Dissertation ETH No. 24134

Limits of linear forcing-feedback frameworks

A thesis submitted to attain the degree of
Doctor of sciences of ETH Zurich
Dr. sc. ETH Zurich
2017

presented by
Maria Almut Amata Rugenstein
Master of Science ETH in Environmental Sciences
born on November 26, 1986, citizen of Germany

accepted on the recommendation of
Dr. Reto Knutti, examiner
Dr. Kyle Armour, co-examiner
Dr. Thomas Frölicher, co-examiner
Dr. Jan Sedláček, co-examiner

Acknowledgments

I am extraordinarily thankful to Reto Knutti for giving me luxurious conditions, trust, freedom, feedback, just the right amount of forcing and guidance, opportunities, unlimited CPU hours and questions.

Ken Caldeira, Jonathan Gregory, Urs Beyerle, and Jan Sedláček shaped my thinking and helped with large parts of this thesis. I am also very thankful to members of the two labs I spent the last three years for copious help, discussions, collaborations, and joyful everyday life. I profited from the larger research community – reviewers, discussions at conferences, other people's work – which I experienced as supportive and creating a very positive, engaging atmosphere.

I am deeply grateful and indebted to Jeremy, my family, and friends and I dedicate this thesis to my father Björn.

Contents

Abstract

In this thesis I analyze large spatial scale energy balances of the climate system as simulated by global climate models. One of the key questions of climate science since it emerged from meteorology and geology is, "Can we predict the sensitivity of the climate system to external or internal perturbations and distinguish these signals from internal variability?"

An intuitive and physically motivated way to analyze this question is to use an energy balance perspective, which differentiates between a forcing (the perturbation), feedbacks (the internal response processes that amplify or dampen the response to that perturbation), and the sensitivity (the overall temperature response). The different chapters zoom into these terms and their relation to each other: Chapter 1 reviews the current state of discussion regarding the global energy balance perspective, different climate sensitivity terms, and the common assumption that feedbacks scale linearly with the global average temperature response to a forcing. Chapter 2 focuses on the forcing term and the very initial, annual time scale, climate system response to the application of a forcing. Chapter 3 deals again with nonlinear global feedbacks and links them to the state of the climate system response, in particular the surface heat flux induced spatial sea surface temperature (SST) patterns. Chapter 4 details the relation between forcing and feedbacks and shows that ocean warming and sea level rise depend nonlinearly on the forcing level. Chapter 5 discusses themes from chapter 1 and 3, especially the reliability of estimates of equilibrium temperature response based on forcing and feedback information of decadal to centennial time scales.

My focus in this thesis is to understand possible climate system behavior. To reduce the complexity of the real world and to separate linear from nonlinear responses, I undertake several idealization steps: First, I use climate models, but no observational data. Second, I use only the forcing agent carbon dioxide and the forcing time history of prescribed step forcings. This implies that I assume the reaction time scale of carbon cycle, ice sheets, and interactive vegetation is slower than millennial time scales or negligible on century time scales. To increase the signal-to-noise ratio, I mostly use strong forcing magnitudes, and either very long – decadal to millennial – time scales or large ensembles.

A large part of this work is "modeling for insight not numbers" (Hamming, 1962) – and the research questions do have a conceptual touch. However, their answers are highly relevant to very practical questions such as decadal to centennial climate predictability, the usage of the widely used terms climate sensitivity and forcing, or committed millennial time scale impacts of atmospheric CO_2 concentrations. Notably, almost every question in climate science, even if not directly related to forcing, feedbacks, and climate sensitivity, touches upon these concepts or uses – often unstated – assumptions, which are questioned in this thesis.

Zusammenfassung

In dieser Dissertation befasse ich mich mit grossskaligen Energiebilanzen von Klimamodellen. Eine der herausragenden Fragen der Klimawissenschaften – die gestellt wird, seitdem sich die Klimawissenschaften aus der Meteorologie und Geologie entwickelt haben – ist, ob man die Reaktion des globalen Klimasystems auf exogene und endogene Störungen vorhersagen kann und wie sich dieses Signal von interner Variabilität unterscheidet.

Intuitiv und physikalisch liegt es nah, diese Fragen über Energiebilanzen zu lösen. Dabei wird üblicherweise unterschieden zwischen der Störungsursache (dem „Forcing"), den Rückkopplungsprozessen, welche die Reaktionen des Klimasystems auf die Störung entweder verstärken oder dämpfen („Feedbacks"), und der Gesamtwirkung, die bezüglich der Temperatur „Klimasensitivität" genannt wird. In den verschiedenen Kapiteln analysiere ich diese drei Terme und ihre Beziehung zueinander: In Kapitel 1 wird der aktuelle Diskurs zur globalen Energiebilanzen ausgewertet und die verbreitete Annahme untersucht, dass Feedbacks linear mit der globalen Temperaturanomalie skalieren. In Kapitel 2 befasse ich mich mit der Reaktion des Klimasystems auf idealisierte CO_2-Störungen auf sehr kurzen Zeitskalen von weniger als zehn Jahren. In Kapitel 3 untersuche ich nichtlineare Feedbacks und zeige, wie sich die Nichtlinearität auf Prozesse im Klimasystem – insbesondere die Muster von Energieflüssen und Ozeanoberflächentemperaturen – zurückführen lassen. In Kapitel 4 analysiere ich die Beziehung verschiedener CO_2-Störungen und der Reaktion der globalen Meeresspiegelerhöhung und zeige, dass diese nicht, wie oft angenommen, linear ist. In Kapitel 5 werden Resultate aus Kapitel 1 und Kapitel 3 herangezogen, um die Vorhersagbarkeit eines Gleichgewichtszustandes des Klimasystems aus nichtstationären Prozessen herzuleiten.

Diese Arbeit soll dazu beitragen, mögliche Reaktionen des Klimasystems besser zu verstehen. Drei idealisierende Annahmen wurden getroffen, um Komplexität zu reduzieren und lineare von nichtlinearen Prozessen zu unterscheiden: Erstens benutze ich ausschliesslich Klimamodelle und keine Beobachtungsdaten. Zweitens analysiere ich lediglich CO_2-induzierte Erwärmung und Abkühlung und schreibe idealisierte Szenarien vor. Letzteres impliziert, dass die Reaktionszeit des Kohlenstoffkreislaufs, der Landeismassen und der

Pflanzen auf einer Zeitskala von tausend Jahren als vernachlässigbar angesehen werden. Um das Rauschsignal zu verbessern, verwende ich drittens vor allem deutliche Störungen und Signalintegrationszeiten von Jahrzehnten bis Jahrtausenden oder Mittelwerte verschiedener Prozessrealisationen.

Ein Grossteil der vorliegenden Arbeit ist „Modellierung um der Erkenntnis, nicht der Zahlen willen"(Hamming, 1962) und die Ergebnisse sind eher konzeptuell als klimapolitisch relevant. Allerdings tangieren einige der Diskussionen und Resultate der Arbeit auch akute Fragen, wie die nach dekadischer Vorhersagbarkeit, den Gebrauch von strittigen Begriffen, wie den der Klimasensitivität und des Forcings, und unvermeidbare Klimaauswirkungen eines geringen CO_2-Anstiegs auf Zeitskalen von Jahrtausenden. Beinahe jede Frage, die heutzutage in den Klimawissenschaften gestellt wird, bezieht sich direkt oder indirekt auf die Energiebilanzperspektive. Weitverbreitete oft unkritisch getroffene Annahmen werden in der vorliegenden Arbeit in Frage gestellt. Im Sinne der Begriffspräzisierung und Vorhersagbarkeitsanalyse ist diese Dissertation also für zahlreiche Probleme der Klimawissenschaft von Bedeutung.

Overview

In the following, I show how the chapters fit together and change the way to think about simple energy balance models and feedback concepts. I summarize each chapter with an emphasis on the method, the novelty of the results, and their broader implications, some of which are not discussed in the papers themselves.

Chapter 1 Feedbacks, climate sensitivity, and the limits of linear models

We review the achievements, limitations, and necessity of the concept of climate sensitivity. From a historical perspective, we speculate why the concept is popular and successful even though it characterizes the climate system only partially and is an emerging property of the system which does not have a direct measurable proxy. We summarize literature on the limitations of linear feedback assumptions (sketched in Figure 1) concerning initial conditions and tropospheric adjustments, feedbacks on annual to decadal time scales being activated by surface warming patterns, temperature dependent feedbacks, forcing agent and forcing magnitude dependent feedbacks, global – tectonic, ocean gateway, ice sheets – state dependent feedbacks, and the long time scale processes as weathering or oceanic carbon sequestration, which are captured in the concept of Earth System Sensitivity.

We use large ensembles of one complex coupled climate model and one intermediate complexity model to show how the timescales of equilibration depend on the forcing magnitude and sign.

The novelty of this work is the method of the "moving bin regression". Ideally, but not necessarily, done with a large initial condition ensemble, this method depicts for the first time the widely discussed time, state, or temperature dependent feedback parameter as the local tangent in temperature–TOA space. The resulting continuously – annual to centennial – changing feedback parameter emphasizes that a two line fit as done in the two layer model including ocean heat uptake efficacy is better than a simple linear regression, but does not capture the actual characteristic evolution in time of the system.

We conclude that state- and forcing-dependency of feedbacks are not con-

sidered adequately in many previous studies, especially when combining feedback estimates from climate models with observationally based temperatures of energy flux imbalances.

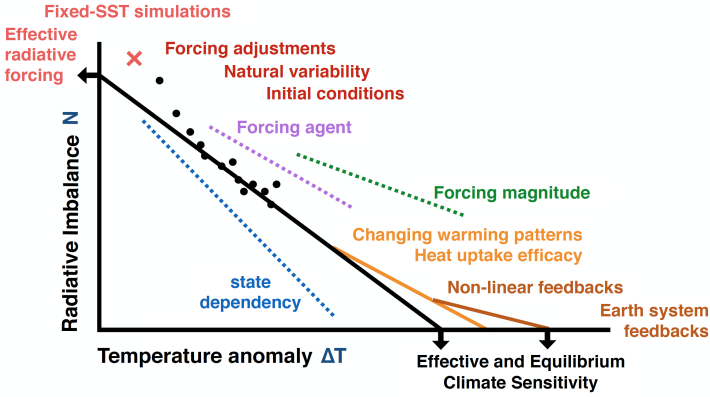

Figure 1: *Sketch to indicate known issues (colors) around the constant feedback parameter assumption, in which the temperature response would follow the straight black line and approach a zero radiative imbalance at equilibrium. Black dots illustrate annual mean model output. References for each issue are noted in the first chapter.*

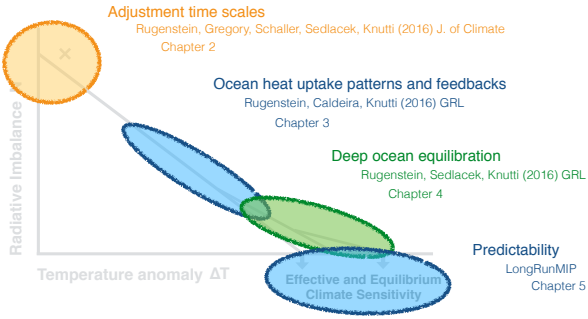

Figure 2: *Sketch to position the following chapters in the reference frame of Fig. 1.*

Chapter 2 Multiannual Ocean–Atmosphere Adjustments to Radiative Forcing

We use a 120 member ensemble of abrupt4xCO_2 simulations to analyze the first few years after the forcing is applied. While the concept of stratospheric adjustment has been around since the early days of climate science, tropospheric adjustment was recognized only a few years ago, probably due to the increased resolution and complexity of climate models, especially their bounday layer and cloud schemes (e.g., Shine et al., 1990; Lahellec and Dufresne, 2014). Tropospheric adjustment takes place because the surface is less out of balance than the top-of-the-atmosphere just after the forcing is applied, leading to circulation and flux responses of opposite sign from the long-term or feedback response. Tropospheric adjustment is thought of as happening within months, certainly the first year after the forcing is applied (e.g., Bala et al., 2010). However, Fig. 3, which depicts each surface heat flux component, shows that the short wave and latent heat fluxes over the oceans take several years to reach even the sign of their long term response (dark red in panel a and green in panel b). The latent heat flux, which dominates the total heat flux, takes about 10 years to become negligible over land (difference between light and dark green line in panel b and brown line in panel c). The difference between TOA and surface is 1 Wm^{-2} in the first year, but still 0.5 Wm^{-2} in the second year and larger than 0.1 Wm^{-2} a decade after the forcing.

In the paper we show that the adjustment responses have a time scale of several years. We ascribe this rather long time scale to an adjustment of the ocean, which is forced with the anomalous surface fluxes and wind stresses caused by the tropospheric adjustment and, in turn, sets the lower boundary conditions for the troposphere to respond to the overall warming. We show with an illustrative model that the adjustment processes follow a characteristic time scale of a few years, and do not act as temperature dependent feedbacks.

Thus, the novelty of this work is the idea that the adjustment problem should be regarded and studied as a coupled one. This is also a profound methodological critique, since many previous adjustment studies are done with fixed-SST and slab oceans. This implies that the separation of forcing and feedbacks is even more complicated than assumed. As one rudimentary solution, we suggest the new forcing definition of "virtual forcing", which includes adjustment processes. In the outlook chapter I discuss the motivations and strength of this concept in greater detail.

Chapter 3 Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes

In this study we show that ocean heat flux induced SST pattern changes substantially influence atmospheric radiative feedbacks. This has been shown in a few studies before and is the idea behind the concept of ocean heat uptake

efficacy. However, the existing studies use idealized setups or argue that the ocean heat uptake efficacy is a constant, comprising the shift of heat uptake from a global homogeneous to high latitude amplified pattern. We bridge an important gap in model hierarchy and reproduce time slices of a coupled model with a slab ocean model. We further show that the entire magnitude of feedback parameter reduction during equilibration time in a coupled model can be attributed to the changing surface heat flux patterns. In other words, ocean heat uptake efficacy is time dependent and potentially also spatially rather complex.

Methodological, we introduce a new way to quantify feedbacks: the surface or mixed layer heat flux patterns from a transient state of a coupled model are scaled up and down in magnitude to force a series of slab ocean models, implying different TOA imbalances. This procedure imitates constant ocean heat uptake patterns and shows that for a constant pattern, the feedbacks parameter is indeed linear. Although computationally elaborate, this method circumvents problems with kernel methods, such as testing states far away from where the kernel was developed for.

Overall, we find that conclusions from idealized settings carry over to more realistic settings, meaning the same amount of heat taken up in high latitudes more efficiently changes global temperatures than the same amount of heat taken up in low latitudes. The realistic patterns are more complex than implied in idealized studies though and different feedback components explain the overall effect.

As a sideline implication, this work helps to understand the question of local linear versus non-linear feedbacks. While local feedbacks in idealized and realistic setups depend on local and distant heat flux patterns – i.e., are not constant in time under a changing pattern (see also Rose et al., 2014; Kang and Xie, 2014; Andrews et al., 2015) –, they change substantially less for realistic surface heat flux pattern. This leaves the possibility open to describe local feedbacks on short enough time scales and large enough areal integrals as constant (e.g., Armour et al., 2013; Roe et al., 2015). Fig. 4 shows for the short wave cloud radiative effect as an example that this holds when feedbacks are defined relative to the local or global temperature.

The most pressing implication of this chapter is that in global climate models a priority should be given to correctly represent sea surface temperature patterns, ocean circulation, and ocean-atmosphere interactions. Another implication is that an inherent limit of predictability of surface climate may exist on decadal time scales, due to the difficulty to understand SST pattern formation and internal modes of variability. The outlook discusses questions this work raised.

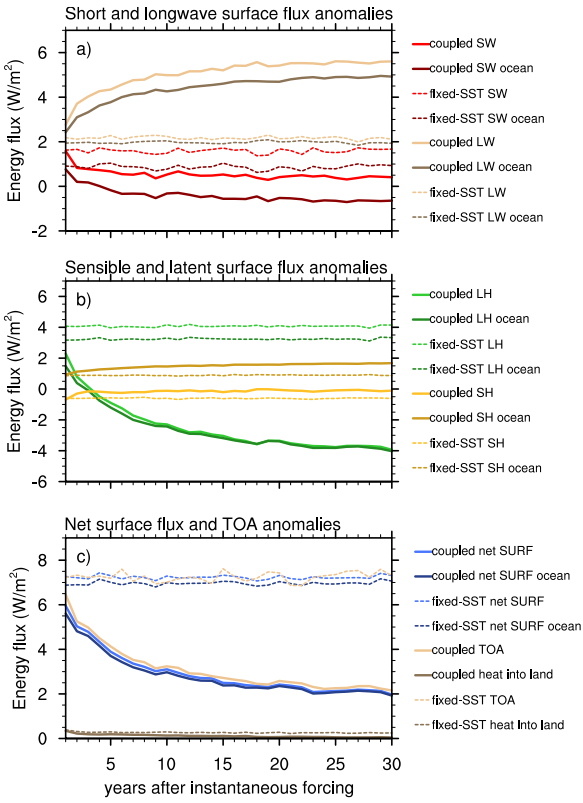

Figure 3: *Surface heat flux component responses to an abrupt4xCO_2 forcing. Dashed lines are the averages of four fixed-SST ensembles, solid lines are the average of 12 coupled simulation ensemble members. In the paper we argue that adjustment is a coupled process, precisely because it includes surface heat fluxes, and thus, should not be studied with fixed-SST simulations.*

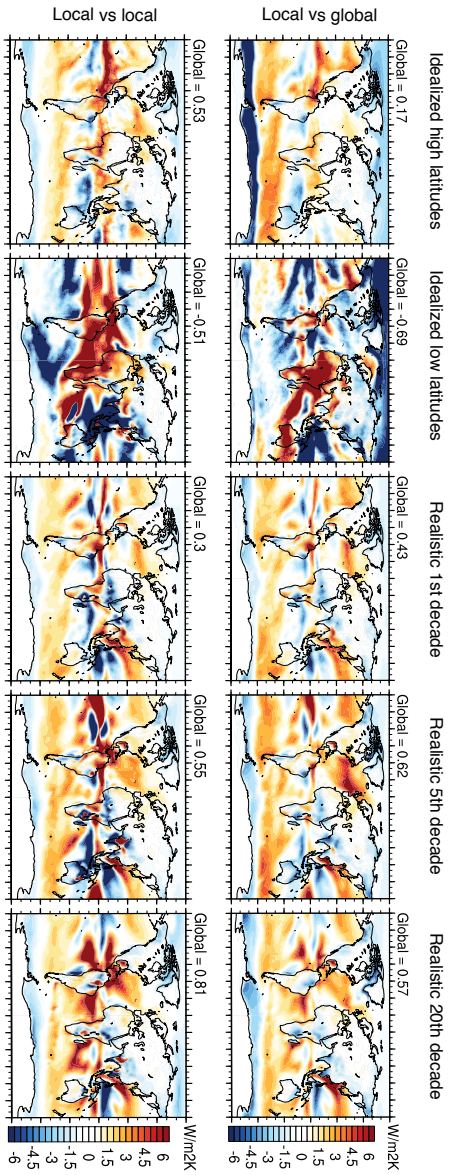

Figure 4: *Short wave cloud radiative effect as local SW CRE against global temperature (as in Fig. 3.2, upper row) and local SW CRE against local temperature (lower row).*

Chapter 4 Nonlinearities in patterns of long term ocean warming

In this paper we deal with how the TOA imbalance is distributed within the climate system through time and in the new equilibrium through space. We explore the centennial to millennial scale patterns of ocean warming and their dependence on time and on forcing level induced ocean circulation changes. Since the sub-surface ocean takes centuries to millennials to equilibrate even after the surface ocean is in quasi-equilibrium with the atmosphere (see Fig. 5) it is less studied and strong assumptions are made to extrapolate for example centennial to millennia scale behavior. In general, little is known on ocean circulation in this time frame, but the interpretation of paleo proxies sensitively depends on assumptions made around mixing or the representativeness of specific places and time scales. We show two common assumptions to be invalid: (a) mid and long-term ocean expansion is not proportional to surface warming and (b) long-term ocean warming is not homogeneously distributed.

