterozoic (1 Ga), eukaryotes were probably present alongside cyanobacteria in terrestrial mats, but whether these were algae is unclear." The soil-crust mats occur on modern midlatitude deserts that are seasonally above freezing, as in Utah and Nevada. Similar environments may therefore have offered a habitat for mixed prokaryotic/eukaryotic life in the deserts of Snowball Earth wherever the soil temperatures were above freezing seasonally, as also proposed by Retallack (2023). The locations would have to be in deserts (PE>P) to avoid burial by land-glaciers, but they could have received water from the local sparse precipitation, or from runoff modulated through topography.

2.5 Discussion and Conclusions

In our simulations we used the albedo of firn rather than glacier ice (Table 2.1) to represent the snow-free parts of the frozen ocean, biasing the climate colder. Actual sea glaciers

should have had a slightly darker surface, allowing for a warmer climate. Yet, despite our conservative choice of a brighter ice surface, our results suggest that a "hard" snowball climate with ice extending to the equator could have allowed some locations to sustain the surface liquid water needed to host photosynthetic life, despite extremely cold global-mean temperatures. Our global annual mean surface temperatures $\overline{T}$ are considerably colder than those of other GCMs simulating the hard-snowball climate. Abbot et al. (2013) reviewed six GCMs; we can estimate $\overline{T}$ from their Figure 2.1a for 100 ppm CO_2: for five GCMs, $\overline{T} \approx -38$°C; the cold outlier (the FOAM GCM) had $\overline{T} \approx 46$°C. Our finding of above-freezing locations even with our extremely cold global mean surface temperatures, all in the range -61 to -69°C (Table 2.1), thus argues strongly for refugia on or near ocean bays.

Although we find strong evidence to support the potential for refugia, the distribution and type of refugia (ice-surface or open-water) would be sensitive to the actual bare-land albedo. The albedo of bare land surfaces during the Neoproterozoic may have been even darker than our darkest case (albedo 0.2), as land plants had not yet evolved, which would limit the erosion of rocks into smaller grain sizes, so that stony deserts, which have broad-band albedo 0.10-0.15, would be more likely than modern deserts of soil or sand. If Neoproterozoic land surfaces were indeed dark like stony deserts, this lower land albedo would result in more refugia than we have simulated here.

We find that modern nearly-enclosed bays, resulting from continental rifting (e.g., the Red Sea), are especially habitable and could support seasonal refugia depending on the land surface albedo. Since the dynamics of sea-glacier invasions into nearly-enclosed bays occur below the resolution of the model simulations employed here (Campbell et al., 2011, 2014), our results quantify the maximum potential for refugia within a hard-snowball climate with modern continents; but the exact distribution of such nearly-enclosed bays on Neoproterozoic continents would determine true habitability. A constriction at the entrance to the bay helps to slow the sea glacier (Campbell et al., 2014), but the entrance must not be too shallow because ocean water needs a path below the ice to reach the refugium. The strait at the entrance to the Red Sea (Bab el Mandeb) is only 137 m deep (Siddall et al., 2002), so a sea glacier would likely become grounded there.

Since Neoproterozoic continental reconstructions are constrained primarily by paleomag-

netism, which constrains the latitude but not the longitude, there remains large uncertainty in the likelihood of large continental interiors, which we expect to strongly influence our results (Merdith et al., 2021). The total land area was probably only slightly smaller than today's (Hawkesworth et al., 2019), but the tropical bias in land distribution and degree of continental "clumping", as well as the location and height of mountain ranges, could influence our results (Laguë et al., 2023).

Our conclusion is that the ends of narrow oceanic bays were likely to serve as refugia for photosynthetic eukaryotes, even during the coldest early phase of a snowball event.

Chapter 3

IDENTIFYING THE INFLUENCE OF CARBON CYCLE REPRESENTATION ON SINK SENSITIVITY, TOTAL WARMING, AND GLOBAL CARBON BUDGETS IN SIMPLE CLIMATE MODELS

Abstract

Complex models of the Earth system are increasingly able to represent details of the carbon cycle that determine how much carbon is taken up by the Earth's land and ocean, but a variety of approaches exist to simplify the representation of these processes in reduced-complexity models, or simple climate models (SCMs). Such models are useful for building conceptual understanding as well as informing policy decisions intended to mitigate global warming. We perform a comparison of four SCMs with varying carbon cycle structures in the historical period and emulate global mean temperature response to idealized emissions trajectories. We find that carbon cycle representation influences the magnitude and uncertainty of land carbon uptake in decarbonization scenarios. These findings suggest that using one model in isolation to guide decarbonization decision-making may hide large uncertainties in how land and ocean sinks may behave in a decarbonzation or post-net-zero emissions landscape.

3.1 Introduction

3.1.1 Quantifying structural uncertainty in the coupled Earth systems

Identifying and quantifying uncertainty in the coupled climate-society system is important for understanding the total outcome of any human policy, as well as for developing climate-oriented benchmarks, such as the remaining carbon budget (Rogelj et al., 2019).

The Earth's carbon cycle contains many important feedback processes that together influence how the Earth's climate responds to increased or decreased net emissions of carbon dioxide (Friedlingstein et al., 2014; Arora et al., 2020).

In Earth System Models (ESMs) biogeochemical processes on land contribute significant uncertainty to the global climate response to emissions (Friedlingstein et al., 2022). Modeling the coupled carbon-water cycle on land requires representing complex biological and ecological processes occurring on small scales (plant stomata open and close on the micrometer scale) by using poorly-constrained parametrizations operating at resolutions that are at least 10 orders of magnitude coarser.

The challenge of how to represent processes within the terrestrial carbon cycle is not likely to be resolved soon since it reflects an incomplete understanding of how to represent them at global scales, and the number of parametrizations of different biological processes is more likely to increase than stabilize or decrease. Wieder et al. (2015) illustrate this by comparing two equally plausible formulations for biological nitrogen fixation (one calculated from net primary productivity and one calculated from evapotranspiration) in CESM2. They find that while both formulations lead to similar carbon fluxes and pools in mean stead-state pre-industrial pools, they diverge under a high-emissions scenario (RCP8.5) by about 30% in the northern hemisphere.

While these differences in carbon cycle representation can lead to different answers, having a diversity of ways of representing processes unveils the uncertainty stemming from incomplete representation. Comparing outcomes across different model architectures enables us to demarcate how much of our total uncertainty is due to this source.

3.1.2 Simple Climate Models

While Earth system models are capable of representing details of the climate system, the computational burden prevents them from being used for assessing a wide range of pathways. Simple climate models (SCMs) are used by a range of actors from scientists for intuition, global leaders to vet policy, and companies to scope climate-oriented projects. They are even used in the IPCC WG3 to scope climate solutions.

SCMs can be built to emulate ESMs. While ESMs may contain representations of carbon cycle processes that are more mechanistic and rely on empirical or mechanistic representations of observable mechanisms, leading to deterministic simulations, the structure of SCMs

varies.

One benefit of SCMs is their computational efficiency. Because of their reduced complexity, SCMs can simulate thousands of scenarios on one CPU in a matter of seconds or minutes, allowing users to sample across both internal variability and scenario uncertainty with ease. This usability make them far more attractive and ubiquitous tools for policy- or technology-scoping use cases.

Since CMIP-class ESMs are too computationally expensive to explore the full parameter space of structural uncertainties and future scenarios, we can use reduced complexity or "simple" models, which emulate CMIP-class GCMs to test the range of temperature response we get from decreasing CO_2 emissions.

Emissions-driven climate model emulators are physically based emulators that can estimate ESM dynamics and offer insight into dominant carbon cycle processes. They are used to assess and plan decarbonization pathways, scope CMIP scenarios for ESMs, emulate the climate for economic models, and offer insight into the dominant carbon cycle processes at the global scale.

For this reason, they currently play a powerful role in scenario development and in particular for Working Group III Mitigation Assessment of the IPCC Working Group II (2022), a set of three climate model emulators were used to vet and classify a large database of scenarios generated from economic models (Kikstra et al., 2022). The emulators were used to classify mitigation scenarios into different warming categories, with those limiting warming to below 2°C being the most desirable. In the report, pathways that are able to limit warming to 2°C (with high likelihood), projected CO_2 emissions to be reduced between 2019 and 2050 by around 49% for energy demand, 97% for energy supply, and 136% for agriculture, forestry, and other land use. The implication is that successful mitigation strategies entail dramatic modifications to the existing land carbon sink. While three models are run, FaIRv1.6.2, MAGICCv7.5.3, and CICERO-SCM, only MAGICC's results are reported in the final assessment since they capture the range of the other models.

At the international scale, multiple models are run to characterize potential mitigation scenarios. At marginally smaller scales, it is common to run only one emulator, varying parameters to emulate ESMs to sample across uncertainty in the forced response to emis-

sions. As an example, the independent research and analysis think tank, Rhodium Group, which provides research insights for economic and policy decision-makers in the public and privates sectors, recently published its Rhodium Climate Outlook (RCO) report intended to provide an assessment of likely future climate risk given projected future global emissions. To translate global emissions to global temperature, the authors run FaIR. The very first conclusion of the report states: "The world is very likely on track to exceed 2°C above pre-industrial levels, but we've avoided the most catastrophic projections" (Larsen et al., 2023).

McKinsey & Company published a similar energy outlook report in January 2024: "Global Energy Perspective 2023: CO_2 emissions outlook" (Nivard et al., 2024). In it the authors conclude that projected emissions will likely cause global temperatures to exceed 1.5°C by 2035. To reach this conclusion, they employ MAGICC (v7.5.3) and have since provided recommendations for land-based carbon removal strategies that align with the emissions reductions targets set in a 2021 report: "Nature and Net Zero" (Adams et al., 2021).

As another example, CarbonPlan, an independent climate- and energy-focused research non-profit, investigated the comparative timescales of carbon removal that result from different carbon dioxide removal (CDR) technologies, specifically the difference between direct air capture (DAC) and direct ocean removal (DOR) (Zarakas et al., 2023). Their assessment informs policy-makers on the nuanced ways in which changes to CO_2 concentration depend on multiple processes coming to equilibrium, which take effect over overlapping timescales and gives recommendations for reconciling differences in these approaches. In their assessment, they run HECTOR, a reduced-complexity model which contains a process-based carbon cycle model to illustrate the comparative delayed in carbon removal that might be felt if DOR were employed versus DAC, and caution decision-makers to consider the influence of carbon cycle feedbacks fairly when comparing technologies (i.e., suggesting the use of emissions-driven assessments of technologies in models that represent the carbon cycle feedback from both technologies). This assessment exists as an ideal example of the explanatory power of SCMs. HECTOR's structural simplicity and consequent computational slightness make it an excellent tool for achieving insight into the varying timescales of the

global carbon cycle but also for facilitating the connection between those biogeochemical insights and decarbonization policy.

Taken a step further, a final example of the ubiquity of simple climate models and emulators in the decision-making space is the planned use of MESMER, a spatially-resolved climate model emulator, for EU policy assessment as part of the Horizon Europe SPARCCLE project (Beusch et al., 2020). MESMER has been built to replicate ESM climate impact variables, including fire weather, soil moisture, and annual maximum temperature, without the process-representation and incumbent computational cost. The project works in collaboration with the European Commission's Joint Research Centre, which provides direct scientific support for EU policy-making.

Since SCMs are built to produce the climate response to radiative forcing, they inherently include a representation of the carbon cycle. The SCMs vary in how implicit or explicit that representation is. An important obstacle in modeling any kind of solution, even with an ESM is that decarbonization emissions scenarios are necessarily new territory for our observed climate system and our climate models. The ubiquity of their use in modern climate decision-making, their diversity in structure, and the inherent uncertainty of an imminent untested emissions regime, together make SCMs important and exciting subjects for inter-comparison for the purposes of assessing structural uncertainty.

3.1.3 Metrics of climate sensitivity in a decarbonization regime: Transient Climate Response to Cumulative CO_2 Emissions (TCRE) and the Zero-Emissions Commitment (ZEC)

We use two metrics that can be used in conjunction with one another to quantify the remaining carbon budget for climate stabilization. The first is the Transient Climate Response to Cumulative CO_2 Emissions (TCRE) (MacDougall, 2016) and the second is the Zero-Emissions Commitment (ZEC). The TCRE can be defined as the modeled proportionality (or slope) of transient global mean warming to cumulative emissions. ESMs consistently exhibit this emergent response to cumulative emissions across a wide range of cumulative emissions pathways (Koven et al., 2022). In scenarios reaching net-zero (or net-negative)

emissions, some of the committed temperature response to emissions is realized after emissions reach zero, which is quantified as the ZEC, evaluated at some designated time after emissions cease. ZEC can be seen as a measure of the relative strength of lagged warming to lagged CO_2 uptake operating over longer timescales when emissions reach net-zero or net-negative (Koven et al., 2022). The use of the two metrics together allows for the quantification of a total CO_2 remaining carbon budget international warming threshold targets of 1.5 or 2°C (Rogelj et al., 2019; Jones and Friedlingstein, 2020).

We can use TCRE and ZEC to estimate a remaining carbon budget using the following inputs: (1) anthropogenic warming to date (2) cumulative historical CO_2 emissions and (3) current non-CO_2 fraction of total anthropogenic forcing

In theory, ZEC can be quantified as the temperature change taking place after net-zero, though in practice, when emissions decrease progressively (as opposed to falling to zero rapidly), some of the ZEC can be realized before emissions cease (Koven et al., 2023), meaning that ZEC may be more accurately quantified as the temperature change with respect to the expected linear TCRE.

Quantifying and understanding the underlying dynamics controlling ZEC is essential for predicting peak-CO_2-driven warming, the relative timing between peak emissions and peak warming, and the remaining carbon budgets (Koven et al., 2023).

3.1.4 Research Questions

1. How does the carbon cycle structure in an SCM influence the magnitude and timing of warming in an decarbonization scenarios?

2. How do different model structures allocate carbon to land and ocean sinks, and how do the allocations differ depending on emissions mode (increasing or decreasing)?

3.2 Methods

3.2.1 Models

We compare four models of varying structures, levels complexity, and applications. In Table 3.1, we describe the components of how the carbon cycle is represented in each model.

Model	Carbon cycle structure	Reference
FaIR	4 fixed-allocation timescales, $\alpha\tau$, of carbon removal.	Millar et al. (2017); Smith et al. (2018); Leach et al. (2021)
HECTOR	3-box terrestrial model with LUC 4-box ocean model	Hartin et al. (2015, 2016); Dorheim et al. (2023)
MAGICC	3-box terrestrial model Modified impulse response ocean model	Meinshausen et al. (2011)
FaIR_bgc Box Model	9-box terrestrial model 7-box ocean model	Parton et al. (1987, 1988); Swann (2010) Toggweiler (1999)

Table 3.1: Carbon Cycle Approach by Model

In Figure 3.1 each model's carbon cycle structure is portrayed as a schematic diagram, illustrating the variation in structure that we explore.

FaIR

At the lowest level of complexity, we run the Finite-amplitude Impulse Response Model, version 2.1.3, or FaIR (Leach et al., 2021; Smith et al., 2018; Millar et al., 2017). FaIR's carbon cycle is a modified version of the four-timescale impulse response function for carbon dioxide derived by Joos et al. (2013). In practice, FaIR's carbon cycle is composed of four reservoirs with unique timescales of carbon removal from the atmosphere, $\alpha\tau_i$, such that τ_i is the characteristic atmospheric lifetime, or decay timescale, for reservoir, i, and α is an adjustment factor, computed at each time step and applied to all characteristic timescales, shown in Equation 3.7. Each timescale of decay, τ_i operates on a specified fraction of the emissions to the atmosphere, a_i, (approximately one fourth for each pool). The evolution

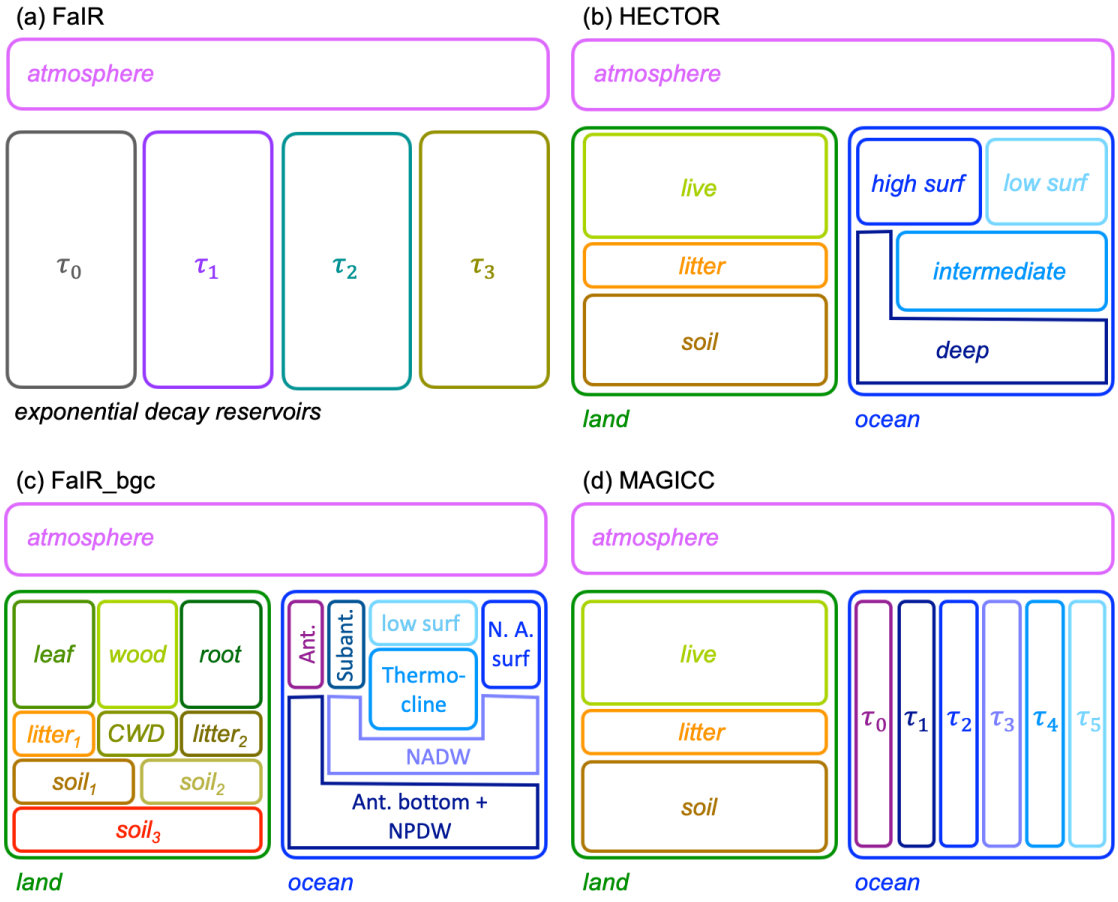


Figure 3.1: **Carbon cycle structures for each SCM**. In each schematic, a pink horizontal rectangle at the top represents the SCM's atmosphere carbon pool. HECTOR, FaIR_bgc and MAGICC represent land and ocean carbon cycles separately and explicitly and have been colored green and blue to indicate terrestrial and ocean carbon representations, respectively. HECTOR and FaIR_bgc contain representations of ocean circulation, which influences carbon draw-down from the atmosphere to the ocean. FaIR_bgc includes live, litter, and soil organic matter pools, each of which have three sub-pools.

of total atmospheric CO_2 in the atmosphere, $C(t)$ can therefore be described as the sum of four emissions reservoirs in the atmosphere, R_i with $i = [0, 1, 2, 3]$, plus preindustrial CO_2 concentration such that

$$C(t) = C_0 + \sum_{i=0}^{3} R_i(t), \tag{3.1}$$

where C_0 is the preindustrial CO_2 concentration, and $C(t) - C_0$ is the atmospheric burden above preindustrial, $G_a(t)$. $G_a(t)$ is calculated as a sum of the atmospheric pools as in Eq. 3.2. The total sink at any time is the difference between cumulative emissions and atmospheric burden as in Eq. 3.3, where $\sum_{s=t_0}^{t} E(s)$ is cumulative emissions since the simulation start, t_0.

$$G_a(t) = \sum_{i=0}^{3} R_i(t) \tag{3.2}$$

$$G_u(t) = \sum_{s=t_0}^{t} E(s) - G_a(t) \tag{3.3}$$

We calculate the evolution of each individual sink reservoir, $G_{u,i}(t)$ as the difference between the emissions allocated to each atmospheric pool and the respective atmospheric pool concentration:

$$G_{u,i}(t) = \sum_{s=t_0}^{t} a_i E(s) - R_i(t) \tag{3.4}$$

The evolution of CO_2 concentration in each atmospheric pool, $R_i(t)$ is given by Eq. 3.5.

$$\frac{dR_i(t)}{dt} = a_i E(t) - \frac{R_i(t)}{\alpha(t)\ \tau_i} \tag{3.5}$$

In Figure 3.1, each sink reservoir is labeled as τ_i. While values for each τ_i are fixed. The lifetime adjustment parameter, $\alpha(t)$ allows accumulated sink uptake, R_u, and global temperature, T, to feed back on each lifetime Smith et al. (2018):

$$\sum_{i=0}^{3} \alpha a_i \tau_i \Big[1 - \exp\big(\frac{-100}{\alpha \tau_i}\big)\Big] = r_0 + r_u G_u + r_T T \tag{3.6}$$

In earlier versions of FaIR, the unique root α is computed at each time step. Beginning with FaIR version 2.0.0, the solution is approximated with the exponential form in Eq. 3.7, where g_0 and g_1 are constants that align both solutions when $\alpha = 1$ (Leach et al., 2021). r_0 can be interpreted as the strength of preindustrial uptake from the atmosphere; r_u is the sink sensitivity to cumulative sink uptake; r_T is the sink sensitivity to temperature. While α is computed at each time step, the value is the same for each reservoir (i.e., there is no α_i).

$$\alpha(t) = g_0 \ \exp\big(\frac{r_0 + r_u G_u(t) + r_T T(t)}{g_1}\big) \tag{3.7}$$

The lifetime parameter, α, is state-dependent for only CO_2 and methane, CH_4. For CH_4, an additional sensitivity to atmospheric concentration, $r_a G_a(t)$, is included in Eq. 3.7 (Leach et al., 2021). For other greenhouse gases, α is constant.

Each of the characteristic timescales can be interpreted as the rate at which carbon is drawn out of the atmosphere and into a sink reservoir. At each time step, the four reservoirs are computed and used to determine the new atmospheric concentration (together with the pre-industrial concentration). Millar et al. (2017) interpret the four timescales of uptake as geological re-reabsorption ($\tau_0 = 1 \times 10^9$ years), deep ocean invasion/equilibration ($\tau_1 = 394.4$ years), biospheric uptake/ocean thermocline invasion ($\tau_2 = 36.54$ years), and rapid biospheric uptake/ocean mixed-layer invasion ($\tau_3 = 4.304$) years), and the computed adjustment factor, $\alpha(t)$ provides uniform biogeochemical feedbacks on all atmospheric lifetimes based on the current time step's levels of accumulated emissions (Gu) and global temperature (T).

HECTOR

At the next level of complexity, we run HECTOR version 3.2.0, an open-source, process-based, carbon-climate model that calculates the globally-resolved flow of carbon and energy between the atmosphere, ocean and terrestrial biosphere, as described by Hartin et al. (2015, 2016). Anthropogenic emissions are emitted from a geologic carbon pool, `earth_c`, and LUC emissions are emitted as fractions of the land carbon pools (`veg_c`, `detritus_c`,

and `soil_c`).

In HECTOR's terrestrial carbon cycle component, atmospheric carbon flows to the vegetation pool through net primary productivity (NPP). The rate of NPP varies with atmospheric carbon, modified by a user-set fertilization parameter, β. Carbon flows from live vegetation to detritus and soil; detritus flows to soil, and both detritus and soil carbon return to the atmosphere through respiration. Respiration is modified by a Q_{10} factor. The CO_2 fertilization feedback and respiration response to temperature that β and Q_{10} represent, respectively, modify the flux of carbon between land and atmosphere pools in opposing directions. NPP increases with higher CO_2, and respiration also increases with higher CO_2 driven by higher temperatures.

HECTOR's ocean carbon cycle component is composed of four ocean boxes – two surface (one cold high-latitude and one warm low-latitude box), one intermediate, and one deep ocean box. The flux of carbon between the atmosphere and each surface ocean box is a function of the gradient in pCO_2 between the atmosphere and ocean box at each time step (Equation 11 of Hartin et al., 2015). Within each surface box, pCO_2 is calculated as a function of temperature, salinity, and pH. Temperatures of the surface ocean boxes are linearly related to atmospheric temperature (allowing the cold high-latitude box to be a stronger sink). Carbon is transported between ocean boxes in a prescribed circulation. The total ocean sink at each time step is the sum of all ocean fluxes.

MAGICC

MAGICC is the primary simple climate model used by the IPCC. We run MAGICCv6.0, a hemispherically-averaged upwelling-diffusion ocean coupled to an atmosphere layer and a globally averaged carbon cycle model (Meinshausen et al., 2011). While many aspects of MAGICC are highly complex, the carbon cycle model is distinct from but of similar complexity to HECTOR and the other models included in this analysis.

MAGICC's terrestrial carbon cycle component is similar to HECTOR's terrestrial carbon cycle with three pools (live vegetation, detritus, and soil pools) representing carbon flow between each pool. It can also represent deforestation explicitly.

The ocean carbon cycle component uses a modified impulse-response function to represent the carbon sink (Joos et al., 1996; Meinshausen et al., 2011).

FaIR_bgc

FaIR_bgc has the highest-complexity carbon cycle model in its structure of both the land and ocean. Similarly to HECTOR, atmospheric CO_2 is computed as the difference between emissions sources and land and ocean carbon sinks.

For FaIR_bgc's terrestrial carbon cycle, a reproduction of the CENTURY land carbon cycle model is used (Parton et al., 1987, 1988). The model consists of nine carbon pools, with fluxes between them. These include three live pools, three detritus pools and three soil pools - with their own turnover times. Each carbon pool has a timescale, τ, associated with it. For the live pools (leaves, roots, and wood) the timescale represents the average lifetime before that part of the plant will die (mortality). For the dead pools it represents the average time carbon spends in that pool before either being respired or transferred to another dead pool. The timescale of a given carbon pool depends on the environment and ecosystem in which it exists. The FaIR_bgc model also represents CO_2 fertilization and the response of respiration to temperature as the β and Q_{10} as user-set parameters respectively.

The ocean model in FaIR_bgc is a seven-box model of the ocean overturning circulation with nutrient and carbon cycles adapted from Toggweiler (1999). The seven boxes include four surface, one thermocline/intermediate, one mid-depth, and one deep ocean box. Like HECTOR, it contains a chemistry submodule to compute pCO_2 from salinity and temperature.

We couple this combined global carbon cycle model to FaIR, so that at annual time steps, FaIR computes a temperature from atmospheric CO_2 concentration, using a 3-layer energy balance model, which is then passed to the carbon cycle model at the next time step.

AR5-IR

For the idealized simulations, we also include CO_2 concentration output from the AR5-IR, which is FaIR's predecessor and used in the IPCC-AR5 (Myhre et al., 2013; Millar

et al., 2017). Its atmospheric response to emissions is similar to FaIR's (with four pools of atmospheric emissions corresponding to four fixed, characteristic timescales of exponential carbon removal) except that it does not include a lifetime adjustment parameter, α, and therefore does not include allow CO_2 uptake or temperature to feedback on the sink rates. We include this model only for context in figures and do not include it in analysis.

3.2.2 Simulations

We sample across a range of decarbonization scenarios. Emissions trajectories for each of the four scenarios is shown in Figure 3.2. The pulse experiment provides a clean way to compare model behavior in a setting that has been used to calibrate simple models (Joos et al., 2013; Schwarber et al., 2019). The Flat10 and Bell Curve experiments are ideal experiments for calculating the fundamental climate mitigation metrics (discussed in **Methods**) The historical emissions trajectory is also useful for assessing necessary mitigation; it also allows us to compare the models to observations, and provides a direct comparison how models behave in a regime against which they are tuned when assessing these models in observed regimes. Together, these scenarios allow us comprehensively compare the climate predictions from these SCMs under idealized mitigation and trace the origin of any difference.

Historical: Global Carbon Budget

We drive a historical simulation for each model using emissions of CO_2-FFI and CO_2-LULUCF beginning in 1850 from Forster et al. (2023). Emissions are dropped to zero after the record ends in 2021, and simulations are continued until 1000 years of simulation are complete.

Observations of global mean CO_2 concentration and growth rate are from the NOAA Global Monitoring Network (Lan et al., 2024).

Idealized

We run a set of idealized simulations that are designed to diagnose the response of the carbon cycle to a decarbonization regime. These include a Flat10 emissions scenario followed by a

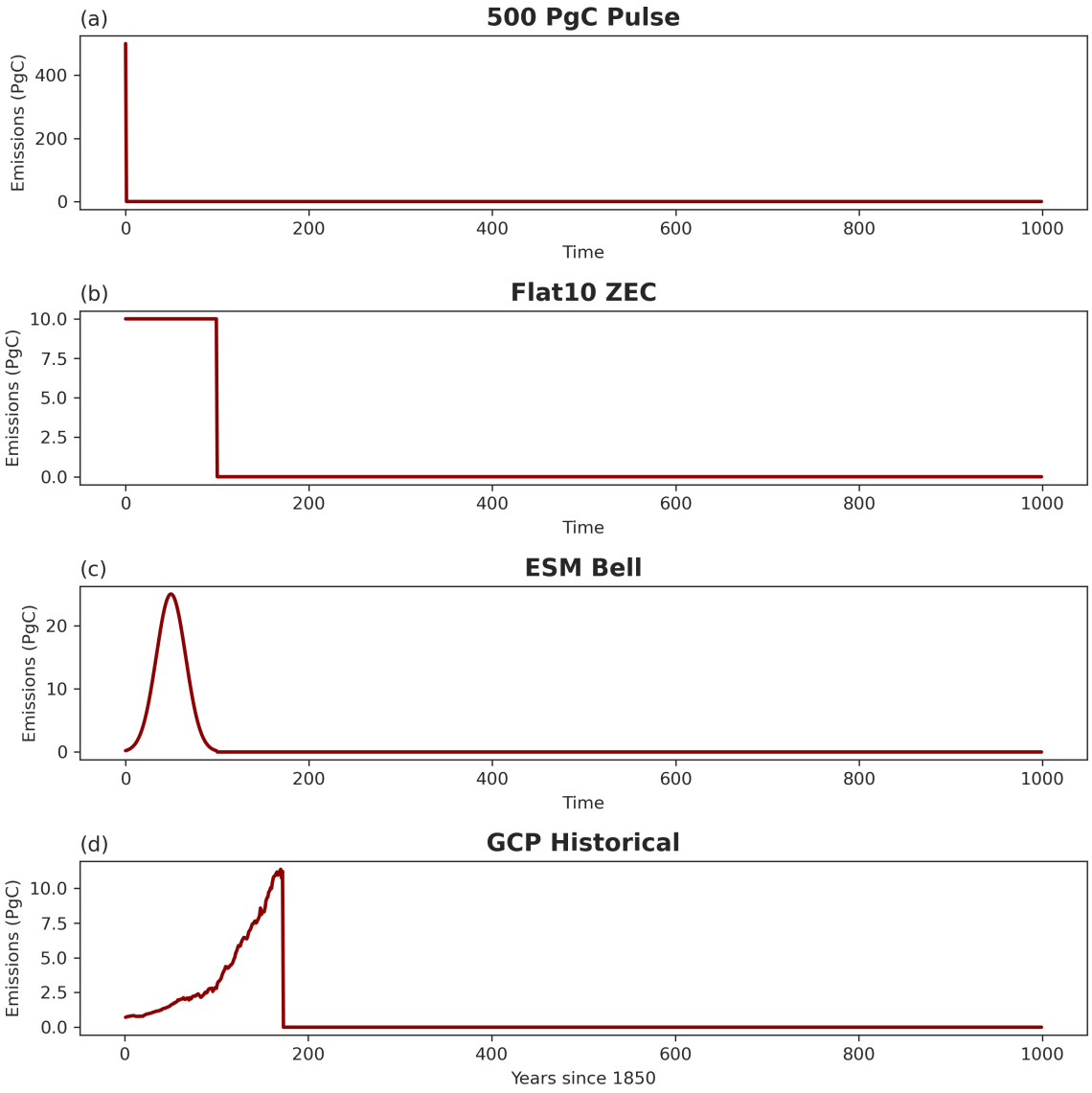


Figure 3.2: **Emissions scenarios.** Emissions scenarios for each of the three idealized scenarios (**a** - **c**) as well as historical emissions from Forster et al. (2023) (**d**).

drop to zero emissions after 100 years (referenced hereafter as *esm-flat10-zec*), a Gaussian emissions scenario, which we refer to as the *esm-bell1000PgC* scenario, and a 500 Pg C pulse emissions scenario. In the *esm-flat10-zec* scenario, 10 Pg C is emitted every year for 100 years, immediately followed by zero emissions. In *esm-bell1000PgC*, we use the emissions prescribed in the ZECMIP protocol, in which a Gaussian bell-shaped curve of emissions is followed, which reaches 1,000 Pg C in year 100, and smoothly transitions to zero emissions at year 100 (Jones et al., 2019). Both the *esm-flat10-zec* and *esm-bell1000PgC* scenarios reach 1,000 Pg C of CO_2 emissions before reaching net-zero emissions in year 100. In the *pulse-500PgC* experiments, a pulse of 500 Pg C is emitted followed by zero emissions.

We run all simulations for a total runtime of 1000 years for each emissions scenario. In the historical simulation, after 2021, emissions are dropped to zero until 1000 years of simulation have passed.

3.2.3 Configuration

Since SCMs can be tuned to produce virtually any climate response, model developers and have established calibration and constraining processes that restrict model behavior to an acceptable and reliable response. To tune these models to their default parameter values, developers largely use observations of present-day trends in global mean temperature (Smith et al., 2018; Meinshausen et al., 2011; Dorheim et al., 2020). Additional calibration approaches may be used, but each model is tuned independently. By running MAGICC, FaIR, and HECTOR with their default, "out-of-the-box" settings, we attempt to compare

Unlike MAGICC, FaIR, and HECTOR, FaIR_bgc carbon cycle model is not formally tuned in its default configuration. We perform minor hand-tuning to *only FaIR_bgc* to calibrate its model parameters to align with historical observations of temperature. Specifically, we adjust parameters that influence the biogeochemical response to accumulated CO_2 and temperature by carbon cycling processes. These include the plant CO_2 fertilization rate, the Q_{10} heterotrophic respiration response to temperature, and scaling parameters on ocean surface gas exchange, circulation, and circulation's response to temperature. These specific parameters are discussed further in Ch.4. All of the parameter values were adjusted within

ranges supported by observations of these values.

Outside of this adjustment, the SCMs are run in their default configuration in order to diagnose the initial inter-model spread and in order to bring into relief any potential bias that would be assumed in running a single model on its own.

We run an ensemble in FaIR using its published calibrated, constrained ensemble, version 1.2.0 (Smith, 2023), which is an AR6-consistent calibration of FaIRv2.1.1. It consists of 1,001 parameter combinations, largely constrained from historical observations of temperature and ESM temperature metrics. More description of the ensemble is available with the published dataset.

3.2.4 Computing a Remaining Carbon Budget (RCB)

The remaining carbon budget (RCB) for climate stabilization is a critical metric used by the IPCC to direct climate mitigation policy (IPCC Working Group I, 2021b). It is defined as the net amount of carbon that can be emitted (starting today) while staying below a chosen global warming limit (Rogelj et al., 2019). It is often taken with respect to 1.5 or 2°C above preindustrial temperature, which is the stated goal of the UN Paris Agreement (UNFCCC, 2015). In this study, we do not restrict the RCB to positive values. However, we would interpret a negative RCB as the degree to which we have exceeded the global total carbon budget rather than a recommended necessary negative emissions budget.

Rogelj et al. (2019) define the RCB for a given temperature change limit ($\mathrm{T_{lim}}$) as a function of the historical anthropogenic warming to date (T_{hist}), the non-CO_2 contribution to future temperature rise, $T_{\mathrm{nonCO_2}}$, the Zero Emissions Commitment, given as T_{ZEC} in Equation 3.8 and TCRE, and finally an adjustment factor for sources of unrepresented Earth system feedbacks (E_{Esfb}).

$$\mathrm{RCB} = \frac{T_{\mathrm{lim}} - T_{\mathrm{hist}} - T_{\mathrm{nonCO_2}} - T_{\mathrm{ZEC}}}{\mathrm{TCRE}} - E_{\mathrm{Esfb}} \tag{3.8}$$

We set a limit of 1.5°C for T_{lim}. T_{hist} can be estimated from observations with respect to an 1850-1900 baseline. To do so, we use the Forster et al. (2023) combined temperature dataset: anthropogenic warming in 2021 is 1.11°C warmer than the baseline. For cumulative

CO_2 emissions from fossil fuels and land use (from 1870 to the end of 2021), we use the Global Carbon Project global emissions dataset starting in 1870. In 2023, cumulative emissions is 675.31 Pg C.

To determine T_{nonCO_2} we use our results from the Flat10_ZEC following the example of Damon Matthews et al. (2021) given in Equation 3.9:

$$RCB = E \times \left(\left(\frac{\Delta T_{lim} - \Delta T_{ZEC}}{\Delta T_{anth}}\right)\left(\frac{1 - f^*_{nc}}{1 - f_{nc}}\right) - 1\right) \tag{3.9}$$

Here, a distinction is drawn between the $\text{TCRE}_{\text{CO}_2}$ and the observed temperature response to anthropogenic emissions as a way to embed non-CO_2 forcing, wherein:

$$\text{TCRE}_{CO_2} = \frac{\Delta T_{anth}}{E} \times (1 - f_{nc}) \tag{3.10}$$

We diagnose a CO_2-driven TCRE from our modeled historical simulation, which we can relate to the observed temperature response to historical emissions to estimate f_{nc}, or the historical non-CO_2 forcing fraction. We calculate f_{nc} as the mean ratio of non-CO_2 to total anthropogenic radiative forcing as shown in Equation 3.10.

Damon Matthews et al. (2021) estimate the historical non-CO_2 forcing fraction, f_{nc}, as a linear function of f^*_{nc}.

We calculate the TCRE from the flat10_ZEC experiment for the CO_2-driven warming. We then add the expected non-CO_2-driven warming as $(\text{TCRE}_{\text{flat10\_ZEC}} \times 1000 \text{ Pg C}) \times f_{nc}$.

We diagnose observed warming to cumulative CO_2 since pre-industrial baseline as the slope in temperature to cumulative emissions since 1950.

3.3 Results

3.3.1 SCMs produce historical CO_2 concentration record well.

All models reproduce historical observations of CO_2 concentration in their default "out-of-the-box" configuration. We compare the each model's airborne CO_2 to observations of global mean CO_2 concentration from the NOAA Global Monitoring Network (Lan et al., 2024). Historical airborne fraction in response to historical emissions is consistent across

models and is in line with observations. Figure 3.3 shows the fidelity of each model to the observed CO_2 record.

These SCMs were have been built and calibrated to reproduce historical warming trends from historical emissions. In the cases of FaIR, HECTOR, and MAGICC, they have been additionally tuned to reproduce ESM responses to CMIP scenarios (Meinshausen et al., 2011; Smith et al., 2018; Dorheim et al., 2020) and consequently real-world CO_2 concentrations from emissions records. With this in mind, the models' high fidelity to observed CO_2 concentration is relatively unsurprising. However, it does demonstrate the SCMs' ability to produce a comparable response to the observed climate and validates their performance in increasing-emissions regime. As a result, we can use this result to validate that SCMs are able to capture the principle dynamics of the global carbon cycle at play in an increasing-emissions regime of the historical period.

3.3.2 Wide range of climate paths from the same emissions trajectory.

We examine the spread in modeled CO_2 concentration and global temperature within each emissions scenarios with a specific focus on before and after the date of net-zero emissions.

Scenario: Historical emissions

While the response to historical emissions was not surprising given that all models have been built to reproduce historical emissions-to-concentrations, what is surprising is the response to emissions starting in 2022 when emissions are dropped to net-zero. Figure 3.4 shows atmospheric CO_2 concentration in the historical (pre-2022) and after a hypothetical drop to zero emissions (post-2022). The spread in CO_2 concentration across SCMs reaches a maximum of 57.7 ppm 49 years following cessation of emissions (compared to a maximum spread of 23.4 ppm in 2021, the last year of historical emissions).

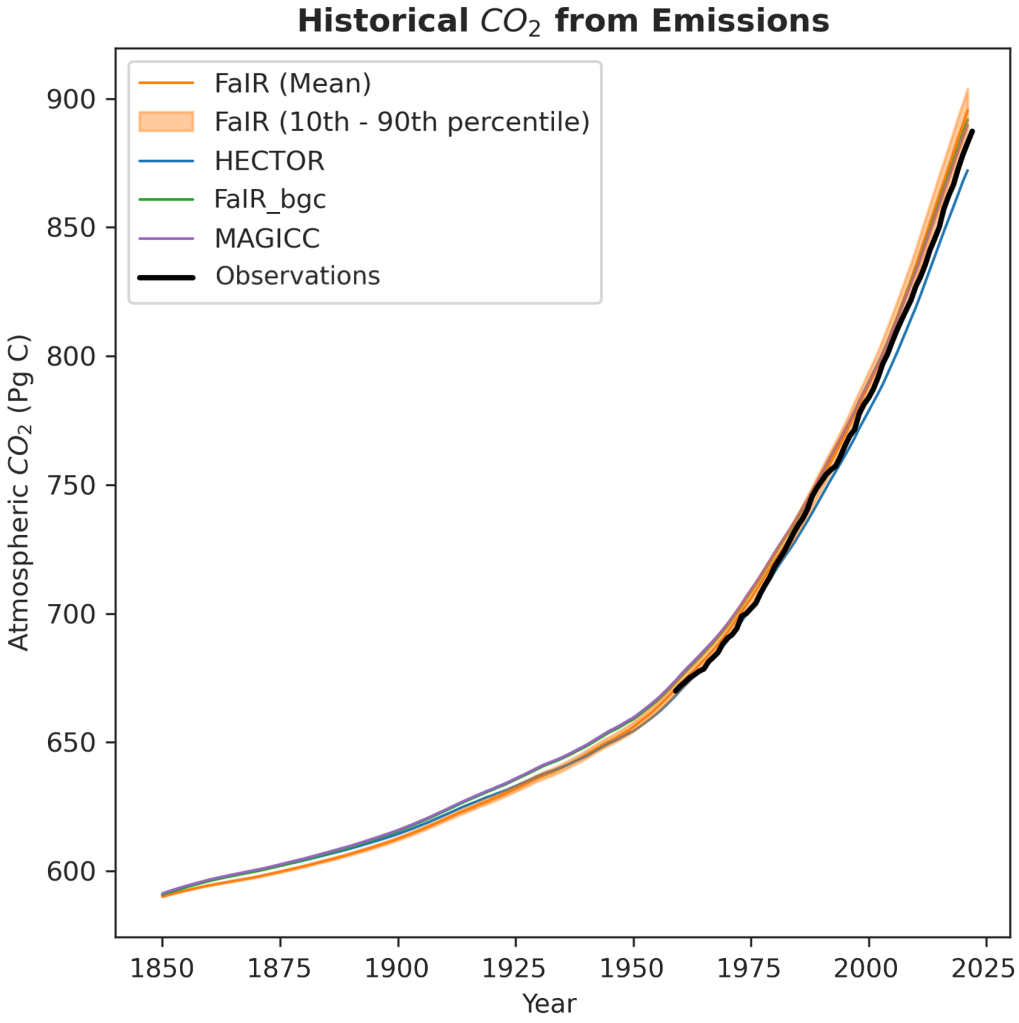


Figure 3.3: **Historical: CO_2 concentration for all models.** Modeled atmospheric CO_2 from historical emissions compared to global observations of CO_2 concentration from Lan et al. (2024).

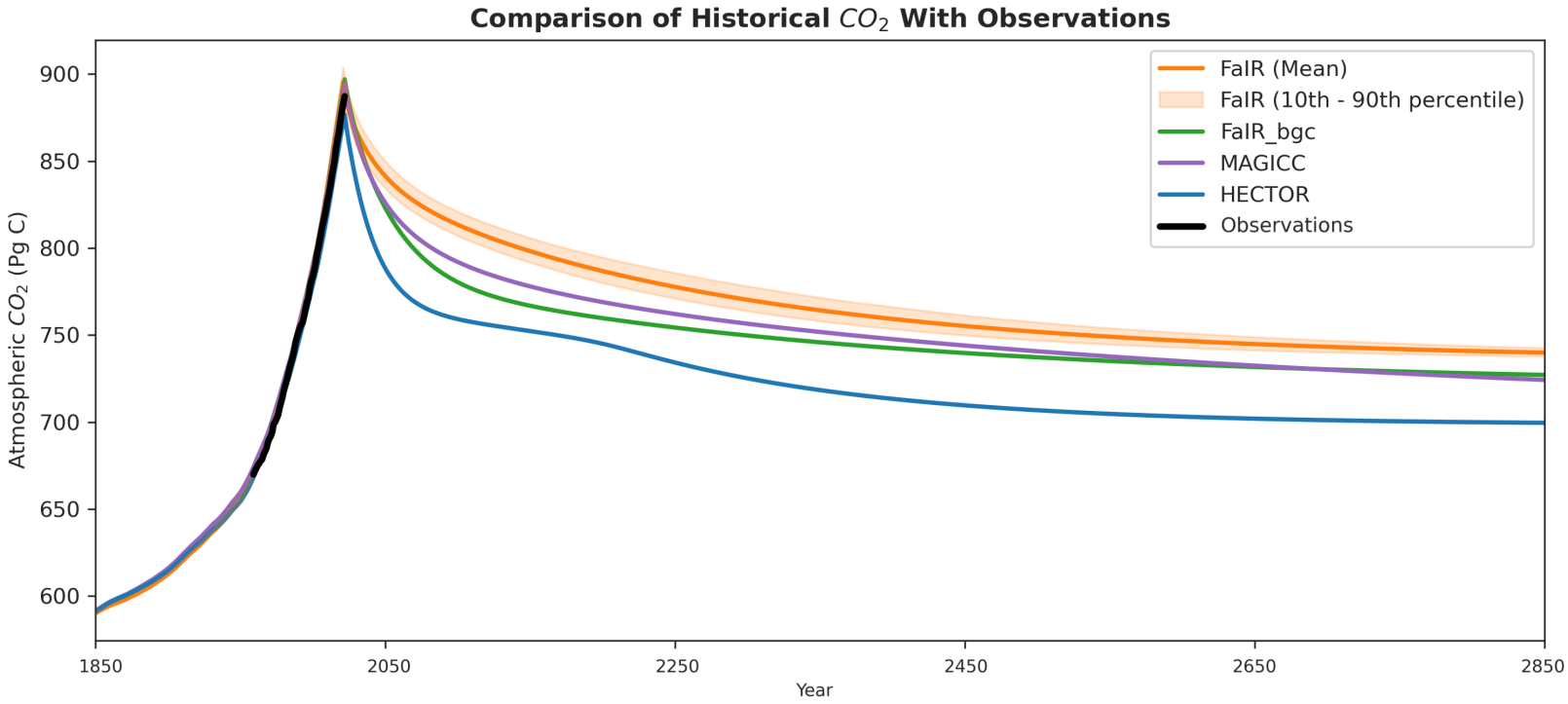


Figure 3.4: **Historical followed by flattened emissions: CO_2 concentration for all models.** Evolution of atmospheric CO_2 in each model when driven with historical emissions until 2021. In 2022, CO_2 emissions drop to 0 Pg C/year until the end of the simulation.

Scenario: 500 Pg C emissions pulse

In response to the 500 Pg C pulse, a large initial spread across models of more than 150 ppm is driven by a sharp decrease in airborne fraction due to FaIR's shortest carbon uptake timescale, which causes FaIR to show a sharp decrease in atmospheric CO_2 during the first four years of the scenario. After this initial divergence, the models see a secondary maximum spread 68 years following net-zero of 53.6 ppm (between FaIR and HECTOR) driven by a more rapid decline in atmospheric CO_2 in HECTOR compared to FaIR. The spread in modeled CO_2 concentration only slightly declines after this point.

From this pulse experiment, we can detect that long-term differences in atmospheric CO_2 can be traced back to differences in rates of carbon removal in the first 200 years following the decrease in emissions to net-zero.

The spread in atmospheric CO_2 concentration across models leads to a corresponding spread in the temperature response in the *pulse_500PgC* emissions scenario. Temperature differences persist and increase slightly after year 200, leading to a 0.23K range in mean global temperature anomaly after 1000 years.

1000 PgC Scenarios: esm-bell1000PgC and esm-flat10-zec

In both the *esm-bell-curve* (Figure 3.6) and the *esm-flat10-zec* (Figure 3.7) emissions scenarios, the SCMs develop a spread in atmospheric CO_2 in the first 200 years following the beginning of decarbonization (years 50-100 for the *esm-bell1000PgC* scenario and year 100 for the *esm-flat10-zec*). Similarly to the *pulse_500PgC* scenario, the spread in CO_2 concentration stabilizes after the first 200 years and persists until the end of the simulation. The concomitant spread in global mean temperature increases and persists, largely driven by the unique evolution of HECTOR's CO_2 concentration. The timing of HECTOR's carbon sinks lead to a consistent kink in atmospheric CO_2 approximately 80 years after net-zero in all scenarios.

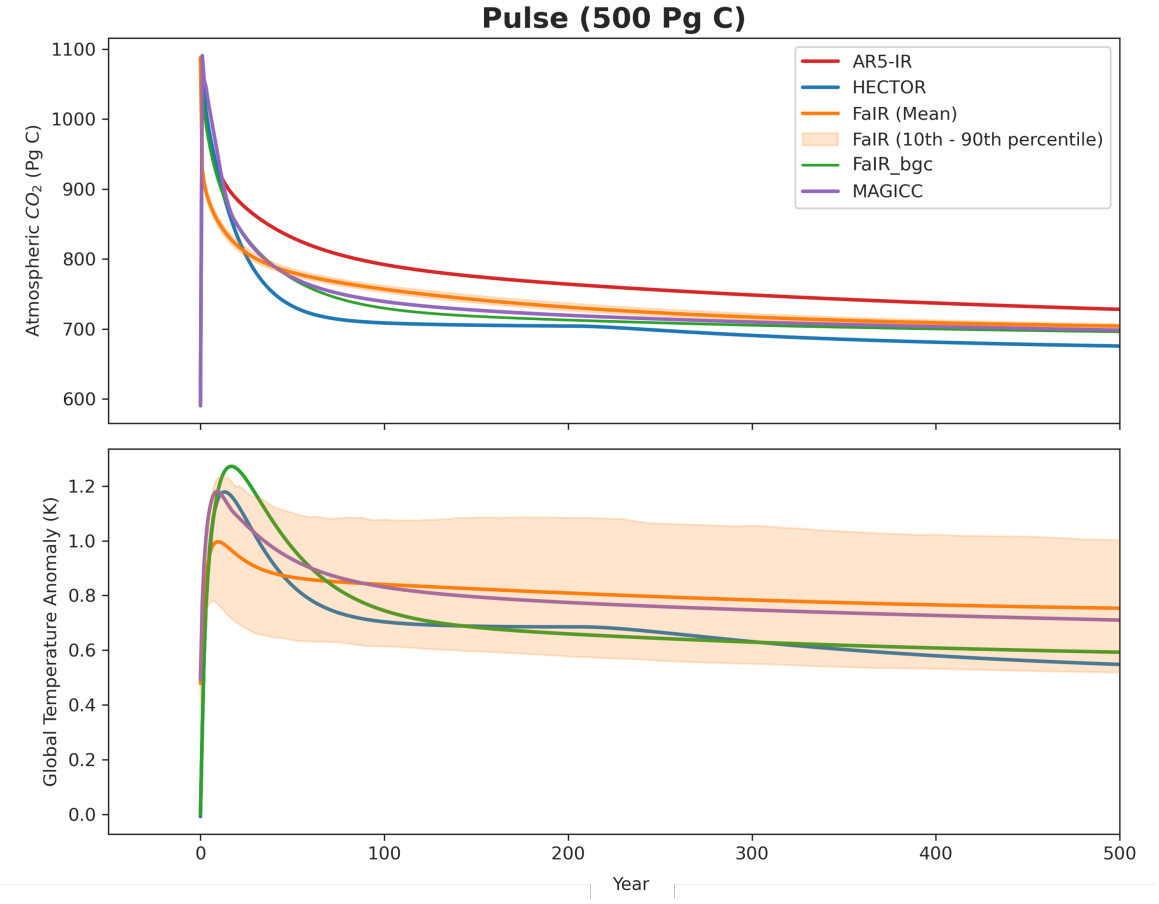


Figure 3.5: **500 Pg C Pulse: CO_2 and temperature.** Modeled atmospheric CO_2 and global mean surface temperature for all models in response to a 500 Pg C pulse of CO_2 in year 0.

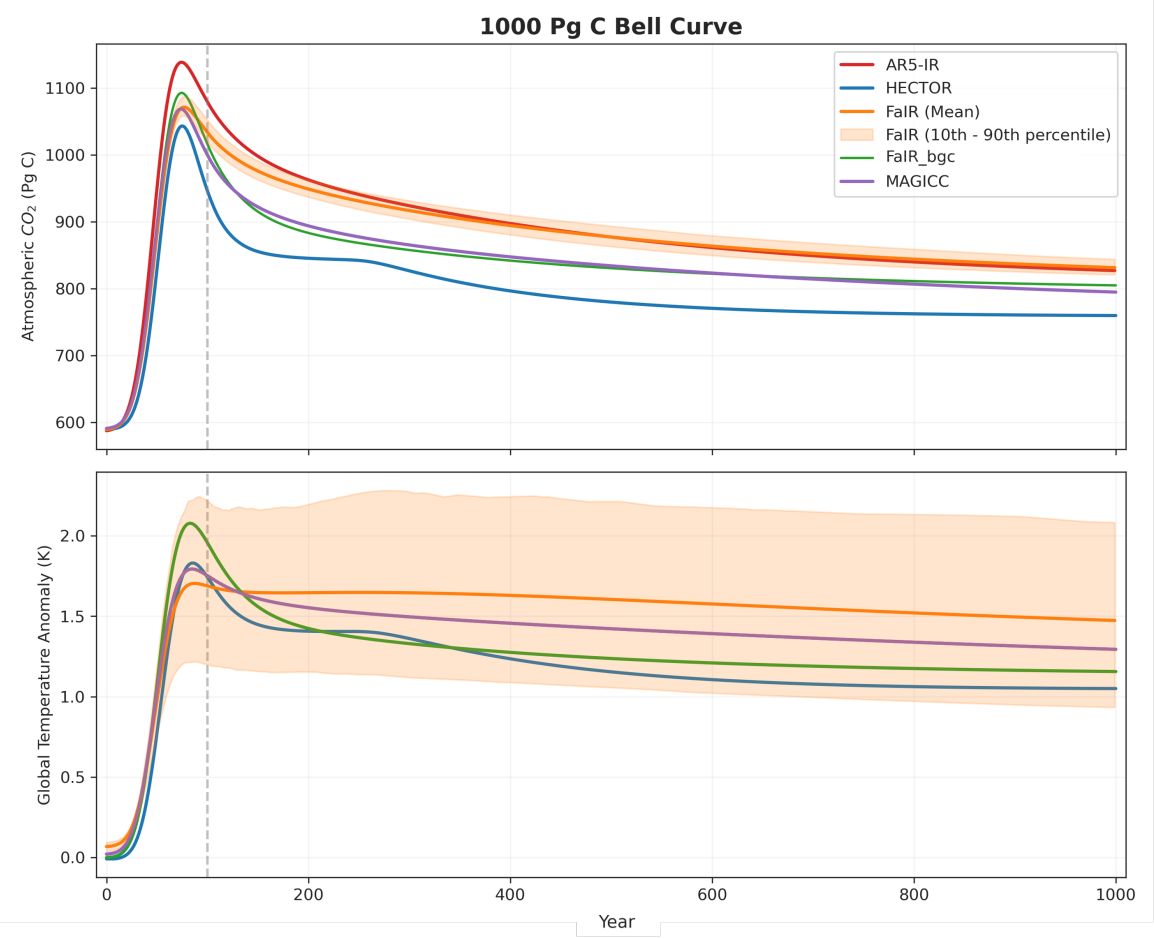


Figure 3.6: **ESM Bell Curve: CO_2 concentration and temperature.** Modeled atmospheric CO_2 and global mean surface temperature for all models in response to the *esm-bell1000PgC* emissions scenario

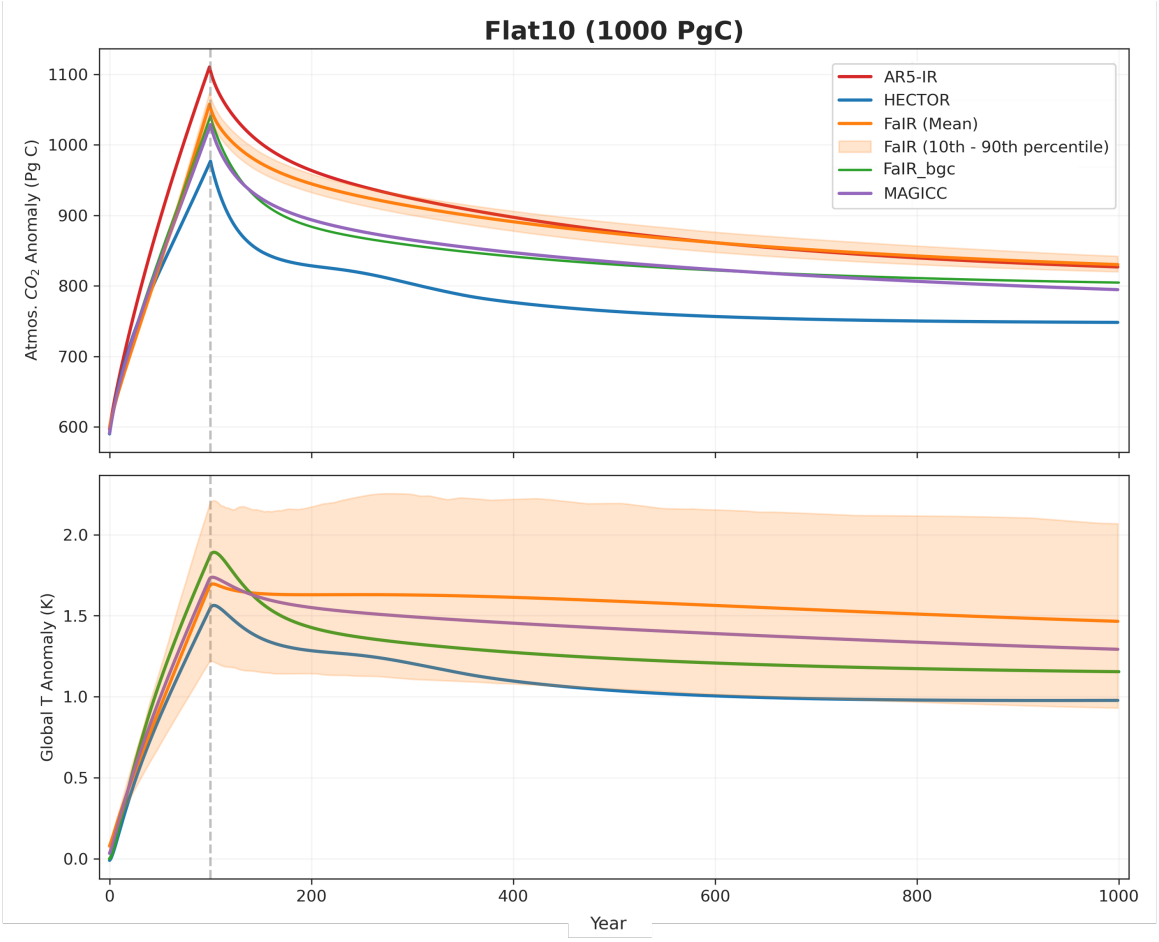


Figure 3.7: **Flat10: CO_2 concentration and temperature** Modeled atmospheric CO_2 and global mean surface temperature for all models in response to the *esm-flat10-zec* emissions scenario

3.3.3 Models disagree on long-term global temperature response.

Overall, in all scenarios, the spread across SCMs in CO_2 concentration and global temperature anomaly increases following net-zero emissions. The spread increases to a maximum between 50 and 200 years after net-zero with minimal decrease in spread afterward.

Figure 3.7 shows the carbon and climate response across models in the *esm-flat10-zec* scenario. While some differences emerge in the emissions phase (before year 100), which can be seen in Figure 3.8, the difference in ZEC, or the temperature evolution after decarbonization can be compared in Figure 3.9, in which all models begin with respect to their own TCRE at year 100 and the evolution of temperature with respect to TCRE1000 is show. This illustrates that the behavior of the long-term sinks varies from each other, leading to a widening spread in ZEC over the first 200-300 years, and allow the spread to stabilize but not shrink over the following 600 years.

Figures 3.8 shows the comparison of TCRE+ZEC50 in SCMs (squares). While SCMs exhibit differences in TCRE and appear to converge after 50 years, this convergence is only temporary, marking the approximate point where the temperature anomalies of FaIR, FaIR_bgc, and MAGICC shift position. If the clock were stopped at ZEC50, it might look like the sum of TCRE and ZEC is consistent across these three models, but over time, the change in temperature driven by ZEC outweighs the TCRE, and the models display a spread of about 0.5°C by year 1000 of simulation.

In Figure 3.8, are the TCRE+ZEC50s calculated from the same *esm-flat10-zec* emissions scenario for a selection of ESMs participating in a Flat10 intercomparison project for context. Fifty years after net zero, ESMs have a range in total temperature response of over 1°C.