We analyze CMIP5 model behavior for different forcing scenarios but also use a model of intermediate complexity, since it allows us to study millennia scale time scales for a large ensemble and several forcing levels. One hypothesis is that an almost undetectable perturbation (of less then 30 ppm CO_2 increase) will distribute homogeneously with depth since the ocean circulation pattern will not respond too much. To detect this small perturbation we run 90 ensemble members with different initial states, and also 90 corresponding control simulations to reduce internal variability. Overall, we find a strong dependence on the forcing level, which not only sets the initial ocean stratification after the forcing is applied (as can be seen in Fig. 5), but also sensitively and non-linearly affects circulation response over several centuries. In turn, this circulation response determines where heat gets stored in the equilibrium and how much its expansion contributes to sea level rise.

This more conceptual work has a lot of implications. Centuries after a forcing is applied, the climate system still adjusts, potentially in difficult to predict terms. Our understanding of the century time scale processes all rely on comparatively low resolution models of intermediate complexity, which potentially may not display the real world sensitivity of ocean meridional overturning circulation (and certainly mixing).

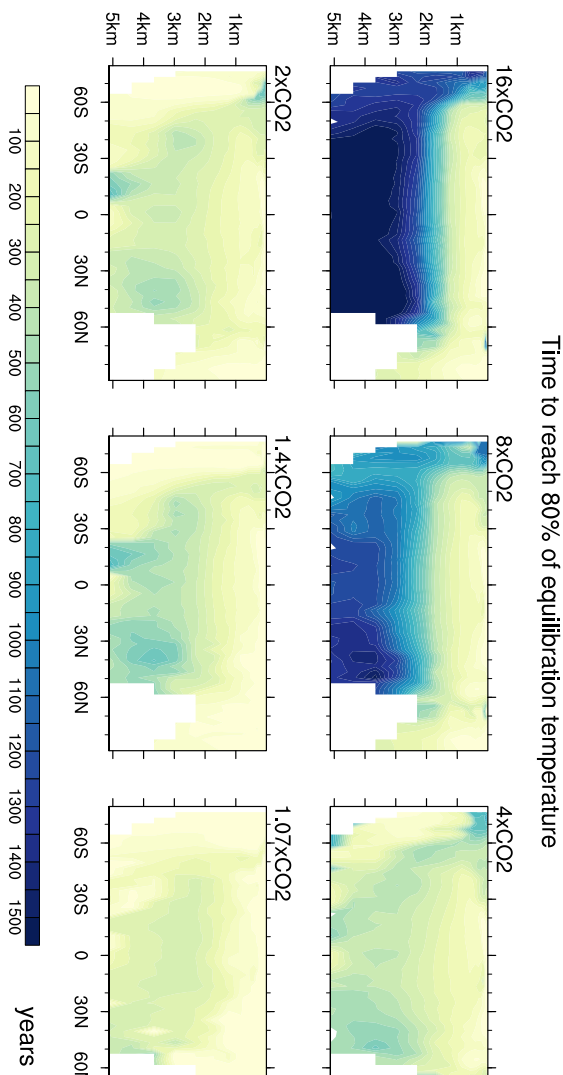

Figure 5: *Zonally averaged equilibration time scales for different forcing levels of EC-Bilt CLIO.*

In short – end of the century – terms, the clearest implication of this work is not to scale patterns in the ocean to extrapolate to a few decades beyond 2100 or between scenarios. This also means that the committed impact of climate change depends on the path of emissions and their potential reduction. Additionally, paleo proxy interpretations should be cautious about their assumptions on forcing history and equilibration state. Our analysis suggests that the common assumption that anomalies in intermediate to deep ocean temperatures represent anomalies of global mean or high latitude temperatures is probably entirely invalid.

Chapter 5 Effective and Equilibrium Climate Sensitivity

Motivated by some of the shortcomings of the other chapters and the inherent limitations of CMIP5, we initiate a Model Intercomparison Project of millennia scale coupled complex climate model simulations ("LongRunMIP"). We do not develop new methods, but collect and homogenize more than 14TB of data of 38 scenario simulations of 14 different models originating from nine modelling centers, each between 1,000 and 6,000 years long. A small number of existing publications discuss single model integrations of century to millennial time scales, but comparing model behavior is fruitful in understanding not only processes in single models, but their robustness or potential biases. The project has the potential to contribute to a range of ongoing discussions in climate science and in paleoclimate, as discussed in chapter 5.6 and 6.

The first aim is to test energy balance models, which predict equilibrium from transient behavior and study the assumptions and limitations of linear feedback analysis. As a first result, we show that the two most wide-spread models of extrapolating transient conditions fail to predict equilibrium conditions for GCMs. I discuss preliminary results on surface temperature patterns, the oceanic Meridional Overturning Circulation, and scaling the response to different forcing levels. I finish by posing a number of pressing questions on which this MIP will, hopefully, provide answers.

1 Feedbacks, climate sensitivity, and the limits of linear models

published by Reto Knutti and Maria Rugenstein, in Philosophical Transactions of the Royal Society, Special Issue on "Feedbacks on climate in the Earth system" (2015) doi: 10.1098/rsta.2015.0146

The term "feedback" is used ubiquitously in climate research, but implies varied meanings in different contexts. From a specific process that locally affects a quantity, to a formal framework that attempts to determine a global response to a forcing, researchers use this term to separate, simplify, and quantify parts of the complex Earth system. We combine new model results with a historical and educational perspective to organize existing ideas around feedbacks and linear models. Our results suggest that the state- and forcing-dependency of feedbacks are probably not appreciated enough, and not considered appropriately in many studies. A non-constant feedback parameter likely explains some of the differences in estimates of equilibrium climate sensitivity from different methods and types of data. Clarifying the value and applicability of the linear forcing feedback framework and a better quantification of feedbacks on various timescales and spatial scales remains a high priority in order to better understand past and predict future changes in the climate system.

1.1 Introduction

Partly originating from control theory, the analysis of feedbacks is a powerful tool to study dynamical systems, in which one quantity affects another, thereby attenuating or amplifying the original signal (for a review Roe, 2009). For example, warmer temperatures lead to melting of snow and ice, which exposes a darker surface that absorbs rather than reflects incoming solar radiation, which leads to more warming and melting than would have occurred if the snow cover area had been fixed. In simple systems with few components and interactions, such feedback frameworks can separate cause and effect, and allow for a mathematical description of a dynamical system. However, there are difficulties in applying it to the global climate system, which is not closed, and where the interplay of different feedbacks and forcings complicate the description. Some feedbacks may only become relevant in the future, or may no longer be relevant (e.g., if there is no snow and ice left), while some changes may be nonlinear, abrupt, or irreversible. For instance, systems like the El Niño-Southern Oscillation could potentially show regime shifts, invalidating simple linear feedback formulations and potentially making feedback analysis less relevant for both understanding the past and predicting the future. Yet despite all these potential complexities, the construction of linear feedback frameworks has been helpful in the past, if applied carefully to parts of

the whole climate system, and within certain bounds on timescales and climate states that we discuss below.

The perspective provided here, focusing on the global forcing feedback framework, emerged from an overview talk presented at the Royal Society Meeting "Feedbacks on climate in the Earth system". We attempt to provide an extended context and perspective to the more detailed papers in this special issue. As a consequence, some conceptual material presented here is not novel, though we hope to stimulate potential avenues of future research.

1.2 The case for forcing feedback frameworks

A specific forcing might affect the climate system response on a large range of timescales. In the usual forward thinking and modeling chain, shown in Fig. 1.1a, the use of fossil fuels leads to greenhouse gas emissions and an increase in their atmospheric concentrations, a change in radiative forcing, which causes a climate response. In the more detailed view in Fig. 1.1b, the change in the CO_2 concentration causes an instantaneous forcing, which – after being adjusted for very fast responses – becomes an effective radiative forcing, defined as the change in the top of atmosphere radiative balance before the surface temperature responds (see Knutti and Hegerl (2008) for an overview). By warming, the surface restores the radiative balance by increasing the radiation to space, but this warming causes water vapor, lapse rate, albedo, clouds, vegetation, ice sheets, permafrost, and/or atmospheric chemistry to change. Those changes – directly or indirectly – affect the Earth's radiation budget, and amplify or damp the temperature response.

Equilibrium climate sensitivity (ECS) is an attempt to combine many of these changes in a tractable manner, and is one of several key numbers that are used to characterize the temperature response of the Earth to a change in forcing or the CO_2 concentration. ECS is usually defined as the equilibrium global average surface warming in response to the radiative forcing from an atmospheric CO_2 doubling, and includes the changes in water vapor, lapse rate, surface albedo and clouds (see magenta box in Fig. 1.1). By definition, in equilibrium the ocean heat uptake is zero, but in a transient climate it damps the warming. The transient climate response (TCR) characterizes the warming at the time of CO_2 doubling after a 1%/yr increase in the CO_2 concentration (see violet box). The Transient Climate Response to cumulative carbon Emissions (TCRE, light blue box) characterizes the warming as a function of the total emitted CO_2, and is relevant to estimate the carbon budgets and emission reductions required for stabilizing global temperature (see Knutti and Rogelj (2015) for an overview). In some sense, the definition of ECS is arbitrary and has survived only because of historical development, convenience in modeling, and the lack of better alternatives. The early generations of climate models included only the water vapor, lapse rate, albedo and cloud feedbacks, and had no appropriate representation of land ice, vegetation, chemistry, or biogeochemical cycles, nor did they include a dynamical ocean component. Doubling the atmospheric CO_2 concentration for a few decades in such a model was therefore a benchmark to characterize the overall temperature response to a well-defined forcing, and a measure of the total feedback on timescales of decades to centuries.

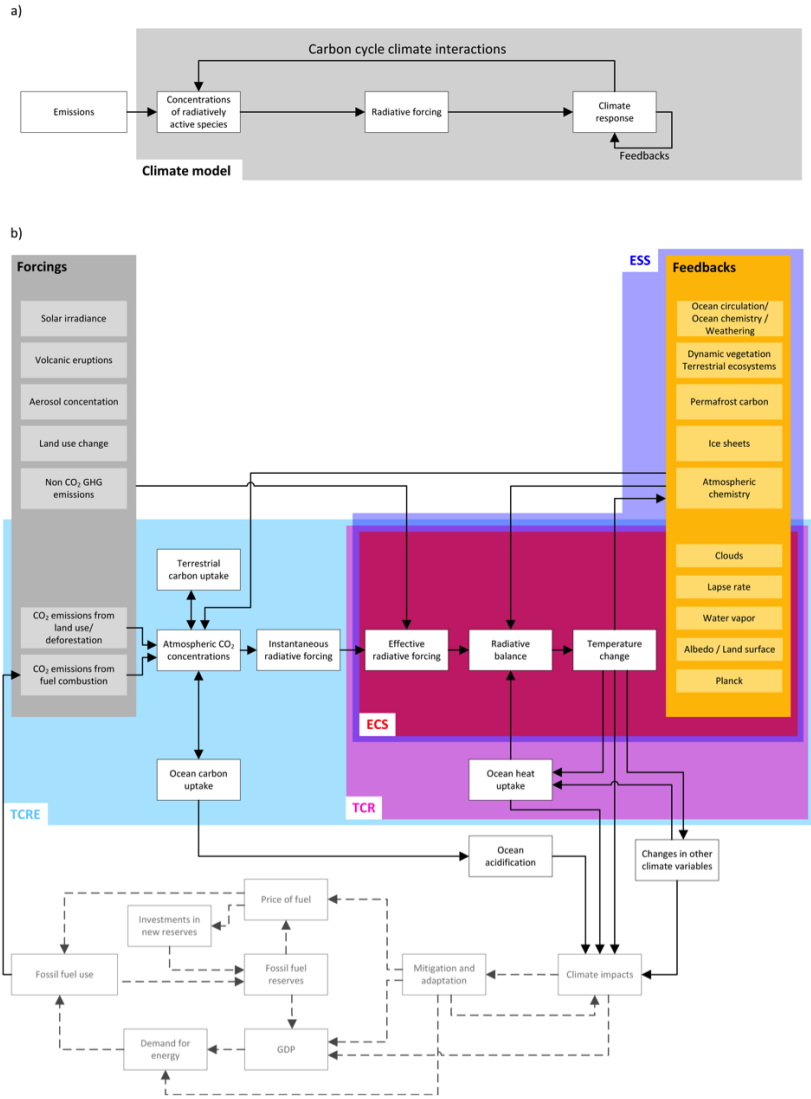
a)
Carbon cycle climate interactions
Emissions
Concentrations of radiatively active species
Radiative forcing
Climate response
Feedbacks
Climate model
b)
Forcings
Solar irradiance
Volcanic eruptions
Aerosol concentation
Land use change
Non CO_2 GHG emissions
CO_2 emissions from land use/ deforestation
CO_2 emissions from fuel combustion
Terrestrial carbon uptake
Atmospheric CO_2 concentrations
Instantaneous radiative forcing
Effective radiative forcing
Radiative balance
Temperature change
ESS
Feedbacks
Ocean circulation/ Ocean chemistry / Weathering
Dynamic vegetation Terrestrial ecosystems
Permafrost carbon
Ice sheets
Atmospheric chemistry
Clouds
Lapse rate
Water vapor
Albedo / Land surface
Planck
ECS
Ocean carbon uptake
Ocean heat uptake
TCRE
TCR
Ocean acidification
Changes in other climate variables
Price of fuel
Investments in new reserves
Fossil fuel use
Fossil fuel reserves
Mitigation and adaptation
Climate impacts
Demand for energy
GDP

Figure 1.1 *(previous page)*: *a) Simplistic and generalized modeling chain (adapted from ref. IPCC, 2007, Fig. 10.1) and b) more refined distinction between feedbacks acting in and on the climate system. Greenhouse gas emissions perturb the radiative balance, which force the temperature to respond. Temperature change is causing various feedbacks (yellow box, interactions between the feedbacks are not marked) to act back onto the radiative balance, which again causes the temperature to adjust. The Equilibrium Climate Sensitivity (ECS, magenta box) covers only some of the feedbacks. In a transient reference framework (Transient climate response TCR, violet box), the rate of ocean heat uptake affects the radiative balance and temperature change in return. The Transient Climate Response to cumulative Carbon Emissions (TCRE, light blue box) characterizes the temperature response to emissions and includes carbon emissions, uptake, and release of the land biosphere and the ocean. The Earth System Sensitivitity (ESS, blue box) includes more feedbacks, generally but not exclusively acting on longer than century time scales. The separation of forcings (gray box) and feedbacks (yellow box) is in some sense arbitrary and has to be defined for each problem. Climate change – through temperature and other variables' change will impact socio-economic systems, which finally will feed back on emissions. See text for further discussion. The feedback loops sketched act on different time scales.*

From Fig. 1.1, it becomes clear that ECS and TCR are rather limited characterizations of a much larger and interactive system. Other feedbacks such as vegetation, chemistry, or land ice are now included in some climate models as their relevance is better understood. Some feedbacks operate on very long timescales that are determined by the internal dynamics of the system, and their response is not proportional to temperature. Thus, a more recent concept is an equilibrium Earth System Sensitivity (ESS, dark blue box) which encompasses all climate (but not human) feedbacks. The separation of ECS and ESS is often made along timescales, with the argument that those feedbacks included in ECS essentially scale with surface temperature, while the others in ESS partly have their intrinsic (and often slower) timescales. However, this does not apply to atmospheric chemistry which responds quickly. Here the reason is a historic one, as the early climate models simply did not simulate interactive chemistry. This supports the argument that the separation of ECS and ESS is somewhat arbitrary in the real world where a lot of processes interact.

How would we go about estimate ECS in the real world? The Earth today is not in equilibrium, and other, non-greenhouse-gas forcings (aerosols, dust, land use, solar, and volcanic), which are smaller than greenhouse gas forcings, are still important locally. Attempting to capture the importance of these other forcings, scenarios of future climate change now prescribe emissions of many gases (Moss et al., 2010). Climate change is also expressed in other variables – from ranges of species' habitats to hail grain size – including their variance and extremes. The state of the changed climate system causes impacts, leading to adaptation and mitigation, which in turn influence the economy and fossil fuel exploitation and use (gray in Fig. 1.1, sketched only roughly to indicate the incompleteness of process understanding), which further influence greenhouse gas emissions. It is tempting to broaden the definition of ECS to include more feedbacks to simplify the comparison with the real world. Even impacts and the human response in terms of adaptation and mitigation could be included in a broader concept of sensitivity (Previdi et al., 2013), encompassing most or all relations shown in Fig. 1.1. However, the decision to incorporate an additional process into "sensitivity" must consider the need to reduce complexity, in order to have a tractable system that is useful for understanding. The human component is a hyper-complex interaction of nature and societies. Humans as biological systems may in theory be described by the laws of physics and chemistry, and could be parameterized similar to other ecosystems, but human de-

cisions, ideas, and inventions can (and have done in history) literally change the course of the world, and thus introduce a problem of predictability of the first kind (sensitivity of the outcome to initial conditions). If climate sensitivity is defined in such broad terms to include human behaviour, it is apt to be unpredictable and fails to provide insight into the climate system.

The idea of the feedback framework in climate science is to break down complex processes and quantify their sensitivities. For long term warming, ECS or TCR may be useful numbers and they explain the largest fraction of uncertainty (Knutti and Hegerl, 2008), but for adaptation purposes, global temperature is of very limited value. For regional change and changes other than temperature, the feedbacks and processes that matter most may be different (e.g., soil moisture, vegetation, or air pollution) from the ones that are most important for TCR or ECS.

Why would we stick to an arguably narrow framework of climate sensitivity, which describes only a limited number of the feedbacks in the real world? A number of reasons partly explain why we have done so for a long time. Firstly, many changes in climatic variables approximately scale with temperature (Tebaldi and Arblaster, 2014; Herger et al., 2015). As a result, global temperature is probably the best proxy for aggregated impacts, even though the relation is likely nonlinear. Global temperature is relatively easy to measure, records extend further back than measurements of most other climate variables, and temperature is more straightforward to reconstruct from paleo data. Together, this provides a way of comparing current and future climate with the climate that would have been without anthropogenic emissions. If we had to reduce climate change to a single aggregate number, for example to agree on a single climate target, global temperature is an obvious choice. Secondly, in the global forcing feedback framework, the radiative forcings and their responses are assumed to be additive, as discussed further in section 1.3. This is key for the relevance of the radiative forcing definition, as it means that $1\,\mathrm{Wm}^{-2}$ of a forcing can be 'traded' against $1\,\mathrm{Wm}^{-2}$ from a different forcing when designing policies towards a climate target, and the total warming is proportional to the total forcing. This additivity is also a key assumption for detection and attribution studies, to break down the observed changes into parts caused by different forcings. Thirdly, many earlier studies (partly based on slab ocean rather than dynamical ocean models) indicated that the global feedback parameter (the inverse of the equilibrium warming per unit forcing) is roughly constant for various forcings and climate states. This is equivalent to a description of a Taylor expansion neglecting higher order terms, as shown in section 1.3. To the degree that this is justified, the global feedback can be used in simple energy balance models to estimate the future warming from future emissions or forcings, or in integrated assessments models. Such models, where the forcing is seen as the cause, and warming as the effect, are known to be a simplification of the real world, but have been crucial for understanding how models of various degrees of complexity respond to perturbations, and to which degree past and future climate change can to first order be described as an energy balance problem (Huber and Knutti, 2012).

Fig. 1.1 does not indicate typical process time scales, but it is obvious that cloud droplet formation acts on different temporal and spatial scales than weathering of rocks or land use changes. Climate sensitivity was defined with a century timescale in mind and, as such, can only characterize certain processes. Fig. 1.2 compares the most common process time scales. The direction – warming or cooling, positive or negative feedback – is not taken into account in this representation, and some processes have different sensitivities for warming than for cooling (see discussion in section 1.3). The colored ellipses indicate different methods to define sensitivity in broad terms. Climate sensitivity is not a quantity that can be measured, and it characterizes only a part of the relevant processes and feedbacks, but it is an emerging property of the system. From past climate it can be approximated by relating equilibrium warming to radiative forcing. In Global Climate Models (GCM), climate sensitivity is normally not tuned, but

it results from aggregating or parameterizing small-scale processes and ignoring long term ones (red ellipse in Fig. 1.1). GCM based estimated of TCR and ECS ignore certain processes even within the time frames they consider (gray bars within the red ellipse).

On short time scales (green ellipse), the observed surface warming, ocean heat uptake, and an estimate of radiative forcing, provide an estimate of the anthropogenic contribution to the observed warming and the global feedbacks (Huber and Knutti, 2012), and therefore ECS and TCR (e.g., Aldrin et al., 2012; Andronova and Schlesinger, 2001; Forest et al., 2002; Forster and Gregory, 2006; Frame et al., 2005; Gregory et al., 2002; Johansson et al., 2015; Knutti and Tomassini, 2008; Knutti et al., 2002; Lewis, 2013; Lewis and Curry, 2015; Otto et al., 2013; Skeie et al., 2014; Tomassini et al., 2007). Those again can be used for probabilistic projections that are conditional on, i.e. constrained by, past warming (Allen et al., 2009; Knutti et al., 2002, 2003; Meinshausen et al., 2009; Rogelj et al., 2012, 2014; Sokolov et al., 2009; Stott and Kettleborough, 2002; Tomassini et al., 2010). Observations and simulations of the response to natural external forcings (volcanic or solar) (Boer et al., 2007; Soden et al., 2002; Wigley, 2005; Yokohata et al., 2005) or unforced climate variations on short, or very long timescales (green and yellow ellipse in Fig. 1.2), or the climatology and seasonal cycle may provide information on feedbacks (Dessler, 2010; Hegerl et al., 2006; Knutti et al., 2006), but the inferred numbers in ($Wm^{-2}K^{-1}$) may differ from those on the century timescale. Both the short term and proxy methods are often called "observational", but it is important to note that they rely on models and assumptions as much as GCMs. Their radiative forcing is derived from a GCM, the magnitude and timescales of internal climate variability often come from climate model control runs or statistical models, and in many cases, strong assumptions about linearity and spatial aggregation are made, as discussed in the next sections. Information from paleoclimate combined with models (e.g., Braconnot et al., 2012; Caballero and Huber, 2013; Hargreaves et al., 2007, 2012; Köhler et al., 2010; Martinez-Boti et al., 2015; Masson-Delmotte et al., 2006; Rohling et al., 2012; Schmidt et al., 2014; Schneider von Deimling et al., 2006) provides further support for an ECS value in the consensus range of 1.5-4.5 °C, but also highlights that feedbacks for warmer or colder states and on longer timescales may differ from those today.

Several pressing questions become clear from Fig. 1.2. The first is why different lines of evidence point to different ECS values. Specifically, some but not all recent studies on the 20^{th} century warming find rather low ECS values (median at or below 2 °C) (Aldrin et al., 2012; Lewis, 2013; Lewis and Curry, 2015; Otto et al., 2013). Climate models show a large spread in ECS, with the spread half as big as the median value. The highest uncertainty can be attributed to the cloud feedbacks (traceable to certain cloud types and regions), and the lapse rate feedback (Andrews et al., 2015, 2012; Bony et al., 2006; Vial et al., 2013). But all comprehensive climate models indicate sensitivities above 2 °C, and those that simulate the present day climate best (Fasullo and Trenberth, 2012; Huber and Knutti, 2012; Sanderson et al., 2015; Sherwood et al., 2014b) even point to a best estimate of ECS in the range of 3-4.5 °C. The second question is how to infer present-day ECS from the climate sensitivity in warmer or colder states, from shorter or longer time scales, or for a non CO_2 perturbation ("mapping"). Both questions are partly rooted in the use of simple linear forcing feedback models with a constant feedback parameter, discussed in depth in the following sections.

1.3 Climate sensitivity, timescales, and commitment

1.3.1 General concepts

In equilibrium, the global radiation budget, R, – the sum of net incoming solar shortwave and outgoing terrestrial longwave radiation – is closed ($R = 0$). The degree of

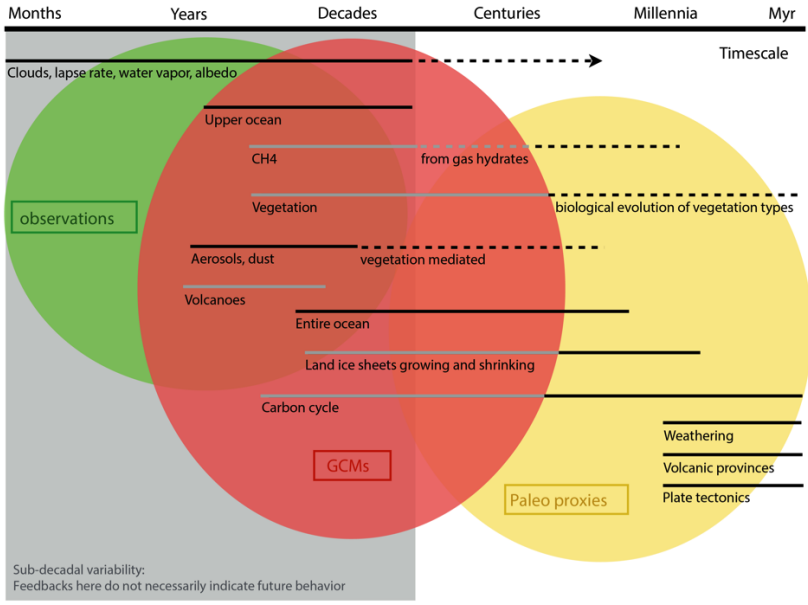

Figure 1.2: *Timescales of climate relevant processes. Light gray bars indicate processes that act on timescales that a global climate model (GCMs) can resolve, but are usually assumed to be (partly) inactive or non-existent. Dashed lines indicate timescales where specific feedbacks are weaker or only operate under certain circumstances. The arrow for clouds, lapse rate, water vapor, and albedo indicates that those feedbacks operate on short timescales, but because the surface warming takes centuries or more to equilibrate, these feedbacks continue to change and affect the overall response of the systems up to millennia. This can apply similarly to other feedbacks which respond quickly but continue to change over long timescales in response to other feedbacks. The colored ellipses each cover different methods used to estimate climate sensitivity. The vertical ordering of the feedbacks is arbitrary. Models of intermediate complexity (EMICs) can bridge the gap between GCMs and paleo proxies, e.g., by including carbon cycles, weathering and ice sheets. Usually there are trade offs between simulating very long time scales and the level of detail of short time scale processes.*

imbalance ($R \neq 0$) at some time following a perturbation can be ascribed to the temperature response itself (ΔT), and changes induced by the temperature response, the radiative response ($\alpha\Delta$T), thus R = R(ΔT, α(T)) (e.g., Roe, 2009; Zaliapin and Ghil, 2010). To study how a small change in the radiation budget ΔR is related to the temperature response, one can use the Taylor expansion of R, in T and α(T):

$$\Delta \mathrm{R} = \frac{\partial \mathrm{R}}{\partial \mathrm{T}} \Delta \mathrm{T} + \frac{\partial \mathrm{R}}{\partial \alpha} \frac{\partial \alpha}{\partial \mathrm{T}} \Delta \mathrm{T} + \mathcal{O}((\Delta \mathrm{T})^2) \tag{1.1}$$

The perturbation of the radiation budget is the effective radiative forcing F minus the heat flux or TOA radiative imbalance N, which is non-zero as long as the system is not in equilibrium. The reference height of the heat flux is usually the tropopause. Over time scales longer than a year this is the same as the heat flux into the ocean, ice, and land. The first term on the right-hand side describes the strongest negative feedback, sometimes called the Planck feedback. Increased temperatures lead to increased top of the atmosphere outgoing longwave radiation. Other feedbacks would have to be stronger than the Planck feedback to lead to a runaway climate. The second term on the right-hand side describes the sum of the feedbacks, which scale with the temperature response:

$$\frac{\partial \mathrm{R}}{\partial \alpha} \frac{\partial \alpha}{\partial \mathrm{T}} = \sum \frac{\partial \mathrm{R}}{\partial \alpha_i} \frac{\partial \alpha_i}{\partial \mathrm{T}},$$

with i = water vapor, lapse rate, albedo, and cloud feedback. These are the common physical feedbacks analyzed in CMIP5-type climate models (see the violet TCR box in Fig. 1). The feedbacks can be positive (e.g., water vapor) or negative (e.g., lapse rate) and sometimes difficult to determine (e.g., for the cloud feedbacks). Processes that involve several of the feedbacks can lead to correlations between them. For example, the sum of the water vapor and lapse rate feedback is better constrained than the individual parts (Soden and Held, 2006). Finally, the last term on the right-hand side of Eq. 1.1 is the sum of all higher order terms of the Taylor expansion, representing the non-linearities of individual process and the interaction between the different feedbacks.

The linear approximation generally neglects the last term since the temperature response from interactions between the feedbacks is usually small. Focusing on the linear term helps to distinguish and quantify the single feedbacks' influence on the final response (Klocke et al., 2013). However, it is not clear what a "small perturbation" comprises and when higher order terms should be taken into account, such as for high emission scenarios or paleoclimate studies with large perturbations or additional active feedbacks (Fig. 1.2). Another limitation arises because the climate system may include thresholds and tipping points, where the linear assumptions are not justified (Zaliapin and Ghil, 2010). As discussed in section 1.2, part of why the linear approximation is so widely used is its simplicity, convenience, and lack of alternatives; its validity is not in all cases examined. Studies investigating limitations of the linearization would help to strengthen trust in the findings obtained within the linear framework.

All terms in Eq. 1.1 are globally defined and hold for large temporal integrated scales. To analyze feedbacks on a local scale a heat-flux divergence term has to be added (Armour et al., 2013; Roe et al., 2015). The meridional structure of feedbacks tends to compensate for local non-linearites (Feldl and Roe, 2013).

As shown, the climate feedbacks are treated as relative contributions to the response compared to the strongly negative Planck feedback. One can define a reference temperature (increase T_0) caused by the Planck feedback (about 1.1 °C for a doubling of the atmospheric CO_2 concentration). The additional temperature response caused by the feedbacks can then be described by $\Delta \mathrm{T} = \Delta \mathrm{T}_0/(1 - \mathrm{f})$ with $f = \partial \mathrm{T}/\partial \mathrm{R}(\partial \mathrm{R}/\partial \alpha \partial \alpha/\partial \mathrm{T})$ the feedback factor. For an ECS value of approximately 3 °C, this implies that more than half of the warming is caused by feedbacks in the climate system, and less than half is a direct Planck response to forcing.

Accepting the linear assumption and adopting the naming conventions mentioned above, one can rewrite Eq. 1.1 as

$$F - N = \lambda \Delta T \tag{1.2}$$

The linearization leads to the assumption that the feedback parameter λ is constant, meaning the net feedback strength is independent of the climate state ΔT and the forcing F (Gregory et al., 2004). It is assumed that the real world climate system has an *a priori* unknown λ and climate models can help finding the value of that λ and then project ΔT into the future. When the system settles into the new equilibrium, the net heat flux, N, at the top of the atmosphere is zero, and the temperature change necessary to reach the new equilibrium $\Delta T = F/\lambda$, is – by convention and as defined in section 1.2 – , the Equilibrium Climate Sensitivity (ECS), if the forcing is a doubling of the preindustrial CO_2 concentration. The less efficient the Earth is at emitting energy to space (smaller λ), the higher temperature increase ΔT is necessary to restore the balance. By incorporating the heat uptake as a measure of the transient climate response, the global feedback (and thus ECS) can be inferred from $\Delta T = (F - N)/\lambda$. The transient response can be approximated from the ratio between temperature change and forcing, and is smaller than ECS. As a consequence, keeping F fixed at a certain time during a warming simulation would result in further surface warming for several centuries. This is the commitment warming or "warming in the pipeline" (Hansen et al., 2005a; Meehl et al., 2005). The magnitude of the commitment warming depends on ECS, because the response timescale is longer, and therefore the fraction of realized equilibrium warming (discussed later in Fig. 1.3) is smaller, for higher ECS. In other words, if ECS is high, the current temperature (expressed as a fraction) is further away from the equilibrium temperature for that forcing. As a consequence, TCR becomes less sensitive to ECS for high ECS (i.e. a high and very high ECS are difficult to separate in their short term response as indeed in many other observables), which often results in probability density functions with fat tails to high values (Knutti et al., 2005; Roe and Baker, 2007; Baker and Roe, 2009).

The description of equilibration in Eq. 1.2 – as N approaches 0 – is a choice of a reference framework and might be more helpful for certain questions than for others. While the global energy balance has to be closed of course, the ability of Eq. 1.2 to physically explain different time scales is limited. There is no physical necessity that the response scales with the global mean surface temperature change, although many variables do (see discussion in section 1.2).

To study the validity of the assumptions discussed above, and to analyze different processes and time scales, step forcing experiments are useful. The forcing, F, doesn't vary in time (as it does in reality), but is prescribed as an instantaneous increase or decrease and then held constant, to let the system approach a new equilibrium. A fully equilibrated state is never reached in the real world, because boundary conditions (e.g., orbital forcing, tectonics) always change, and some feedbacks have very long response timescales. Nevertheless, these experiments are the cleanest method of studying the time scales of different processes involved in the radiative restoration or equilibration. Climate model intercomparisons reveal a large spread in time scales for a certain responses (Andrews et al., 2012; Caldeira and Myhrvold, 2013). This indicates a large uncertainty when analyzing climate change impacts and risks. Step experiments can be further used to predict the response to a more realistic time varying CO_2 forcings (Geoffroy et al., 2013b,a; Good et al., 2011, 2013).