3.3.4 Structural differences between SCMs lead to different carbon cycle outcomes and climate responses

HECTOR, MAGICC, and FaIR_bgc allocate carbon to physical pools corresponding to the terrestrial and ocean carbon cycles. Figure 3.10 shows how those pools are allocated in the *esm-flat10-zec* scenario. Notably the largest differences are in long timescale carbon

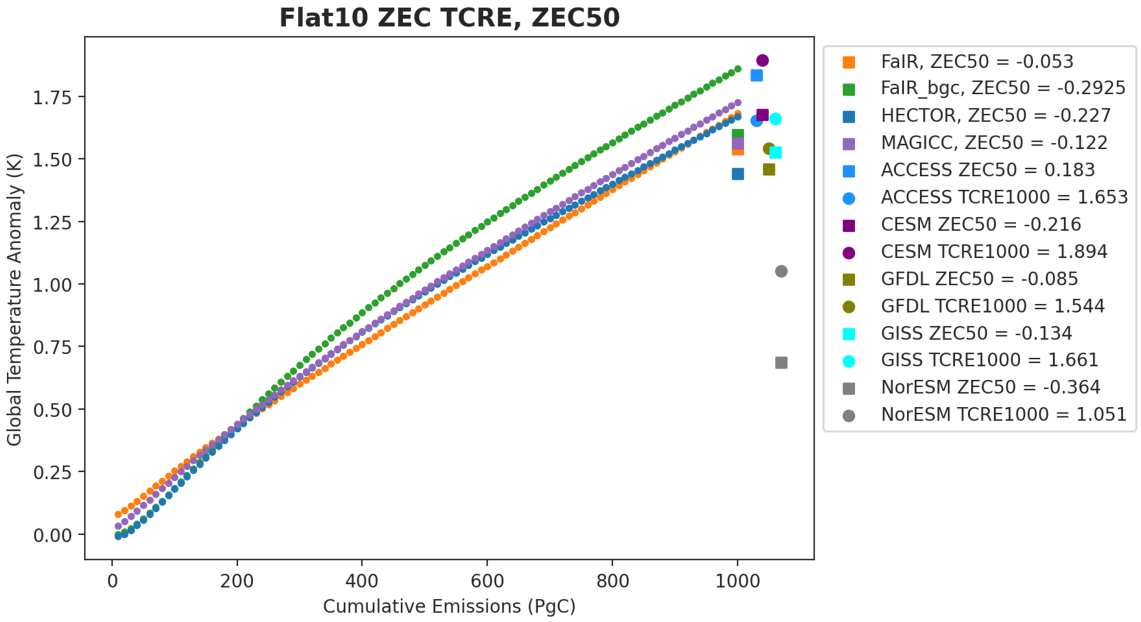


Figure 3.8: **TCRE from *esm-flat10-zec*.** For each year in the *esm-flat10-zec* simulation, we calculate the slope of the TCRE (using all preceding years' emissions). We multiply each year's TCRE by its cumulative emissions to obtain the inferred global temperature anomaly in Kelvin. Shown in dots are each SCM's inferred global temperature anomaly calculated from the TCRE vs. cumulative emissions. Corresponding squares are plotted showing the ZEC50 relative to the TCRE1000 (rightmost circle represents for each SCM). The difference between each model's square and last circle is the value of ZEC50. We display the TCRE and ZEC relative to each other this way in order to illustrate the relative contributions to total temperature change. We also include TCRE1000 (circles) and corresponding ZEC50 (squares) in the same way for several ESMs that have run the same *esm-flat10-zec* experiment.

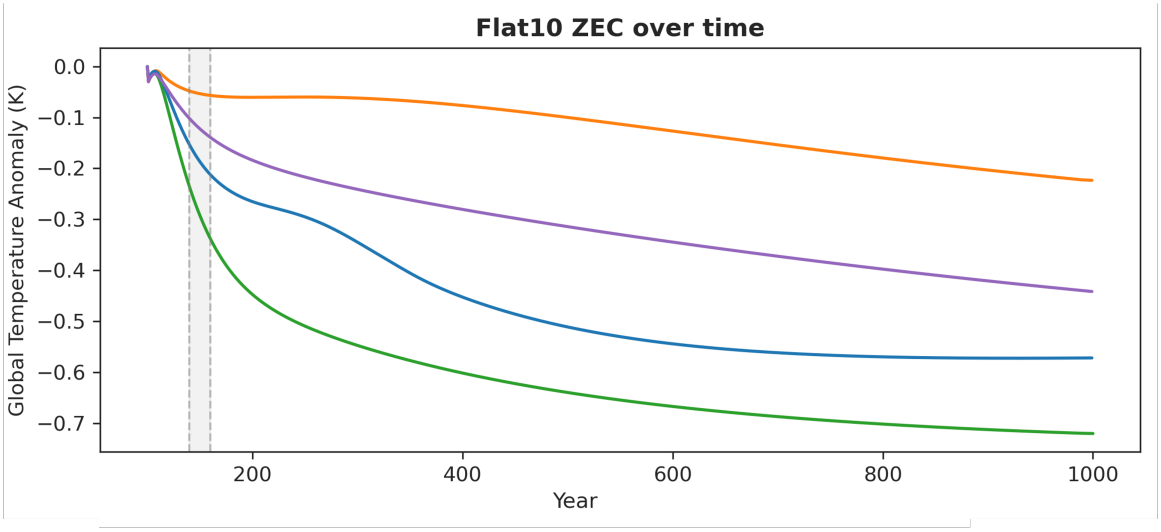


Figure 3.9: **ZEC over time in *esm-flat10-zec* scenario.** Global mean temperature anomaly relative to TCRE1000 for each year, starting in year 100 (time of net-zero). FaIR_default is shown in orange, MAGICC in purple, HECTOR in blue, and FaIR_bgc in green. Grey shading indicates years over which ZEC50 is averaged when reported.

pools – ocean and soil pools. During the period with the largest differences in atmospheric CO_2, HECTOR allocates more carbon to soil pools than to detritus or live vegetation pools put together. Whereas MAGICC and FaIR_bgc allocate larger amounts of carbon to live vegetation (leaves, wood, and roots). FaIR_bgc allocates the least amount of carbon to ocean pools.

Finally, we compute an RCB for each SCM, shown in Figure 3.11a, using a warming threshold of 1.5°C and a ZEC100 from the *esm-flat10-zec*. We compare the rates of required decarbonzation across models by expressing the RCB as an linear ramp-down in emissions necessary to keep warming at or below 1.5°C. With a linear decrease in emissions starting in 2021, we find that FaIR should achieve net-zero by 2026, FaIR_bgc by 2074, HECTOR by 2035, and MAGICC by 2043. A range of about 250 Pg C corresponds to a range of 50 years in net-zero date.

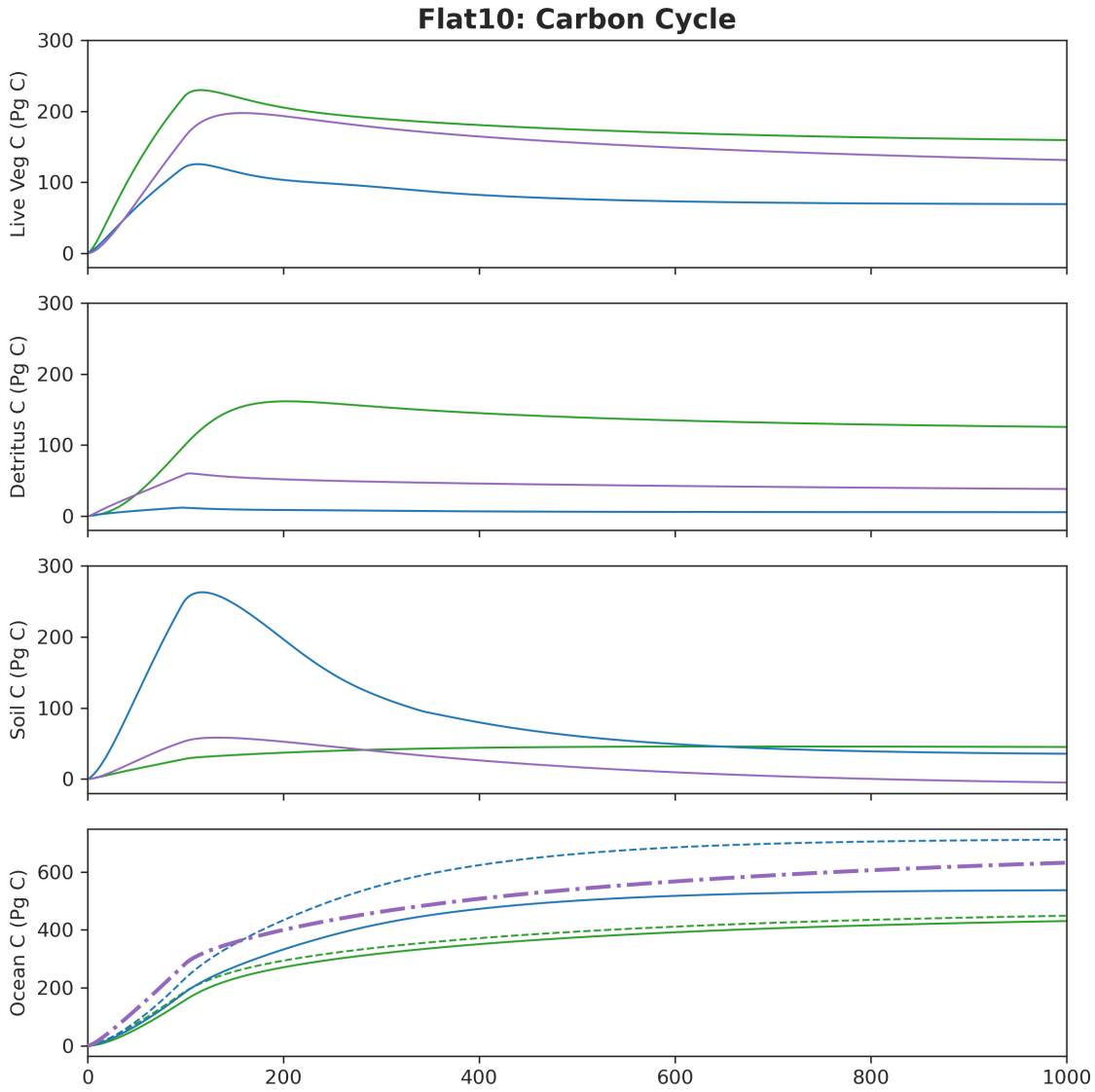


Figure 3.10: **Change in carbon pools for *esm-flat10-zec*.** For FaIR_bgc (green), HECTOR (blue), and MAGICC (purple), time evolution of carbon pools for the *esm-flat10-zec* experiment relative to year 0. Solid lines in the ocean pools indicate the sum of the non-surface pools; dashed lines represent the total ocean pool size. MAGICC's total ocean carbon is plotted as the purple dash-dot line.

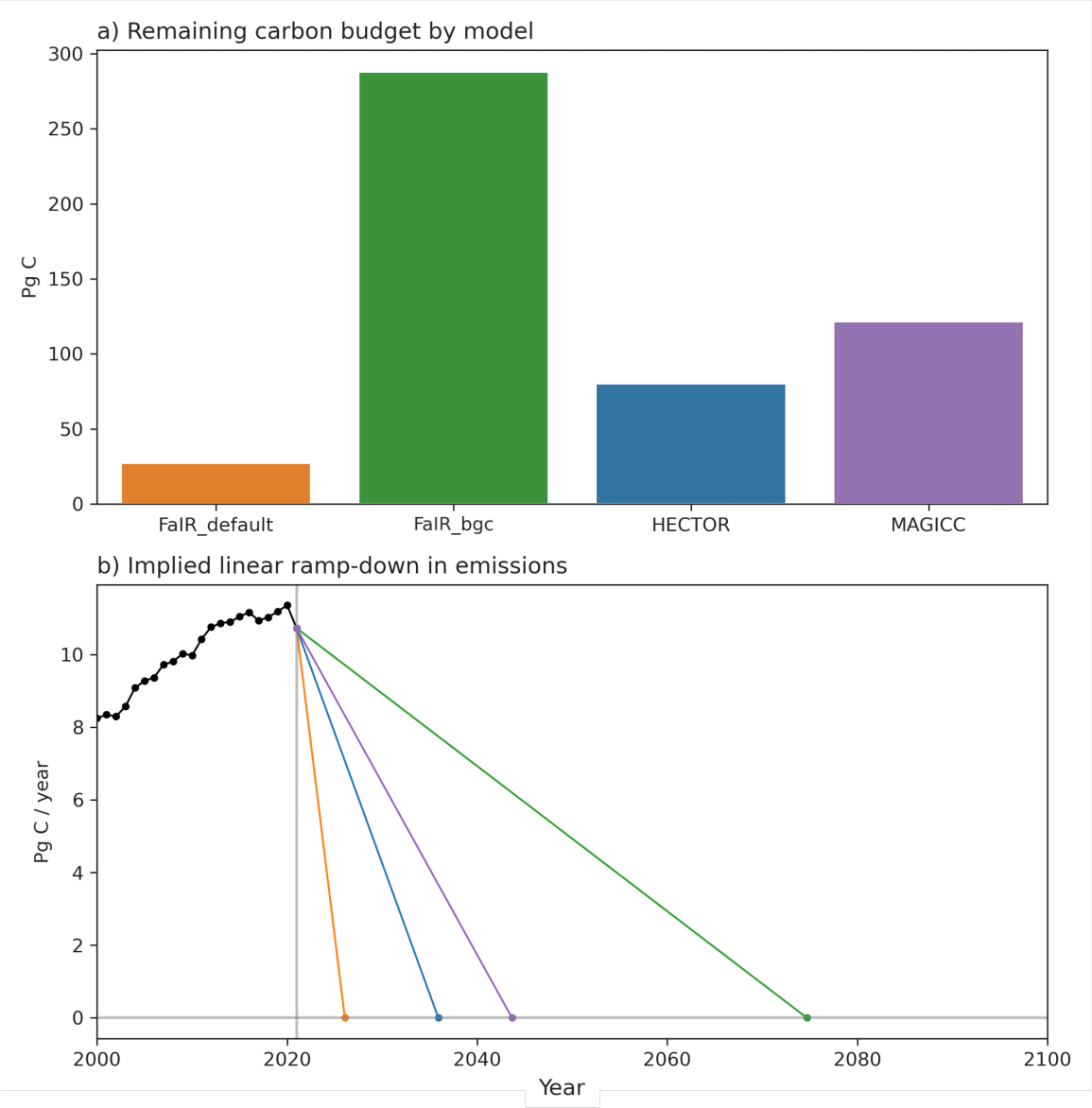


Figure 3.11: **Remaining carbon budgets across models. a** Remaining carbon budget in the default configuration for each model, given a historical warming rate beginning in 1950. **b** Implied necessary ramp-down in emissions beginning in 2021 to keep warming below a warming threshold of 1.5°C above pre-industrial temperature.

3.4 Discussion & Conclusions

How does the carbon cycle structure in an SCM influence the magnitude and timing of warming in decarbonization scenarios? In the historical simulation, spread in the CO_2 concentration response to emissions emerges after emissions cease, suggesting that structural uncertainty in the carbon-climate system may be largest after emissions stop. In the idealized simulations, the spread in global mean temperature emerges in the 50-200 years following emissions cessation. If we compare the *esm-flat10-zec* TCRE1000 (Figure 3.8) and the corresponding long-term evolution of ZEC (Figure 3.9), we see a larger spread in temperature change across models in the long-term ZEC than in the TCRE1000. However, differences in ZEC alone cannot account for the differences in RCB. FaIR_bgc consistently has the most negative ZEC (strongest post-net-zero cooling), but it also exhibits the strongest TCRE (strongest pre-net-zero warming). HECTOR consistently has the second most negative ZEC, but its RCB is smaller than MAGICC's. HECTOR also has the smallest TCRE1000 (slightly smaller than FaIR's). As a result, for the same CO_2 emission, more historical warming may be attributed to non-CO_2 forcing in the RCB calculation, leading to a decrease in the RCB. In this way, while the ZEC may have direct RCB consequences, small differences in non-CO_2 warming implied by differences in the TCRE contribute to differences in the RCB. Previous work has shown that non-CO_2 emissions are a strong source of uncertainty in the RCB (Rogelj et al., 2016; Tokarska et al., 2018; Lamboll et al., 2023); however, in our experimental setup in each of these models, only differences in the CO_2 cycling drive these implied effects.

We can begin to explain differences modeled TCRE and ZEC by examining how each model allocates carbon across its different pools. From Figure 3.10, which shows the evolution of land and ocean carbon sink reservoirs in FaIR_bgc, HECTOR, and MAGICC from the *esm-flat10-zec* scenario, we can see that FaIR_bgc allocates the least amount of carbon to its terrestrial soil carbon pools and the most carbon to its live pools. HECTOR, by contrast, allocates more carbon to its soil pool and less to its live pool. In the ocean, by year 200 (100 years after emissions cease), FaIR_bgc has allocated over 100 Pg C less carbon to its total ocean carbon pool. Additionally, HECTOR's soil pool shows a strong response to

emissions, increasing rapidly during the first 100 years and decreasing rapidly afterward. By contrast, HECTOR's ocean carbon pool continues to increase even after emissions cease, which contributes to its negative ZEC. Generally, MAGICC sits between FaIR_bgc and HECTOR, but in the ocean, MAGICC initially sees the strongest growth in total ocean carbon and is overtaken by HECTOR only after emissions cease. While HECTOR's larger ocean carbon pools are compensated by smaller land carbon pools in the *esm-flat10-zec* experiment, HECTOR generally shows a stronger total carbon sink during the emissions phase (because of the ocean) and the lower atmospheric burden during these first 100 years contributes indirectly to HECTOR's low RCB. FaIR_bgc achieves the strongest TCRE and the most negative ZEC by allocating more carbon to short timescale land carbon pools, and as a result of its strong ZEC, FaIR_bgc has the largest RCB.

In this study, we perform a first-step model intercomparison project (MIP) with particular focus on carbon cycle structure in the types of SCMs being used to scope decarbonization pathways. We find that differences in carbon cycle structure can meaningfully influence the timing and amount of expected warming during decarbonization scenarios without changing historical airborne fraction. This underscores an inherent lack of knowledge about the principle dynamics at play in decarbonization scenarios.

Our main conclusion from this work is that structural diversity in carbon cycle representation in these simple climate models leads to a difference in warming response due to differences in how the carbon is allocated across the Earth system. As emissions are increasing, these differences may not influence the overall trend in atmospheric CO_2, but in decarbonization scenarios it changes the climate response. This can be seen most clearly in the *esm-flat10-zec* experiment and validated in other emissions scenarios, like the *esm-bell1000PgC*, we see the same divergence when emissions begin to decrease.

This work highlights the need for emissions-forced decarbonization runs in ESM that can help us train emulators to get the right answer for the right reasons (building reliability and insight).

Secondly, it highlights the danger in using only one SCM to project carbon cycle changes in response to mitigation scenarios. A better approach is to use at least two with different approaches to emulation.

Finally, it motivates the following questions, which will be addressed in the next chapter: (1) What are the timescales of carbon uptake in the long-term, and how different are they from each other in each carbon cycle representation? (2) How sensitive are these uptake timescales to choice of parameterization?

In this chapter, we calibrate the analyzed models loosely to the historical record and as a result produce a wide range of climate outcomes. This choice is made because there are many ways to calibrate these simple models. By running these models in their default configuration with minimal calibration to the historical period, we illustrate the risk of running these models in isolation, as they often are. While these models are built to be able to be calibrated to various ESMs, choice of calibration can be subjective and lead to different climate outcomes (Sanderson, 2020). By using the default configuration, we attempt to be agnostic to this choice, but are also subject to varying biases in calibration. For that reason, in the following chapter, we take a secondary approach to calibration in order to minimize the influence of this source of spread.

Chapter 4

CHARACTERIZING CARBON SINK RESPONSES TO DECARBONIZATION ACROSS MODEL STRUCTURES

Abstract

Complex models of the Earth system, are increasingly able to represent details of the carbon-climate system, but a variety of simple climate models (SCMs) use simplified, parameterized representations of the Earth system, which make them easily deployed tools for climate mitigation assessment and accessible tools for conceptual understanding. However, these SCMs vary in their approach to simplifying the Earth system, especially in their representation of the carbon cycle. In this chapter, we begin by examining how two different carbon cycle structures within one SCM, FaIR, differ in their projections of future climate under idealized decarbonization. We find that differences in carbon cycle structure lead to differences in the timing of carbon sinks especially at long timescales but that these differences do not lead to or explain differences in warming under the same decarbonization emissions scenario. Differences in the metrics of warming are instead primarily controlled by assumptions about climate feedbacks and non-carbon cycle forcing, which are parameterized separately from carbon cycling. However, when we introduce a physically-motivated link reflecting the connection between ocean circulation and energy balance, we see a change in the set of climate feedbacks necessary to explain our observed carbon-climate system. As a result, projections for future warming are shifted, and the measures necessary to mitigate climate impacts are also changed.

4.1 Introduction

The IPCC's bases its assessment of Earth's habitability for future human populations on a framework that uses indices of vulnerability, severity, and persistence with respect to climate hazards. In its Special Report on Global Warming of 1.5°C, the authors find that

a change in near-surface air temperatures between 1.5 and 2°C leads to higher risks due to climate change across these indices (IPCC SR1.5, 2018).

To limit warming to habitable levels, it it will be necessary to cease emissions entirely (Matthews and Caldeira, 2008; Allen et al., 2009; Solomon et al., 2009). As global technologies and economies transition to meet this goal, methods have been developed to place an upper limit on the amount of CO_2 that can be emitted without exceeding international warming limits (Rogelj et al., 2019; MacDougall et al., 2020; Jones and Friedlingstein, 2020; Damon Matthews et al., 2021). The CMIP process explores possible futures using ESMs, but as discussed in Chapter 3 simple climate models (SCMs) offer an opportunity to explore many more possible options due to their computational efficiency. In Chapter 3, we show that SCMs differ in their projection of climate under decreasing emissions.

It may not be surprising that SCMs disagree on how the climate will respond to decreasing emissions since our understanding of the carbon-climate system has been almost entirely validated in an increasing-emission regime. In this chapter, we investigate exactly which processes in our understanding of the carbon-climate system are responsible for the spread in projected climate response to decreased emissions. We focus on one simple climate model, FaIR, described in more detail in Chapter 3.

In FaIR, differences in carbon cycling influence global mean temperature through changes in the atmosphere burden, which is then used to compute radiative forcing and temperature by the energy balance model. Carbon cycling is represented by four pools with characteristic timescales of carbon removal from the atmosphere. In Chapter 3 we saw that assumptions about the structure of the carbon cycle, including how carbon is passed between reservoirs within the land or ocean, influence the behavior of carbon removal from the atmosphere. We also know that many of these processes are temperature-dependent in some way (i.e. through changes in soil respiration or ocean stratification). Within FaIR, temperature also feeds back on the timescales of carbon removal from the atmosphere through parameters in the carbon cycle module that represent sensitivities to atmospheric temperature (Smith et al., 2018; Leach et al., 2021).

In Chapter 3, we introduced two metrics that can be used to to characterize the biogeochemical response to CO_2 emissions, the transient response to cumulative carbon emissions

(TCRE) and the Zero Emissions Commitment (ZEC). Both quantities are measurements of temperature change in response to emissions and constitute the most important components of the remaining carbon budget (RCB) (MacDougall, 2016; Rogelj et al., 2019; Damon Matthews et al., 2021). The ZEC is the amount of global mean temperature change predicted to occur after emissions completely cease, and the TCRE describes the proportionality between CO_2-induced warming and cumulative CO_2 emissions up to the same point. By definition, the TCRE does not account for the delayed warming response to CO_2 after emissions reach zero, and the ZEC only accounts for this delayed response. Taken together, the TCRE and the ZEC can describe the full climate response of an emissions pathway, when calculated from the same scenario (Krasting et al., 2014; MacDougall, 2019; Koven et al., 2022).

Previous work constraining the TCRE has found that Earth system models generally exhibit higher TCREs than best estimates from observations. This discrepancy has been attributed to strong negative radiative forcing from aerosol emissions, relatively low diagnosed cumulative emissions in models (from concentration-driven experiments), in addition to strong CO_2-induced and non-CO_2 warming in models (Millar and Friedlingstein, 2018). The role of the ZEC has only recently received more attention related to its contribution to the RCB. This is partly due to the approximation that continued warming after emissions cease would be balanced by continued removal of atmospheric CO_2 by carbon sinks (IPCC SR1.5, 2018).

The fact that the ZEC is an empirical quantity that emerges from the coupling of climate and the carbon cycle gives it powerful applicability, but makes its physical meaning difficult to (a) interpret and (b) distinguish from other "committed" warming metrics such as the "constant emissions" and the "constant composition" commitments and (c) characterize with meaningful uncertainty (Palazzo Corner et al., 2023; Damon Matthews et al., 2021). Using simulations in the UVic ESCM, Ehlert and Zickfeld (2017) show that because thermal and biogeochemical processes are connected, the ZEC can be interpreted as a balance between the timescales of biogeochemical relaxation and thermal equilibration. They find that thermal equilibration has a stronger influence on the ZEC and that the rate of emissions reduction toward zero has a strong control on its magnitude. Previous quantifica-

tion of the ZEC relied on experimental setups that included unrealistic termination shocks (in atmospheric CO2 concentration), and differences in inferred emissions pathways (IPCC Working Group I, 2021a). In response, recent efforts have been made to stress the importance using idealized emissions-forced experiments to quantify the ZEC (Koven et al., 2022; Jones and Friedlingstein, 2020; Sanderson et al., 2023). New idealized model experiments show that the ZEC can be partially realized before net-zero emissions in a ramp-down emissions scenario, illustrating that the ZEC can be understood to be the deviation from the path-independence of the TCRE relationship than as a measure of the committed warming after reaching net zero (Koven et al., 2023). Finally, a recent review of the processes that might influence the ZEC shows that processes that influence both our ability to constrain future ocean heat uptake, the ocean biogeochemical cycle, the land carbon cycle, and physical climate feedbacks will have bearing on both TCRE and ZEC, with implications for the climate mitigation planning Palazzo Corner et al. (2023).

It can be seen both in the broader context and from the results of this analysis that while the TCRE and ZEC may be quantified separately, they are very entangled parts of a total climate response to cumulative emissions.

In this chapter, we seek to understand which aspects of our representation of the carbon-climate system within simple climate models exert control on carbon sinks and how they subsequently influence carbon budgets for climate mitigation. To do this we focus on two versions of the FaIR model, the first is FaIR as from the published literature (FaIR_default), and the second is a modified version of FaIR that we created called FaIR_bgc. FaIR_bgc has an identical energy balance component to FaIR_default, but the simple gas cycle model from FaIR has been replaced with a 9-box terrestrial carbon cycle and a 7-box ocean carbon cycle (Parton et al., 1987, 1988; Toggweiler, 1999) (see further discussion in Chapter 3). We compare these two SCMs, and investigate the role of both carbon cycle representation and links between carbon cycling and heat uptake of the ocean.

In its default version, FaIR removes carbon from the atmosphere according to four timescales, $\alpha\tau_i$. If the fixed characteristic timescales of removal, τ_i, were representative of different carbon sinks, we could compare the allocation carbon between the different pools. Smith et al. (2018) offer some interpretation of the four timescales: the longest timescale

($\tau_0 = 10^9$ years) can be interpreted as a carbon sink from rock weathering; the next timescale ($\tau_1 = 394.4$ years) as carbon taken up into the deep ocean; the next ($\tau_2 = 36.54$ years) as carbon taken up into the terrestrial biosphere; the shortest ($\tau = 4.304$ years) as carbon uptake into the surface ocean reservoir. However, as described by Leach et al. (2021), these timescales of uptake do not formally correspond to representations of carbon cycle reservoirs, but were diagnosed from a series of pulse experiments that contained carbon cycle representations and are emergent timescales. Neither do these pools correspond to comparable land or ocean sink reservoirs in comparison to the FaIR_bgc reservoirs. In contrast FaIR_bgc contains these reservoirs explicitly, and the processes that control the cycling of carbon between the atmosphere, land, and ocean reservoir (including photosynthesis, respiration, gas exchange in the ocean, and circulation), are explicitly represented through parameterizations.

This comparative analysis is done because the construction of FaIR is such that energy balance, and by consequence the model's temperature evolution, is represented separately from carbon cycling and concomitant carbon sink evolution. This divided construction, described below in *Methods*, is not unique to FaIR. Other SCMs, including HECTOR and MAGICC, also use separate modules with independent parameterizations of carbon cycling and energy balance to represent the carbon-climate system.

In this chapter we first compare carbon cycle processes to thermal processes in their influence on carbon sink rates. We then examine how differences in carbon sink rates lead to differences in climate in an idealized decarbonization scenario. As a final step, we examine how differences in the biogeochemical representation lead to differences in long-term climate.

4.2 Methods

We use the *esm-flat10-zec* decarbonization scenario to quantify the TCRE and ZEC using a perturbed parameter experimental approach. For each model, a unique perturbed parameter ensemble (PPE) is generated according to the structure of the carbon cycle. Constructing the PPE allows us to sample across the range of climate responses within each model structure and to further investigate which aspects of the carbon cycle exert the most influence on the carbon cycle response.

After identifying any differences in the long-term temperature response to emissions, we investigate differences in the carbon cycle representations themselves. To do so, we fit exponential memory functions that reproduce the timescales of each models' total carbon sink. This approach allows us to compare across model structure and systematically identify which elements of carbon cycle timescales contribute most to differences in carbon cycle response to decarbonization.

Finally, we would like to identify to what extent the differences in carbon sink timescales influence long-term temperature evolution response to decarbonization and if there are differences between models.

4.2.1 Models

We use FaIR_default and FaIR_bgc in this chapter.

FaIR_default contains an energy balance model for computing global temperature and a so-called "gas cycle" model for computing the evolution of greenhouse gas and aerosol constituents in the atmosphere. The models are linked through atmospheric concentration – effective radiative forcing (ERF) relationships for each constituent. We modify the default version of FaIR (FaIR_default) to create FaIR_bgc by removing its carbon cycle module and replacing it with a more complex carbon cycle module.

Briefly, FaIR_default represents total carbon uptake by the global sink as exponential decay using four timescales of carbon removal operating on emissions equally (see further description in Chapter 3). FaIR_bgc represents carbon uptake from the atmosphere by land with a nine-box carbon cycle model and by ocean with a seven-box carbon cycle model (see more description in Chapter 3).

We devote this section to describing the thermal component of FaIR (which is used in both FaIR_default and FaIR_bgc) and the parameters that are perturbed within both models.

To compute surface temperature from ERF, it uses a three-layer energy balance model described in detail in Cummins et al. (2020). These models are linked through forcing from atmospheric CO_2 concentration (as well as other greenhouse gases).

Each of the three vertically stacked layers has a temperature, T, and heat capacity, C. The transfer of heat from Layer 1 to Layer 2 and subsequently to Layer 3 represents the represents the thermal inertia of the mixed layer ocean and deep ocean and as a result determine the timescales of warming in each layer (see Figure 1 of Leach et al., 2021). The parameters within the energy balance model control the transfer of energy between the three layers of the energy balance model. They include `clim_kappa1`, which is the "climate feedback" parameter, denoted as κ_1 or λ more often and is the functional inverse of climate sensitivity, such that an equilibrium response to radiative forcing, F can be written as the linear relation between GMST anomaly, T_1 and TOA net downward radiative flux change, N: $N(t) = F(t) - \kappa_1\ T_1(t)$ (Gregory et al., 2004; Cummins et al., 2020; Leach et al., 2021).