1.3.2 Coupled model results

We use two models to illustrate some of above concepts, and to highlight the limitations of the linear forcing feedback framework. First, the Community Earth System Model

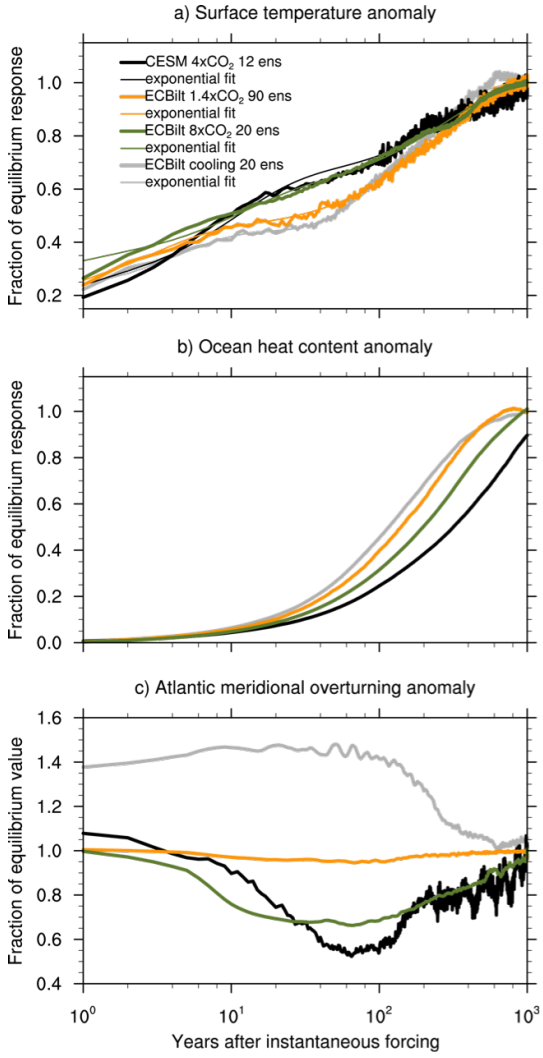

Figure 1.3: *Response timescales, expressed as fraction of the realized equilibrium response, for the global surface temperature (a), the global ocean heat content (b), and fraction of the final equilibrium value of the maximum Atlantic Meridional Overturning Circulation at 30 °N (c) for different models (CESM and ECBilt-CLIO) and forcing levels of 1.4 to 8xCO_2 and cooling to 100ppm. For CESM only the 12 longest runs are used here. The number of ensemble members is noted in the label (ens).*

(CESM1.0.4), a comprehensive ocean-atmosphere-land-sea ice model, is used with fixed vegetation (Bitz et al., 2012; Danabasoglu et al., 2012; Gent et al., 2011). A set of 120 ensemble members branched off from different control run years - thus different in their initial oceanic, atmospheric, and sea ice state - are forced with an instantaneous quadrupling of the CO_2 concentration from the preindustrial value. All members are run for two years, twelve for 100 years, six for 250 years, and one member for 1300 years. Its final state is regarded as being equilibrated to calculate the fraction of equilibration shown in Fig. 1.3a, although the deep Southern Ocean is still adjusting. The novel result here is that the forced response (show here as the anomaly to the control run) and therefore the changes in the global feedback can be estimated on all timescales due to the many ensemble members. Most GCM studies using the energy balance equation (Eq. 2) are done with a 150 year time series and one or a few simulations for each model (e.g., Andrews et al., 2012; Armour et al., 2013; Good et al., 2015). The second model is ECBilt-CLIO, a model of intermediate complexity, with a three level quasi-geostrophic atmosphere with simple parameterizations for the diabatic processes and a free-surface ocean general circulation model coupled to a thermodynamic-dynamic sea-ice model (Goosse and Fichefet, 1999; Opsteegh et al., 2011). We conducted five step forcing experiments composed of instantaneously increasing the CO_2 concentration 1.4, 2, 4, 8, and 16 times above the preindustrial concentration, and one step experiment with reduced forcing. In this cooling scenario, the CO_2 concentration is instantaneously set to 100 ppm, thus 0.35 times the preindustrial value of 280 ppm. For each of the six ECBilt-CLIO experiments we simulate - depending on the signal to noise ratio – 10 to 90 realizations of the same forcing from different initial conditions, all of which are run for 1000 years. One member per experiment is run for 10,000 years until equilibrium. These simulations provide insight into how the global feedback changes with different forcing levels, and from transient to equilibrium.

Fig. 1.3a shows the realized temperature response at a certain time (relative to equilibrium) for the ensemble average of four different experiments: The 4xCO_2 CESM (black), 1.4xCO_2 ECBilt-CLIO (orange), 8xCO_2 ECBilt-CLIO (green), and the cooling ECBilt-CLIO (grey) all in thick lines. The assumption that λ is independent of forcing level and climate state or temperature implies that at all times the fraction of equilibration is the same in all experiments, which is not the case. There are roughly three time scales that all experiments have in common: A short timescale lasting up to a few years, a decadal timescale, and a century timescale, consistent with processes operating on different timescales as shown in Fig. 1.2. Despite the instantaneous forcing, the realized warming is only 30-50% after a decade, and 60-80% after a century, confirming the commitment warming idea discussed above and in section 1.3.3.

A minimized-least-squares fit of a sum of three exponentials to the dimensionless temperature response function θ(t) of the form

$$\theta(\mathrm{t}) = 1 - (\theta_0 e^{-t/\tau_0} + \theta_1 e^{-t/\tau_1} + \theta_2 e^{-t/\tau_2})$$

is shown as thin lines in Fig. 1.3a. The choice of the exponential function is arbitrary – a sum of two exponentials, a fit to a heat diffusion equation, or a transfer function might be valuable for certain purposes as well (Caldeira and Myhrvold, 2013; MacKay and Ko, 1997). The time scales reveal the differences: τ_0 ranges from 0.4 to 4.7 years, τ_1 from 2.6 to 50 years, and τ_2 from 194 to 310 years. The models differ most on decadal time scales, with the weakest forcing case (orange) having a small warming initially (relative to equilibrium) and an increased rate of warming after one hundred years, while the stronger forcings (green and black) lead to initially stronger warming and a slower increase on the century time scale. The amount of realized warming at a given time differs up to 15% between the experiments. The rate of temperature change involved when approaching a cooler state is initially smaller, but after some decades it is larger than in the warming situation.

Panel b shows the oceanic time scales, which are of course much longer, leading to

small fraction of realized warming. Models with initially large atmospheric warming have a delayed oceanic response. The spread of realized warming or cooling is up to 30% around year 400. One reason for the differences is that stronger warming leads to a higher ocean stratification, which reduces diffusive heat uptake (Stouffer, 2004; Stouffer et al., 2006).

Finally, panel c shows one of several reasons why the oceanic heat uptake efficiency changes over time. The Atlantic Meridional Overturning Circulation decreases due to the freshwater and heat flux forcing, but reaches its control run strength after around 1000 years. It responds within decades and a decreased AMOC on decadal time scales leads to an increased heat uptake (panel b) (Knutti and Stocker, 2000) and reduced surface warming. The magnitude of the AMOC reduction depends on the magnitude of warming. In the cooling case, after strengthening for a decade, the AMOC reduces by a few Sverdrups and stays at its new state without re-strengthening as it does in the warming case.

1.3.3 The limits of linear models

So far, we have shown that not only different models show different time scales of equilibration, but also that within one model the response time scales depend on the forcing magnitude and sign. To analyze the constancy of λ, Fig. 1.4a shows top of the atmosphere (TOA) radiative imbalance (N in Eq.2) versus the surface temperature anomaly (ΔT) for all experiments. The slope of the regression line through the points of one experiment corresponds to λ, and it should be a straight line (Gregory et al., 2004). The annual averages of each ensemble member are depicted by small dots, while the large dots are initial condition ensemble averages. Annual averages are shown until year 150, after which decadal averages reduce the large internal variability, which dominates over the small forced signal close to the equilibrium. The standard way to estimate the climate feedback parameter, effective climate sensitivity (the intersect of the regression line with the horizontal axis), is to linearly regress annual averages of year 1 to 150 of one realization of a 2xCO_2 step forcing simulation per model (e.g., Andrews et al., 2015, 2012).

There are several known issues with this regression method and the linear assumptions described in section 1.3.1. It is unclear how much of the signal in the first year is impacted by the initial conditions and by the tropospheric adjustment to the application of the forcing (Andrews and Forster, 2008; Bala et al., 2010; Lambert and Faull, 2007; Sherwood and Fu, 2014). Fig. 1.4a shows a very large spread of responses for the first years. For example, the 120 ensemble members CESM (black) differ by more than 2 Wm^{-2} and by 1 K for the same forcing in the first year. A deviation from a constant λ has been found in earlier studies not only for the annual time scale but the first two decades (Armour et al., 2013; Williams et al., 2008; Winton et al., 2010, 2014) and is treated so far inconsistently, by cutting off a few years before regressing N against ΔT, leading to an ambiguous definition of the effective radiative forcing and effective climate sensitivity. Efficacy factors are used for different forcing agents (to account for the different spatial forcing distribution, shortwave vs. longwave and top of atmosphere vs. surface forcings) and the ability of the ocean to cool the atmosphere by taking up heat (Hansen et al., 2005b; Winton et al., 2010). A dependency of λ on century time scales has been studied in just a few models (Meraner et al., 2013; Senior and Mitchell, 2000) and can be ascribed to the cloud, albedo and water vapor feedback depending non-linearly on temperature. Closely related is the dependency on the forcing level, i.e. the temperature dependency not only over equilibration time – thus, temperature – but also as climate base state (e.g., that surface albedo feedbacks will be weaker in a much warmer world without snow and ice) (Colman and McAvaney, 2009; Good et al., 2015; Jonko et al., 2013, 2012). State dependency also applies to paleo studies (Caballero and

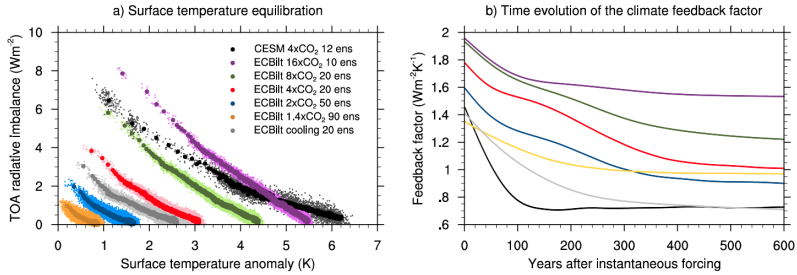

Figure 1.4: *(a) Surface temperature equilibration (imbalance N versus temperature change ΔT) for different forcing levels (colors) and models (CESM and ECBilt-CLIO) where each ensemble member annual averages is a small dot and the ensemble mean a large dot (annual until year 150, decadal averages until equilibration). The number of ensemble members is noted in the label (ens). For the first two years 120 CESM ensemble members are used, afterwards only 12. (b) Time evolution of the climate feedback parameter λ, according to the "moving bin regression", thus ΔN/ΔT. For the cooling case (gray) absolute values are shown.*

Huber, 2013; von der Heydt et al., 2014; Zeebe, 2013). Finally, fully coupled GCMs, with a deep ocean can amplify feedback magnitudes of lapse rate and short wave cloud feedbacks compared to their slab ocean version (Boer and Yu, 2003b; Danabasoglu and Gent, 2009; Jonko et al., 2012). Recently it has been suggested that the non-constancy in the global λ is caused by the evolving spatial surface temperature pattern, which (through ΔT) enhances certain local feedbacks at different times (Armour et al., 2013). Further, it has been shown that the evolving sea surface temperature pattern alone could explain the time or state dependency of λ (Andrews et al., 2015; Gregory et al., 2015).

To quantify the dependency of λ on the forcing level and the temperature or integration time, we calculate the local derivative (ΔN/ΔT) in each point. The radiative imbalance is regressed against temperature for all ensemble members of each model in a certain temperature bin – a few Kelvins wide – which is moved in small steps throughout the temperature range. Different bin widths are used for each simulation, according to the level of forcing and density of points in the N-T-space. This "moving bin regression" circumvents the common problem of either putting more weight to later years when using annual averages, or not addressing the first years when averaging over decades before regressing N against Δ T. The evolution of λ over the temperature range obtained by this moving bin regression is then transferred back to the time domain, shown in Fig. 1.4b. The apparent time dependence is a temperature dependence. Time, in our case, is characterized by how close a state is to the equilibrium state. After year 600, all model simulations show a near-constant λ (cut off in panel b). The feedback parameter decreases especially strongly within the first hundred years. For CESM, λ reduces from 1.5 to 0.7 Wm^{-2}K^{-1}. The CMIP5 model mean value obtained with the standard regression method is 1 ± 0.5 Wm^{-2}K^{-1}. Accordingly, the effective climate sensitivity increases, in CESM, from 4.2 to 6.8 K for 4xCO$_2$. In the runs with a strong CO$_2$ forcing, the time it takes to reach a roughly constant λ level is several hundred years shorter, and the absolute value is higher, than for the lower CO$_2$ forcing levels. Even after several hundred years λ has a small trend. Using five year instead of annual averages lead to the same result on time scales longer than 10 years.

1.4 Are the current concepts of feedbacks and climate sensitivity still useful?

1.4.1 What have we learned from simple models?

Describing a complex system like the climate with a very simple model inevitably means that many factors are ignored, or assumed to be constant. The results above show that the global temperature response to different forcing magnitudes and timescales cannot be fully described with the assumption of a constant feedback parameter λ even in models that ignore long-term Earth system feedbacks (ice sheets, dynamic vegetation, permafrost), non-CO_2 forcings, chemistry, and land used change. In our models, the feedback parameter varies by about 50% or more between different forcing magnitudes and over time as the system approaches equilibrium. The concept of a universal constant climate sensitivity as a fundamental climate system property is very likely wrong, even when ignoring many feedbacks and forcings. This could be an explanation – next to model biases in feedback strength – for the questions outlined in section 1.2 (Fig. 1.2). The inconsistency of ECS estimates based on the observed warming and those based on GCMs with freely evolving SST evolution could be partly caused by the assumption of a constant λ. The estimates based on the observed warming, which use an effective radiative forcing estimated from GCMs together with the assumption of a constant λ, would be biased low, if λ would in fact not be constant but time or temperature dependent, as shown in Fig. 1.4b. In the same way, a state and temperature dependency of λ makes the mapping of GCM, paleo-proxy, and short term observational estimated sensitivities a lot more difficult.

Does this imply the zero-order linear energy balance model is useless? A model is always wrong with regard to reality in a strict sense, but the constant feedback parameter model may still be an adequate approximation for some purposes. As an example, in our case, running the CESM model for 200 years and ignoring the first 150 years for the regression of N against ΔT, would allow us to predict the further evolution of the model. We argue that the quote "modeling for insight, not numbers" makes an essential point here (Bloomfield and Updegrove, 1981). We have to conclude that the global linear forcing feedback model may be of limited value to estimate quantities like the ECS of the real world, or at least we have to be more careful in understanding and quantifying in which range of forcings, timescales and climate states a simple model with a constant feedback parameter can be adequately used. But irrespective of whether the numbers tell us much about the real world, such simple models are, and have been, valuable tools to understand fundamental properties of the system (Held, 2005).

For example, the fact that the transient response simulated in models (or observed, e.g., as the 20^{th} century warming), particularly on short timescales, becomes less sensitive to ECS at high sensitivities, and that it is therefore harder to constrain the upper bound on ECS (Allen, 2006; Knutti et al., 2005; Roe and Baker, 2007; Baker and Roe, 2009) has in fact been noted decades ago with simple energy balance models. Wigley and Raper (1990) pointed out "that the response of the climate system to high-frequency forcings such as volcanic eruptions and the seasonal insolation cycle must be virtually independent of the sensitivity. High frequency information is therefore of little value in trying to estimate, empirically, the climate sensitivity. This is an obvious, but little appreciated result." Wigley and Schlesinger (1985) wrote that the "the observed global warming over the past 100 years can be shown to be compatible with a wide range of CO_2-doubling temperature changes ECS", and as a consequence "it may be very difficult to determine ΔT_{2x} (ECS) from observational data". Recent evidence from observations and models that the climate system will continue to warm for a constant forcing, the commitment warming (Hansen et al., 2005a; Meehl et al., 2005), can be traced back to Siegenthaler and Oeschger (1984), and Wigley and Schlesinger (1985)

who noted that "at any given time, the climate system may be quite far removed from its equilibrium with the prevailing CO_2 level" and Schlesinger (1986), who wrote that "sequestering of heat into the ocean's interior is responsible for the concomitant warming being only about half that which would have occurred in the absence of the ocean. These studies also indicate that the climate system will continue to warm towards its yet unrealized equilibrium temperature change, even if there is no further increase in the CO_2 concentration." These same authors also demonstrated the causes, shown in Fig. 1.3, namely that the characteristic timescales to reach equilibrium range from decades to centuries. These response timescales, and as a consequence the commitment warming, depend on the feedback strength and sensitivity of the model. Hansen et al. (1985) noted that "the response times are particularly sensitive to (i) the amount that the climate response is amplified by feedbacks and (ii) the representation of ocean mixing. If equilibrium climate sensitivity is 3 °C or greater for a doubling of the carbon dioxide concentration, then most of the expected warming attributable to trace gases added to the atmosphere by man probably has not yet occurred."

The basic ideas of additive feedbacks enhancing the Planck response also go back to work by Hansen et al. (1984), and earlier pioneering work, both conceptual and based on climate models (Budyko, 1969; Manabe and Wetherald, 1967; North et al., 1981; Sellers, 1969). All of those old insights are qualitatively still correct, and helpful as thought experiments. More recent work has helped to clarify some of the concepts and point to their limitations (see discussion in section 1.2-3). As GCMs become more complex and include more feedback processes, simple models are necessary to aggregate, approximate, and understand the complex models (Bouttes et al., 2013; Gregory et al., 2015; Held, 2005).

1.4.2 Have we made progress?

ECS was initially used as a model benchmark that was simple to calculate and well defined, an overall measure of the response to increased atmospheric CO_2. It is neither a characterization of all aspects of climate change, nor the most relevant number for policy for all questions. The anchoring on ECS as the holy grail of climate science since the early report by Charney (1979) is not helpful. Some feedbacks like clouds were challenging back then (Cess et al., 1989) and still are (Bony and Dufresne, 2005; Bony et al., 2015; Fasullo and Trenberth, 2012; Sherwood et al., 2014b; Stevens and Bony, 2013), and as a result the uncertainty in climate projections has not decreased much (Knutti and Sedlacek, 2013). But observations and models have greatly improved, paleoclimate has given us a substantially improved view of what has and could happen, we know how to model many processes more realistically, and we have a better understanding of the robust results and key uncertainties. Charney based his conclusions on essentially two global climate models, citing five sources, of which a single one was actually published (Manabe and Wetherald, 1975), the other sources were in press, submitted, or labeled as personal communication. The published model by Manabe dates back to 1969 and has a limited computational domain with equal areas of land and ocean, an idealized topography, no heat transport by ocean currents, and fixed cloudiness. Thus, the fact that the range for climate sensitivity today is similar as was guessed by Charney over three decades ago based on sketchy evidence should not be interpreted as a lack of progress, and using the range of ECS as a measure of success for climate research fails to characterize the state of research.