The heat transfer coefficients, κ_2 and κ_3 control the transfer of energy between the first and second and the second and third layers, respectively.

$$C_1\ \frac{dT_1}{dt} = F(t) - \kappa_1\ T_1(t) - \kappa_1\ (T_1 - T_2) \tag{4.1}$$

Another parameter within the energy balance model that contributes to the variance in the TCRE and ZEC is ϵ, the "efficacy factor" parameter, which adjusts the temporal variation in the relationship between T_1 and N. This efficacy factor is based on Held et al. (2010).

$$N(t) = F(t) - \kappa_1\ T_1 + (1-\epsilon)\kappa_3(T_2 - T_3) \tag{4.2}$$

$$C_2\ \frac{dT_2}{dt} = \kappa_2(T_1 - T_2) - \kappa_3(1-\epsilon)(T_2 - T_3) \tag{4.3}$$

In this chapter, as in Ch. 3, we replace the FaIR_default carbon cycle module with a more complex carbon cycle model. We refer to this coupled construction in this chapter as FaIR_bgc, to indicate the explicit representation of biogeochemical processes as its parameters. Briefly, FaIR_bgc cycles carbon through its 9-box terrestrial carbon cycle box model and its 7-box ocean carbon cycle model. The land model is based on the CENTURY land carbon cycle model (Parton et al., 1988) and uses parameterizations of photosynthesis by leaves and respiration by litter and soil to exchange carbon with the atmosphere. The ocean model is a reproduction of Toggweiler (1999), with gas exchange at four surface boxes, and

a prescribed circulation and transport of carbon between boxes (see further description in Chapter 3).

Perturbed energy balance and carbon cycle parameters in FaIR_default

We use FaIR's calibrated, constrained ensemble, version 1.2.0 (Smith, 2023) to determine the bounds of eleven parameters that determine the model's climate response to emissions. These include eight parameters that are used in the FaIR's three-layer energy balance model and three parameters used in FaIR's CO_2 gas cycle. Three additional energy balance model parameters are also perturbed in both ensembles but are used only when the stochastic variability component is activated in FaIR. We do not use stochastic variability in our forcing, and as a result, these parameter changes have no significant effect. We thus do not include them in our experimental descriptions.

The three carbon cycle parameters, denoted with *cc_* prefixes, are used to calculate the lifetime adjustment parameter, α, at each time step as described by Leach et al. (2021). The parameters we modify include `cc_r0`, which is the strength of preindustrial uptake from the atmosphere, `cc_ru`, the sensitivity of uptake to the cumulative sink of CO_2, and `cc_rT`, the sensitivity of uptake to temperature change since the beginning of the simulation. We sample across values for all parameters that are used in the non-stochastic, three-layer energy balance model, described in Cummins et al. (2020) with the exception of preindustrial CO_2 concentration. The bounds of each perturbed parameter is given in Table 4.1.

We generate a 1,000-member ensemble of parameter sets sampled across these three carbon cycle parameters as well as the eight energy balance parameters using a Latin Hypercube random sampling technique.

Perturbed carbon cycle parameters in FaIR_bgc

In FaIR_bgc, in addition to the energy balance parameters described above we perturb five parameters that characterize important carbon cycle feedbacks. They include two parameters that characterize biogeochemical feedbacks in the terrestrial carbon cycle (`beta_550` and `q10_resp`) and three in the ocean carbon cycle (`kwScalar`, `PsiScalar`, and `dPsidb`).

Parameter, variable (units)	**Min. value**	**Max. value**
FaIR energy balance parameters (see Cummins et al. (2020))		
`clim_c1`, C_1 (W yr m^{-2} K^{-1})	2.02	7.77
`clim_c2`, C_2 (W yr m^{-2} K^{-1})	7.16	66.5
`clim_c3`, C_3 (W yr m^{-2} K^{-1})	26.3	371
`clim_kappa1`, κ_1 (m^{-2} K^{-1})	0.581	2.60
`clim_kappa2`, κ_2 (m^{-2} K^{-1})	0.605	8.64
`clim_kappa3`, κ_2 (m^{-2} K^{-1})	0.00933	2.80
`clim_epsilon`, ϵ (unitless)	0.0391	1.93
`clim_F_4xCO2`, $F_{4\times CO_2}$ (W m^{-2})	4.29	10.8
FaIR_default carbon cycle parameters (see Leach et al. (2021))		
`cc_r0`, r_0 (unitless)	26.9	41.7
`cc_ru`, r_U ((Pg CO$_2$)$^{-1}$)	-0.00643	0.0118
`cc_rT` r_T (K^{-1})	-2.85	8.04
FaIR_bgc carbon cycle parameters		
`beta_550`, β_{550} (unitless)	1.48	1.73
`Q10_resp`, Q_{10} (unitless)	1.51	2.48
`kwScalar`, S_{kw} (unitless)	0.531	1.98
`PsiScalar`, S_Ψ (unitless)	0.505	1.97
`dPsidb`, $S_{d\psi/db}$ (unitless)	0.204	0.974

Table 4.1: Parameter bounds for FaIR and FaIR_bgc

We describe each parameter and how it is used in the FaIR_bgc carbon cycle below:

`beta_550`, β_{550}: CO_2 fertilization is parameterized as β_{550}, a rate ratio of NPP to change in atmospheric CO_2, where

$$\beta_{fert} = \frac{\beta_{550} - 1}{\ln(550/365)} \tag{4.4}$$

$$\text{NPP}(t) = \text{NPP}_0 \times \text{NPPfac} \times (1 + \beta_{fert} \times \ln(\frac{C(t)}{C_0})) \tag{4.5}$$

The scaling factor, NPPfac, accounts for the fact that within FaIR_bgc, a seasonal cycle is interpolated. As β_{550} increases, β_{fert} also increases, increasing the response of NPP to atmospheric carbon.

`q10_resp`, Q_{10}: rate ratio of respiration of terrestrial vegetation taking place at temperatures differing by 10 °C or K.

`kwScalar`, S_{kw}: Scaling factor for piston velocity in units of meters per day, regulating the flux of CO_2 between the ocean and the atmosphere at the surface. Piston velocity in the FaIR_bgc ocean model operates on surface ocean boxes scaled by the fraction without ice, which is 0.9 for all boxes except the low-latitude surface box, which has an open-water fraction of 1.

For each surface box, i, effective kw is calculated from the scaling factor, the fractional ice coverage, F (0.1 in all surface boxes except the low latitude box, where it is 0), and the default piston velocity, kw_d, which is 3 m/s.

$$kw_i = S_{kw}\ kw_d\ (1 - F_i) \tag{4.6}$$

$$F_{gasx,i} = kw_i\ A_i\ K_{sol}\ (C_{w,i}\ -\ \alpha\ C_a) \tag{4.7}$$

F_{gasx} is the air-sea gas exchange flux, and A is the surface area of each ocean box.

`PsiScalar`, S_ψ: Scaling factor for circulation. Carbon is circulated between ocean boxes at transport rates representing large-scale circulation patterns representative of the North Atlantic Deep Water (NADW) overturning, Antarctic bottom water overturning (AABW), Thermocline overturning, and convective mixing. Each as specified overturning rates, and all scaled uniformly by S_ψ.

`dPsidb`, $S_{d\psi/db}$: Climate response of ocean circulation. This sensitivity parameter scales overturning in the ocean to the total buoyancy gradient (between the Antarctic bottom water and low-latitude surface boxes).

It should be noted that while the carbon cycle model contains boxes with spatial relations as well as ice coverage, the atmosphere is uniform, and when passing information to the energy balance model, no spatial information is included. Thermal calculations are performed separately and do not involve any considerations except for ERF from CO_2 concentration.

Additional perturbed parameter ensembles

In addition to the standard perturbed parameter ensembles for FaIR_default and FaIR_bgc we create two modified parameter ensembles for FaIR_bgc.

First, we explore the behavior of FaIR_bgc if the parameter controlling heat transfer between the surface ocean and deep ocean, denoted by κ_3 is held fixed, while all other parameters are varied. We remove variation in FaIR's κ_i parameters by setting a constant value for each κ parameter, taken as the median value from the top 100 ensemble members as constrained by only CO_2 concentration. For all other parameters in FaIR_bgc, we construct a second Latin Hypercube perturbed parameter ensemble. This experiment is denoted as 'cnst_k'. We run both the historical and *esm-flat10-zec* experiments and calculate three-timescale exponential memory functions for each ensemble member constrained to both historical CO_2 and historical temperature.

Second, we explore the behavior of FaIR_bgc if the parameter controlling heat transfer between the surface ocean and deep ocean, denoted by κ_3, is correlated with the ocean circulation rate, S_Ψ. To do so, we conduct a second experiment in which we correlate the variation in `PsiScalar` parameter and κ_3 parameter values. This experiment is denoted as 'corr_kp'.

To construct the correlated ensemble, we remove the `PsiScalar` parameter from the Latin Hypercube parameter sampling and generate a set of `clim_kappa3`. We use the slope computed from the upper and lower bounds for each parameter (see Table 4.1) to set a relationship between the two parameters. Once we have a range of parameter values for

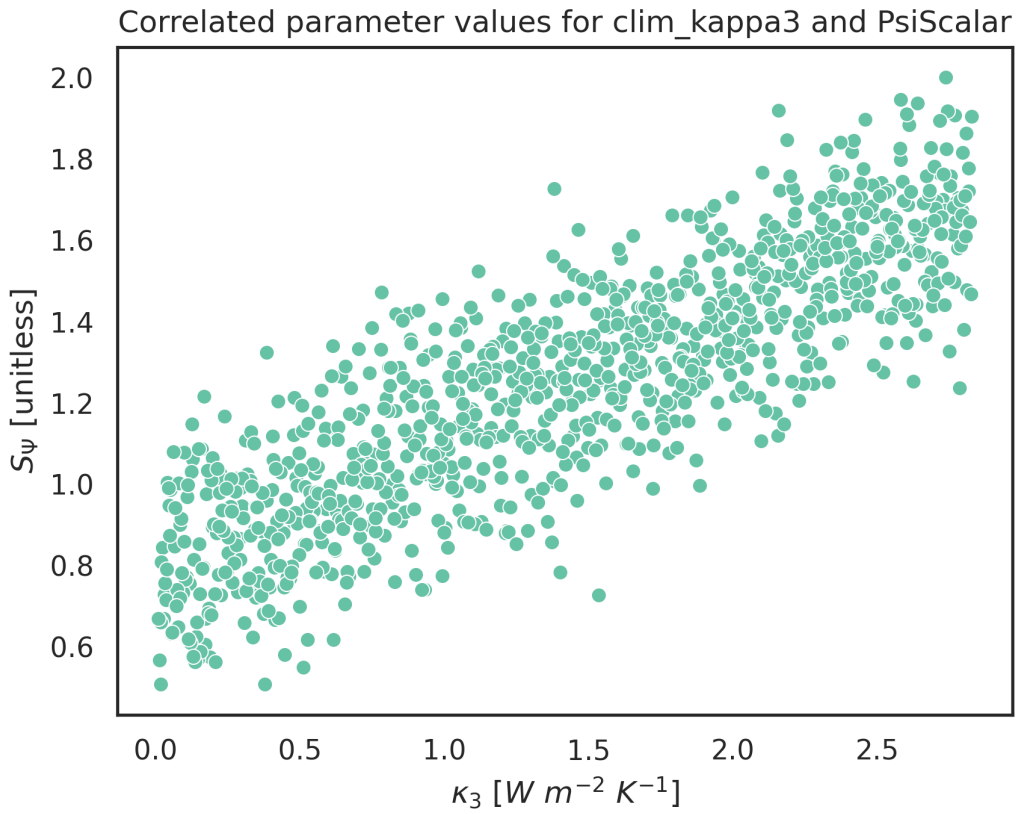


Figure 4.1: **Correlated parameter values.** Parameter values within correlated `PsiScalar` and κ_3 parameter ensemble used in FaIR_bgc-corr_kp experiment.

`clim_kappa3` from the Latin Hypercube, in order for `PsiScalar`'s values to stay within the parameter bounds, we create a temporary `clim_kappa3` array by normalizing to a smaller range (+/- one standard deviation of the total `clim_kappa` range). From this normalized array, we calculate `PsiScalar`'s values using the computed slope and add noise to the `PsiScalar` values.

The correlation in parameter values can be seen in Figure 4.1. We run historical and *esm-flat10-zec* experiments, and again calculate three-timescale exponential memory functions for each ensemble member constrained by both CO_2 and temperature.

In all experiments we include constant concentrations of CH_4 (808.24 ppb) and N_2O

(273.02 ppb).

4.2.2 Calibrating and constraining the ensembles

There are many ways to constrain models to observations, and as models advance, the methods used to calibrate and constrain them will evolve. We take a straightforward approach, and for each of the ensembles, we tested three observations to constrain our ensemble to parameter combinations that match the historical period.

Constraining by global mean temperature trends

First, we constrain the ensembles using historical trends in global mean surface temperature. We use the observed temperature record from Forster et al. (2023). As is done with FaIR (Smith et al., 2018; Smith, 2023), we select the model-driven temperatures that fall in observational uncertainty following Thompson et al. (2015), which compares the regression slope of the modeled temperature to the slope of observed temperatures using a 90% confidence interval. The method includes autocorrelation between years. Observed temperature change is also driven by non-CO_2 forcing, but in our simulations, we hold CH_4 and N_2O concentrations constant and do not drive the historical simulations with any other emissions forcing except CO_2. Thus, while we do not expect the modeled temperature change to match up completely to observations, to account for non-CO_2-driven warming, we constrain for climate feedback parameters that produce climate responses that match our record of CO_2 emissions with observed warming. Using this method, 341 FaIR_default and 411 FaIR_bgc ensemble members fall within this temperature trend constraint.

Constraining by atmospheric CO_2 growth rate

Second, we constrain the ensembles to a range of uncertainty in records of atmospheric growth rate (GR). We use 2.5 times the reported uncertainty in CO_2 GR from NOAA GML between 1959 and 2021. Reported ranges of uncertainty are higher for the period between 1959 and 1979, but we compare our total uncertainty to total uncertainty throughout the period of record. 418 FaIR_default and 603 FaIR_bgc ensemble members fall within this

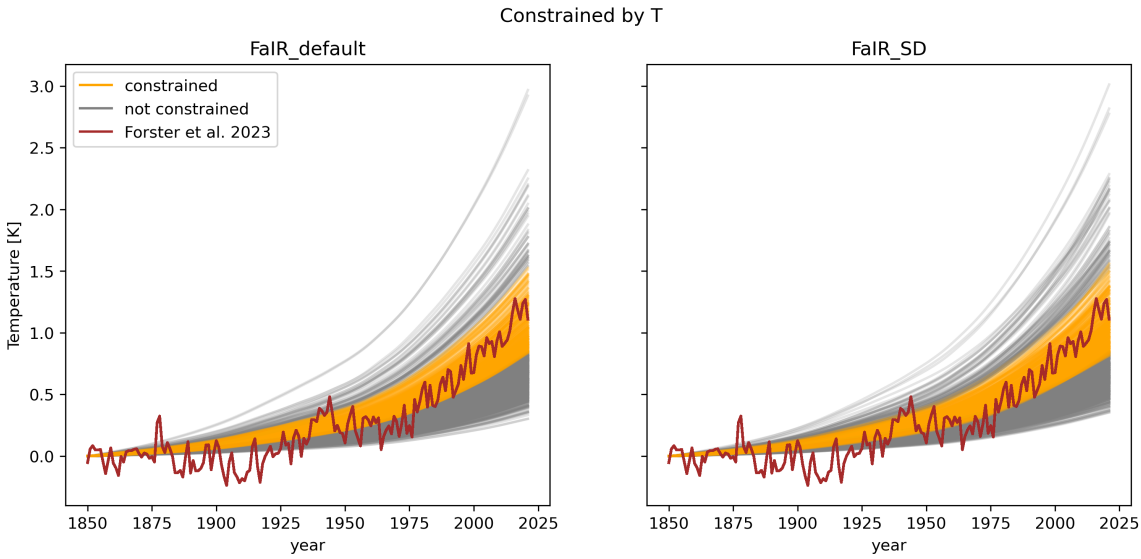


Figure 4.2: **Constraining to observed temperature.** Timeseries of historical emissions-driven temperature change (relative to 1850) for FaIR and FaIR_bgc during the period of record, 1850 to 2021. For both models, ensemble members that fall within a 90% confidence interval around the mean regression slope are indicated in orange, and ensemble members that fall outside this range are shown in grey.

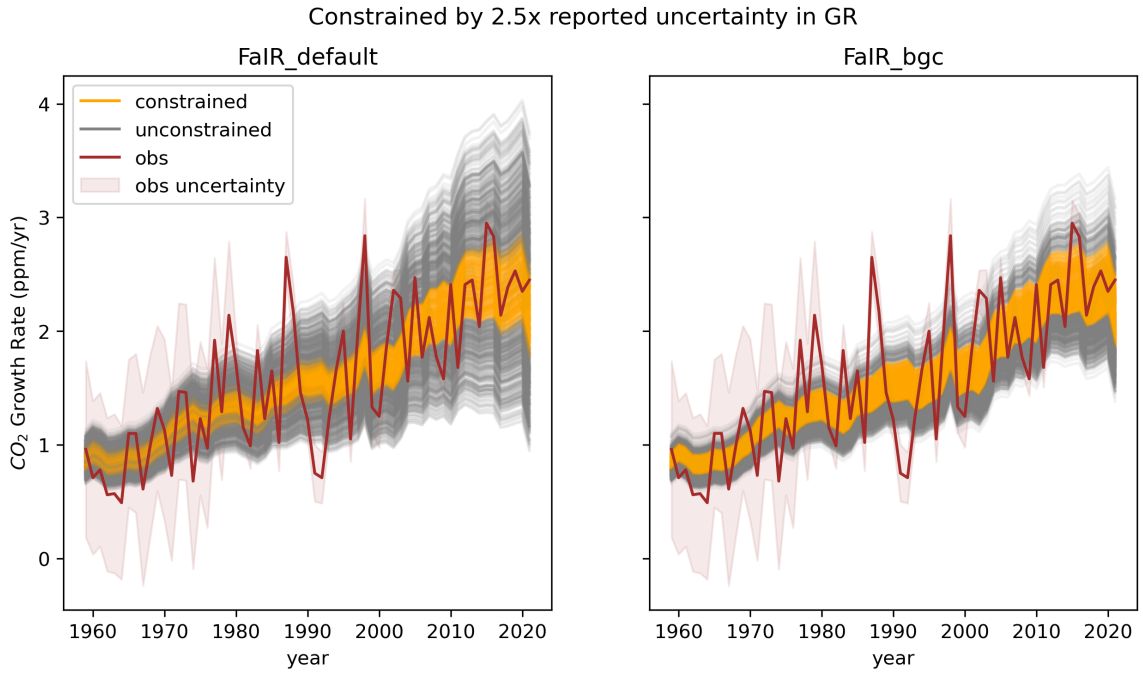


Figure 4.3: **Constraining to observed CO_2 growth rate.** Timeseries of historical emissions-driven growth rate of atmospheric CO_2 for FaIR and FaIR_bgc during the period of observation, 1959 to 2021. For both models, ensemble members that fall within a constraint of 2.5 $\times$ the reported uncertainty in CO_2 GR are shown in orange, and ensemble members that fall outside this constraint are shown in grey.

CO_2 growth rate constraint.

Constraining by atmospheric CO_2

Third, we constrain the ensemble to within a range of uncertainty in records of atmospheric CO_2 mixing ratio as reported by NOAA's Global Monitoring Laboratory (GML) as driven by historical emissions.

While the atmospheric CO_2 record is known to high precision, records of global CO_2 emissions have larger levels of uncertainty. As in chapter 3, we use historical emissions as reported by the Global Carbon Project between 1979 and 2021 (Friedlingstein et al., 2022;

Forster et al., 2023). Additionally, we drive each of our models using only emissions of CO_2 from fossil fuels and land use change, the real atmospheric CO_2 record will be influenced by more than just CO_2 emissions. As a result, we constrain our ensemble to within 100 times the uncertainty in NOAA GML reported global mean CO_2, a somewhat arbitrary limit. The annual average reported uncertainty is 0.06 ppm. Our limit allows for an annual error of about 6 ppm. Despite the inflated uncertainty bounds CO_2 mixing ratio was the most restrictive of our constraints. Each ensemble begins with 1,000 members. 310 FaIR_default ensemble members and 452 FaIR_bgc ensemble members ensemble members fall within the CO_2 concentration constraint.

For our analysis we want to match the historical behavior of both energy cycling and carbon cycling. Therefore we select the ensemble members that satisfy both the temperature and the CO_2 concentration constraints, which results in 111 ensemble members for FaIR_default, and 177 ensemble members for FaIR_bgc. This is a departure from what has done in prior work, where CO_2 has not been used as a constraint on parameters Smith et al. (2018); Meinshausen et al. (2009). More recent approaches to calibration of FaIR now incorporate a CO_2 constraint (Smith et al., 2024).

4.2.3 Simulations

As described above, we perform a historical simulation, using global emissions data between 1850 and 2021 from Friedlingstein et al. (2022) for both fossil fuel (FF) and land use change (LUC) to constrain the ensembles.

After selecting parameter combinations which fit within our tolerance of historical observations of temperature trends and CO_2 concentration, we conduct decarbonization simulations using the *esm-flat10-zec* protocol that prescribes a constant rate annual emissions at 10 Pg C/year (as FF) for 100 years, followed by 1000 years at 0 Pg C/year. The *esm-flat10-zec* experiment provides an elegant experimental design for quantifying the TCRE since the first 100 years produce a linear increase in cumulative emissions.

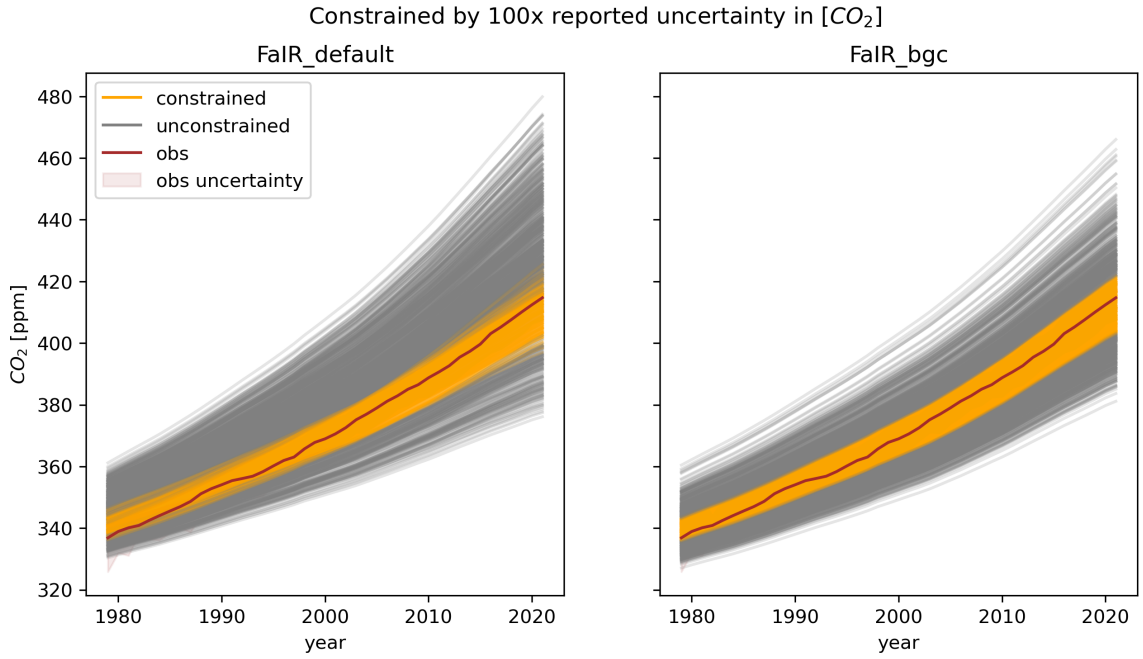


Figure 4.4: **Constraining to observed CO_2 concentration.** Timeseries of historical emissions-driven atmospheric CO_2 concentration for FaIR and FaIR_bgc during the period of CO_2 concentration observation, 1979 to 2021. For both models, ensemble members that fall within a constraint of 100 $\times$ the reported uncertainty in CO_2 concentration are shown in orange, and ensemble members that fall outside this constraint are shown in grey.

4.2.4 Analysis

Single exponential memory function

In order to compare carbon sinks across models, we define a set of simple exponential memory functions that relate sink strength to prior emissions. To first order, we define a single exponential memory function as:

$$S(t) = \int_0^t E(t_e)\, a\, e^{\frac{-(t-t_e)}{\tau}}\, dt_e \tag{4.8}$$

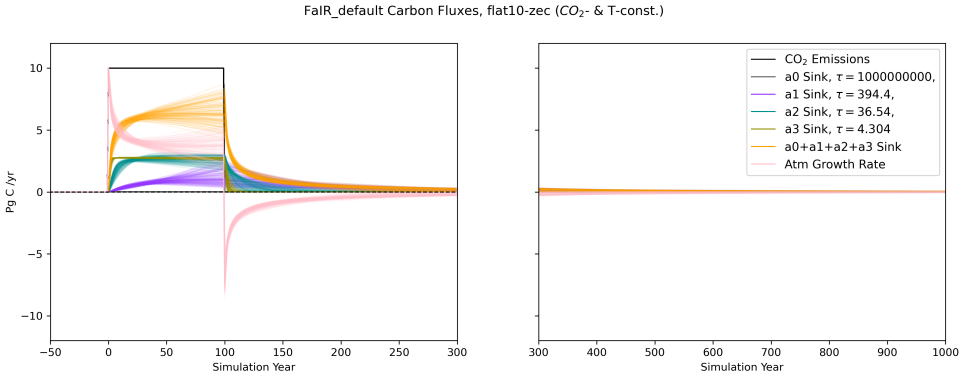
We use a non-linear least squares optimization to fit the memory function to a single ensemble member and yield τ_i and a_i parameters for the function in Equation 4.8. Our optimization method uses the Levenberg–Marquardt algorithm to fit data to non-linear function using a damped least-squares minimization (Virtanen et al., 2020; Vugrin et al., 2007). This method allows us to characterize individual ensemble members' total carbon sink timescales, and we apply it to ensemble members constrained by atmospheric CO_2 concentration and temperature change for each model.

Diagnosed exponential decay timescales are used to describe each model's total sink

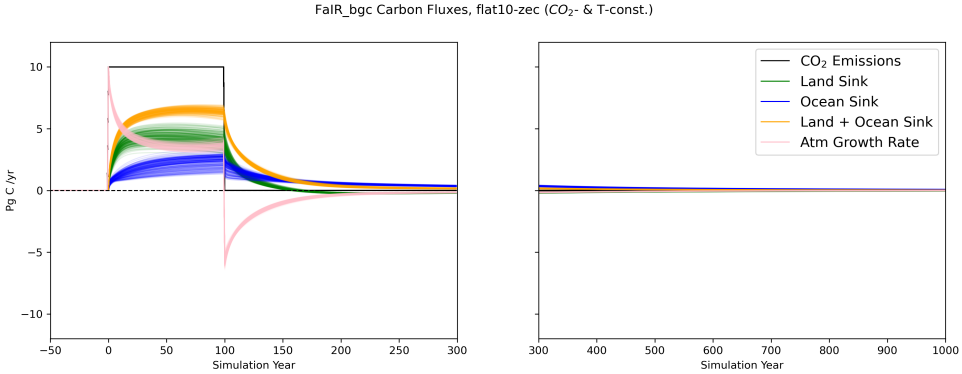
Using the same optimization, we fit additional predictions for total carbon sink using a linear sum of exponential functions. We use a 3-timescale memory models fit to *esm-flat10-zec* sinks given in Equation 4.9.

$$S(t) = \int_0^t E(t_e) \left(a_1\, e^{\frac{-(t-t_e)}{\tau_1}} + a_2\, e^{\frac{-(t-t_e)}{\tau_2}} + a_3\, e^{\frac{-(t-t_e)}{\tau_3}} \right) dt_e \tag{4.9}$$

We find that a 3-timescale memory function is sufficient to reasonably reproduce the timeseries of total carbon removal for both FaIR and FaIR_bgc both in the emissions phase and decarbonization phase (after year 100). Figure 4.10 shows the comparison between various exponential fits. While FaIR itself uses a four-timescale removal of carbon, the longest timescale removes little carbon from the atmosphere even over 1000 years. The fit triple exponential timescales have functionally the same form as the original FaIR structure



(a) FaIR carbon fluxes for 111 ensemble members constrained by CO_2 and temperature.



(b) FaIR_bgc carbon fluxes for 177 ensemble members constrained by CO_2 and temperature.

Figure 4.5: **Carbon fluxes in *esm-flat10-zec* emissions scenario.** Timeseries of CO_2- and temperature-constrained *esm-flat10-zec* carbon fluxes for each model's reservoir are shown. FaIR was modified to compute these fluxes, since only the total sink is computed by default. The x-axis is divided at 300 years to expand the first 200 years.

cycle if the diagnosed sensitivity parameters sum to approximately 3/4, i.e. $a_i \approx 0.25$ in Equation 4.9.

TCRE and ZEC

To calculate the TCRE, we regress annual mean temperature on cumulative emissions over the first 100 years of the *esm-flat10-zec* simulation, which are the only emissions years. In this chapter we report the TCRE in most cases as the TCRE1000 (K), which is warming expected from 1,000 Pg C (i.e. slope multiplied by 1,000 PgC). Quantifying the TCRE this way allows us to more effectively compare the changes in temperature due to the TCRE and the ZEC as temperature changes due to the same emissions.

The ZEC is the projected global mean temperature change relative to the temperature at net-zero emissions and can be diagnosed at different points in time following the cessation of emissions. The ZEC at year X is calculated as the average temperature change over the 20 years surrounding year X relative to the TCRE-implied warming. For example, to calculate the ZEC50 from the *esm-flat10-zec* experiment, we take the average temperature from 40 to 60 years following the year of net-zero emissions (100 years after simulation start) and subtract the TCRE1000. Though not expected to be a large factor in this experimental design, this formulation accounts for any expression of the ZEC before emissions reach cessation (Koven et al., 2023).

4.3 Results & Discussion

4.3.1 Carbon pools

We examine the differences between how models allocate carbon to their various carbon sinks over the historical period. We first focus on the two primary models FaIR_default and FaIR_bgc. In section 4.3.6 we discuss the variants with altered parameter distributions.

To compare carbon fluxes and total pools across models, we calculate the inferred reservoir allocation from for each $\alpha\tau_i$ within FaIR_default, and if we consider FaIR_default's shortest timescale to correspond to an ocean reservoir, $\alpha\tau_3$, the FaIR_default's total ocean carbon flux would be larger than FaIR_bgc by about 50 Pg C by 2021; whereas if the fast

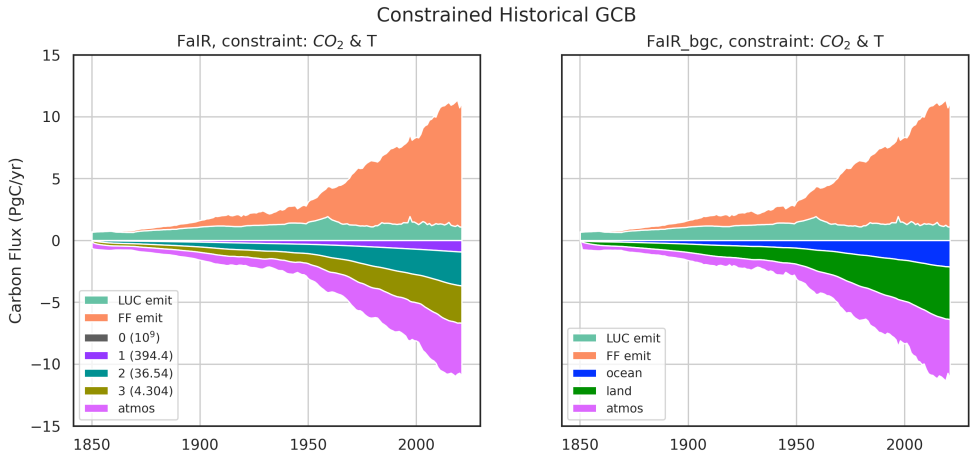


Figure 4.6: **Comparison of historical global carbon budgets between FaIR_default and FaIR_bgc.** Carbon sources and sinks for historical simulations as annual carbon fluxes in Pg C/yr. Positive values are carbon sources and negative values are carbon sinks. Both models produce realistic airborne fractions of emissions (shown in pink) as an outcome of different individual sink rates.

timescale is interpreted as being driven by land processes, FaIR_default's total land reservoir would be similarly larger than the land pools in FaIR_bgc.

Figure 4.6 shows annual carbon fluxes for the mean of each model's CO_2- and T-constrained ensembles. ESMs generally estimate about a third of historical carbon emissions have been taken up by the land sink and a third by the ocean sink (Friedlingstein et al., 2023). Recent model assessments suggest that ESMs are limited in their ability to accurately represent processes that govern the land carbon sink especially, identifying observational uncertainties, as GPP and respiration-related biases that exist across models (Seiler et al., 2022). While the models are able to reproduce the airborne fraction of CO_2 (and are constrained to do so in this case), they differ in how carbon is removed from the atmosphere. FaIR_bgc estimates a stronger land carbon sink than ocean sink in ensemble mean, with land making up about 67% of the total integrated sink of carbon since the preindustrial time period.

4.3.2 TCRE and ZEC computed from esm-flat10-zec

We estimate the TCRE and ZEC from a set of *esm-flat10-zec* simulations using parameter combinations constrained to match CO_2 and trends in temperature over the historical period.

Our estimated TCRE is consistent across models with both FaIR_default and FaIR_bgc displaying a mean TCRE of 1.2 K/1000 PgC with positive skew. We find little to no significant difference between models in TCRE (Figure 4.8).

The TCRE is a metric of the climate response to increasing cumulative emissions, an emissions regime that is part of the observed historical period. These models have been developed and tuned in the context of increasing emissions and tuned to reproduce observations of the real world carbon-climate system in an increasing-emissions regime. Therefore it might be unsurprising to see that these simple models produce a similar TCRE. What is less certain is how models behave in a decarbonization regime, as measured by the ZEC, which is so far not part of our observational record.

Since the ZEC is intended to characterize the evolution of temperature after decarbonization, we diagnose it at several points after emissions reach net-zero: 20-, 50-, and 100-years afterward, as shown in Figure 4.9. The two models agree less on the value of the ZEC, exhibiting consistent differences at the 95% confidence level using a Student's t-test.

FaIR_bgc shows a mean negative ZEC, while FaIR_default displays a mean positive ZEC, indicating slight warming after emissions have ceased. The distributions of the ZEC diagnosed from each model diverge more strongly at longer timescales (Figure 4.9). Using a Student's t-test, we determine that the two constrained distributions are significantly different from each other at the 95% confidence level starting just 50 years after achieving net-zero (ZEC50).

4.3.3 Timescales of carbon sinks

To explain the differences in carbon sinks, we fit a set of increasing order exponential memory functions to the different models' carbon cycle responses, which allows us to compare the various timescales of carbon removal between models.

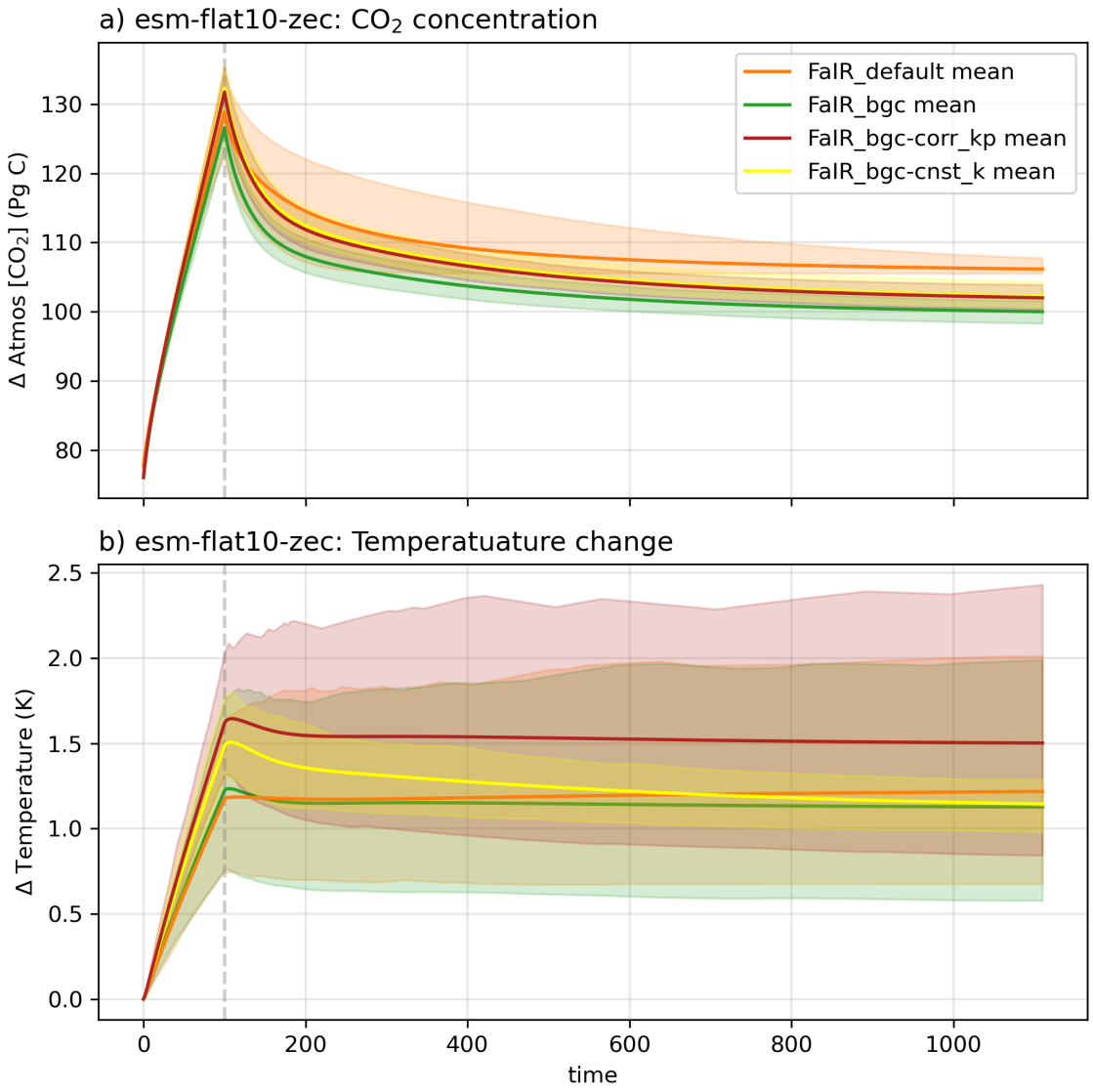


Figure 4.7: **FaIR_default and FaIR_bgc CO_2 and temperature in *esm-flat10-zec* scenario.** Top panel shows the evolution total atmospsheric CO_2 for FaIR_default and FaIR_bgc for each experimental configuration. Lower panel shows the evolution of global mean surface temperature.

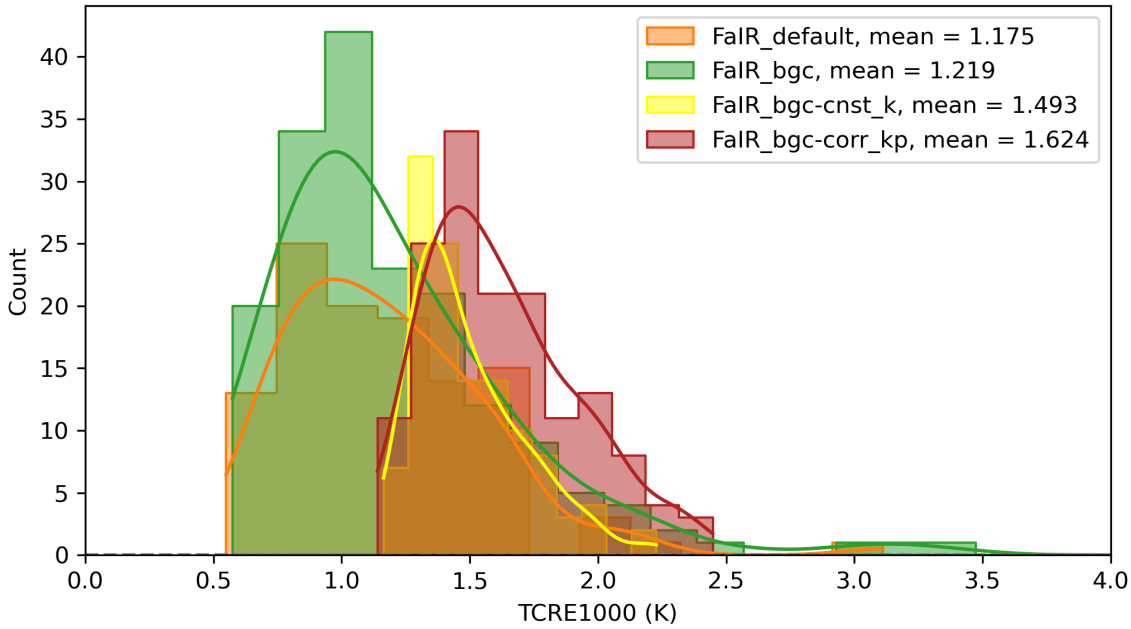


Figure 4.8: **Distributions of TCRE × 1,000 PgC cumulative emissions.** Distribution of TCRE1000 diagnosed from the *esm-flat10-zec* models and experiments, all constrained to both CO_2 and temperature change. FaIR_default is shown in orange; FaIR_bgc is shown in green; FaIR_bgc run with invariant κ_i is shown in yellow; FaIR_bgc run with correlated κ_3 and S_{Psi} is shown in red.

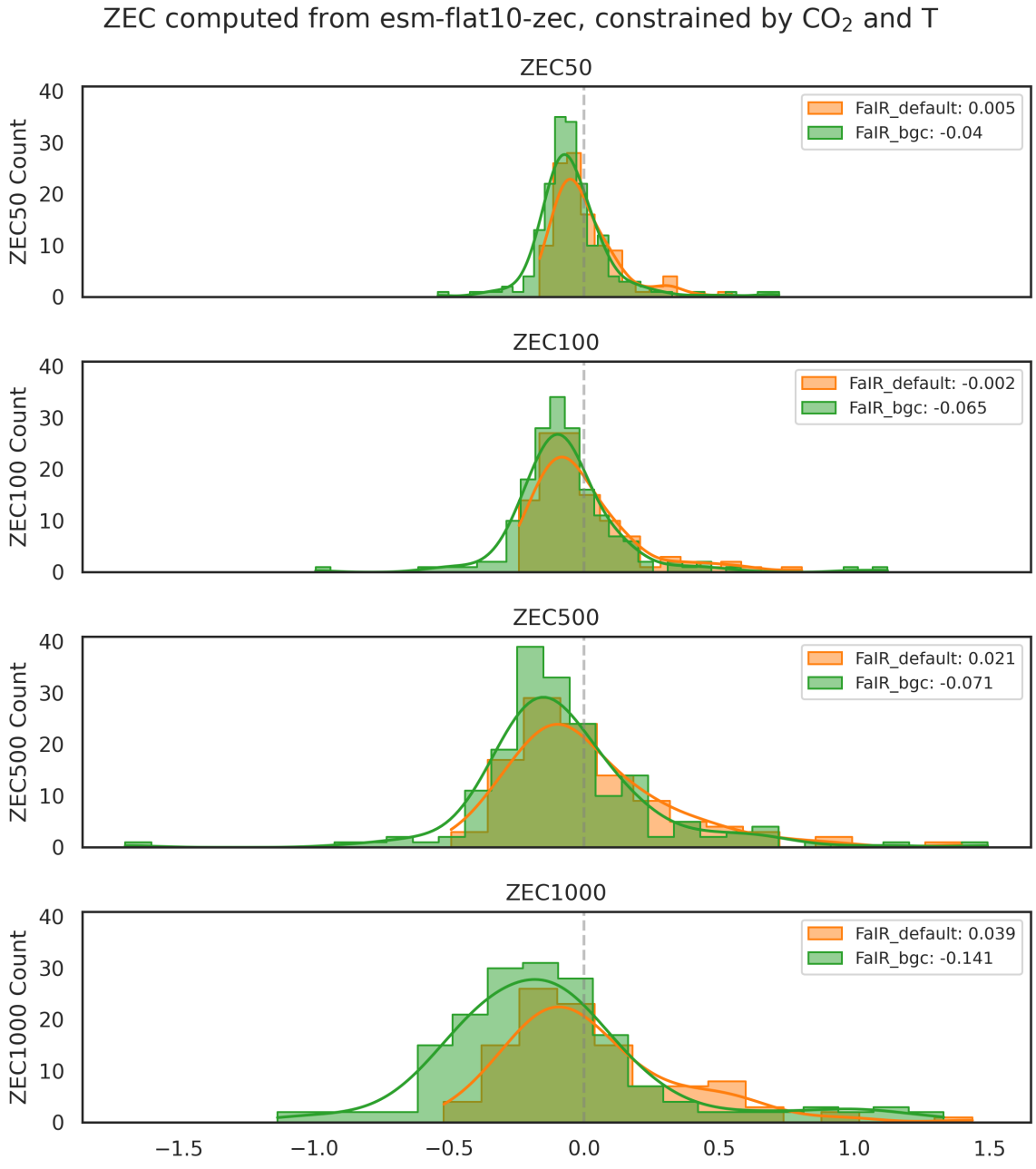


Figure 4.9: **FaIR and FaIR_bgc ZEC over time.** Panels show the distributions of ZEC (K) for each model at 50, 100, 500, and 1000 years following net-zero in the *esm-flat10-zec* emissions scenario. The legends includes the mean values for each model.

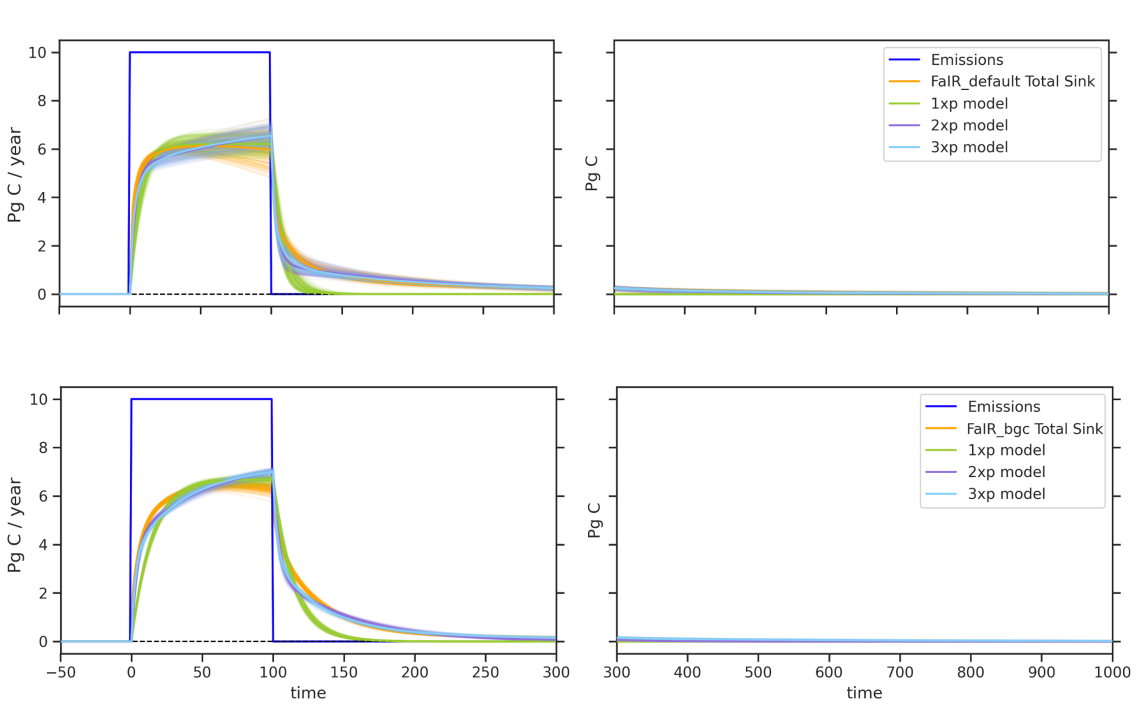


Figure 4.10: **Predicted carbon fluxes for *esm-flat10-zec* emissions scenaro using exponential curve fits.** Timeseries of total carbon sink fluxes with single-, double-, and triple-exponential curve fits using emissions and sink responses over the entire simulation time period for FaIR_default (top panel) and FaIR_bgc (lower panel).

Total sink responses in FaIR_bgc

We examine differences between diagnosing total sink and aggregated individual land + ocean timescales. While diagnosing timescales of uptake using land and ocean separately does improve the curve fit to the modeled ensemble slightly, using the total sink and fitting a memory function using the full simulation period allows us to capture the model behavior better, especially using multiple timescales of exponential decay.

We also examine the effect of applying the least-squares optimization using only the first 100 years (when emissions are nonzero) versus the whole emissions period (1110 years) to attain parameters for the exponential functions. We find that fitting only emissions-phase SCM output to emissions produces a better representation of emissions-phase sinks, but

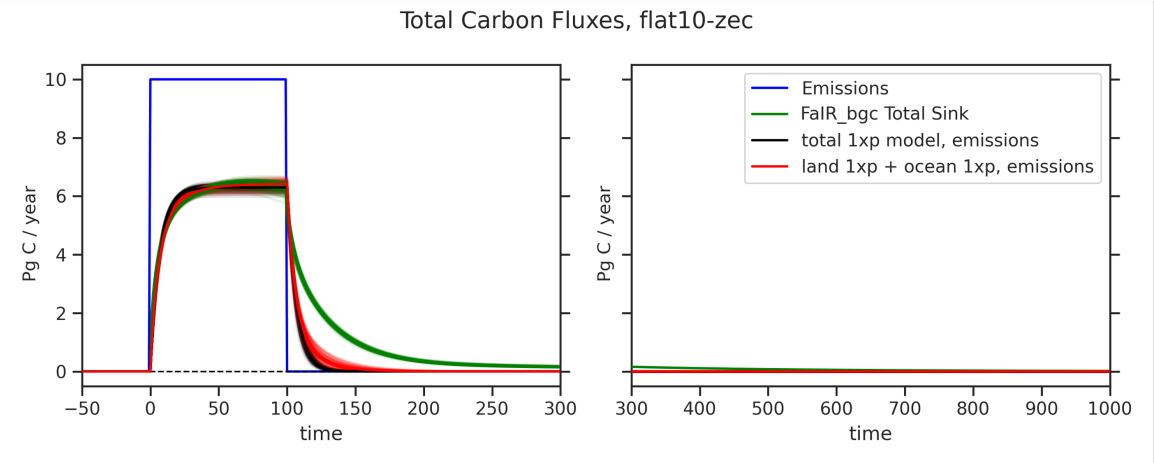


Figure 4.11: **FaIR_bgc total carbon sink fluxes fit from emissions years only**. Single Exponential fit to CO_2-constrained ensemble members using emissions and sink response over first 100 years of simulation

including the entire 1110 year time period produces a better fit during the decarbonization phase. This suggests that similar approaches to timescale diagnostics may underestimate the importance of post-net-zero sink behavior if only a short amount of emissions-phase sink data is used (as in pulse-emissions scenarios). This can be seen in Figures 4.11 and 4.12, which show the comparative fit for single-exponential fits to the top 100 ensemble members when constrained only by CO_2.

Comparison of total sink timescales between FaIR and FaIR_bgc

We estimate τ parameters from exponential curve fits for both FaIR_default and FaIR_bgc.

FaIR_default is built using timescales of removal, and we compute the α lifetime modifier as an output at each time step. We then compare the mean effective timescales of removal in FaIR_default to the curve-fit timescales. The emergent lifetimes ($\bar{\alpha}\tau_i$) generally agree with the predicted timescales of removal (Figure 4.13). In particular, our predicted τ_3 is most aligned with the longest mean emergent lifetime, ($\bar{\alpha}\tau_3$).

Figure 4.14 shows the distributions of the diagnosed timescales, τ, of removal for the

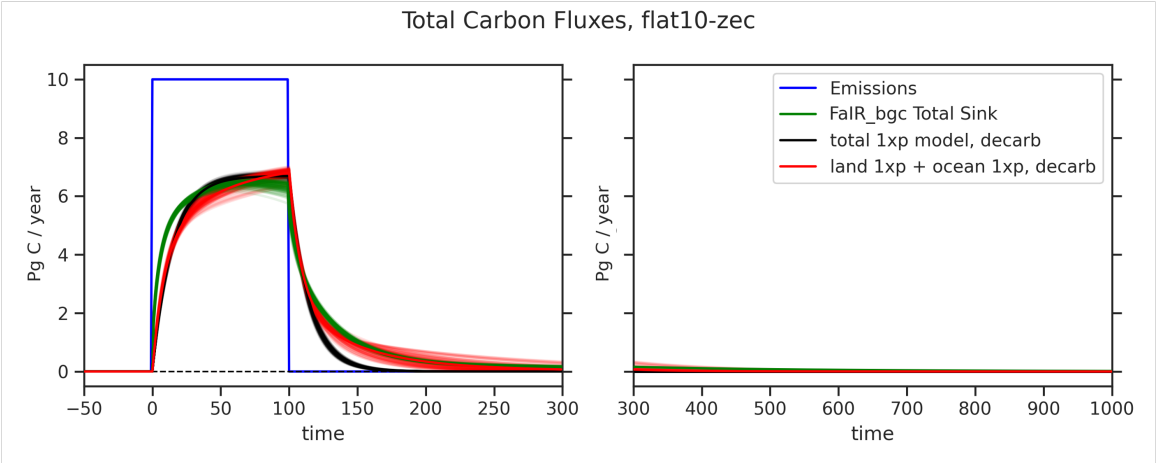


Figure 4.12: **FaIR_bgc total carbon sink fluxes fit from full simulation length**. Single Exponential fit to CO_2-constrained ensemble members using emissions and sink response over entire length of simulation

total sink. The two different carbon cycle structures embedded within FaIR_default and FaIR_bgc, even when constrained to historical CO_2 and temperature trends, produce differing distributions of values for all τs, and especially for the longest timescale of removal, τ_3.

However, overall the τ timescales are not entirely dissimilar, and we diagnose consistent ranges of carbon lifetimes in the atmosphere from two very different carbon cycle structures. This result highlights that FaIR_default's gas cycle construction may be sufficient to capture the dominant timescales that emerge from in FaIR_bgc's more complex structure.

4.3.4 Identifying controls on timing of carbon sinks and timing of warming

Our next step is to identify how each of the parameters we modified in our perturbed parameter ensembles influences the diagnosed timescales of carbon uptake in each model. Figure 4.15 shows the variance in τs explained by a linear model for each carbon cycle parameter in FaIR_bgc, and we find that these parameters are able to influence short- and long-term timescales of uptake in intuitive ways. For example, the parameter `kwScalar` adjusts a fast

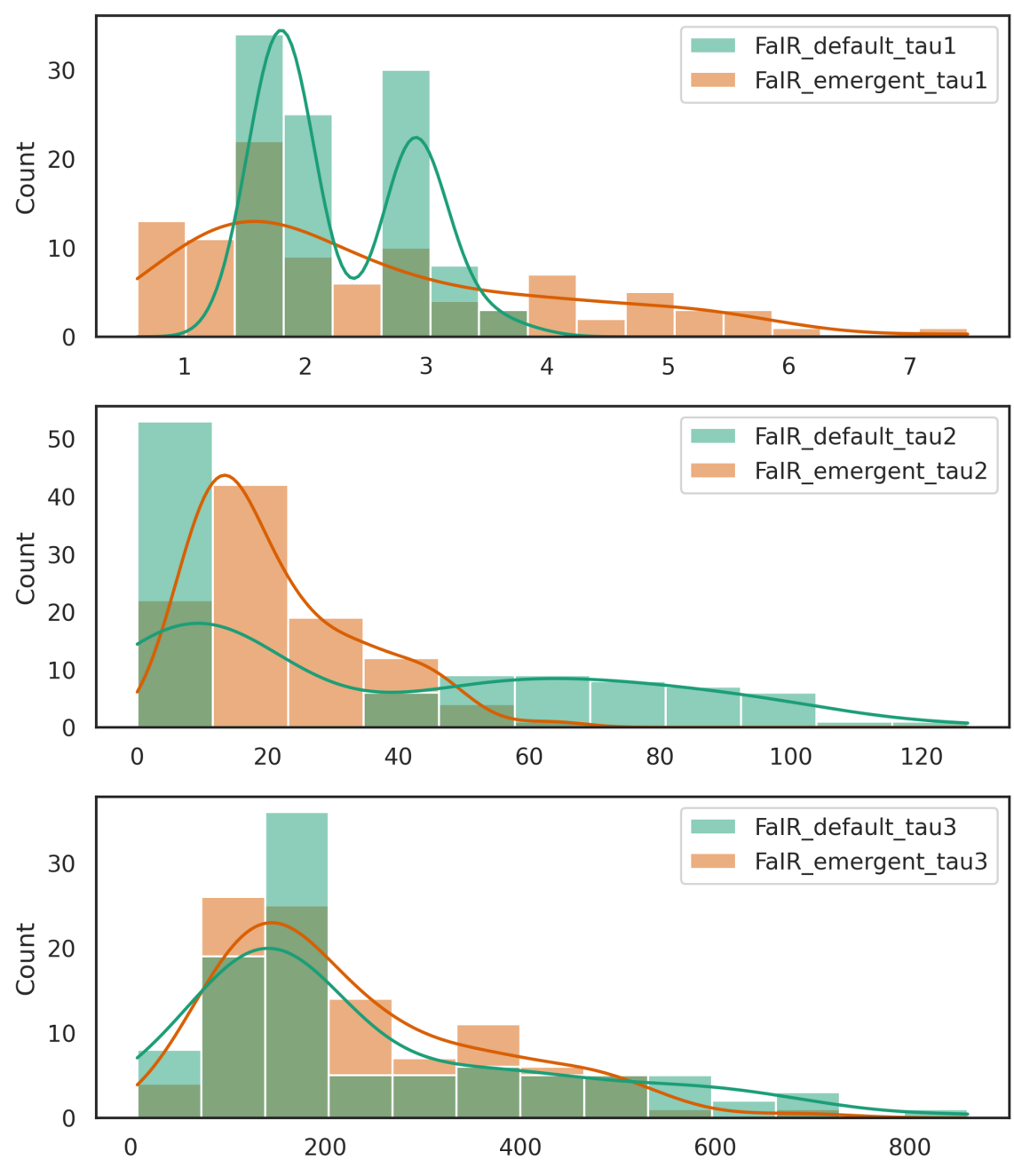


Figure 4.13: **Validation of predicted FaIR carbon sink timescales against mean calculated carbon sink timescales.** Comparison of FaIR_default predicted timescales (years) to emergent timescales (years) as calculated from the mean lifetime modifier, α, times the characteristic timescale, τ_i.

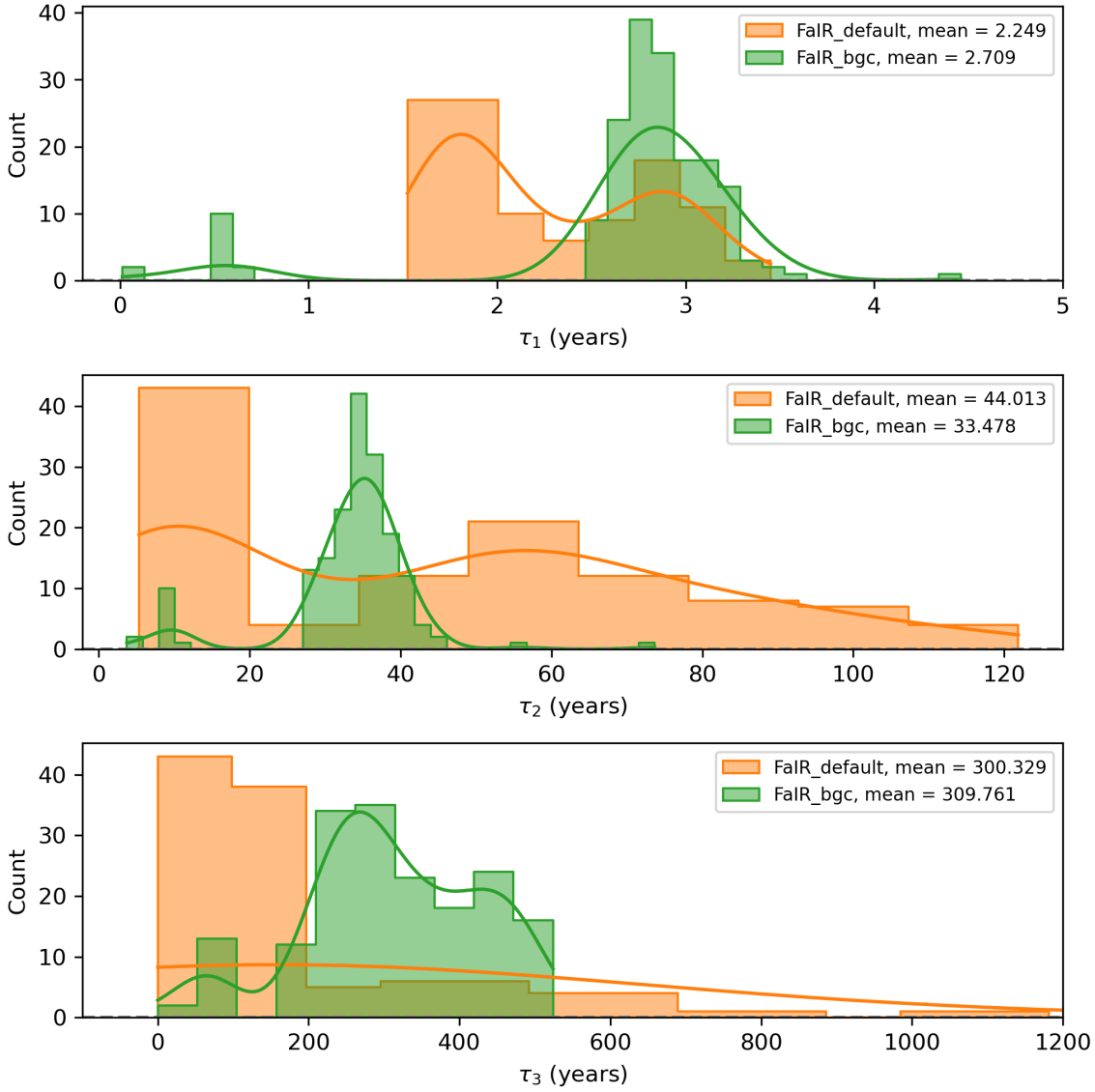


Figure 4.14: **Diagnosed timescales of carbon removal for three-timescale memory functions for both FaIR_default and FaIR_bgc.** Top panel: distribution of τ_1 with both models showing a mean of less than 5 years. Middle panel: distribution of τ_2, which have a decadal range. Bottom panel: distribution of τ_3, which have centennial range .

process, gas-exchange between the atmosphere and the ocean surface. Correspondingly, changes to its parameter value are felt most strongly in the shortest lifetime. By contrast, changes to the parameter, `PsiScalar`, explains the most variance in the longest predicted timescale. The CO_2 fertilization parameter, `beta_550` exerts the strongest influence on the longest timescale of uptake as well.