1.4.3 Possible ways forward

There are 'top down methods', in which a global feedback is inferred from a global energy balance equation, and 'bottom up methods', in which the total feedback is an

emerging property of the myriad of processes that we try to model quantitatively based on insight into each process and data obtained to constrain or parameterize it. And there are, of course, methods in between that combine elements. All methods have in common that they are a fusion of models and observations, and there is no pure observational constraint on ECS. Either we define a simple conceptual model like an energy balance model, aggregate the inputs and constrain ECS, such as relating forcing to cooling in the Last Glacial Maximum. We then use complex models to argue that the simple model is correct and consequently use simple models to predict future warming. Alternatively, we use a complex model directly and relate whatever observations we have straight to model quantities, and use a constrained set of models for prediction (Annan and Hargreaves, 2015; Hargreaves et al., 2012; Rowlands et al., 2012; Schmidt et al., 2014). In this case, the mapping of a paleoclimate sensitivity to a modern ECS is not prescribed, but is implicit in the GCM by the fact that feedback changes spatially and as a function of the climate state in the GCM. In all of those questions, the treatment of uncertainties is key. In an energy balance approach, the uncertainties for different time periods are dominated by either uncertainties in radiative forcing, feedback, ocean heat uptake, or natural variability. For paleoclimate, the perturbations are large and the response is close to equilibrium, but forcing and response are uncertain. The strength of the feedbacks may differ, and additional feedbacks may become relevant, as discussed with the difficulties in defining ECS vs. ESS in section 1.2. For short timescales and forcings other than CO_2, the feedbacks are different and variability is large. For climatological constraints, the problem is that climate models have common biases pointing to common problems in representing key feedbacks, because many relevant processes are not resolved but parameterized. Therefore, all methods have uncertainties in the climate models, the observations, the forcings, in structural and statistical assumptions (e.g., priors in Bayesian methods, or assumptions about constant feedback parameters), and in how the estimated sensitivity relates to the present day ECS in which we are interested.

All methods, but in particular the 'bottom up' which attempt to simulate each individual process accurately, of course require a detailed process understanding to ensure that no important feedbacks are overlooked. This again requires high quality long term and spatially resolved observations, and larger computing capacity to improve (and at some point eliminate where possible) parameterizations of key processes in climate models. New approaches in data assimilation and bridging the gap between numerical weather prediction and climate modeling could be important steps in that direction (Palmer et al., 2008; Rodwell and Palmer, 2007).

The understanding of single feedbacks has increased dramatically in the last few years. The focus has moved to understanding the effect of the temperature pattern $\Delta T(\text{lat, lon, time})$ that acts on local feedback processes and their aggregation to the global ΔT term. Analyzing local scales complicates feedback analyses since the skill of GCMs in simulating regional and local processes is reduced and model comparisons are more difficult. Trying to understand local feedbacks also includes the evolution of the pattern of ocean heat uptake, heat convergence, and TOA imbalance, and research on this subject has barely begun. Understanding regional changes though is more relevant for impact and risk assessments and might bridge the gap between the understanding of global energy budget constraints and localized impact studies. The structural problem of separating individual feedbacks in models – e.g., by keeping parts of the model fixed, or by regression, radiative kernel, or partial radiative perturbation – and comparing them to observations – in which partial derivatives are impossible – persists (e.g., Bony et al., 2006; Klocke et al., 2013). Next to the evaluation of the full blown feedback processes in the models, a key challenge is to study the limits of using the linear framework discussed in this paper. How far can one push a GCM into being very sensitive or very insensitive to explore the range of plausible magnitudes of feedbacks and their rate of change? Do cloud, convection, and aerosol parameterizations bias GCMs to be too

sensitive, or not sensitive enough? For which purposes can we safely use the effective radiative forcing estimates of the linear regression methods? Over which time frames is the assumption of a constant λ justified? Can GCMs serve as a perfect model test bed for simple frameworks, as shown in Fig. 1.4 above? For which climatic base states, feedbacks, and their interaction would it be wise to include non-linear descriptions? For which temperatures, forcing scenarios, and locations does the rate of change of the feedback term matter? When is using a certain fit to estimate the global or regional temperature response justified? How does the coupling of ocean, atmosphere, and sea ice determine the evolution of surface temperature patterns enhancing different feedback processes? How can we understand uncertainty propagation in nonlinear systems, with correlated uncertainties, and using computationally expensive climate models? In the light of all these questions, we argue to further explore various uses of feedback frameworks rather than squeezing them into a one-fits-all-concept, and to carefully explore the applicability and predictive capacity of each concept for a range of purposes.

2 Multiannual ocean–atmosphere adjustments to radiative forcing

published by Maria Rugenstein, Jonathan Gregory, Nathalie Schaller, Jan Sedlá-ček, and Reto Knutti in Journal of Climate (2016) doi: 10.1175/JCLI-D-16-0312.1

In radiative forcing and climate feedback frameworks, the initial stratospheric and tropospheric adjustments to a forcing agent can be treated as part of the forcing and not as a feedback, as long as the average global surface temperature response is negligible. Here, we analyze with a very large initial condition ensemble of the Community Earth System Model how the ocean shapes the fast response to radiative forcing. We show that not only the stratosphere and troposphere but also the ocean adjusts. This oceanic adjustment includes meridional ocean heat transport convergence anomalies, which are locally as large as the surface heat flux anomalies, and an increase of the Atlantic Meridional Overturning Circulation. These oceanic adjustments set the lower boundary condition for the atmospheric response of the first few years, in particular, the short wave cloud radiative effect. This cloud adjustment causes a non-linear relationship between global energy imbalance and temperature. It proceeds with a characteristic time scale of a few years in response to the forcing rather than scaling non-linearly with global mean temperature anomaly. We propose that even very short time scales are treated as a fully coupled problem and encourage other modeling groups to investigate whether our description also suits their models' behavior. We introduce a definition of the forcing term ("virtual forcing"), which includes the oceanic adjustment processes, and serves as interpretive idea for longer time scales.

2.1 Introduction and tropospheric adjustment

The response of the global energy budget to an external perturbation of the energy content can be described by the heat uptake of ocean, ice, and land (N), the perturbation or radiative forcing (F), and the feedback response (λT), with the climate feedback parameter λ and temperature anomaly T:

$$C\frac{dT}{dt} = N = F - \lambda T \tag{1}$$

with the heat capacity of the climate system, C. Changes that are mediated by the climate system's response to the perturbation are called feedback responses. In contrast,

changes that depend on the nature of the perturbation, before the global temperature response happens, are termed adjustments. While the differentiation between forcing and feedbacks is seemingly a nominal problem, their clear separation is important to compare global climate models (GCMs), calibrate models which do not represent radiation and feedbacks dynamically, and to determine the widely used equilibrium climate sensitivity (ECS) from both models and observations (Gregory et al., 2004; Rogelj et al., 2011; Geoffroy et al., 2013b; Long and Collins, 2013). In this chapter, we put forward the idea of multi-annual coupled atmosphere-ocean adjustments. We use a large abrupt4xCO_2 ensemble to robustly detect this adjustment and ascribe it to the oceanic response to radiative forcing. We argue with a new conceptual modification of Eq. 1 that these processes are indeed time dependent adjustments to the forcing and not temperature dependent feedback responses. We suggest that the use of a modified forcing term –"virtual forcing", which includes the multi-annual adjustments – is a useful interpretative idea for longer timescales. We now first discuss tropospheric adjustment mechanisms in detail.

In climate models, adjustments in the stratosphere have long been accounted for when determining radiative forcings (Shine et al., 1990). The concept of tropospheric adjustment emerged only recently. Tropospheric effects have been called "fast feedbacks" or "the initial fast features" (Lahellec and Dufresne, 2014), "semi-direct effects" (Andrews and Forster, 2008), "fast responses", "rapid adjustments" (Bala et al., 2010), or "direct response to CO_2" (e.g., Merlis, 2015). The separation between fast tropospheric adjustments and feedbacks partly arises from the approximation of the global radiative response by a feedback term that depends linearly on global temperature anomaly (λT in Eq. 1). Tropospheric adjustments, which have short timescales, are included in the effective radiative forcing (ERF, e.g., Boucher et al. (2013); Forster et al. (2013)).

From a process point of view, tropospheric adjustment for CO_2 happens because directly after the forcing is applied, the radiative imbalance at the Earth's surface is smaller than at the top of the atmosphere (TOA). The middle and lower troposphere warm before the surface temperatures increase, causing increased stability, and reduced evaporation, convection, and precipitation over oceans (Cao et al., 2012; Andrews et al., 2012; Kamae et al., 2015). The tropical upward velocities weaken over the oceans and strengthen over land in all models (Bony et al., 2013). The circulation also weakens – independently of the land sea warming contrast – due to the spatial pattern of the CO_2 radiative forcing acting on the climatological distribution of clouds and humidity (Merlis, 2015). Trace gases alone force the stratosphere and upper troposphere to increase the eddy momentum flux, to accelerate stratospheric westerlies, and to displace the tropospheric jets poleward without any sea surface temperature response (Wu et al., 2011; Grise and Polvani, 2014a; Staten et al., 2014). Either the reduced surface latent heat flux or the reduced relative humidity at the top of the boundary layer leads to a reduction in low level cloud cover (Colman and McAvaney, 2011; Kamae and Watanabe, 2013; Wyant et al., 2012; Watanabe et al., 2012; Tomassini et al., 2013; Zelinka et al., 2013). The cloud response can be attributed to both the aforementioned dynamic and thermodynamic component. Other adjustment effects are the increased transport of heat from the land to the ocean due to enhanced land-ocean heating contrast (Williams et al., 2008; Dong et al., 2009; Webb et al., 2013) and the CO_2 physiological effect which enhances the land warming and moisture transport onto land (Cao et al., 2011; Kravitz et al., 2013; Doutriaux-Boucher et al., 2009; Kamae and Watanabe, 2013). Different forcing agents, such as solar or CO_2 forcing, show different adjustment processes (Lambert and Faull, 2007; Bala et al., 2010; Andrews et al., 2010; Schaller et al., 2013). Overall, the tropospheric CO_2 adjustment is also defined as sum of all processes which happen before the net forcing at the TOA and surface are equal, as a measure of the equilibration of the troposphere with the surface (Bala et al., 2010; Lahellec and Dufresne, 2014).

Some of the tropospheric adjustment effects are consistent across models, while others are model dependent, or even model version or cloud scheme dependent (e.g., Chung and Soden, 2015a). Studies disagree as to how important the different tropospheric adjustment processes are compared to long term or equilibrium responses. Williams et al. (2008); Andrews and Forster (2008); Gregory and Webb (2008); Bala et al. (2010); Webb et al. (2013); Bony et al. (2013) and Lahellec and Dufresne (2014) find that the ECS or the overall uncertainty to external forcing is influenced, whereas Tomassini et al. (2013); Vial et al. (2013) and Grise and Polvani (2014b) find no statistically significant effect of (their) adjustment processes on the feedback strengths or ECS. Ringer et al. (2014) shows a correlation between forcing and feedback, which further complicates the distinction between the two and is discussed further in Section 2.5.

There is no a priori reason why all adjustment processes should be fast, so the distinction between adjustment and surface temperature mediated response is not clear (Williams et al., 2008; Caldeira and Myhrvold, 2013; Zelinka et al., 2013; Sherwood et al., 2014a). Some forcing-dependent processes may take place over months or years after the imposition of the forcing, during which the temperature may increase by several degrees, so adjustment and feedback may be taking place simultaneously. An adjustment process influences the local temperatures and heat fluxes, but has little impact on the global mean surface temperature (T in Eq. 1). We define the system which is being forced and which experiences adjustments and feedbacks as that which determines T.

Technically, the spatial pattern of a tropospheric adjustment response is determined by either fixed sea surface temperature (SST) runs, in which a climatological SST field is prescribed (used e.g. in Hansen et al., 2005b; Bala et al., 2010; Andrews et al., 2012; Zelinka et al., 2013; Meraner et al., 2013), by defining the adjustment as the first year of a 4xCO2 simulation of a coupled model (used e.g. in Kravitz et al., 2013; Bony et al., 2013), or a slab ocean model (used e.g. in Bony et al., 2013; Grise and Polvani, 2014b), or by regressing each grid boxes' variable evolution against the global mean surface air temperature (Gregory and Webb, 2008; Andrews et al., 2015). Attempts and problems using radiative kernels to define the tropospheric adjustment response are discussed in Chung and Soden (2015a), Block and Mauritsen (2013) and Larson and Portmann (2015). Forcings are determined by either fixing the global surface temperature (Shine et al., 2003), or – more commonly – fixed-SST simulations, a regression method, which uses Eq. 1 and regresses N, the Top of the Atmosphere flux imbalance, against T, the surface temperature response (Figure 2.1), or radiative kernels Chung and Soden (2015b, for a comparison of the different definitions of radiative forcings).

In this chapter, we want to challenge the common understanding of sub-annual tropospheric adjustment. We find robust multi-annual adjustment responses in a large ensemble of coupled simulations introduced in Section 2.2. This response includes mechanisms described above, but we also find the ocean to strongly adjust to the radiative forcing, shaping in turn the tropospheric adjustment processes. We introduce a conceptual framework to show that this response is indeed better described as a time dependent forcing adjustment than a temperature dependent feedback (Section 2.3). We then show in the large ensemble that the multi-annual adjustment is caused by short wave cloud radiative effect over the oceans and argue that this is due to the spatial pattern formation of sea surface temperature, ocean heat uptake and loss, and the meridional ocean heat transport (Section 2.4). Finally, since the conventional definitions of radiative forcing only includes the adjustment of sub-annual tropospheric processes, we discuss the consequences of oceanic adjustment for the definition of radiative forcing (Section 2.5).

2.2 Model and experimental setup

We generate a large initial condition ensemble of abrupt4xCO_2 simulations – in which CO_2 is quadrupled at the start and then held constant – with the coupled atmosphere-ocean-sea ice-land Community Earth System Model (CESM1.0.4 with a resolution of roughly two degrees in the atmosphere and land and one degree in the ocean and sea ice component, Gent et al. (2011); Bitz et al. (2012); Danabasoglu et al. (2011)). Abrupt forcing allows us to study different timescales, and the conclusions also apply for more realistic gradual forcing scenario simulations, which can be thought of as a convolution of infinitesimal abrupt forcing changes (Good et al., 2011, 2013; Geoffroy et al., 2013b). From a several century long control run, each January an ensemble member is branched off. In total, 121 different combinations of ocean, sea ice, and atmospheric states are used as initial conditions for the abrupt CO_2 forcing. This nearly eliminates internal variability when considering the ensemble average. Simulations starting in months other than January were conducted, but yield similar results. All ensemble members are run for two years, 13 members for 100 years, 6 members for 250 years, and one member is run for 1300 years. All atmospheric data shown here are annual anomalies of each ensemble member with respect to the stable annual averaged control run. All oceanic anomalies are the difference of the forced ensemble member and the corresponding years following the control run branch off years, which are up to 100 years apart, to account for the small drift in the control run deep ocean. Ensemble averages are shown, except where noted. Results for other CMIP5 models with a similar experimental setup, but only a few ensemble members are shown by Andrews et al. (2012); Kravitz et al. (2013); Vial et al. (2013); Flato et al. (2013); Chung and Soden (2015a).

To obtain the forcing, an atmosphere only control simulation with climatological fixed prescribed SST derived from the coupled control run was run for several decades. Four 30 year long quadrupling CO_2 simulations are branched off from different initial conditions. The last ten years of each simulations' averaged TOA imbalance are depicted as a red cross in Figure 2.1. Further, for illustration purpose only, we conduct two more 150 year step function simulations with 2x and 8xCO_2.

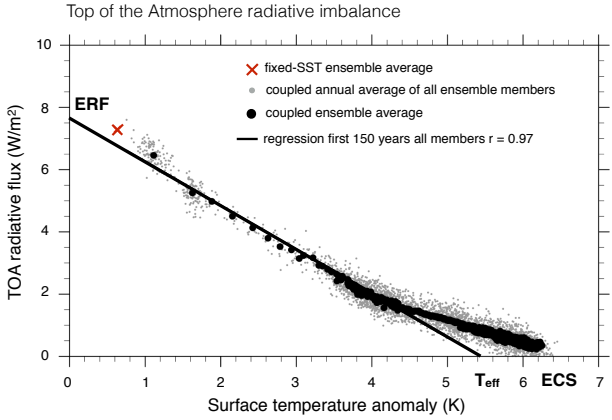

Figure 2.1: *Global mean net TOA downward radiative flux evolution against global average surface air temperature change of the large ensemble 4xCO_2. The regression covers the first 150 years (black line). Small dots depict the 121 ensemble member annual averages, while large dots are ensemble averages – annual until year 100, at ca. 4.4 K temperature increase, and decadal afterwards.*

2.3 Forcing adjustment versus time dependent feedbacks

Figure 2.1 shows the time evolution of the TOA radiative imbalance, N, against the surface temperature anomaly for all ensemble members of the 4xCO_2 step forcing simulations. While little gray dots are annual averages of individual ensemble members, the thick black dots are annual and ensemble means. Starting in the upper left with a temperature increase of 1.1K in the first year, the climate system evolves towards the equilibrium in the lower right. The linear regression of N against T for the first 150 years (treating the years as independent, black line) leads to the definition of effective radiative forcing (ERF, N at T=0) and effective climate sensitivity (T_{eff}, intersect of regression line with the horizontal axis, e.g., Boucher et al. (2013)). T_{eff} is substantially smaller than the equilibrium climate sensitivity (ECS), defined as the intersect of points with horizontal axis ($N_{t\to\infty} = 0$, e.g., Senior and Mitchell (2000); Gregory et al. (2004); Williams et al. (2008); Bitz et al. (2012); Li et al. (2013b); Andrews et al. (2015)). Originally, ECS and ERF are defined for doubling of CO_2 concentration from preindustrial values. Throughout this chapter we show values for the quadrupling simulations, which can be divided by two to get approximately the standard values for ECS, ERF, and T_{eff}. Deviations from the linear regression imply that a global average λ in Eq. 1 is not constant. This seems to be the case not only on the century time scales, but also over very short time scales of the first few years. To analyze the time evolution of N versus T we discuss two ways to adapt Eq. 1, in both cases by making a first-order perturbation for simplicity:
(A) The feedback term could be described as temperature dependent (Armour et al. (2013); Andrews et al. (2015); Gregory et al. (2015)),

$$C\frac{dT}{dt} = N = F - \lambda T(1 - f_A(T)) \tag{2A}$$

or (B) one could treat the climate feedback parameter as constant and adjust the forcing F for processes with a time scale longer than a year, analogously to the tropospheric adjustment due to processes on shorter time scales

$$C\frac{dT}{dt} = N = F(1 - f_B(t)) - \lambda T. \tag{2B}$$

$f_A(T)$ and $f_B(t)$ are unknown functions. However, by comparing Eq. 2A and 2B for different forcing levels, we can assess which formulation describes the CESM output better. Assume different step forcing levels F_1, F_2, F_3, e.g., $2xCO_2$, $4xCO_2$, $8xCO_2$, do relate $F_1 = n \cdot F_2 = m \cdot F_3$.