We perform the same exercise for FaIR_default's carbon cycle parameters (Figure 4.16) but gain little insight from these relationships because the lifetime coefficient operates on each characteristic τ_i uniformly, and therefore do not project strongly onto particular timescales.

We next examine the variance across our perturbed parameter ensembles in TCRE, ZEC50, and ZEC500 that can be explained by a linear relationship between each temperature metric and each carbon cycle response sink rates (τ_1, τ_2, or τ_3). We find that diagnosed timescales of carbon removal for both FaIR_default and FaIR_bgc explain little to none of the variance in warming metrics (Figure 4.17).

Rather, much of the variance in temperature metrics can be explained by parameters that are used in FaIR's energy balance model. In particular the climate feedback parameter, `clim_kappa1`, which is often denoted by λ, exhibits the strongest influence on the TCRE and ZEC (Figure 4.18).

4.3.5 Connecting FaIR's energy balance and carbon cycle modules

The three parameters included in Figure 4.18 are used in FaIR's 3-layer energy balance model (see further description in *Methods*) and together explain almost all of the variation in temperature on long timescales (as measured by the TCRE and ZEC) in both FaIR_default and FaIR_bgc. The τ parameters that we diagnose describe the rate of CO_2 removal from the atmosphere by carbon sinks, and removal of CO_2 controls the concentration of CO_2 in the atmosphere and thus radiative forcing. Although the role of heat fluxes and climate feedback rate have long been understood to play a substantial role in setting the temperature response of global climate to forcing, it surprising that the timing of CO_2 removal from the atmosphere does not have much impact on temperature metrics. In Figure 4.17, the

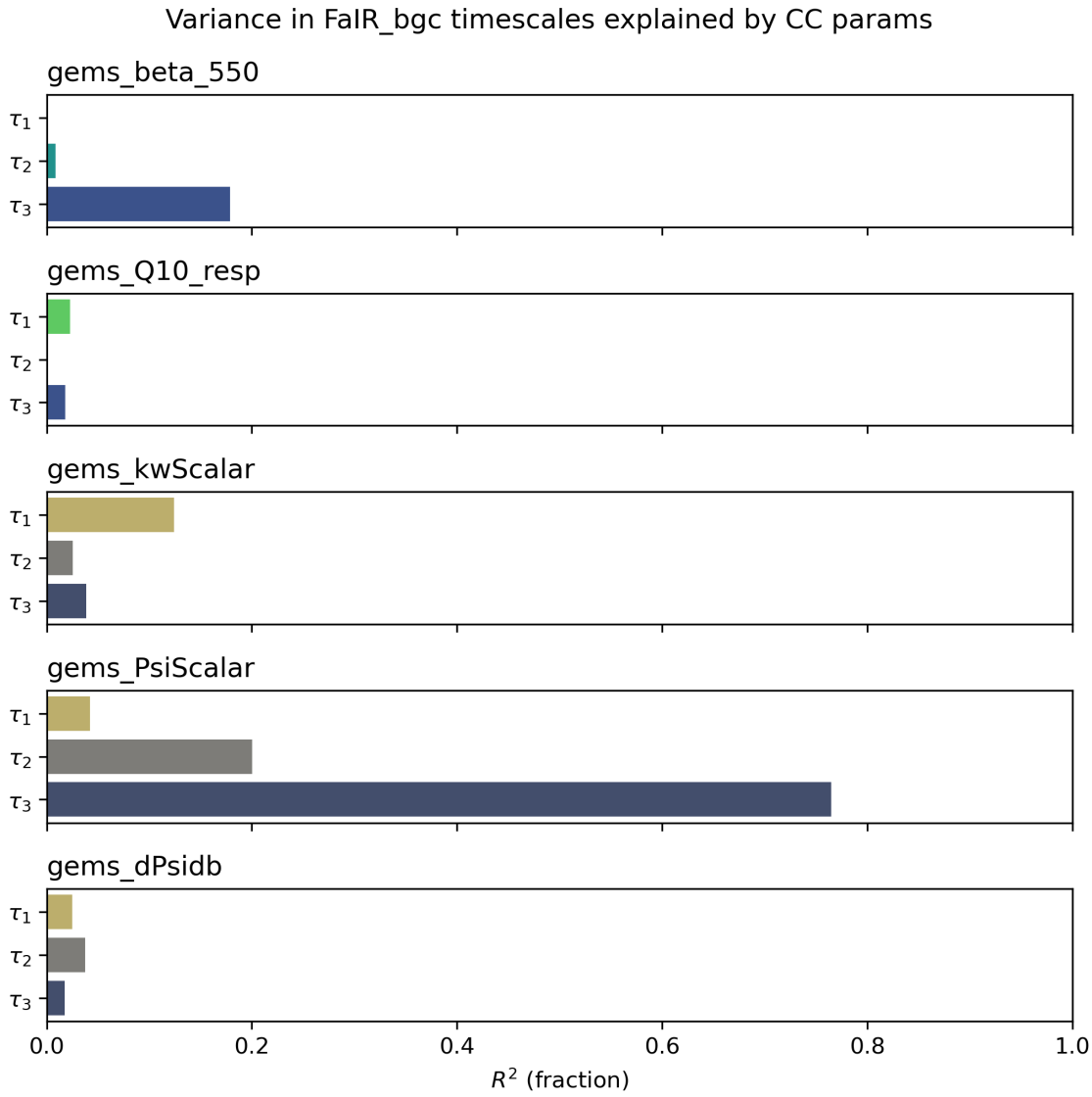


Figure 4.15: **FaIR_bgc carbon cycle parameter influence on diagnosed carbon sinks.** Variance in each timescale of carbon removal explained (R^2) by each FaIR_bgc carbon cycle parameter.

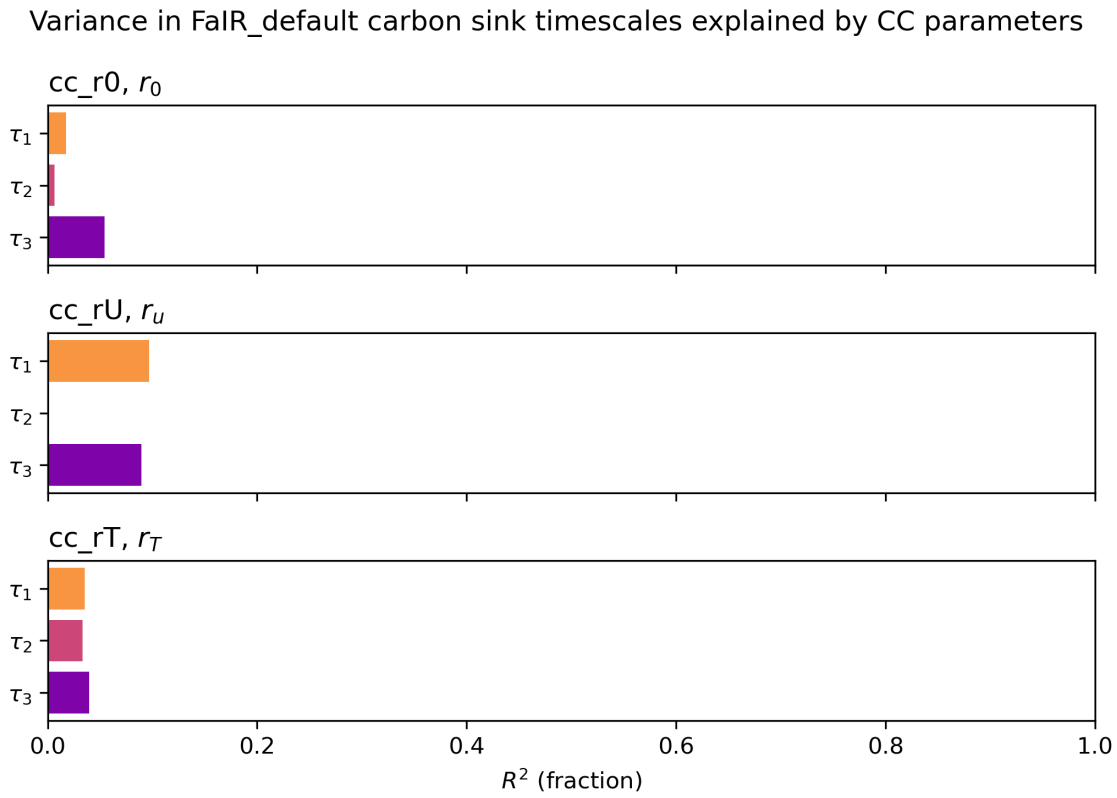


Figure 4.16: **FaIR_defalut carbon cycle parameter influence on diagnosed carbon sinks.** Variance in timescale of carbon removal explained (R^2) by each FaIR_default carbon cycle parameter.

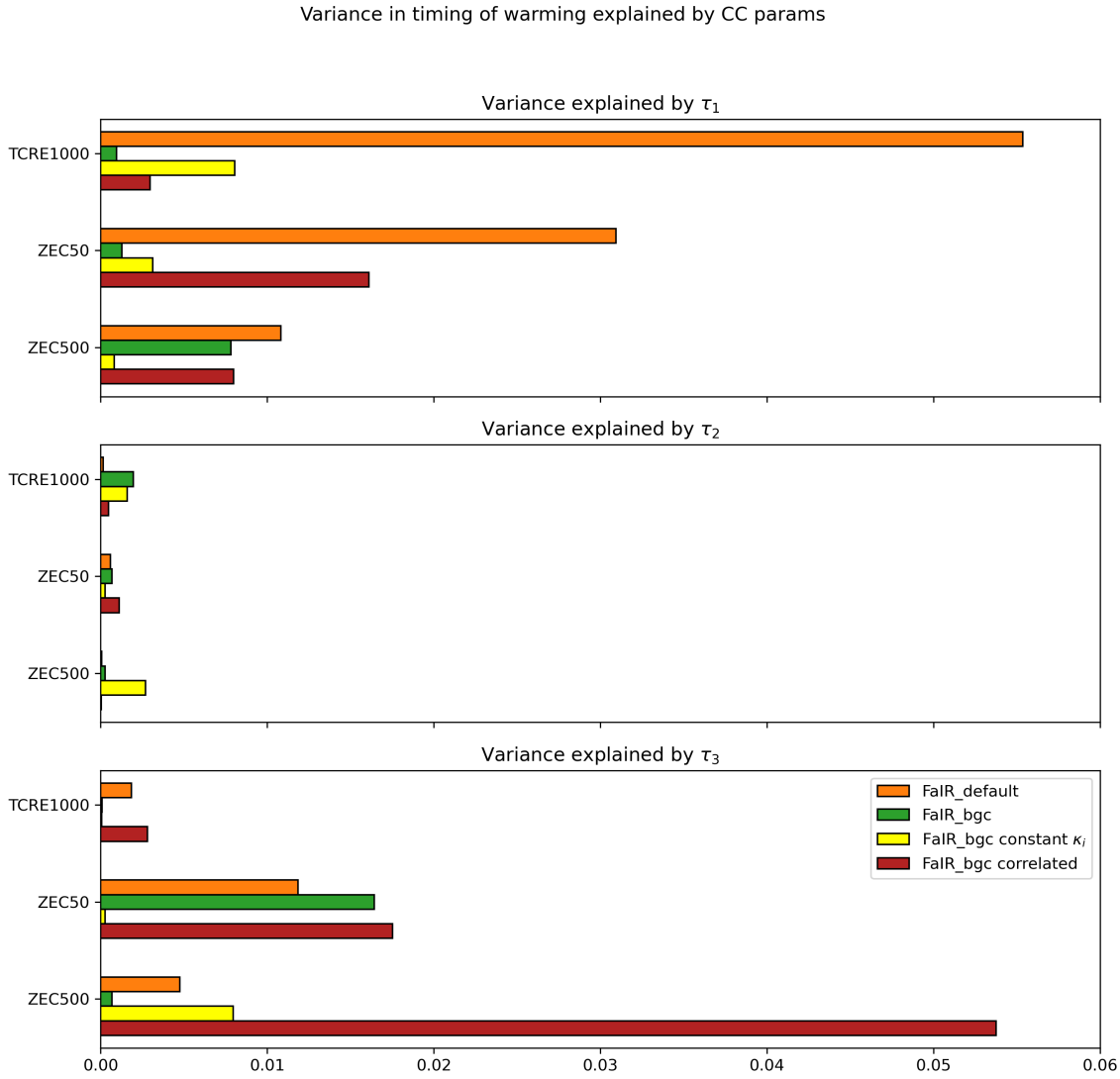


Figure 4.17: **Influence of diagnosed carbon sink timescales on warming metrics across models.** Variance in diagnosed warming metrics TCRE1000, ZEC50, ZEC500 explained by carbon cycle parameters for each experimental configuration, as measured by the linear coefficient of determination, R^2.

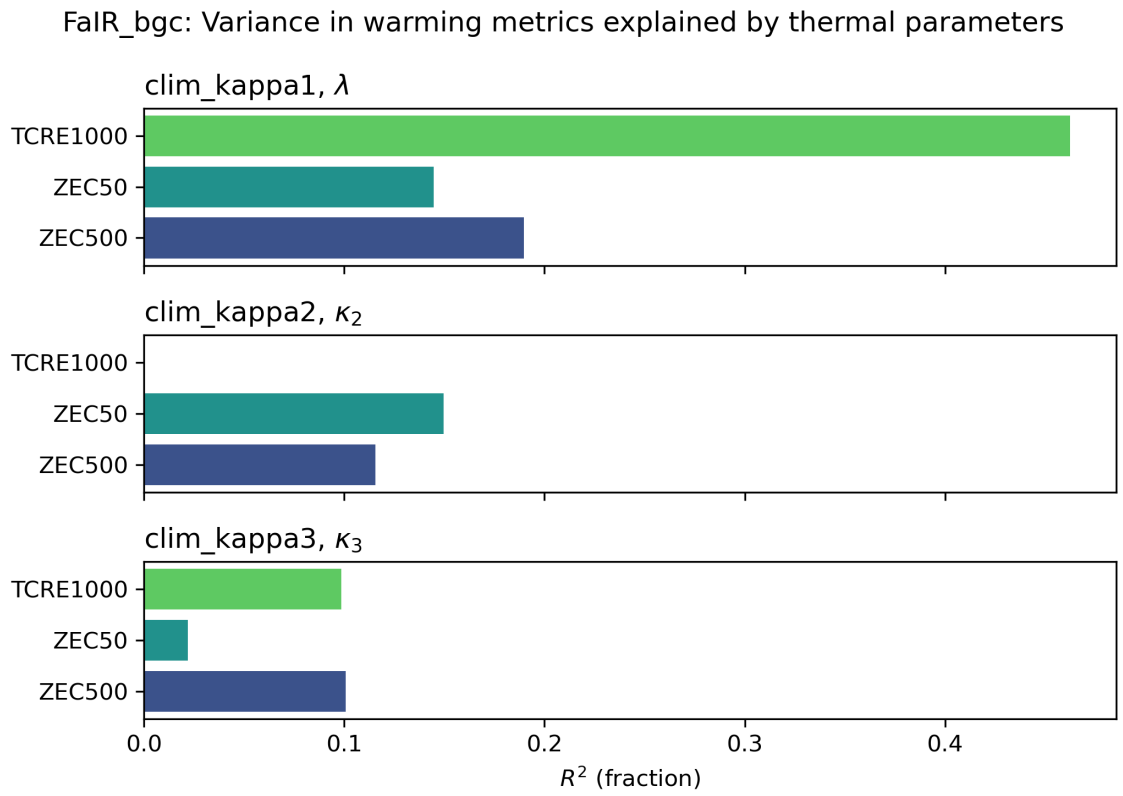


Figure 4.18: **Influence of energy balance model parameters on warming metrics in FaIR_bgc.** Variance explained by FaIR_bgc energy balance model parameters used to characterize ocean heat transfer, κ_1, the climate feedback parameter (also known as λ), κ_2, the heat transfer coefficient between the atmosphere and the mixed layer, and κ_3, the heat transfer coefficient between the mixed layer and the deep ocean.

diagnosed timescales of CO_2 uptake in FaIR_default (orange bars) and FaIR_bgc (green bars) explain only up to 5% of the variance in warming metrics, TCRE, ZE50, and ZEC500. We further explore the importance of carbon cycle processes in setting the temperature response of the climate system by linking the energy balance model and the carbon cycle model within FaIR.

What we would like to know is: is FaIR's temperature response to emissions truly insensitive the behavior of the carbon sink (Figure 4.17), or is there something about the structure of FaIR_default that dictates this response?

As we have shown above, carbon sink timescales have little influence on long-term temperature evolution in either FaIR_default or FaIR_bgc. In the following sections, we perform two sensitivity tests to identify the response of FaIR's temperature calculation to changes in carbon cycling when variability in energy balance parameters is muted and when ocean heat flux is influenced by ocean circulation rates.

We offer three explanations for why variability in carbon sink rates does not influence long-term temperature variability in the original FaIR. First, since a Latin Hypercube sampling framework samples near-randomly from the parameter space and presumes no correlation between parameters, the total range of variation in carbon cycle parameters may be overwhelmed by the total range of variation in energy balance model parameters, given the bounds we have selected for each parameter. Second, the structural separation between FaIR's energy balance model and its carbon cycle model may predetermine an artificially reduced explanatory power of carbon cycle in predicting temperature since, in reality, processes that govern biogeochemistry in land and ocean are intimately linked to processes that govern heat transport, and are not only connected through changes to atmospheric CO_2 forcing. Finally, it may be the case that FaIR's TCRE and ZEC insensitivity to carbon cycle structure illuminates the potential insight that biogeochemical processes have little bearing on the TCRE, the ZEC, and therefore the RCB.

4.3.6 FaIR TCRE and ZEC sensitivity to removing and correlating variability in heat transport parameters

We can begin to distinguish between the first and second explanation by conducting two sensitivity tests. In our first sensitivity tests, the energy balance parameters controlling heat transfer between the surface ocean and deep ocean, denoted by κ_i, are held fixed, while all other parameters are varied (cnst_k).

In the second experiment, the energy balance parameter specifically controlling heat transfer between the mixed layer and deep ocean, denoted by κ_3, is correlated with the ocean circulation rate, S_Ψ, a parameter with in the ocean carbon cycle model (corr_kp).

The justification for this correlated ensemble (corr_kp) is that large-scale ocean circulation rates should be expected to influence vertical ocean heat transport. The global overturning circulation controls the rate at which excess heat absorbed by the mixed layer is transported to depth. Since the total ocean circulation in the carbon cycle model of FaIR_bgc is scaled by a scaling parameter, `PsiScalar`, it follows that variations in this parameter should correlated with variations in the rate of heat transport in a coupled model. By correlating heat transport coefficients and ocean circulation scaling parameter, we connect the carbon-climate system and reveal connections between carbon cycle processes and long-term climate responses to emissions as measured by the TCRE and ZEC. This explanatory power emerges from the downstream counteraction between increases in long-timescale ocean carbon uptake and long-timescale ocean heat uptake, playing out a compensating mechanism that is used to explain a mean value of zero for the ZEC across ESMs (IPCC Working Group I, 2021a).

When the three energy balance parameters are held constant, we see little change in the explanatory power of carbon sink timescales on temperature metrics, as shown in Figure 4.17, which displays the coefficients of determination for a linear regression between each diagnosed timescale of total carbon uptake (τs) and TCRE, ZEC50, and ZEC500. This does not change when we add correlation between `PsiScalar` and `clim_kappa3`.

However, while we do not see a larger explanatory power of carbon removal timescales on temperature metrics, we do see shifts in the diagnosed distributions of parameter values after

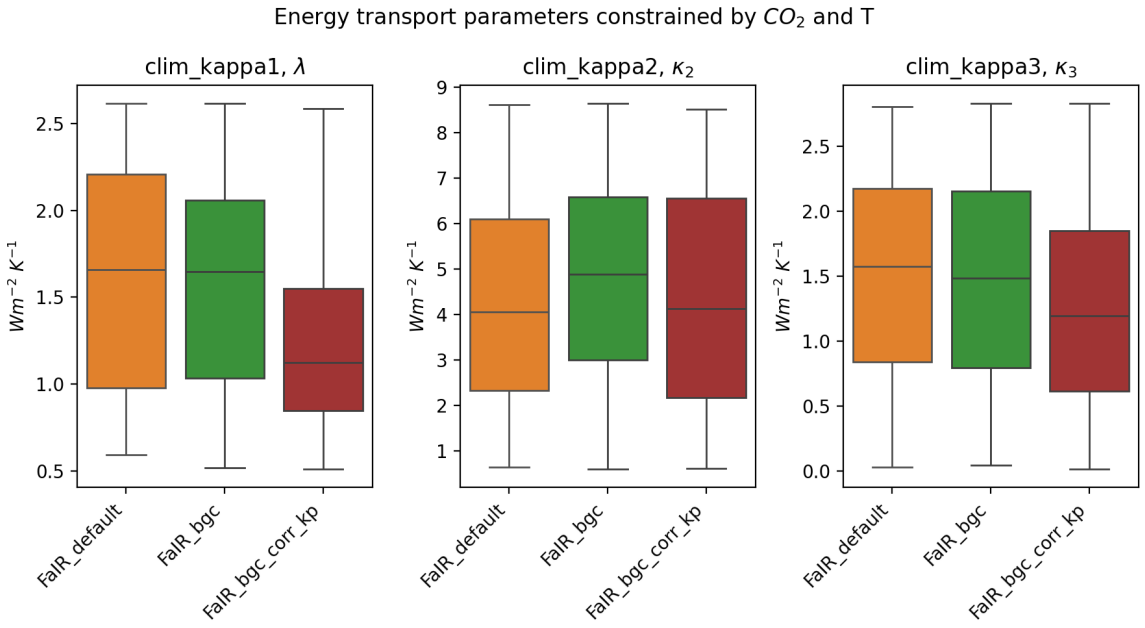


Figure 4.19: **Constrained thermal parameter distributions.** Comparison of diagnosed energy transport parameter values among FaIR_default, FaIR_bgc, and FaIR_corr-kp when constrained to CO_2 and temperature change. Each panel represents one energy balance model parameter.

constraining for ensemble members that satisfy our CO_2 concentration and temperature trend thresholds. Specifically, when `PsiScalar` is correlated with `clim_kappa3`, we see a significant shift in the TCRE toward higher temperature responses to cumulative CO_2 (Figure 4.8). This shift can be explained by a shift in the mean constrained climate feedback parameter, λ, toward lower values when we constrain the ensemble to the observed climate (Figure 4.19). For each ensemble we end up selecting for values in the energy balance parameters that compensate for these carbon-driven differences in climate.

When a high value for `PsiScalar` is correlated with high values for ocean heat uptake, a stronger ocean carbon sink is correlated with a stronger ocean heat sink. As a result, when constrained to observations, the constrained value for the climate feedback parameter shifts toward lower values (Figure 4.19), producing a mean climate that has a higher temperature

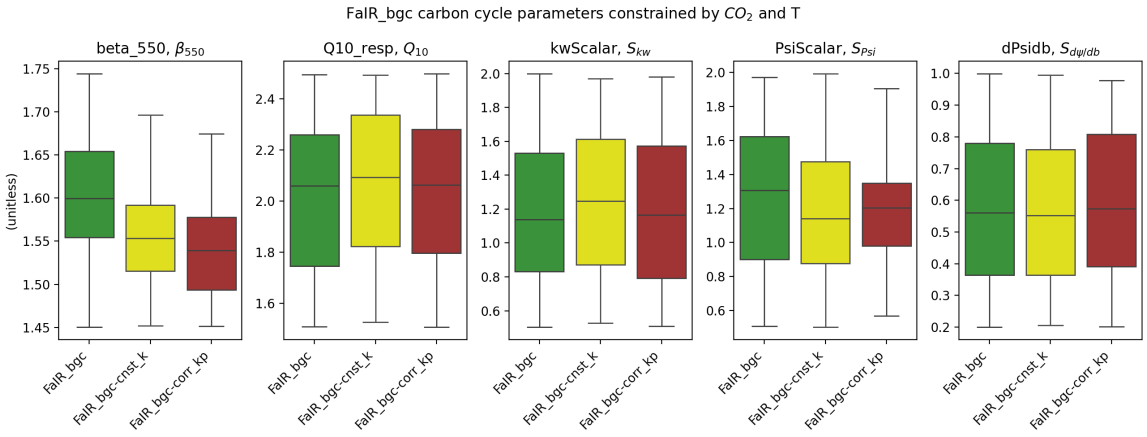


Figure 4.20: **Constrained carbon cycle parameter distributions.** Comparison of diagnosed carbon cycle parameter values when constrained to CO_2 and temperature change. Each panel represents one carbon cycle parameter.

sensitivity to radiative forcing. This can be seen in the shift toward higher TCRE values in the experiment with correlated parameters, denoted as "corr-kp", in Figure 4.8. As a corollary, the selected values for other parameters that regulate other carbon feedbacks also see a shift. In particular, the constrained values for `beta_550` shift toward lower values, indicating a weaker land sink.

Additionally, when we examine the selection of energy balance model parameters when constrained to CO_2 and temperature, in addition to the shift in the distribution the climate feedback, most notably the climate feedback parameter, we also see a shift in the influence of κ_3 on long-term temperature.

In the original FaIR_bgc experiment, increasing κ_3 leads to a decrease in global temperature, consistent with more efficient transport of heat to the deep ocean. However, in the correlated ensemble, increasing this same parameter leads to an increase in long-term temperature, despite the efficient transport of heat to the deep ocean. Figure 4.21 shows timeseries of temperature for two subsets of each ensemble.

The change in temperature response to identical emissions is due to the shift in overall climate feedbacks, which leads to the dramatic shift toward higher TCRE, overwhelming

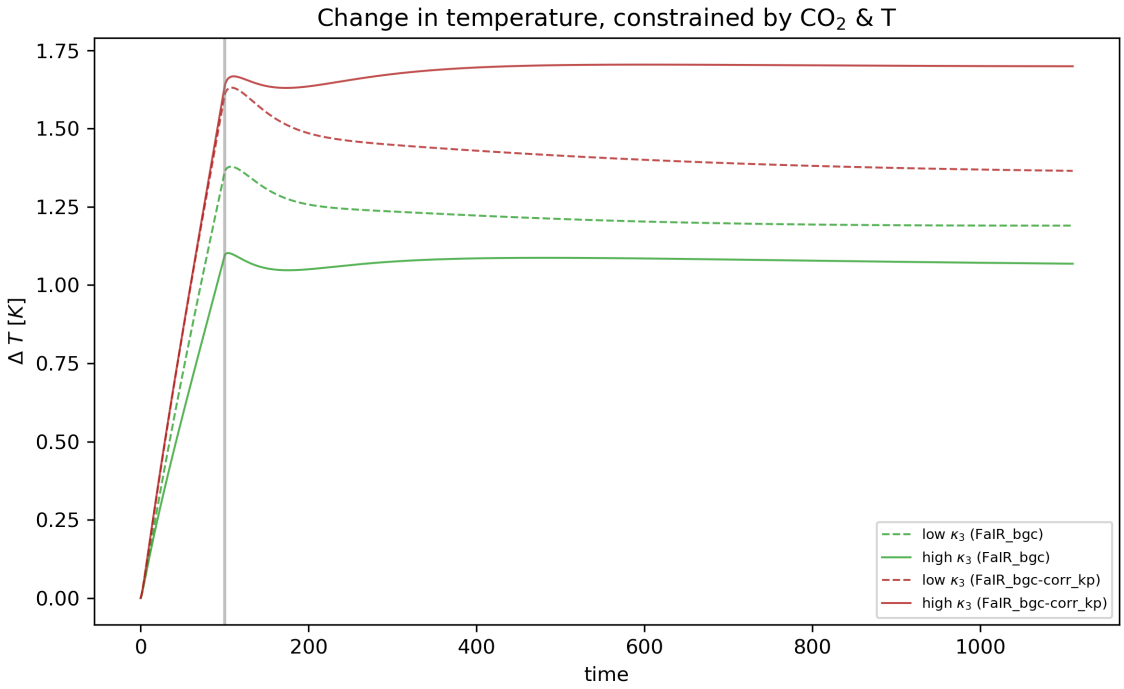


Figure 4.21: **Relative effect of increasing κ_3.** Comparative climate influence of changes to κ_3 between FaIR_bgc and FaIR_bgc_ckp under the *esm-flat10-zec* scenario. Dashed lines show the mean of all ensemble members (constrained by CO_2 and temperature) that have values for κ_3 below the parameter median value; solid lines show the mean of high κ_3 ensemble members for each experiment.

the influence of κ_3 on the ZEC which alone would tend to decrease temperature. The role of TCRE is to bring the system to a warmer temperature at the point of zero emissions, and adjustments to temperature through the ZEC are too small to compensate for this different initial temperature at the point of decarbonization.

As a result, we see a shift in what we expect for future climate in the same emissions scenario. Figure 4.22 shows the consequence of the remaining carbon budget (RCB) for the correlated ensemble (FaIR_bgc-corr_kp) in comparison to FaIR_bgc when constrained to CO_2 concentration and temperature as well as the other models assessed in Chapter 3.

The separation between energy balance and carbon cycling in SCMs is not unique to

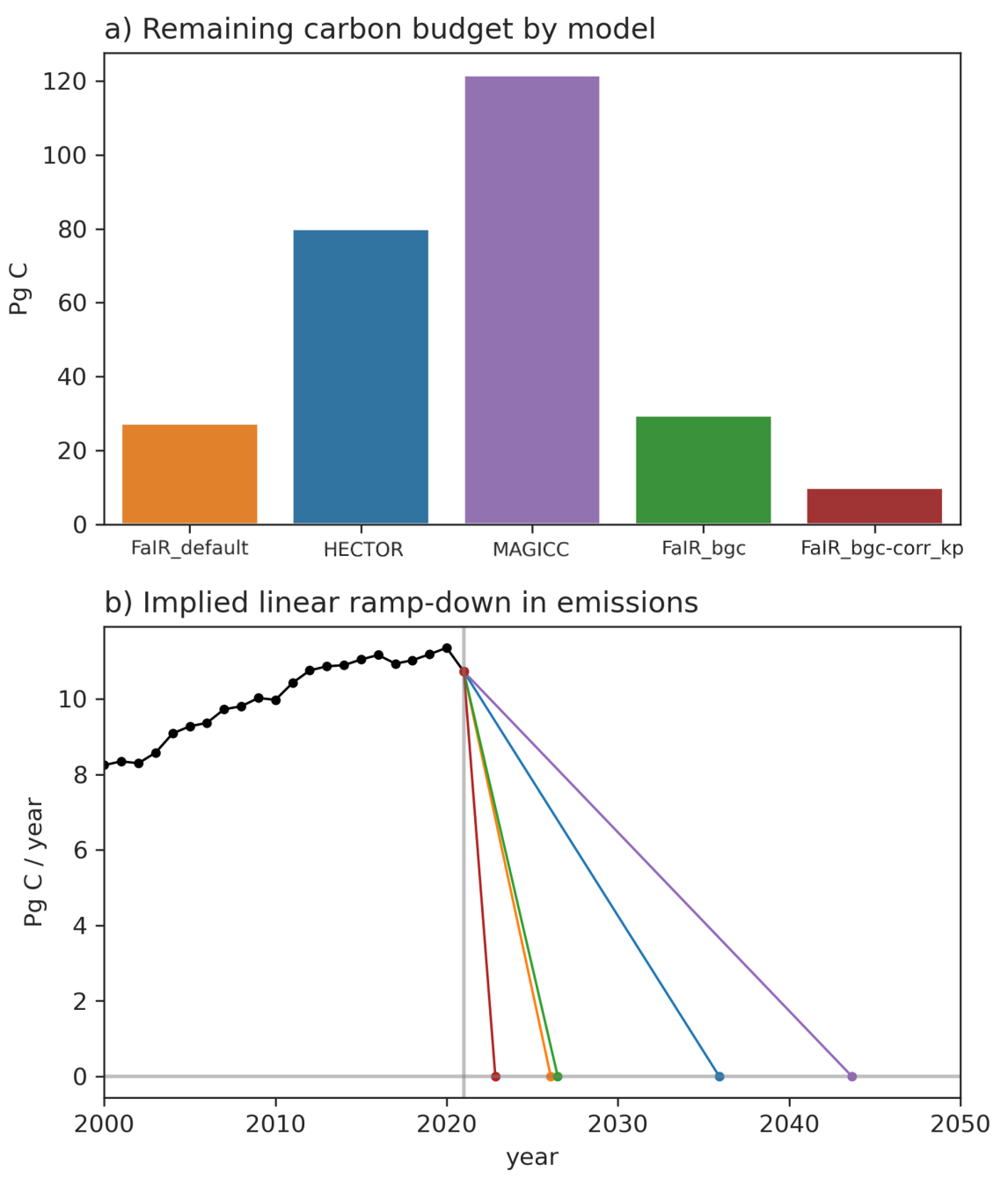


Figure 4.22: **Remaining carbon budget (RCB) for different SCMs.** **a** Remaining carbon budget as of 2021 to limit warming to 1.5°C above preindustrial temperatures. **b** Implied linear ramp-down to net-zero necessary to achieve RCB limit.

FaIR. In fact, we find that none of the common models we investigated across these chapters (FaIR, HECTOR, and MAGICC) have connections between carbon cycle processes and heat uptake processes. Connecting physical carbon-climate processes with heat uptake rates reveals existing trade-offs in our interpretation of the climate system, illustrating the need for ESM-driven constraints on processes within SCMs as well as outcomes.

Connecting the carbon cycle to energy balance through the physical dynamics of ocean circulation changes our implied climate feedbacks. Importantly, both parameter ensembles are able to reproduce the observed carbon-climate system, but contain implicit differences in thermal characteristics that lead to different behavior of the system in the future.

4.4 Conclusions & Implications

When constrained to observed CO_2 and temperature change, differences in carbon cycle structure alone lead to insignificant changes in the TCRE (orange and green curves in Figure 4.8). However, they lead to small, significant changes in the ZEC that increase over time (Figure 4.9), suggesting that differences in biogeochemical feedbacks emerge after net-zero emissions is achieved.

When fit to exponential memory functions with multiple timescales of decay, SCMs exhibit the largest differences in longer (decade- and century-scale) timescales, with sensitivities that correspond to real-world processes that we know influence the exchange of carbon between chemical reservoirs.

However, our experimental design uses a set of parameters that are intentionally uncorrelated with one another, created with a Latin Hypercube sampling method. In the physical climate system, some carbon and climate processes are effectively governed by the same structure, in particular through ocean circulation. This is the case for the processes that govern the ZEC in the real world. In FaIR, since the flow of energy between Earth systems is calculated separately from the flow of carbon, biogeochemical feedbacks may not have the same influence we expect them to, but this is due to the structure of the model inhibiting connections which exist in the real system.

In the FaIR_bgc simulation, differences in biogeochemical feedbacks influence the rate of carbon uptake (Figure 4.15). Variance in diagnosed carbon sink timescales can be explained

by changes to the parameters that define biogeophysical feedbacks in ways that correspond to the functionality of those changes. However, these carbon cycle parameter perturbations do not explain any changes in long-term warming (Figure 4.18). A possible explanation is that our perturbations to energy balance model parameters produce spread in the temperature response that overwhelms the temperature change due to the perturbations to carbon cycle parameters.

By generating a modified parameter distribution with logically correlated parameters, we test whether separating carbon cycle and energy balance processes masks the influence of carbon cycle structure on ZEC. We find that by connecting parameters that represent connected processes, we shift our expectation for future climate even while staying consistent with observed climate, in particular to higher climate sensitivity and larger response of photosynthesis to CO_2. This result has critical implications for how we continue to model mitigation scenarios in SCMs. It suggests that changes to our parameter space that do not actively account for processes that affect both carbon cycling and energy balance can lead to overestimation of climate feedbacks and non-CO_2 warming and overestimation of the historical land carbon sink. We see a 0.5 Pg C/year drop in our estimate of carbon flux to the land in 2021 (from about 4.5 Pg C/year to 4 Pg C/year). Since many proposed mitigation scenarios (both public and private) rely on modifications to the land sink, it will be critical to evaluate and constrain the physical assumptions implicit in SCMs' models structures, especially as they affect climate sensitivity and land carbon feedbacks.

To address this problem, modelers can restrict the SCM parameter space to physically-justified ranges using emissions-driven decarbonization experiments in ESMs. While certainly an ongoing part of model development (Smith et al., 2024), the task of perpetually constraining parameters can be a daunting challenge, especially for SCMs; increasing constraints on parameter choices for SCMs reduces their ability to span the range of climates simulated by very different ESMs. Another approach would be for SCM developers introduce structural links between SCM components that reflect physical connections between energy balance and the carbon cycle.

Both approaches require repeating these experiments with ESMs in emissions-driven configurations, since concentration-driven experiments, which have been used to develop and

calibrate FaIR and other SCMs until today, are not able to fully capture the biogeochemical feedbacks in response to emissions (or decrease emissions) that occur in mitigation scenarios (Sanderson et al., 2023; Melnikova et al., 2023). With CMIP7, we will begin to learn more about which processes in ESMs determine the long-term evolution of carbon.

This work raises an additional question about methods for diagnosing timescales going forward when comparing emissions and decarbonization regimes. We cannot fairly compare the response of emissions-regime carbon sinks to decarbonization-regime carbon sinks because the longer timescales of change do not emerge if only 100 years is given. (This is also a limitation of the Joos et al. (2013) modeling study that only uses 300 years of simulation for some model runs, which are then used to determine input lifetimes in FaIR.) Future scenarios that compare, for example, a 1000 PgC pulse of carbon to the *esm-flat10-zec* and a 1 PgC/yr for 1000 years emissions scenario to investigate differences between sink responses during the emissions-phase vs. decarbonization-phase.

We have shown that SCMs may be missing key linkages between processes that have significant impact on the diagnosed behavior of our climate system to warming. One positive take on this conclusion is that in many ways, the range of SCM structure has wider diversity than ESM model structure, and as a result, identifying key drivers of differences in the influence of carbon cycle feedbacks on temperature across simple models could aid in identifying drivers of differences between ESMs. As the two types of models continue to evolve, this work may constitute just one way in which they may continue to accelerate each other's advancement.

Chapter 5

CONCLUSIONS AND FUTURE WORK

Conclusions

In two divergent climate settings, we show that habitability is sensitive to the behavior of processes at the interface between the atmosphere and surface. In particular, the specific response of processes on land to atmospheric forcing exerts a strong control on the overall climate state.

Since ESM components reflect fields within Earth Science, the challenge of improving coupling between model components can be as difficult as interdisciplinary research. Approaching and understanding these boundaries allows us to push fundamental boundaries in our understanding and ability to predict unobserved climates. This dissertation is an attempt to better understand both past and future climate using (together) simple and complex representations of our current Earth system, and in doing so, understand the hinges that open and close the door to habitability.

We motivate each case with the goal of defining "habitability" because it forms a clear use case for the model experiment, and a result dictates the specific model components necessary to answer the question.

In Chapter 2, we crossed many model component boundaries and to find that in the context of very different *ocean* states, *land* properties matter greatly for past *atmospheric* moisture demand, which might allow for ocean-based life to survive at the boundary between land and ocean because of important *ice* dynamics.

Performing these experiments required the use of a particular kind of land model capable of resolving properties of the land surface that matter for the atmosphere, namely surface albedo, evaporative resistance, and roughness. Beneath these properties lie more complex components of soil type and composition, but for the purposes of passing information to the atmosphere, we can simplify the land system, and the result is a more clear understanding

of land's role in overall climate. In this way, Laguë et al. (2019) cleanly illustrates the ideas laid out in Held's 2005 essay on the importance of maintaining climate model hierarchies.

At present, the most crucial boundary that thoughtful and creative model design might address is the disconnect between human influence on the terrestrial biosphere and climate. While Integrated Assessment Models (IAMs) can simulate human action as it influences carbon emissions and consequently inputs to climate models, it is difficult to identify a single climate model (simple or complex) that is able to produce reliable outputs that are specifically relevant for human decision-making, which is perhaps the most critical element influencing the global climate system.

In chapter 3, we show that models designed to emulate the Earth system with computational efficiency, which removes the barrier of high and iterative sampling needed for human decision-making, disagree in their predictions of the Earth system response to decarbonization. In chapter 4, we found that by structurally separating carbon cycling from energy balance, SCMs may exaggerate the degrees of freedom in the carbon-climate system, leading to unconstrained counteracting feedbacks and unphysical representations of historical climate. These representations may be consistent with historical observations but yield meaningful differences for future climate habitability under decarbonization.

To address this uncertainty, ESMs may be used to further tune the SCM parameter space to account for connected processes. In addition, we may also be able introduce structural changes to FaIR and other SCMs as these models continue to develop, that enforce physical constraints from connections across the carbon-climate system.

Finally, this work motivates continued examination and characterization of ESM carbon cycle dynamics under decarbonization. This is especially urgent in the context of the terrestrial carbon cycle as wildfires and other disturbances exacerbated by climate change drive increased uncertainty in projections of forest ecosystems and consequent above-grown carbon storage (Wu et al., 2023).

Future Work

Further work is planned for both refinement of Snowball Earth habitability and exploration of the influence of decarbonization on the carbon-climate system.

We plan to test the sensitivity of our precipitation minus evaporation and temperature thresholds to continental configuration in CESM, prescribing Neoproterozoic continents with the same experimental design. Further work is needed to investigate changes to atmospheric circulation due to the orography of ice sheets and their consequent influence on habitability.

An crucial difference between the definitions of habitability during Snowball Earth and future climate is the inclusion human influence change on the climate state through fossil fuel and land use-driven emissions.

The addition of the *esm-flat10* experiments in the CMIP7 Fast Track experiment list is an exciting advancement. The results of these simulations will allow us to include uncertainty in carbon cycle feedbacks in our assessment of mitigation policies. Importantly, *esm-flat10-zec* and *esm-flat10-cdr* are decarbonization pathways, allowing for key metrics for warming abatement to be calculated in the context of decreasing emissions. These experiments are designed specifically to inform IPCC Assessment Reports, which are subsequently used in guiding global climate agreements and local climate policy.

While concentration-driven runs are useful for allowing a larger diversity of models to participate in CMIP, emissions that are used to drive the SCMs and compose RCPs differ from the inferred emissions (Liddicoat et al., 2021; Sanderson et al., 2023). This inconsistency poses a challenge when interpreting model results for policy decision-making. In the case of land-use change for carbon uptake and hypothetical decarbonization technologies, which exist in IAMs but not in ESMs, interred emissions cannot be calculated. An additional next step for emissions-driven Earth System Models will be to include land use-driven emissions ad technologies such as Bioenergy with Carbon Capture and Storage (BECCS), which makes use of the terrestrial biosphere to capture CO_2, as an endogenous process. Doing so would allow us for coherence in the entire carbon cycle and the process of quantifying uncertainty in the carbon-climate system.

Finally, this work illustrates the need for sensitivity tests within and across ESMs to parameter choices that influence carbon cycle processes and feedbacks, again, in decarbonization scenarios. While carbon cycle PPEs may be a computational challenge, we are already making decisions about decarbonization that assume untested characteristics of the Earth system, and the challenge of reversing or regretting those decisions is even larger.

The need to incorporate human-centric metrics into climate model design has another dimension beyond emissions and policy, which is climate impacts as they are felt by human beings. As part of the EU-funded SPARCCLE project, I will be working to extend the climate model emulator, MESMER (the Modular Earth System Model Emulator with spatially Resolved output) (Beusch et al., 2020, 2022) to include biophysical impact indicators for the purposes of climate hazard attribution as well as uncertainty quantification. These indicators include heatwave duration, heat stress, heat index, and runoff, and will be built using results from the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) (Warszawski et al., 2014; ISIMIP3, 2024). ISIMIP provides a protocol for models to assess climate impacts as they relate to specific economic sectors and regions, giving policymakers usable information to plan climate mitigation as well as adaptation.

My hope is this work will help to cross new disciplinary and model boundaries in our coupled climate system, advancing our understanding of the human-Earth system through the lens of habitability in a new way.

BIBLIOGRAPHY

Dorian S. Abbot, Aiko Voigt, and Daniel Koll. The jormungand global climate state and implications for neoproterozoic glaciations. *Journal of Geophysical Research: Atmospheres*, 116(D18):D18103, 2011.

Dorian S. Abbot, Aiko Voigt, Dawei Li, Guillaume Le Hir, Raymond T. Pierrehumbert, Mark Branson, David Pollard, and Daniel D. B. Koll. Robust elements of snowball earth atmospheric circulation and oases for life. *Journal of Geophysical Research: Atmospheres*, 118(12):6017–6027, 2013.

Jordan T. Abell, Alex Pullen, Zachary J. Lebo, Paul Kapp, Lucas Gloege, Andrew R. Metcalf, Junsheng Nie, and Gisela Winckler. A wind-albedo-wind feedback driven by landscape evolution. *Nature Communications*, 2020a. doi: 10.1038/s41467-019-13661-w.

Jordan T. Abell, Stefan R. Rahimi, Alex Pullen, Zachary J. Lebo, Dehai Zhang, Paul Kapp, Lucas Gloege, Sean Ridge, Junsheng Nie, and Gisela Winckler. A quantitative model-based assessment of stony desert landscape evolution in the hami basin, china: Implications for plio-pleistocene dust production in eastern asia. *Geophysical Research Letters*, 47(20):e2020GL090064, 2020b.

Justin Adams, Teresa Hartmann, Rebecca King, Dominic Kailash Nath Waughray, Daniel Aminetzah, Emily Birch, Julien Claes, Joshua Katz, Peter Manion, Dickon Pinner, and Christopher Neill. Nature and net zero. Technical report, World Economic Forum, McKinsey & Company, 2021. URL $\texttt{https://www3.weforum.org/docs/WEF}_{C}onsultation_{N}ature_{a}nd_{N}et_{Z}ero_{2}021.pdf$.

Eric Agol, Caroline Dorn, Simon L. Grimm, Martin Turbet, Elsa Ducrot, Laetitia Delrez, Michaël Gillon, Brice-Olivier Demory, Artem Burdanov, Khalid Barkaoui, Zouhair Benkhaldoun, Emeline Bolmont, Adam Burgasser, Sean Carey, Julien de Wit, Daniel

Fabrycky, Daniel Foreman-Mackey, Jonas Haldemann, David M. Hernandez, James Ingalls, Emmanuel Jehin, Zachary Langford, Jérémy Leconte, Susan M. Lederer, Rodrigo Luger, Renu Malhotra, Victoria S. Meadows, Brett M. Morris, Francisco J. Pozuelos, Didier Queloz, Sean N. Raymond, Franck Selsis, Marko Sestovic, Amaury H. M. J. Triaud, and Valerie Van Grootel. Refining the transit-timing and photometric analysis of trappist-1: Masses, radii, densities, dynamics, and ephemerides. *The Planetary Science Journal*, 2(1):1, jan 2021. doi: 10.3847/PSJ/abd022.

Frédérique Aït-Touati and Susan Emanuel. *Fictions of the cosmos : science and literature in the seventeenth century*. University of Chicago Press, Chicago, 2011.

Myles R. Allen, David J. Frame, Chris Huntingford, Chris D. Jones, Jason A. Lowe, Malte Meinshausen, and Nicolai Meinshausen. Warming caused by cumulative carbon emissions towards the trillionth tonne. *Nature*, 458(7242):1163–1166, 2009. doi: 10.1038/nature08019.

V. K. Arora, A. Katavouta, R. G. Williams, C. D. Jones, V. Brovkin, P. Friedlingstein, J. Schwinger, L. Bopp, O. Boucher, P. Cadule, M. A. Chamberlain, J. R. Christian, C. Delire, R. A. Fisher, T. Hajima, T. Ilyina, E. Joetzjer, M. Kawamiya, C. D. Koven, J. P. Krasting, R. M. Law, D. M. Lawrence, A. Lenton, K. Lindsay, J. Pongratz, T. Raddatz, R. Séférian, K. Tachiiri, J. F. Tjiputra, A. Wiltshire, T. Wu, and T. Ziehn. Carbon–concentration and carbon–climate feedbacks in cmip6 models and their comparison to cmip5 models. *Biogeosciences*, 17(16):4173–4222, 2020. doi: 10.5194/bg-17-4173-2020.

L. Beusch, L. Gudmundsson, and S. I. Seneviratne. Emulating earth system model temperatures with mesmer: from global mean temperature trajectories to grid-point-level realizations on land. *Earth System Dynamics*, 11(1):139–159, 2020. doi: 10.5194/esd-11-139-2020.

L. Beusch, Z. Nicholls, L. Gudmundsson, M. Hauser, M. Meinshausen, and S. I. Seneviratne. From emission scenarios to spatially resolved projections with a chain of computationally efficient emulators: coupling of magicc (v7.5.1) and mesmer (v0.8.3). *Geoscientific Model Development*, 15(5):2085–2103, 2022. doi: 10.5194/gmd-15-2085-2022.

C. M. Bitz, K. M. Shell, P. R. Gent, D. A. Bailey, G. Danabasoglu, K. C. Armour, M. M. Holland, and J. T. Kiehl. Climate sensitivity in the community climate system model version 4. *J. Climate*, 25:3053–3070, 2012. doi: 10.1175/JCLI-D-11-00290.1.

C.M. Bitz and W.H. Lipscomb. An energy-conserving thermodynamic model of sea ice. *Journal of Geophysical Research: Oceans*, 104(C7):15669–15677, 1999.

Gordon Bonan. *Ecological Climatology: Concepts and Applications*. Cambridge University Press, 3 edition, 2015.

Gordon B. Bonan, F. Stuart Chapin, and Starley L. Thompson. Boreal forest and tundra ecosystems as components of the climate system. *Climatic Change*, 29(2):145–167, 1995. doi: 10.1007/BF01094014.

C. Braun, J. Hörner, A. Voigt, and J.G. Pinto. Ice-free tropical waterbelt for snowball earth events questioned by uncertain clouds. *Nature Geoscience*, 15(6):489–493, 2022.

M.I. Budyko. The effect of solar radiation variations on the climate of the earth. *Tellus*, 21:611–619, 1969.

C.E. Bøggild, R.E. Brandt, K.J. Brown, and S.G. Warren. The ablation zone in northeast greenland: ice types, albedos and impurities. *Journal of Glaciology*, 56(195):101–113, 2010.

A.J. Campbell, E.D. Waddington, and S.G. Warren. Refugium for surface life on snowball earth in a nearly-enclosed sea? a first simple model for sea-glacier invasion. *Geophysical Research Letters*, 38(19):L19502, 2011. doi: 10.1029/2011GL048846.

A.J. Campbell, E.D. Waddington, and S.G. Warren. Refugium for surface life on snowball earth in a nearly-enclosed sea? a numerical solution for sea-glacier invasion through a narrow strait. *J. Geophys. Res.*, 119:2679–2690, 2014. doi: 10.1002/2013JC009703.

R. D. Cess, G. L. Potter, J. P. Blanchet, G. J. Boer, A. D. Del Genio, M. Déqué, V. Dymnikov, V. Galin, W. L. Gates, S. J. Ghan, J. T. Kiehl, A. A. Lacis, H. Le Treut, Z.-X. Li, X.-Z. Liang, B. J. McAvaney, V. P. Meleshko, J. F. B. Mitchell, J.-J. Morcrette,

D. A. Randall, L. Rikus, E. Roeckner, J. F. Royer, U. Schlese, D. A. Sheinin, A. Slingo, A. P. Sokolov, K. E. Taylor, W. M. Washington, R. T. Wetherald, I. Yagai, and M.-H. Zhang. Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. *Journal of Geophysical Research: Atmospheres*, 95 (D10):16601–16615, 1990.

A.J. Conley, J.-F. Lamarque, F. Vitt, W.D. Collins, and J. Kiehl. Port, a cesm tool for the diagnosis of radiative forcing. *Geoscientific Model Development*, 6(2):469–476, 2013. doi: 10.5194/gmd-6-469-2013.

T.J. Crowley and S.K. Baum. Effect of decreased solar luminosity on late precambrian ice extent. *Journal of Geophysical Research: Atmospheres*, 98(D9):16723–16732, 1993.

Donald P. Cummins, David B. Stephenson, and Peter A. Stott. Optimal estimation of stochastic energy balance model parameters. *Journal of Climate*, 33(18):7909–7926, 2020. doi: 10.1175/JCLI-D-19-0589.1.

R. Dadic, P.C. Mullen, M. Schneebeli, R.E. Brandt, and S.G. Warren. Effects of bubbles, cracks, and volcanic tephra on the spectral albedo of bare ice near the transantarctic mountains: Implications for sea glaciers on snowball earth. *Journal of Geophysical Research: Earth Surface*, 118(3):1658–1676, 2013.

H. Damon Matthews, Katarzyna B. Tokarska, Joeri Rogelj, Christopher J. Smith, Andrew H. MacDougall, Karsten Haustein, Nadine Mengis, Sebastian Sippel, Piers M. Forster, and Reto Knutti. An integrated approach to quantifying uncertainties in the remaining carbon budget. *Communications Earth & Environment*, 2(1):7, 2021. doi: 10.1038/s43247-020-00064-9.

G. Danabasoglu, J.-F. Lamarque, J. Bacmeister, D. A. Bailey, A. K. DuVivier, J. Edwards, L. K. Emmons, J. Fasullo, R. Garcia, A. Gettelman, C. Hannay, M. M. Holland, W. G. Large, P. H. Lauritzen, D. M. Lawrence, J. T. M. Lenaerts, K. Lindsay, W. H. Lipscomb, M. J. Mills, R. Neale, K. W. Oleson, B. Otto-Bliesner, A. S. Phillips, W. Sacks, S. Tilmes, L. van Kampenhout, M. Vertenstein, A. Bertini, J. Dennis, C. Deser, C. Fischer, B. Fox-

Kemper, J. E. Kay, D. Kinnison, P. J. Kushner, V. E. Larson, M. C. Long, S. Mickelson, J. K. Moore, E. Nienhouse, L. Polvani, P. J. Rasch, and W. G. Strand. The community earth system model version 2 (cesm2). *Journal of Advances in Modeling Earth Systems*, 12(2):e2019MS001916, 2020. doi: 10.1029/2019MS001916.

Y. Donnadieu, F. Fluteau, G. Ramstein, C. Ritz, and J. Besse. Is there a conflict between the neoproterozoic glacial deposits and the snowball earth interpretation: an improved understanding with numerical modeling. *Earth and Planetary Science Letters*, 208(1): 101–112, 2003.

K. Dorheim, S. Gering, R. Gieseke, C. Hartin, L. Pressburger, A. N. Shiklomanov, S. J. Smith, C. Tebaldi, D. Woodard, and B. Bond-Lamberty. Hector v3.1.1: functionality and performance of a reduced-complexity climate model. *EGUsphere*, 2023:1–20, 2023. doi: 10.5194/egusphere-2023-1477.

Kalyn Dorheim, Robert Link, Corinne Hartin, Ben Kravitz, and Abigail Snyder. Calibrating simple climate models to individual earth system models: Lessons learned from calibrating hector. *Earth and Space Science*, 7(11):e2019EA000980, 2020. doi: https://doi.org/10.1029/2019EA000980.

Dana Ehlert and Kirsten Zickfeld. What determines the warming commitment after cessation of co2 emissions? *Environmental Research Letters*, 12(1):015002, 2017. doi: 10.1088/1748-9326/aa564a.

D.A.D. Evans. Stratigraphic, geochronological, and paleomagnetic constraints upon the neoproterozoic climatic paradox. *American Journal of Science*, 300(5):347–433, 2000.

Bernard Le Bovier Fontenelle. *Entretiens sur la pluralité des mondes.* C. Blageart, 1686.

P. M. Forster, C. J. Smith, T. Walsh, W. F. Lamb, R. Lamboll, M. Hauser, A. Ribes, D. Rosen, N. Gillett, M. D. Palmer, J. Rogelj, K. von Schuckmann, S. I. Seneviratne, B. Trewin, X. Zhang, M. Allen, R. Andrew, A. Birt, A. Borger, T. Boyer, J. A. Broersma, L. Cheng, F. Dentener, P. Friedlingstein, J. M. Gutiérrez, J. Gütschow, B. Hall, M. Ishii, S. Jenkins, X. Lan, J.-Y. Lee, C. Morice, C. Kadow, J. Kennedy, R. Killick, J. C. Minx,

V. Naik, G. P. Peters, A. Pirani, J. Pongratz, C.-F. Schleussner, S. Szopa, P. Thorne, R. Rohde, M. Rojas Corradi, D. Schumacher, R. Vose, K. Zickfeld, V. Masson-Delmotte, and P. Zhai. Indicators of global climate change 2022: annual update of large-scale indicators of the state of the climate system and human influence. *Earth System Science Data*, 15(6):2295–2327, 2023. doi: 10.5194/essd-15-2295-2023.

P. Friedlingstein, R. M. Andrew, J. Rogelj, G. P. Peters, J. G. Canadell, R. Knutti, G. Luderer, M. R. Raupach, M. Schaeffer, D. P. van Vuuren, and C. Le Quéré. Persistent growth of co2 emissions and implications for reaching climate targets. *Nature Geoscience*, 7(10): 709–715, 2014. doi: 10.1038/ngeo2248.

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

P. Friedlingstein, M. O'Sullivan, M. W. Jones, R. M. Andrew, D. C. E. Bakker, J. Hauck, P. Landschützer, C. Le Quéré, I. T. Luijkx, G. P. Peters, W. Peters, J. Pongratz, C. Schwingshackl, S. Sitch, J. G. Canadell, P. Ciais, R. B. Jackson, S. R. Alin, P. An-

thoni, L. Barbero, N. R. Bates, M. Becker, N. Bellouin, B. Decharme, L. Bopp, I. B. M. Brasika, P. Cadule, M. A. Chamberlain, N. Chandra, T.-T.-T. Chau, F. Chevallier, L. P. Chini, M. Cronin, X. Dou, K. Enyo, W. Evans, S. Falk, R. A. Feely, L. Feng, D. J. Ford, T. Gasser, J. Ghattas, T. Gkritzalis, G. Grassi, L. Gregor, N. Gruber, Ö. Gürses, I. Harris, M. Hefner, J. Heinke, R. A. Houghton, G. C. Hurtt, Y. Iida, T. Ilyina, A. R. Jacobson, A. Jain, T. Jarníková, A. Jersild, F. Jiang, Z. Jin, F. Joos, E. Kato, R. F. Keeling, D. Kennedy, K. Klein Goldewijk, J. Knauer, J. I. Korsbakken, A. Körtzinger, X. Lan, N. Lefèvre, H. Li, J. Liu, Z. Liu, L. Ma, G. Marland, N. Mayot, P. C. McGuire, G. A. McKinley, G. Meyer, E. J. Morgan, D. R. Munro, S.-I. Nakaoka, Y. Niwa, K. M. O'Brien, A. Olsen, A. M. Omar, T. Ono, M. Paulsen, D. Pierrot, K. Pocock, B. Poulter, C. M. Powis, G. Rehder, L. Resplandy, E. Robertson, C. Rödenbeck, T. M. Rosan, J. Schwinger, R. Séférian, T. L. Smallman, S. M. Smith, R. Sospedra-Alfonso, Q. Sun, A. J. Sutton, C. Sweeney, S. Takao, P. P. Tans, H. Tian, B. Tilbrook, H. Tsujino, F. Tubiello, G. R. van der Werf, E. van Ooijen, R. Wanninkhof, M. Watanabe, C. Wimart-Rousseau, D. Yang, X. Yang, W. Yuan, X. Yue, S. Zaehle, J. Zeng, and B. Zheng. Global carbon budget 2023. *Earth System Science Data*, 15(12):5301–5369, 2023. doi: 10.5194/essd-15-5301-2023.

J.C. Goodman. Through thick and thin: Marine and meteoric ice in a "snowball earth" climate. *Geophysical Research Letters*, 33:L16701, 2006. doi: 10.1029/2006GL026840.

J.C. Goodman and R.T. Pierrehumbert. Glacial flow of floating marine ice in "snowball earth". *Geophys. Res. Oceans*, 108(C10):3308, 2003. doi: 10.1029/2002JC001471.

J. M. Gregory, W. J. Ingram, M. A. Palmer, G. S. Jones, P. A. Stott, R. B. Thorpe, J. A. Lowe, T. C. Johns, and K. D. Williams. A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, 31(3), 2004. doi: https://doi.org/10.1029/2003GL018747.

T.C. Grenfell, S.G. Warren, and P.C. Mullen. Reflection of solar radiation by the antarctic snow surface at ultraviolet, visible, and near-infrared wavelengths. *J. Geophys. Res.*, 99: 18669–18684, 1994.

Nicolas Gruber, Dorothee C. E. Bakker, Tim DeVries, Luke Gregor, Judith Hauck, Peter Landschützer, Galen A. McKinley, and Jens Daniel Müller. Trends and variability in the ocean carbon sink. *Nature Reviews Earth & Environment*, 4(2):119–134, 2023. doi: 10.1038/s43017-022-00381-x.

W.B. Harland. Critical evidence for a great infra-cambrian glaciation. *Geologische Rundschau*, 54(1):45–61, 1964.

C. A. Hartin, P. Patel, A. Schwarber, R. P. Link, and B. P. Bond-Lamberty. A simple object-oriented and open-source model for scientific and policy analyses of the global climate system – hector v1.0. *Geoscientific Model Development*, 8(4):939–955, 2015. doi: 10.5194/gmd-8-939-2015.

C. A. Hartin, B. Bond-Lamberty, P. Patel, and A. Mundra. Ocean acidification over the next three centuries using a simple global climate carbon-cycle model: projections and sensitivities. *Biogeosciences*, 13(15):4329–4342, 2016. doi: 10.5194/bg-13-4329-2016.

C. Hawkesworth, P.A. Cawood, and B. Dhuime. Rates of generation and growth of the continental crust. *Geoscience Frontiers*, 10:165–173, 2019.

Isaac M. Held. The gap between simulation and understanding in climate modeling. *Bulletin of the American Meteorological Society*, 86(11):1609–1614, 2005. doi: 10.1175/BAMS-86-11-1609.