We solve Eq. 2A and 2B for T with $C = 7.3\,W\,yr\,m^{-2}\,K^{-1}$ (Geoffroy et al., 2013b), $\lambda = 1.2\,W\,m^{-2}\,K^{-1}$ for all cases, and F = 3.2, 6.9, and 11.2 $W\,m^{-2}$ for the different forcing levels. The estimates are based on the year 20 to 100 regression of the CESM $2xCO_2$, $4xCO_2$, and $8xCO_2$ simulations. F is not proportional to the log CO_2 (Gregory et al., 2015, and references therein). For illustrative purposes, we choose $f_A(T)$ as $3.5 - \sqrt{T} \cdot 0.8$, to make dN/dT decrease as T increases and $f_B(t)$ as exponential functions decaying from $1.3 \cdot F$ to $0.8 \cdot F$, to relax dN/dT to a long term constant value. Fig. 2.2 shows the solution for case A in the left most column, case B in the middle, and data from the coupled CESM simulation in the right most column. The different forcing levels are depicted in gray, red, and blue.

Because Eq. 2A is not linear in T, for case A, $T_1(t)$ is different from $n \cdot T_2(t)$ or $m \cdot T_3(t)$ with the subscripts denoting the use of F_1, F_2, and F_3 (Fig. 2.2a). The same holds for the TOA radiation imbalance, $N_1(t) \neq n \cdot N_2(t) \neq m \cdot N_3(t)$ (panel d). Eq. 2B is linear in T, however, so for case B, $T_1(t) = n \cdot T_2(t) = m \cdot T_3(t)$ (Fig. 2.2b), and equivalently for N(t) (Fig. 2.2e). For CESM, the scaled T (panel c) and N (panel f) nearly coincide for the three forcing levels. This indicates that case B and Eq. 2B is a good description of the time evolution of N and T.

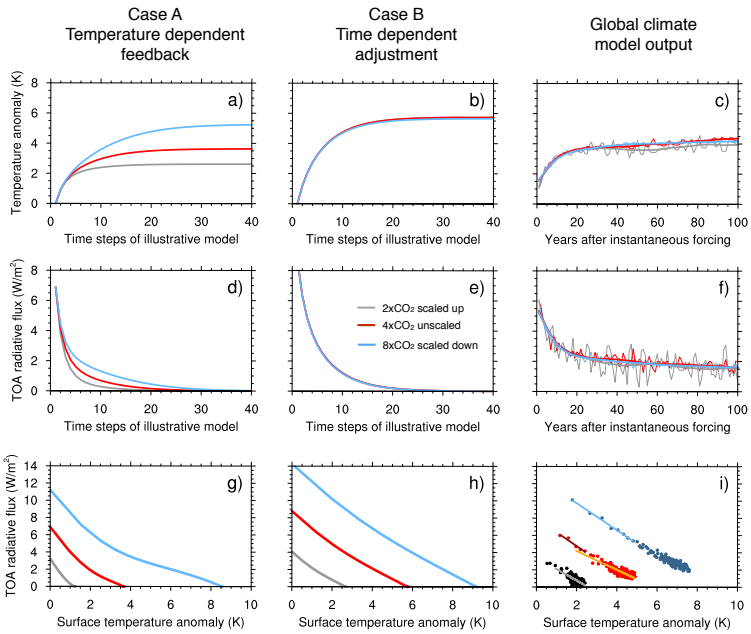

Figure 2.2: *Illustration of case A and B, discussed in the text. Scaled temperature anomaly for the illustrative model (case A panel a, case B panel b, CESM panel c; annual averages thin line, spline fit thick line), scaled TOA radiative imbalance (same setup, panel d-f), unscaled dT/dN (same setup, panel g-i). Gray, red, and light blue depict $2xCO_2$, $4xCO_2$, and $8xCO_2$ forcings, respectively; annual CESM output in panel i is shown in darker colors. The regressions lines – in gray, red, orange, and blue – for different temperature ranges are discussed in the text.*

The evolution of dN/dT behaves similarly: For case A and Eq. 2A, $dN/dT = -\lambda(1 - f_A(T) - T(df_A(T)/dT))$. The right hand side depends only on the temperature T itself and the function $f_A(T)$, assuming λ is constant. This implies that for any given temperature the curves for the different forcing levels have equal dN/dT and are therefore parallel (Fig. 2.2g, unscaled). For case B, it follows from Eq. 2B that, at any given time, $dN_1/dT_1 = dN_2/dT_2 = dN_3/dT_3$ (not shown) and any given temperature $dN_1/dT_1 = n\cdot(dN_2/dT_2) = m\cdot(dN_3/dT_3)$, (Fig. 2.2h, unscaled). Given the good description of case B and CESM data of N and T over time described above, the fit of dN/dT must be the same ((Fig. 2.2i) CESM data, unscaled). Nevertheless, we test the assumption of case A in Fig. 2.2i: For two different temperature ranges – in which the data overlap – the slope is indicated. The first range is 0.8-2.5 K, in which $dN_1/dT_1 = 1.32$ (gray) and $dN_2/dT_2 = 1.59$ (dark red). The second range is 1.6-5 K, in which $dN_2/dT_2 = 1.08$ (orange) and $dN_3/dT_3 = 1.56$ (light blue, all in $W\,m^{-2}\,K^{-1}$). This means the curves are not parallel in the same temperature range (as in Fig. 2.2g) and case A is not a good description for the CESM data.

In summary, we argue that the curvature in the N-T-space for the first few years could be treated as an adjustment problem (case B), rather than a temperature depen-

dent feedback (case A). This does not imply that later on during the equilibration process the feedback parameter has to be constant, that a combination of case A and B might not be a better overall description, or that a spatially dependent feedback parameter might be a helpful description (Armour et al., 2013). We now show that the widely used two box model with an ocean heat uptake efficacy factor conforms to case B, but modifies the forcing on different time scales than $f_B(t)$.

A commonly used refinement of the global model of Eq. 2A and 2B is to consider two layers (Gregory, 2000; Held et al., 2010; Geoffroy et al., 2013b), with T being the temperature of the upper layer and $T_{\text{deep layer}}$ of the deep layer, and a downward heat flux $\gamma(T-T_{\text{deep layer}})$ between them, with γ being a constant coefficient. Thus, $N_{\text{surface}} = F - \lambda T - \gamma(T - T_{\text{deep layer}})$. To model the apparent non-constant behaviour of λ, it has been proposed to introduce an ocean heat uptake efficacy ϵ (Winton et al., 2010; Held et al., 2010; Geoffroy et al., 2013a), so that $N_{\text{surface}} = F - \lambda T - (\epsilon - 1)\gamma(T - T_{\text{deep layer}})$. The overall feedback parameter is initially $\lambda + (\epsilon - 1)\gamma$, but decreases to λ with the time scale of deep ocean equilibration. Thus, the feedback parameter seems time or state dependent (our case A). There is no detailed physical motivation behind this ansatz, but the suggested mechanism is the poleward shift of ocean heat uptake, which modulates the atmospheric feedbacks. However, this model is linear in T and thus conforms to our case B: $N = F(t)^\star - \lambda^\star \cdot T$, with $F(t)^\star = F + (\epsilon - 1)\gamma T(t)_{\text{deep layer}}$ and $\lambda^\star = \lambda + (\epsilon - 1)\gamma$, which is constant in time, as long as ϵ and γ do not change through time. Thus, the scaling argument laid out for case B also describes the two-layer model with ocean heat uptake efficacy.

The timescales on which the curvature develops are different in the two models: Decades to centuries for $F(T(t)_{\text{deep layer}})^\star$ and a few years for $f_B(t)$ of our case B. The two models are physically distinct, although formally similar. We do not analyze multi-decadal time scales here, which motivated the introduction of ocean heat uptake efficacy. The fast time scale of the two-box model is roughly the same as the time scale of $f_B(t)$ (3-4 years, Geoffroy et al., 2013a). In the following, we analyze processes setting this timescale in the large ensemble of CESM. These processes may be related to those which set the fast time scale of the two-box model.

In the next section we show how the short wave cloud radiative response causes the curvature of dN/dT and how the oceanic adjustments – of meridional heat transport, surface heat fluxes, and SST patterns – might be connected to it. Section 2.5 will come back to the questions whether an adaptation of Eq. 1 according to case B might be useful.

2.4 From tropospheric to oceanic adjustment

2.4.1 Short wave cloud radiative response

Most studies point to the short wave cloud response dominating the tropospheric adjustment and the short term non-linearity of feedbacks, mostly focussing on the Tropical West Pacific (Bala et al., 2010; Colman and McAvaney, 2011; Andrews et al., 2012; Zelinka et al., 2013), but also on the Southern Ocean (Grise and Polvani, 2014b). We use the measure of cloud radiative effect (CRE) as a rough indication of the cloud response. The CRE is defined as net TOA all sky minus net clear sky response and we consider only the short wave (SW) component, since the long wave component evolves linearly with temperature and shows cloud masking effects. Its applicability is discussed e.g. by Zelinka et al. (2013); Kamae et al. (2015). On timescales discussed here the SW CRE is dominated by the low latitude response so that potential aliasing errors over sea ice would be too small to qualitatively impact our results. Figure 2.3 shows that in

agreement with some other studies the SW CRE over the oceans is positive in the first year, and by extrapolation that it is non-zero at T=0, indicating tropospheric adjustment (Colman and McAvaney, 2011; Zelinka et al., 2013; Andrews et al., 2015). The SW CRE over the oceans declines as the temperature increases, changes sign after three to four years, and equilibrates after about a decade at a negative value. Models disagree on the temporal evolution and temperature sensitivity of the SW CRE but many models show a different sensitivity during the first decade compared to the century time scale response (Ringer et al., 2014, and gray dots for the global SW CRE response in Figure 2.3). In our case, oceanic SW CRE it is not linearly dependent on global mean temperature, as one would expect of a climate feedback. Global SW CRE varies first strongly and on decadal to centennial timescales only very slightly with global mean temperature. This behavior could be described either as an inconstant climate feedback parameter (case A above), or an adjustment on a longer time scale than a few months (case B). SW CRE over land is also positive in the first year and remains roughly constant, so it can be described as a tropospheric adjustment, with no climate feedback. We will argue below that the SW CRE response comes about not only due to the rapid adjustment to the radiation on monthly time scale, but also due to the oceanic adjustment of heat transport within the first few years after the forcing is applied. We cannot exclude that the surface temperature increase of more than 3K during the first ten years influences the SW CRE response. However, Fig. 2.3 suggests that the SW CRE is not sensitive to temperature anomalies beyond 4K. The four member fixed-SST ensemble average SW CRE (in green and orange for the ocean and land, respectively) lie within the range of the coupled ensembles, i.e. at 0.8 and 0.4 Wm^{-2}, with a global surface temperature increase of 0.7 instead of 1.1K. The standard deviation of the first year's ocean SW CRE is 0.24 Wm^{-2}, which dominates the total TOA ensemble standard deviation of 0.34 Wm^{-2} (Figure 2.1). The spatial SW CRE response pattern is discussed in Section 2.4d.

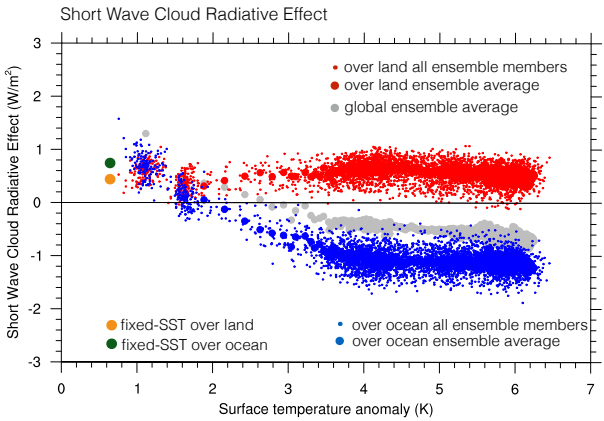

Figure 2.3: *TOA short wave cloud radiative effect (positive downward) over land (red), oceans (blue), and total (gray) for all 121 ensemble members of the abrupt4xCO_2 simulations. Small dots depict individual ensemble annual averages, while large dots are ensemble averages – annual until year 100, at ca. 4.4 K temperature increase, and decadal afterwards.*

2.4.2 Surface temperature, surface heat flux, and ocean heat content patterns

Figure 2.4 shows the anomaly patterns of the surface air temperature (left most column), sea surface temperature (middle left), surface heat flux (middle right, positive downwards), and the rate of change of ocean heat content overlaid by the wind stress (right most column) of the first four years (upper four rows) and the long term average (lowest row). Local values (in K or global W/m^2) are divided by their annual global mean value (lower left box in each panel). The surface air and sea temperature increase includes Arctic amplification, enhanced warming over the Northern Hemispheric continents, and initial cooling in the equatorial Pacific region, especially in the East, which gradually weakens. This La Niña like cooling pattern is attributed to the increase in upwelled cold waters by anomalous surface wind stress forcing (right most column). The deep water was not exposed to the surface warming signal yet and increases the East-West temperature gradient across the Pacific (Clement et al., 1996; Cai et al., 2015). The CMIP5 average shows delayed warming in the east Pacific as well (Andrews et al., 2015). Andrews et al. (2015) showed that the evolving pattern of surface warming is the dominant cause of nonlinearity between N and T in Eq. 1 for the HadCM3 and HadGEM2 atmospheric components and that the pattern of SST changes determines the SW CRE.

Taking this line of thought one step further we show in the remainder of this chapter how not only SSTs but the ocean's heat content and transport respond on short time scales and impacts the tropospheric response. The two right columns of Figure 2.4 show the annual and ensemble average surface heat flux anomaly and the vertically integrated rate of change of ocean heat content. The difference between the two columns is the ocean heat transport convergence. The tropical Atlantic loses heat to the atmosphere (blue in panel k-n) and the subtropical oceans (blue in panel p-s) and so does the whole Eastern and Tropical Pacific. Locally, the time evolution of surface fluxes, heat transport, and surface wind stress is rapidly changing, e.g. in the Indian Ocean, equatorial West Pacific, North Atlantic, or over Eurasia. All responses shown in Figure 2.4 in the first four rows are as high or higher than the inter annual standard deviation of the control run, indicating that they are a forced response. Patterns of surface air and water temperatures in the first years differ strongly from the long term pattern (panel e and j). The increase in surface heat flux patchiness (panel k-o) is dominated by the latent heat flux (not shown). After three years the land does not take up heat. The rate of change of ocean heat content anomaly in year 80-100 (panel t) is very small and multiplied here by 4 to show the distinctively different pattern in all ocean basins, compared with the initial response pattern. Pattern of heat fluxes from the mixed-layer to the deep ocean can influence atmospheric feedbacks and the global surface temperature response through SW CRE in slab ocean aqua planet models (Rose et al., 2014). We argue in the following that meridional ocean heat transport shapes SW CRE on short time scales, on which the mixed-layer is still equilibrating. A direct comparison to Rose et al. (2014) is not possible due to our transient coupled experimental setup. However, their argument that the SST pattern caused by oceanic heat transport influences the SW CRE is the same as laid out here.

We now describe in more detail the oceanic adjustment processes. They are caused by the tropospheric adjustment of the first few months, as well as the the anomalous surface heat fluxes, wind stress pattern, and the land-sea warming contrast of the first few years. Since they respond to the forcing and differ from the surface temperature mitigated ocean patterns of decadal to centennial time scales. Oceanic adjustments last longer than tropospheric adjustments but provides the lower boundary conditions for the short term atmospheric response.

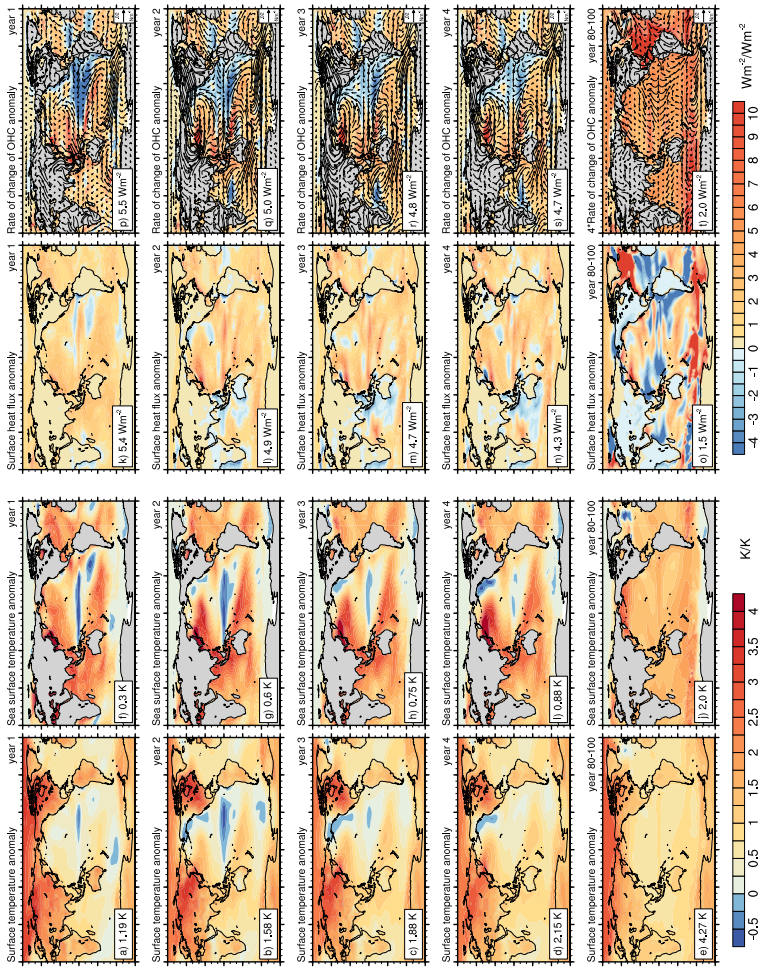

Figure 2.4: *Abrupt4xCO_2 ensemble and annual mean year one to four and 80-100 for surface air (a-e) and sea temperature anomalies (f-j) divided by the global response of that year (lower left corner of each panel, K).*

Figure 2.4 *(previous page)*: *Surface heat flux anomaly (k-o, positive downward) and rate of change of ocean heat content integrated over the whole water column (r-t), again divided by the global value (both W/m^2). The difference between the right most two columns is the ocean heat transport convergence. Surface wind stress anomaly contours overlaid are in the last column (p-t).*

2.4.3 Ocean heat transport and circulation response

Figure 2.5 shows – as a function of latitude and time – the ensemble mean surface heat flux (panel a), the meridional ocean heat transport convergence (b), and the rate of change of ocean heat content (c, all in W per global m^2). As already obvious from Figure 2.4, the zonally integrated anomalous surface heat flux is positive at all latitudes for some years and especially high in the Southern Ocean 50-60°S and the Northern Subtropics around 30°N (panel a). There is anomalous divergence of heat out of the equatorial regions and across the Antarctic circumpolar current (blue in panel b). These two effects lead to a cooling in the equatorial ocean initially, and warming everywhere else, especially in the mid-latitudes, where there is downward wind-driven pumping of heat. The meridional heat transport could be one reason why in the Pacific sea surface temperatures emerge faster than expected from scaling global mean warming (Chadwick et al., 2014). Locally, it can take two to four years until the fluxes reach their highest value, e.g., around 55 °S or 40 °N.

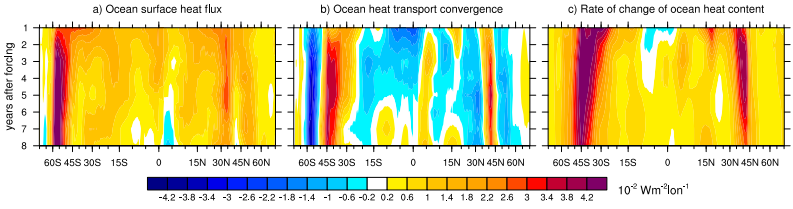

Figure 2.5: *Abrupt4xCO_2 zonally integrated ensemble and annual mean ocean surface heat flux change (a), ocean meridional heat transport convergence (b), and rate of change of ocean heat content (c), all in W per global m^2. Positive values indicate ocean heat uptake (in a), heat accumulation through transport (in b), and an increase in ocean heat content (in c).*

In the North Atlantic the short time scale response in surface fluxes leads to a temporary increasing AMOC strength for two to four years (Fig 2.6a), which is statistically different from the year following the branch off in the control run. Based on our simulations we cannot distinguish whether surface freshwater, surface or lateral freshwater fluxes, the wind field anomalies, sea ice edge, or places of convection and deep water formation cause this AMOC response (e.g., Gregory et al., 2005; Smith et al., 2014). All these fields show anomalous patterns in strength and locality compared with the long term response. The zonally averaged rate of change of ocean heat content in the North Atlantic is much smaller than in other locations during the first years and heat reaches depth later (Figure 2.7 left minus right column). The wind field anomalies change sign within the first two years (Fig. 2.4). The small but robust AMOC response shows that even the deep ocean content can be affected by an anomalous surface forcing within a year and local surface conditions can influence volume fluxes elsewhere (e.g., here at 18 °S where the overturning response is reversed and delayed for a few years, Fig. 2.6b). Local surface fluxes and ocean heat uptake efficiency influence the lower boundary condition for the atmosphere to respond to the radiative forcing (Fig. 2.4) and thus the time scale of tropospheric adjustment and short time scale feedbacks. We trust this surprising result only due to our large ensemble size, with which we can differentiate the response from the control run variability.

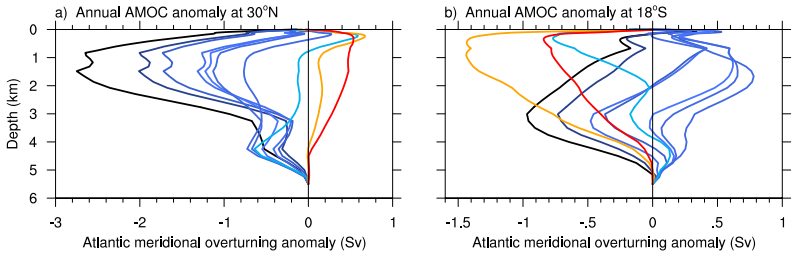

Figure 2.6: *Abrupt4xCO_2 ensemble and annual mean meridional overturning circulation anomaly at 30 °N (a) and 18 °S (b) for year one (red), year two (orange), year three (light blue) etc. to year ten (black). 1 Sv = 10^6 m^3/s*

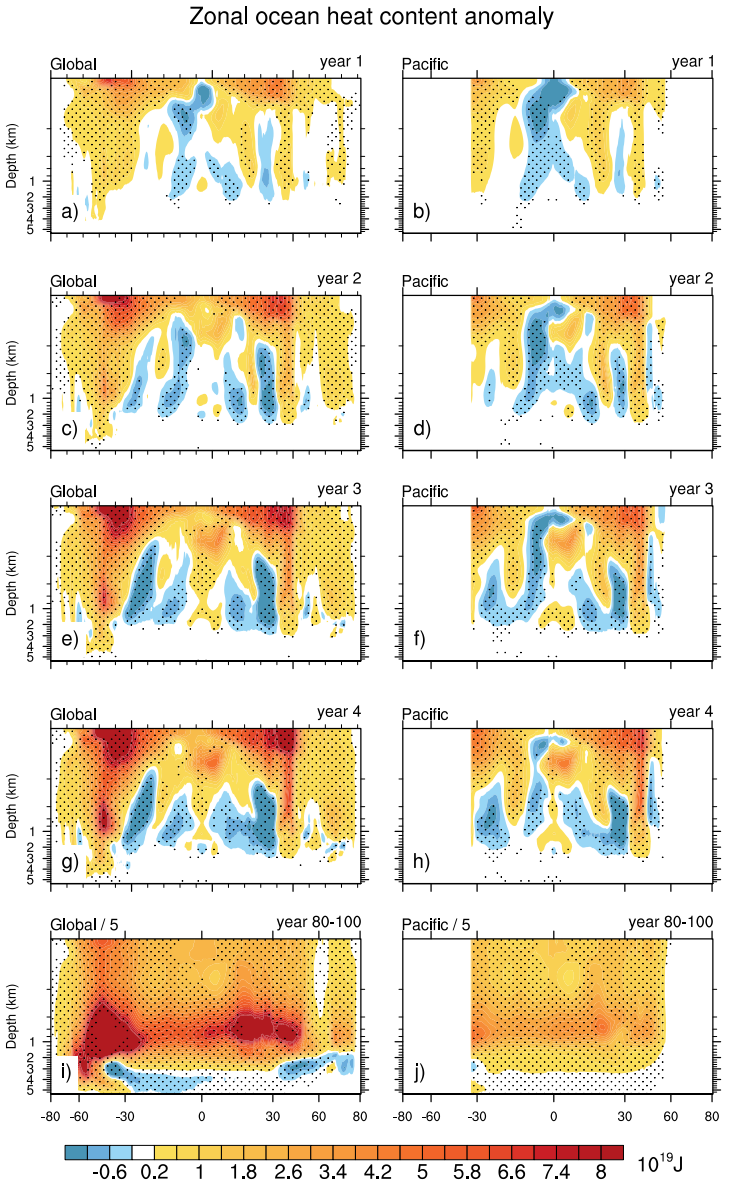
Zonal ocean heat content anomaly
Global
year 1
Pacific
year 1
Depth (km)
a)
b)
Global
year 2
Pacific
year 2
c)
d)
Global
year 3
Pacific
year 3
e)
f)
Global
year 4
Pacific
year 4
g)
h)
Global / 5
year 80-100
Pacific / 5
year 80-100
i)
j)
-80 -60 -30 0 30 60 80
-0.6 0.2 1 1.8 2.6 3.4 4.2 5 5.8 6.6 7.4 8
10^{19}J

Figure 2.7 ***(previous page)*****:** *Abrupt4xCO_2 ensemble mean global (left) and Pacific (right) ocean heat content change for end of year one to four (upper four rows), and average of year 80-100 divided by 5 to fit the scale (last row). Stippling indicates the anomaly is significantly different from the control run variability at the 95% level.*

Changes in the ocean interior temperature result from the changes in ocean heat transport. Figure 2.7 shows the global (left) and Pacific (right) ocean heat content change in joules – including surface heat flux and transport – for the first four years and the long term response (lowest row, scaled to compare the patterns and fit the scale). Stippling indicates where the anomalous heat uptake patterns differs from the control run simulation variability on the 95% level. Next to the Pacific subsurface cooling, the non uniform heating becomes obvious here also with depth. For each location the temperature perturbation reaches depths of a kilometer after the first year and in the Southern Ocean the perturbation reaches a depth of 4 km in the second year.

2.4.4 Evolution of spatial SW CRE patterns

After having sketched out some processes involved in oceanic adjustment, we now show how the SW CRE responds to the oceanic adjustment. We suggest an interpretation along the lines of case B of Section 2.3.

Fig. 2.8 shows the coupled ensemble average of the first and second year's and long term response (panel a, b, d), and the difference of the fixed-SST response to the first year of the coupled response (panel c). The fixed-SST response is averaged over year 10 to 30 of four ensemble members and can be regarded as "year 0". In the global mean, the fixed-SST and coupled values are nearly the same (compare also Fig. 2.3 and discussion in Ringer et al. (2014)). However, the spatial pattern shows locally differences in the order of magnitude of the response itself, also over the land and especially strong over the Southern Ocean. Throughout the ensemble, locally the strongest response in magnitude (both positive and negative) appears in year two (panel b). Panel d indicates that the long term response pattern is very different for the short term response, both in high and low latitudes.

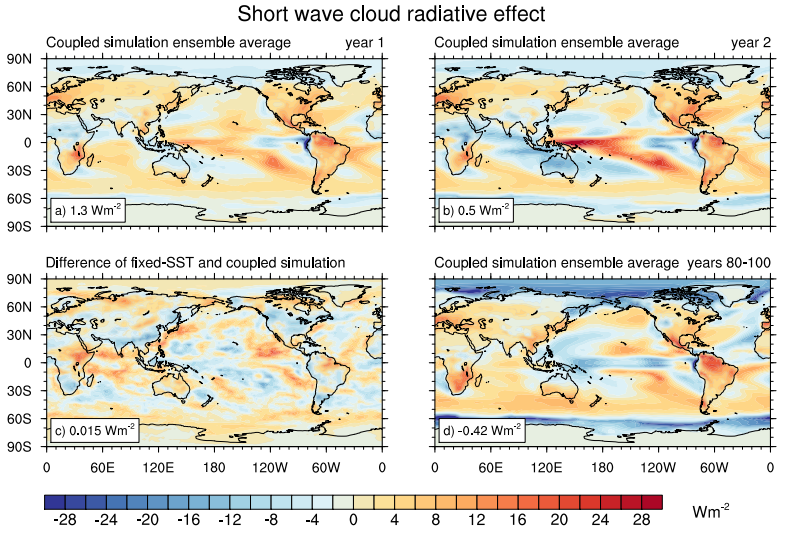

Figure 2.8: *Abrupt4xCO_2 Short Wave Cloud Radiative Effect for the first and second year (a and b), and long term (d) coupled ensemble average. c) Difference between the fixed-SST ensemble average of year 10-30 and the average coupled response of the first year. Global values are indicated in the lower left corner of each panel, all in W/m^2 for the whole global area.*

To measure the time evolution of the spatial pattern and thus, the potential to influence the non-linearity of the feedback term, Fig. 2.9a shows the root mean square difference between each year of each coupled ensemble member and the long term pattern shown in Fig. 2.8d. Each ensemble member is depicted as gray line, the longer ensemble members as colored lines, for the first 20 years of the simulations. The SW CRE pattern differs strongest from the long term pattern in year two in 74% of all simulations. It takes five to ten years to reduce the deviation from the long term pattern by half. Fig. 2.9b shows that the spread between ensemble members for the first year is dominated by the Western equatorial Pacific and Indian Ocean, where the ocean heat transport convergence is also very high and the SST influence the low stratocumulus clouds and thus the SW CRE.

Panel c) and d) show again the root mean square difference to the long term pattern of the rate of ocean heat content change and SSTs, which are similar to the response of SW CRE. To remove the global warming signal, the SST and dOHC/dt patterns are normalized with their global values, as in Fig. 2.4, before the root mean square difference is computed. The SW CRE pattern, however, is not normalized, since it does not scale with global mean temperature anomaly as a feedback would do. The time scale of pattern changes of SW CRE, rate of heat content change, and the SST are similar - most changes happen in the first six years. We interpret this as an indication that the

SW CRE response over the ocean – after being triggered by the application of the forcing – is shaped by the oceanic adjustment of heat transport to the forcing. Perturbed physics experiments fixing the clouds or ocean heat transport would be necessary to understand this relationship in more detail. The analysis here shows that even very short time scales should be studied in coupled atmosphere-ocean instead of fixed-SST or slab ocean frameworks.

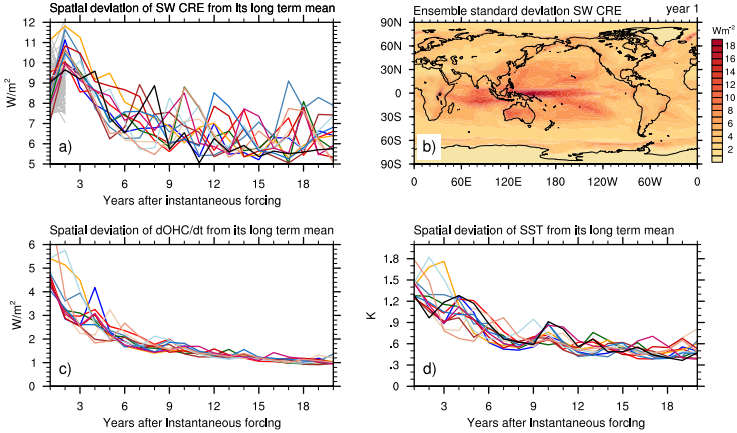

Figure 2.9: *Abrupt4xCO_2 coupled ensemble member (colors and gray): (a) Root mean square deviation of the spatial SW CRE pattern from the long term ensemble average pattern shown in Fig. 2.8d. (b) SW CRE standard deviation of coupled ensemble (c) Root mean square deviation of rate of change of ocean heat content from its long term pattern shown in Fig. 2.4t. (d) Root mean square deviation of rate of sea surface temperature from its long term pattern shown in Fig. 2.4j. Patterns used for panel c) and d) are normalized, as in Fig. 2.4.*

2.5 Virtual radiative forcing

We have shown in Section 2.4 why fixed-SST simulations do not represent the full magnitude of the adjustment processes and discussed which oceanic processes could play a role in the adjustment process. In Section 2.3 we suggested that the forcing term could be modified to capture the non-linear evolution of dN/dT. We now propose a formulation of the forcing term, which includes processes of oceanic adjustments and reattributes part of the N–T curvature to the forcing. If fixed-SST simulation do not represent the full tropospheric adjustment magnitude, which forcing do they represent? Figure 2.1 indicates that the fixed-SST forcing is smaller than the ERF obtained in a regression of the first 150 years. In Flato et al. (2013) five out of ten CMIP5 models, which contributed both the forcing from fixed-SST and the regression method, had smaller or similar forcings obtained by both methods. The brown cross in Fig. 2.10a

and b includes the correction suggested by Hansen et al. (2005b), using the 150 year regression (from Fig. 2.1) to project $F_{fixedSST}$ (red cross) onto the vertical axis. This assumes that the same feedbacks act in a fixed-SST and coupled run and that λ is constant, which is both not the case (e.g., Gregory and Webb, 2008). The orange cross adds the land warming of around $0.4\,Wm^{-2}$ to the fixed-SST forcing. Finally as an example, the green cross is the intersect of the vertical axis with the regression of year 10-150 (red line). This somewhat arbitrary time frame takes into account all adjustment and ocean mixed layer processes of the first ten years.

2.5.1 Alternative method to obtain F and λ: Moving window regression

To examine the non-constancy of F and λ in more detail, we now regress the radiative imbalance not over a certain time, but a limited temperature range. That is equivalent to calculating the local derivative $\Delta N/\Delta T$ and the corresponding axes intersects for that regression. The blue shaded area in Fig. 2.10a is the first temperature window – starting at 0.6 K and ranging up to 2 K. The lower bound is set by the first years' temperature of the coldest ensemble member, while the range of 1.4 K is chosen to be large enough to include at least three years. This prevents regressing members of only one year, while keeping the window small enough to resolve the time or temperature dependence of the feedback parameter. The method is similar to the binned regression of Block and Mauritsen (2013); Ringer et al. (2014) or Andrews et al. (2015), who regress N against T for a certain range of time. While these studies use only time frames (e.g.,year 1-20 versus 20-150), we move the window continuously through the whole temperature range of the first 30 years in 0.1 K steps (gray in panel a), while recording the slope and axis intersects (blue regression line for the first window in blue shading). A centennial perspective of this method is presented by Knutti and Rugenstein (2015). An advantage of the large ensemble is, that each bin has more similar amount of points, so the comparison between regression attributes is more robust than comparing e.g. a regression of 10 versus 140 points. All values are then transferred from the temperature back to the time domain (horizontal axis in panel b). The overall shape of the time dependence in panel b is not sensitive to the regression bin width or the use of annual or decadal averaged data. Panel b shows that the first five to six years contribute most to the effect of T_{eff}, ERF, and λ not being constant on short time scales. After 30 years T_{eff} is still more than 0.8 K away from the approximate 6.6 K ECS. The same holds for λ and ERF, indicated by the arrows and 150 year values on the right axis.