Isaac M. Held, Michael Winton, Ken Takahashi, Thomas Delworth, Fanrong Zeng, and Geoffrey K. Vallis. Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *Journal of Climate*, 23(9):2418–2427, 2010. doi: 10.1175/2009JCLI3466.1.

Paul F. Hoffman, Dorian S. Abbot, Yosef Ashkenazy, Douglas I. Benn, Jochen J. Brocks, Phoebe A. Cohen, Grant M. Cox, Jessica R. Creveling, Yannick Donnadieu, Douglas H. Erwin, Ian J. Fairchild, David Ferreira, Jason C. Goodman, Galen P. Halverson, Malte F. Jansen, Guillaume Le Hir, Gordon D. Love, Francis A. Macdonald, Adam C. Maloof, Camille A. Partin, Gilles Ramstein, Brian E. J. Rose, Catherine V. Rose, Peter M.

Sadler, Eli Tziperman, Aiko Voigt, and Stephen G. Warren. Snowball earth climate dynamics and cryogenian geology-geobiology. *Science Advances*, 3(11):e1600983, 2017. doi: 10.1126/sciadv.1600983.

P.F. Hoffman and D.P. Schrag. Snowball earth. *Scientific American*, 68(1):68–75, 2000.

P.F. Hoffman and D.P. Schrag. The snowball earth hypothesis: Testing the limits of global change. *Terra Nova*, 14:129–155, 2002.

Radley M. Horton, Alex de Sherbinin, David Wrathall, and Michael Oppenheimer. Assessing human habitability and migration. *Science*, 372(6548):1279–1283, 2021. doi: 10.1126/science.abi8603.

S.R. Hudson, S.G. Warren, R.E. Brandt, Grenfell, T.C., and D. Six. Spectral bidirectional reflectance of antarctic snow: Measurements and parameterization. *Journal of Geophysical Research: Atmospheres*, 111:D18106, 2006. doi: 10.1029/2006JD007290.

E. Hunke, W.H. Lipscomb, A. Turner, N. Jeffery, and S. Elliott. Cice: the los alamos sea ice model documentation and software user's manual la-cc-06-012. Technical report, Los Alamos National Laboratory, Los Alamos, New Mexico., 2015.

J. Hörner and A. Voigt. Sea-ice thermodynamics can determine waterbelt scenarios for snowball earth. *EGUSphere*, 2023. doi: 10.5194/egusphere-2023-2073.

IPCC SR1.5. *Mitigation Pathways Compatible with 1.5°C in the Context of Sustainable Development*, pages 93–174. Cambridge University Press, 2018. doi: 10.1017/9781009157940.004.

IPCC Working Group I. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter Future Global Climate: Scenario-Based Projections and Near-Term Information, page 553–672. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021a. doi: 10.1017/9781009157896.006.

IPCC Working Group I. *Chapter 5: Global Warming Levels, Climate Feedbacks, and Impacts*, chapter 5, pages 385–460. Cambridge University Press, 2021b.

IPCC Working Group I. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter The Earth's Energy Budget, Climate Feedbacks and Climate Sensitivity. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021c. doi: 10.1017/9781009157896.009.

IPCC Working Group II. *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK and New York, NY, USA, 2022. doi: 10.1017/9781009157926.

ISIMIP3. Isimip3 simulation protocol. Technical report, 2024. `https://protocol.isimip.org/ISIMIP3a` [Accessed: 18 May 2024].

C. D. Jones, T. L. Frölicher, C. Koven, A. H. MacDougall, H. D. Matthews, K. Zickfeld, J. Rogelj, K. B. Tokarska, N. P. Gillett, T. Ilyina, M. Meinshausen, N. Mengis, R. Séférian, M. Eby, and F. A. Burger. The zero emissions commitment model intercomparison project (zecmip) contribution to c4mip: quantifying committed climate changes following zero carbon emissions. *Geoscientific Model Development*, 12(10):4375–4385, 2019. doi: 10.5194/gmd-12-4375-2019.

Chris D Jones and Pierre Friedlingstein. Quantifying process-level uncertainty contributions to tcre and carbon budgets for meeting paris agreement climate targets. *Environmental Research Letters*, 15(7):074019, jun 2020. doi: 10.1088/1748-9326/ab858a.

F. Joos, R. Roth, J. S. Fuglestvedt, G. P. Peters, I. G. Enting, W. von Bloh, V. Brovkin, E. J. Burke, M. Eby, N. R. Edwards, T. Friedrich, T. L. Frölicher, P. R. Halloran, P. B. Holden, C. Jones, T. Kleinen, F. T. Mackenzie, K. Matsumoto, M. Meinshausen, G.-K. Plattner, A. Reisinger, J. Segschneider, G. Shaffer, M. Steinacher, K. Strassmann, K. Tanaka, A. Timmermann, and A. J. Weaver. Carbon dioxide and climate impulse

response functions for the computation of greenhouse gas metrics: a multi-model analysis. *Atmospheric Chemistry and Physics*, 13(5):2793–2825, 2013. doi: 10.5194/acp-13-2793-2013.

Fortunat Joos, Michele Bruno, Roger Fink, Ulrich Siegenthaler, Thomas F. Stocker, Corinne Le Quéré, and Jorge L. Sarmiento. An efficient and accurate representation of complex oceanic and biospheric models of anthropogenic carbon uptake. *Tellus B: Chemical and Physical Meteorology*, Jan 1996. doi: 10.3402/tellusb.v48i3.15921.

P. Kenrick, C.H. Wellman, H. Schneider, and G.D. Edgecombe. A timeline for terrestrialization: consequences for the carbon cycle in the palaeozoic. *Phil. Trans. Royal Soc. London*, B367(1588):519–536, 2012.

J. S. Kikstra, Z. R. J. Nicholls, C. J. Smith, J. Lewis, R. D. Lamboll, E. Byers, M. Sandstad, M. Meinshausen, M. J. Gidden, J. Rogelj, E. Kriegler, G. P. Peters, J. S. Fuglestvedt, R. B. Skeie, B. H. Samset, L. Wienpahl, D. P. van Vuuren, K.-I. van der Wijst, A. Al Khourdajie, P. M. Forster, A. Reisinger, R. Schaeffer, and K. Riahi. The ipcc sixth assessment report wgiii climate assessment of mitigation pathways: from emissions to global temperatures. *Geoscientific Model Development*, 15(24):9075–9109, 2022. doi: 10.5194/gmd-15-9075-2022.

J.L. Kirschvink. Late proterozoic low-latitude global glaciation: the snowball earth. In J.W. Schopf and C. Klein, editors, *The Proterozoic Biosphere*, pages 51–52. Cambridge Univ. Press, New York, 1992.

A.H. Knoll. The multiple origins of complex multicellularity. *Annual Review of Earth and Planetary Sciences*, 39(1):217–239, 2011.

A.H. Knoll. Paleobiological perspectives on early eukaryotic evolution. *Cold Spring Harbor Perspectives in Biology*, 6(1):a016121, 2014.

C. D. Koven, V. K. Arora, P. Cadule, R. A. Fisher, C. D. Jones, D. M. Lawrence, J. Lewis, K. Lindsay, S. Mathesius, M. Meinshausen, M. Mills, Z. Nicholls, B. M. Sanderson, R. Séférian, N. C. Swart, W. R. Wieder, and K. Zickfeld. Multi-century dynamics of the

climate and carbon cycle under both high and net negative emissions scenarios. *Earth System Dynamics*, 13(2):885–909, 2022. doi: 10.5194/esd-13-885-2022.

Charles D Koven, Benjamin M Sanderson, and Abigail L S Swann. Much of zero emissions commitment occurs before reaching net zero emissions. *Environmental Research Letters*, 18(1):014017, jan 2023. doi: 10.1088/1748-9326/acab1a.

J. P. Krasting, J. P. Dunne, E. Shevliakova, and R. J. Stouffer. Trajectory sensitivity of the transient climate response to cumulative carbon emissions. *Geophysical Research Letters*, 41(7):2520–2527, 2014. doi: 10.1002/2013GL059141.

Eun Young Kwon, François Primeau, and Jorge L. Sarmiento. The impact of remineralization depth on the air–sea carbon balance. *Nature Geoscience*, 2(9):630–635, 2009. doi: 10.1038/ngeo612.

M.M. Laguë, G.B. Bonan, and A.L.S. Swann. Separating the impact of individual land surface properties on the terrestrial surface energy budget in both the coupled and uncoupled land–atmosphere system. *Journal of Climate*, 32(18):5725–5744, 2019.

M.M. Laguë, G.R. Quetin, S. Ragen, and W.R. Boos. Continental configuration controls the base-state water vapor greenhouse effect: lessons from half-land, half-water planets. *Climate Dynamics*, 2023. doi: 10.1007/s00382-023-06857-w.

Robin D. Lamboll, Zebedee R. J. Nicholls, Christopher J. Smith, Jarmo S. Kikstra, Edward Byers, and Joeri Rogelj. Assessing the size and uncertainty of remaining carbon budgets. *Nature Climate Change*, 13(12):1360–1367, 2023. doi: 10.1038/s41558-023-01848-5.

X. Lan, P. Tans, and K.W. Thoning. Trends in globally-averaged co2 determined from noaa global monitoring laboratory measurements version 2024-04. 2024. doi: 10.15138/9n0h-zh07. URL `https://gml.noaa.gov/ccgg/trends/global.html`.

Kate Larsen, Hannah Pitt, Mahmoud Mobir, Shweta Movalia, Alfredo Rivera, Emma Rutkowski, Marie Tamba, Kelly McCusker, and Trevor Houser. Rhodium climate outlook: Probabilistic projections of energy, emissions and

global temperature rise. Technical report, Rhodium Group, 2023. URL https://rhg.com/wp-content/uploads/2023/11/Rhodium-Climate-Outlook$_2$023.*pdf*.

N. J. Leach, S. Jenkins, Z. Nicholls, C. J. Smith, J. Lynch, M. Cain, T. Walsh, B. Wu, J. Tsutsui, and M. R. Allen. Fairv2.0.0: a generalized impulse response model for climate uncertainty and future scenario exploration. *Geoscientific Model Development*, 14(5): 3007–3036, 2021.

T.M. Lenton and S.J. Daines. Matworld – the biogeochemical effects of early life on land. *New Phytologist*, 215:531–537, 2017. doi: 10.1111/nph.14338.

A.J. Lewis, J.P. amd Weaver and M. Eby. Deglaciating the snowball earth: Sensitivity to surface albedo. *Geophysical Research Letters*, 33(23):L23604, 2006. doi: 10.1029/2006GL027774.

D. Li and R.T. Pierrehumbert. Sea glacier flow and dust transport on snowball earth. *Geophysical Research Letters*, 38(17):L17501, 2011.

Spencer K. Liddicoat, Andy J. Wiltshire, Chris D. Jones, Vivek K. Arora, Victor Brovkin, Patricia Cadule, Tomohiro Hajima, David M. Lawrence, Julia Pongratz, Jörg Schwinger, Roland Séférian, Jerry F. Tjiputra, and Tilo Ziehn. Compatible fossil fuel co2 emissions in the cmip6 earth system models' historical and shared socioeconomic pathway experiments of the twenty-first century. *Journal of Climate*, 34(8):2853–2875, 2021. doi: 10.1175/JCLI-D-19-0991.1.

P. Liu, Y. Liu, Y. Peng, J.-F. Lamarque, M. Wang, and Y. Hu. Large influence of dust on the precambrian climate. *Nature Communications*, 11(1):4427, 2020.

Y. Liu, W.R. Peltier, J. Yang, and G. Vettoretti. The initiation of neoproterozoic "snowball" climates in ccsm3: the influence of paleocontinental configuration. *Climate of the Past*, 9 (6):2555–2577, 2013.

Y. Liu, W.R. Peltier, J. Yang, Vettoretti G., and Y. Wang. Strong effects of tropical ice-sheet coverage and thickness on the hard snowball earth bifurcation point. *Climate Dynamics*, 48(11):3459–3474, 2017.

Y. Liu, W.R. Peltier, J. Yang, and Y. Hu. Influence of surface topography on the critical carbon dioxide level required for the formation of a modern snowball earth. *Journal of Climate*, 31(20):8463–8479, 2018.

Y. Liu, P. Liu, D. Li, Y. Peng, and Y. Hu. Influence of dust on the initiation of neoproterozoic snowball earth events. *Journal of Climate*, 34(16):6673–6689, 2021.

James E. Lovelock. Geophysiology: A new look at earth science. *Bulletin of the American Meteorological Society*, 67(4):392–397, 1986. ISSN 00030007, 15200477.

Francis A. Macdonald, Mark D. Schmitz, James L. Crowley, Charles F. Roots, David S. Jones, Adam C. Maloof, Justin V. Strauss, Phoebe A. Cohen, David T. Johnston, and Daniel P. Schrag. Calibrating the cryogenian. *Science*, 327(5970):1241–1243, 2010.

A. H. MacDougall. Limitations of the 1 % experiment as the benchmark idealized experiment for carbon cycle intercomparison in c^{4}mip. *Geoscientific Model Development*, 12(2):597–611, 2019. doi: 10.5194/gmd-12-597-2019.

A. H. MacDougall, T. L. Frölicher, C. D. Jones, J. Rogelj, H. D. Matthews, K. Zickfeld, V. K. Arora, N. J. Barrett, V. Brovkin, F. A. Burger, M. Eby, A. V. Eliseev, T. Hajima, P. B. Holden, A. Jeltsch-Thömmes, C. Koven, N. Mengis, L. Menviel, M. Michou, I. I. Mokhov, A. Oka, J. Schwinger, R. Séférian, G. Shaffer, A. Sokolov, K. Tachiiri, J. Tjiputra, A. Wiltshire, and T. Ziehn. Is there warming in the pipeline? a multi-model analysis of the zero emissions commitment from CO_2. *Biogeosciences*, 17(11) : $2987 - -3016, 2020. doi : 10.5194/bg - 17 - 2987 - 2020$.

Andrew H. MacDougall. The transient response to cumulative co2 emissions: a review. *Current Climate Change Reports*, 2(1):39–47, 2016. doi: 10.1007/s40641-015-0030-6.

H. Damon Matthews and Ken Caldeira. Stabilizing climate requires near-zero emissions. *Geophysical Research Letters*, 35(4), 2008. doi: 10.1029/2007GL032388.

C.P. McKay. Thickness of tropical ice and photosynthesis on a snowball earth. *Geophysical Research Letters*, 27(14):2153–2156, 2000.

M. Meinshausen, S. C. B. Raper, and T. M. L. Wigley. Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, magicc6 – part 1: Model description and calibration. *Atmospheric Chemistry and Physics*, 11(4):1417–1456, 2011. doi: 10.5194/acp-11-1417-2011.

Malte Meinshausen, Nicolai Meinshausen, William Hare, Sarah C. B. Raper, Katja Frieler, Reto Knutti, David J. Frame, and Myles R. Allen. Greenhouse-gas emission targets for limiting global warming to 2 °c. *Nature*, 458(7242):1158–1162, 2009. doi: 10.1038/nature08017.

Irina Melnikova, Philippe Ciais, Olivier Boucher, and Katsumasa Tanaka. Assessing carbon cycle projections from complex and simple models under ssp scenarios. *Climatic Change*, 176(12):168, 2023. doi: 10.1007/s10584-023-03639-5.

Andrew S. Merdith, Simon E. Williams, Alan S. Collins, Michael G. Tetley, Jacob A. Mulder, Morgan L. Blades, Alexander Young, Sheree E. Armistead, John Cannon, Sabin Zahirovic, and R. Dietmar Müller. Extending full-plate tectonic models into deep time: Linking the neoproterozoic and the phanerozoic. *Earth-Science Reviews*, 214:103477, 2021.

R. J. Millar, Z. R. Nicholls, P. Friedlingstein, and M. R. Allen. A modified impulse-response representation of the global near-surface air temperature and atmospheric concentration response to carbon dioxide emissions. *Atmospheric Chemistry and Physics*, 17(11):7213–7228, 2017.

Richard J. Millar and Pierre Friedlingstein. The utility of the historical record for assessing the transient climate response to cumulative emissions. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2119):20160449, 2018. doi: 10.1098/rsta.2016.0449.

Ross N. Mitchell, Thomas M. Gernon, Grant M. Cox, Adam R. Nordsvan, Uwe Kirscher, Chuang Xuan, Yebo Liu, Xu Liu, and Xiaofang He. Orbitally forced ice sheet fluctuations during the marinoan snowball earth glaciation. *Nature Geoscience*, 8(9):704–707, 2015.

J.L. Monteith. Evaporation and surface temperature. *Quart. J. Roy. Meteor. Soc*, 107:1–27, 1981.

Jennifer L. Morris, Mark N. Puttick, James W. Clark, Dianne Edwards, Paul Kenrick, Silvia Pressel, Charles H. Wellman, Ziheng Yang, Harald Schneider, and Philip C. J. Donoghue. The timescale of early land plant evolution. *Proc. Nat. Acad. Sci.*, 115(10):E2274–E2283, 2018. doi: 10.1073/pnas.1719588115.

G. Myhre, D. Shindell, F.-M. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura, and H. Zhang. *Anthropogenic and Natural Radiative Forcing: Supplementary Material*. Intergovernmental Panel on Climate Change, 2013. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Supplementary Material.

R.B. Neale, J.H. Richter, A.J. Conley, S. Park, P.H. Lauritzen, D.L. Gettelman, A. Williamson, P. J. Rasch, S. J. Vavrus, M. A. Taylor, W. D. Collins, M. Zhang, and S.-J. Lin. Description of the ncar community atmosphere model (cam 4.0) (ncar/tn-485+str). Technical report, National Center for Atmospheric Research, Boulder, Colorado, 2010.

Michiel Nivard, Bram Smeets, Christer Tryggestad, Peter van de Giessen, and Rune van der Meijden. Global energy perspective 2023: Co2 emissions outlook. Technical report, McKinsey Energy Solutions, 2024. URL `https://t.ly/wYiEQ`.

Sofia Palazzo Corner, Martin Siegert, Paulo Ceppi, Baylor Fox-Kemper, Thomas L. Frölicher, Angela Gallego-Sala, Joanna Haigh, Gabriele C. Hegerl, Chris D. Jones, Reto Knutti, Charles D. Koven, Andrew H. MacDougall, Malte Meinshausen, Zebedee Nicholls, Jean Baptiste Sallée, Benjamin M. Sanderson, Roland Séférian, Merritt Turetsky, Richard G. Williams, Sönke Zaehle, and Joeri Rogelj. The zero emissions commitment and climate stabilization. *Frontiers in Science*, 1, 2023. doi: 10.3389/fsci.2023.1170744.

W. J. Parton, D. S. Schimel, C. V. Cole, and D. S. Ojima. Analysis of factors controlling soil

organic matter levels in great plains grasslands. *Soil Science Society of America Journal*, 51(5):1173–1179, 1987. doi: 10.2136/sssaj1987.03615995005100050015x.

W. J. Parton, D. S. Schimel, C. V. Cole, and D. S. Ojima. A general model for soil organic matter dynamics: Sensitivity to litter chemistry, texture and management. *Quantitative Modeling of Soil Forming Processes*, pages 147–167, 1988.

H.L. Penman. Natural evaporation from open water, bare soil and grass. *Proc. Roy. Soc. London*, 193A:120–145, 1948.

R. T. Pierrehumbert. Climate dynamics of a hard snowball earth. *J. Geophys. Res. Atmos.*, 110(D01111), 2005. doi: 10.1029/2004JD005162.

D. Pollard and J.F. Kasting. Climate-ice sheet simulations of neoproterozoic glaciation before and after collapse to snowball earth. In Christopher P. McKay, Mark A.S. McMenamin, Linda Sohl, and Gregory S. Jenkins, editors, *The Extreme Proterozoic: Geology, Geochemistry, and Climate, Geophys. Monogr. Ser., vol. 146*, pages 91–105. AGU, Washington, D.C., 2004.

D. Pollard and J.F. Kasting. Snowball earth: A thin-ice solution with flowing sea glaciers. *J. Geophys. Res.*, 110(C07010), 2005. doi: 10.1029/2004JC002525.

D. Pollard, J.F. Kasting, and M.E. Zugger. Snowball earth: Asynchronous coupling of sea-glacier flow with a global climate model. *J. Geophys. Res. Atmos.*, 122(5157-5171), 2017. doi: 10.1002/2017JD026621.

S.M. Porter. The fossil record of early eukaryotic diversification. *The Paleontological Society Papers*, 10:35–50, 2004.

G.J. Retallack. Neoproterozoic snowball earth extent inferred from paleosols in california. *J. Palaeosciences*, 72:9–28, 2023. doi: 10.54991/jop.2023.1851.

Joeri Rogelj, Michiel Schaeffer, Pierre Friedlingstein, Nathan P. Gillett, Detlef P. van Vuuren, Keywan Riahi, Myles Allen, and Reto Knutti. Differences between carbon budget

estimates unravelled. *Nature Climate Change*, 6(3):245–252, 2016. doi: 10.1038/nclimate2868.

Joeri Rogelj, Piers M. Forster, Elmar Kriegler, Christopher J. Smith, and Roland Séférian. Estimating and tracking the remaining carbon budget for stringent climate targets. *Nature*, 571(7765):335–342, 2019. doi: 10.1038/s41586-019-1368-z.

B. Sanderson. The role of prior assumptions in carbon budget calculations. *Earth System Dynamics*, 11(2):563–577, 2020. doi: 10.5194/esd-11-563-2020.

B. M. Sanderson, B. B. B. Booth, J. Dunne, V. Eyring, R. A. Fisher, P. Friedlingstein, M. J. Gidden, T. Hajima, C. D. Jones, C. Jones, A. King, C. D. Koven, D. M. Lawrence, J. Lowe, N. Mengis, G. P. Peters, J. Rogelj, C. Smith, A. C. Snyder, I. R. Simpson, A. L. S. Swann, C. Tebaldi, T. Ilyina, C.-F. Schleussner, R. Seferian, B. H. Samset, D. van Vuuren, and S. Zaehle. The need for carbon emissions-driven climate projections in cmip7. *EGUsphere*, 2023:1–51, 2023. doi: 10.5194/egusphere-2023-2127. URL `https://egusphere.copernicus.org/preprints/2023/egusphere-2023-2127/`.

Jorge Louis Sarmiento and Nicolas Gruber. *Ocean biogeochemical dynamics*. Princeton University Press, Princeton, 2006.

J. Scheff and D.M.W. Frierson. Scaling potential evapotranspiration with greenhouse warming. *Journal of Climate*, 27(4):1539–1558, 2014.

D.P. Schrag, R.A. Berner, P.F. Hoffman, and G.P. Halverson. On the initiation of a snowball earth. *Geochemistry, Geophysics, Geosystems*, 3(6):1–21, 2002. doi: 10.1029/2001GC000219.

A. K. Schwarber, S. J. Smith, C. A. Hartin, B. A. Vega-Westhoff, and R. Sriver. Evaluating climate emulation: fundamental impulse testing of simple climate models. *Earth System Dynamics*, 10(4):729–739, 2019. doi: 10.5194/esd-10-729-2019.

Christian Seiler, Joe R. Melton, Vivek K. Arora, Stephen Sitch, Pierre Friedlingstein, Peter Anthoni, Daniel Goll, Atul K. Jain, Emilie Joetzjer, Sebastian Lienert, Danica Lombardozzi, Sebastiaan Luyssaert, Julia E. M. S. Nabel, Hanqin Tian, Nicolas Vuichard,

Anthony P. Walker, Wenping Yuan, and Sönke Zaehle. Are terrestrial biosphere models fit for simulating the global land carbon sink? *Journal of Advances in Modeling Earth Systems*, 14(5):e2021MS002946, 2022. doi: https://doi.org/10.1029/2021MS002946.

S.A. Sejas, P.C. Taylor, and M. Cai. Unmasking the negative greenhouse effect over the antarctic plateau. *npj Climate and Atmospheric Science*, 1(17), 2018.

W.D. Sellers. A global climatic model based on the energy balance of the earth-atmosphere system. *Journal of Applied Meteorology and Climatology*, 8(3):392–400, 1969.

Denis E. Sergeev, Thomas J. Fauchez, Martin Turbet, Ian A. Boutle, Kostas Tsigaridis, Michael J. Way, Eric T. Wolf, Shawn D. Domagal-Goldman, François Forget, Jacob Haqq-Misra, Ravi K. Kopparapu, F. Hugo Lambert, James Manners, and Nathan J. Mayne. The trappist-1 habitable atmosphere intercomparison (thai). ii. moist cases—the two waterworlds. *The Planetary Science Journal*, 3(9):212, sep 2022. doi: 10.3847/PSJ/ac6cf2.

Greta E. M. Shum, Marysa M. Laguë, Stephanie S. Rushley, and Abigail L. S. Swann. Beautiful days in the neighborhood: Land–atmosphere interactions as drivers of forest expansion. *Earth Interactions*, 27(1):e220017, 2023. doi: 10.1175/EI-D-22-0017.1.

M. Siddall, D. Smeed, S. Matthiesen, and E. Rohling. Modelling the seasonal cycle of the exchange flow in bab el mandab (red sea). *Deep Sea Research Part I*, 49(9):1551–1569, 2002.

C. Smith, D. P. Cummins, H.-B. Fredriksen, Z. Nicholls, M. Meinshausen, M. Allen, S. Jenkins, N. Leach, C. Mathison, and A.-I. Partanen. fair-calibrate v1.4.1: calibration, constraining and validation of the fair simple climate model for reliable future climate projections. *EGUsphere*, 2024:1–36, 2024. doi: 10.5194/egusphere-2024-708.

C. J. Smith, P. M. Forster, M. Allen, N. Leach, R. J. Millar, G. A. Passerello, and L. A. Regayre. Fair v1.3: a simple emissions-based impulse response and carbon cycle model. *Geoscientific Model Development*, 11(6):2273–2297, 2018.

Chris Smith. Fair calibration data (1.2.0). Zenodo, 2023. Version:1.2.0.

E.A. Smith. The structure of the arabian heat low. part i: Surface energy budget. *Monthly Weather Review*, 114(6):1067–1083, 1986.

Susan Solomon, Gian-Kasper Plattner, Reto Knutti, and Pierre Friedlingstein. Irreversible climate change due to carbon dioxide emissions. *Proceedings of the National Academy of Sciences*, 106(6):1704–1709, 2009. doi: 10.1073/pnas.0812721106.

IPCC SR1.5. *Impacts of 1.5°C Global Warming on Natural and Human Systems*, page 175–312. Cambridge University Press, 2018. doi: 10.1017/9781009157940.005.

Abigail L. Swann, Inez Y. Fung, Samuel Levis, Gordon B. Bonan, and Scott C. Doney. Changes in arctic vegetation amplify high-latitude warming through the greenhouse effect. *Proceedings of the National Academy of Sciences*, 107(4):1295–1300, 2010. doi: 10.1073/pnas.0913846107.

Abigail Lynn Segal Swann. *Ecoclimate: Variations, Interactions, and Teleconnections*. eScholarship, University of California, 2010.

David W. J. Thompson, Elizabeth A. Barnes, Clara Deser, William E. Foust, and Adam S. Phillips. Quantifying the role of internal climate variability in future climate trends. *Journal of Climate*, 28(16):6443–6456, 2015. doi: 10.1175/JCLI-D-14-00830.1.

J. R. Toggweiler. Variation of atmospheric co2 by ventilation of the ocean's deepest water. *Paleoceanography*, 14(5):571–588, 1999. doi: 10.1029/1999PA900033.

Katarzyna B Tokarska, Nathan P Gillett, Vivek K Arora, Warren G Lee, and Kirsten Zickfeld. The influence of non-co2 forcings on cumulative carbon emissions budgets. *Environmental Research Letters*, 13(3):034039, mar 2018. doi: 10.1088/1748-9326/aaafdd.

UNFCCC. Paris agreement, 2015. Agreed upon at the 21st Conference of the Parties of the UNFCCC.

W.F. Vincent and C. Howard-Williams. Life on snowball earth. *Science*, 287(5462):2421–2421, 2000.

W.F. Vincent, J.A.E. Gibson, R. Pienitz, V. Villeneuve, P.A. Broady, P.B. Hamilton, and C. Howard-Williams. Ice shelf microbial ecosystems in the high arctic and implications for life on snowball earth. *Naturwissenschaften*, 87(3):137–141, 2000.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

A. Voigt and D.S. Abbot. Sea-ice dynamics strongly promote snowball earth initiation and destabilize tropical sea-ice margins. *Climate of the Past*, 8(6):2079–2092, 2012.

A. Voigt and J. Marotzke. The transition from the present-day climate to a modern snowball earth. *Climate Dynamics*, 35(5):887–905, 2010.

K. W. Vugrin, L. P. Swiler, R. M. Roberts, N. J. Stucky-Mack, and S. P. Sullivan. Confidence region estimation techniques for nonlinear regression in groundwater flow: Three case studies. *Water Resources Research*, 43(3):W03423, March 2007. doi: 10.1029/2005WR004804.

A. Walsh, T. Ball, and D.M. Schultz. Extreme sensitivity in snowball earth formation to mountains on paleoproterozoic supercontinents. *Scientific Reports*, 9(1):2349, 2019.

S.G. Warren and R.E. Brandt. Comment on "snowball earth: A thin-ice solution with flowing sea glaciers" by david pollard and james f. kasting. *Journal of Geophysical Research*, 111(C09016), 2006.

S.G. Warren, R.E. Brandt, T.C. Grenfell, and C.P. McKay. Snowball earth: Ice thickness on the tropical ocean. *J. Geophys. Res.*, 107(C10):3167, 2002.

Lila Warszawski, Katja Frieler, Veronika Huber, Franziska Piontek, Olivia Serdeczny, and Jacob Schewe. The inter-sectoral impact model intercomparison project (isi–mip): Project framework. *Proceedings of the National Academy of Sciences*, 111(9):3228–3232, 2014. doi: 10.1073/pnas.1312330110.

Andrew J. Watson and James E. Lovelock. Biological homeostasis of the global environment: the parable of daisyworld. *Tellus B*, 35B(4):284–289, 1983. doi: 10.1111/j.1600-0889.1983.tb00031.x.

William R Wieder, Cory C Cleveland, David M Lawrence, and Gordon B Bonan. Effects of model structural uncertainty on carbon cycle projections: biological nitrogen fixation as a case study. *Environmental Research Letters*, 10(4):044016, apr 2015.

Eric T. Wolf. Assessing the habitability of the trappist-1 system using a 3d climate model. *The Astrophysical Journal Letters*, 839(1):L1, 2017. doi: 10.3847/2041-8213/aa693a.

Andrew J. Wood, Graeme J. Ackland, James G. Dyke, Hywel T. P. Williams, and Timothy M. Lenton. Daisyworld: A review. *Reviews of Geophysics*, 46(1), 2008. doi: https://doi.org/10.1029/2006RG000217.

Chao Wu, Shane R. Coffield, Michael L. Goulden, James T. Randerson, Anna T. Trugman, and William R. L. Anderegg. Uncertainty in us forest carbon storage potential due to climate risks. *Nature Geoscience*, 16(5):422–429, 2023. doi: 10.1038/s41561-023-01166-7.

J. Yang, W.R. Peltier, and Y. Hu. The initiation of modern soft and hard snowball earth climates in ccsm4. *Climate of the Past*, 8(3):907–918, 2012.

C. Zarakas, G. Badgley, and F Chay. Comparing carbon removal at different timescales. Technical report, CarbonPlan, 2023. URL `https://carbonplan.org/research/cdr-timescale-accounting`.