We suggest to call a forcing not obtained by regressing the first 150 years, but any other time frame virtual forcing. Virtual since there is no real state which experiences that forcing (Andrews et al., 2015), and since it is not only a radiative forcing but includes the whole state of the coupled system. Certain processes are fully, others only partly, included: For example, the virtual forcing of year five ($7.6\,W/m^2$) includes the adjustment of the AMOC and the corresponding influence on ocean heat uptake but only (although the main) part of the SW CRE adjustment.

2.5.2 Applications and limitations

The virtual forcing could be chosen simply as the forcing value that is most suited to maintain a "linear enough" relationship between N and T in a desired range of temperature or time – to answer a certain question. While this approach complicates the definition of the forcing term, it might shift the attention to understanding and comparing processes step by step. If the initial curvature in dN/dT is indeed explained through a modified forcing term, the use of a virtual forcing would help to differentiate responses

caused by the application of the forcing and the surface temperature increase. Ringer et al. (2014) found a correlation between the forcing term and the SW cloud feedback over 150 years. Folding these processes into the forcing term might be helpful to disentangle forcing and feedback. The concept of virtual forcing also might help to compare models with different degrees of adjustment. Assume one model X has a strong sea ice response within the first three years, while model Y has a strong West Pacific SW CRE response within the first two years, while model Z shows a perfectly linear N–T relationship over the first few years. Comparing their virtual forcings might be cleaner than comparing their ERF or fixed-SST forcings, leading to a reduced uncertainty in λ and ECS. Given the evolution of λ, F and ECS term in Fig. 2.10 it is also unclear, which forcing is most suitable to determine ECS from observations (e.g., Otto et al., 2013) and how representative the transient response at any time is not only of equilibrium conditions, but also of any other time frame (Gregory et al., 2015). The virtual radiative forcing has the potential to be more process based than the ERF or fixed-SST forcing. Finally, one might use the virtual forcing for more technical studies: Even in non-coupled simulations one might differentiate with this method between specific atmospheric processes, e.g. by keeping the land surface temperature, certain surface fluxes, or aerosol concentrations fixed and determine their adjustment time scales.

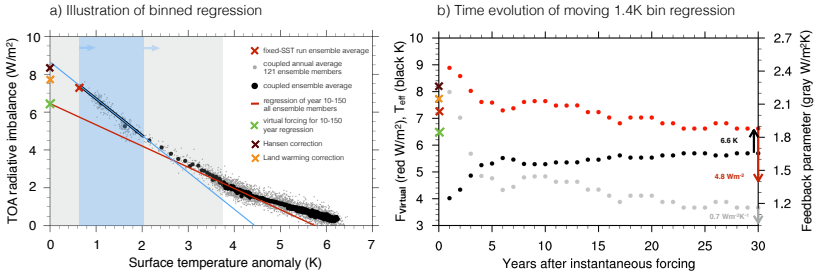

Figure 2.10: *a) Annual global mean TOA net downward radiative flux evolution against global average surface air temperature anomaly, regressed for year 10 to 150 (red), and example of 1.4 K wide window to determine the time dependence of T_{eff}, F_{virtual}, and λ (blue). (b) Time evolution of T_{eff} (black), F_{virtual} (red), and λ (gray) for the time and temperature range indicated by the gray shading in panel a. Values and arrows at the right vertical axis are the values of the same method after 150 years.*

One obvious limitation becomes clear in Fig. 2.10b and by recalling the formulation of case B: The description of the time dependency of the adjustment might be more complicated than the exponential illustrative example in Section 2.3. $f_B(t)$ does not have an obvious formulation and depends on various very likely strongly model dependent processes. In the coupled model reality, temperature dependent processes (either in the form of case A or related to the deep ocean equilibration) can set in while a model is still adjusting. Thus, it is open to which degree the concept described here might be indeed helpful – not only in a model context with clearly defined forcing (here only done for one CO_2 level) and a rough understanding of internal variability – but also concerning observational estimates of surface warming and ocean heat uptake.

2.6 Conclusion

We use a 121 member ensemble of abrupt4xCO_2 simulations to overcome initial conditions dependency and internal variability to explore the heat flux through the coupled system within the first few years after an abrupt forcing. After the forcing is applied, the stratosphere and troposphere adjust within a few months. The resulting anomalous surface flux and wind stress force the ocean to take up and transport heat meridionaly and vertically. Locally, the meridional ocean heat transport convergence can be even stronger than the surface heat flux, leading to a short term tropical Pacific cooling. Circulation adjustments include the Atlantic meridional overturning circulation, which increases for some years. These oceanic adjustments of circulation and heat transport set the conditions for local surface fluxes and thus the atmospheric response of the first few years. The SW CRE over the oceans in particular has an adjustment time scale of several years, after which it does not scale with the global mean atmospheric temperature increase like a feedback would do. Instead, the time scale of pattern formation of SW CRE, SST, and ocean heat transport convergence from their initial homogeneous to a spatially stable pattern changing only in magnitude, is the same. The time scale discussed here is connected to the fast time scales identified by, e.g., Hasselmann et al. (1993); Held et al. (2010); Caldeira and Myhrvold (2013), or Geoffroy et al. (2013a). Simple energy balance models or fits to coupled model output may capture part of the behavior but might be a less helpful framework to understand processes.

We show through scaling an abrupt2xCO_2 and 8xCO_2 simulation that the coupled model output is better described as a forcing adjustment than as a temperature dependent feedback. In other words, the processes causing the curvature of dN/dT during the first few years proceed with a characteristic time scale in response to the forcing (Eq. 2B) and do not scale non-linearly with the global mean temperature (Eq. 2A).

We define a virtual forcing – a variation of the traditional forcing term, which is defined either as effective radiative forcing or fixed-SST forcing. Virtual forcing is the forcing at a time when the feedback parameter is approximately constant in the time range of interest. Folding adjustment processes into the forcing term might help to compare models with different adjustment processes, to circumvent forcing-feedback correlations, or potentially also to better estimate ECS from observed warming and ocean heat uptake (Otto et al., 2013; Knutti and Rugenstein, 2015). However, in principle a model could adjust for a few years through various processes and then either remain linear or display non-linear feedbacks. At this stage it is unclear whether there is a sufficiently robust behavior across models that the concept can be useful. We do not want to argue that this approach is superior to describing the feedbacks as time or state dependent but simply offer one more approach in the recent discussion of the forcing-feedback framework applicability.

3 Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes

published by Maria Rugenstein, Ken Caldeira, and Reto Knutti in Geophysical Research Letters (2016) doi: 10.1002/2016GL070907

In most climate models, after an abrupt increase in radiative forcing the climate feedback parameter magnitude decreases with time. We demonstrate how the evolution of the pattern of ocean heat uptake – moving from a more homogeneous toward a heterogeneous and high latitude enhanced pattern – influences not only regional but also global climate feedbacks. We force a slab ocean model with scaled patterns of ocean heat uptake derived from a coupled ocean-atmosphere general circulation model. Steady-state results from the slab-ocean approximate transient results from the dynamic ocean configuration. Our results indicate that cloud radiative effects play an important role in decreasing the magnitude of the climate feedback parameter. The ocean strongly affects atmospheric temperatures through both heat uptake and through influencing atmospheric feedbacks. This highlights the challenges associated with reliably predicting transient or equilibrated climate system states from shorter-term climate simulations and observed climate variability.

3.1 Motivation

The linear forcing–feedback framework assumes that the warming contribution of globally averaged climate feedbacks depend linearly on the global average near surface temperature response following a radiative forcing, i.e., that the feedbacks are constant. The net global feedbacks are negative, counter-acting the radiative forcing and stabilizing the global mean temperature. Many studies have assumed a constant global climate feedback parameter (e.g., Andrews et al., 2012; Forster et al., 2013; Otto et al., 2013), although some studies show that its magnitude decreases with time following an abruptCO_2 forcing (e.g., Senior and Mitchell, 2000; Gregory et al., 2004; Meraner et al., 2013; Andrews et al., 2015; Knutti and Rugenstein, 2015). Here, we focus on an open question with strong implications for the predictability of global climate evolution on time scales of decades to millennia: To what degree is the global climate feedback parameter dependent on the spatial patterns of surface temperatures and heat fluxes?

Ocean heat uptake, defined as net surface heat flux into the ocean, has a direct cooling effect on the atmosphere, but indirectly affects surface temperatures through

changing the magnitude of local radiative feedbacks. Thus, ocean heat uptake patterns might explain part of the changing magnitude of the global feedback parameter within one model, the spread between models, and the difference between models and observations (Winton et al., 2010; Bitz et al., 2012; Paynter and Frölicher, 2015; Gregory and Andrews, 2016; Rose and Rencurrel, 2016).

Two studies in particular show that atmospheric feedbacks in idealized aquaplanets are very sensitive to the spatial structure of the ocean heat uptake (Rose et al., 2014) or ocean heat release (Kang and Xie, 2014). Not primarily concerned with ocean heat uptake patterns, Andrews et al. (2015) show that 85% of models taking part in the Coupled Model Intercomparison Project Phase 5 (CMIP5) show a significant (22-62%) reduction in magnitude of the global feedback parameter in year 21-150 compared to year 1-20 following an abrupt4xCO_2 forcing. Armour et al. (2013) argues that for one of these models, the reduction of the global feedback parameter comes about through the local structure of warming and ocean heat uptake enhancing constant local feedbacks strength: High latitude warming – where the magnitude of feedbacks are less negative, or positive, thus, destabilizing – leads to a substantially stronger global temperature response than low latitude warming – where stabilizing, negative feedbacks are stronger.

Our aim now is threefold: First, we introduce a new experimental design to quantify transient feedback strengths through forcing a slab ocean model with a series of mixed layer depth heat fluxes derived from a coupled model (Section 3.2, Fig. 3.1). Second, this experimental setup is used to discuss temperature pattern, feedbacks, and the sea ice response for both idealized (similar to Rose et al. (2014) and Kang and Xie (2014), Section 3.3.1) and realistic (i.e. derived from a coupled simulation, Section 3.3.2) heat flux patterns (Fig. 3.2 and 3.3). Thus, we bridge the gap between the highly idealized aquaplanet and more realistic complex coupled simulations presented in recent literature. Third, after concluding with a illustrative summary (Fig. 3.4), we offer an interpretation of the continuously decreasing global feedback parameter in the coupled model through time (Section 3.4, Fig. 3.1b) and discuss broader implications and follow up studies in Section 3.4.

3.2 Models and Method: Generation of Q-flux forcing

We use the fully coupled (ocean-sea-ice-atmosphere-land) and the slab ocean (slab-ocean-sea-ice-atmosphere-land) configuration of the Community Earth System Model 1.0.4 with a finite volume horizontal resolution of 1.9°x2.5° for the atmosphere and roughly 1°x1° for the sea ice and ocean components (Bitz et al., 2012; Hurrell et al., 2013). Fig. 3.1a shows a large initial condition ensemble of the coupled model equilibrating the top-of-the-atmosphere (TOA) and surface temperature imbalances following an abrupt4xCO_2 forcing. The slope of the data points is the global feedback parameter and Fig. 3.1b shows its 60% decrease over 200 years obtained through linearly regressing all annual means in a 1.3 K wide window which is moved in 0.1 K steps through the whole temperature range in Fig. 3.1a (method explained by Rugenstein et al. (2016)). The TOA imbalance caused by the abrupt4xCO_2 forcing is mostly mitigated by ocean heat uptake (Fig. 3.1c), which moves within ten years from a zonally near homogeneous to a heterogeneous pattern. The mixed layer equilibrates in 10-20 years, after which the heat fluxes at the surface and the bottom of the mixed layer are roughly the same. In the slab ocean configuration, the heat flux at the bottom of the mixed layer (Q-flux) is prescribed and on average is the same as the equilibrated surface heat flux, since there is no lateral heat transport within the slab. The partitioning between the shortwave (SW) and longwave (LW), sensible and latent heat fluxes and the sea surface temperatures (SST) evolve freely within the prescribed Q-flux constraint.

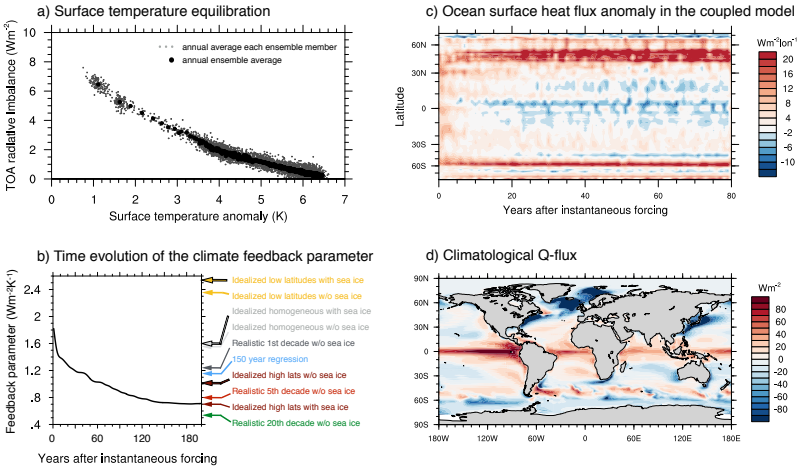

Figure 3.1: *Slab ocean simulations forced with heat flux anomalies derived from a coupled simulation reproduce the temporal evolution of the feedback parameter. a) Global top-of-the-atmosphere imbalance through a 2000 year long equilibration after abrupt $4xCO_2$ forcing for a large initial condition ensemble (Rugenstein et al., 2016). Gray dots are individual ensembles' annual averages, black dots are ensemble mean annual averages (twenty year running mean from year 150 onwards). b) Slope of the regression to the data in panel a). Colored arrows indicate the different magnitudes of feedback parameters we reconstruct with slab ocean simulations, except for the "150 year regression" which is a linear regression over the first 150 years of data in panel a). Arrows with black borders indicate simulations with sea ice, in all other simulations sea ice growth is inhibited. c) Zonal averaged ocean surface heat flux anomaly through time in the abrupt $4xCO_2$ coupled simulation. d) Control simulation's annual average heat flux from the mixed layer to the interior ocean heat flux. Fluxes in c) and d) are positive downward.*

We use an equilibrated 80-year-long coupled control simulation with pre-industrial forcing to generate a climatological reference state Q-flux (Fig. 3.1d), which globally and annually averages to zero. The annual mean climatological Q-flux is dominated by tropical upwelling regions taking up heat and Western boundary currents, the Nordic seas, and the Southern Ocean releasing heat, and compares well with observations (Liang et al., 2015). The reference state for all our experiments is this climatological Q-flux in an equilibrated climate with CO_2 levels of four times its pre-industrial value, leading to 6.4 K warming compared to the control simulation. We then apply the globally non-zero Q-flux pattern as a negative ocean heat uptake forcing to the atmosphere and induce global cooling or – in the presence of a positive CO_2 forcing – reduce the warming. In the supporting information (SI) we discuss details on the choice of our reference state, independence of ocean heat uptake and CO_2 forcing, and the construction of the Q-flux forcing fields (Bitz et al., 2012; Donohoe et al., 2014).

We apply six time invariant ocean heat uptake patterns (Fig. 3.2a-e): sine-shaped bands "Idealized high latitudes" and "Idealized low latitudes" to mimic Rose et al. (2014) and Kang and Xie (2014), "Idealized homogeneous" (not shown, uniform Wm^{-2} anomaly), and three spatially varying "Realistic" patterns reconstructed from the coupled model (see Section 3.3.2). "Realistic" here means imitating the coupled model

behavior, not real world observations. The rationale is to analyze the coupled model's behavior with a series of slab ocean simulations, each imitating one period in time of the coupled simulation (colored arrows in Fig. 3.1b and discussed in Section 3.4). To quantify the feedbacks associated with each ocean heat uptake pattern, we introduce a new approach: We scale the pattern up and down with a constant factor, to produce a global mean ocean heat flux of 1, 2, 3 and 4 Wm^{-2}. Each of these cases is a combination of positive 4xCO_2 and negative ocean heat uptake forcing of varying strength and pattern. All changes in the feedback parameter here are solely due to different ocean heat uptake patterns as opposed to temperature or time dependencies or Earth-system feedbacks (e.g., Senior and Mitchell, 2000; Meraner et al., 2013; Knutti and Rugenstein, 2015; Gregory et al., 2015). Throughout the paper we show the final decade of an equilibrated 40-year-long simulation. The feedbacks are quantified through linearly regressing the top-of-atmosphere imbalance (which in equilibrium is equal to the prescribed net ocean heat uptake) on the near-surface change in air-temperature. Note that there is virtually no internal variability and the uncertainty on the regression is thus very small (Supplemental Table 1 in the Appendix). We also circumvent the problem of regressing points which are not equally spread in the temperature space. The approach is similar to the SST pattern scaling by Andrews et al. (2015), but we – in addition to using a different model – constrain only the net ocean heat uptake, while the SST and heat flux components can evolve freely. Thus, we link the changing feedback parameter to heat uptake and not to changing SST patterns only. Our approach differs from the 2 Wm^{-2} imbalance of Rose et al. (2014) and the 3.3 Wm^{-2} imbalance of Kang and Xie (2014) in that we use realistic topography and patterns of ocean heat uptake, spatially varying mixed layer depth, sea ice, seasonal varying solar insolation, a large range of flux imbalances, a different reference state, and a new method to quantify feedbacks. Our aim here is to explain the typical coupled climate model behavior and potentially real world climate evolution. We analyze to what extent idealized setups mimic realistic behavior and thus close an important gap in the model hierarchy.

To compare to the idealized studies done without sea ice and due to slab ocean configuration technicalities, we simulate all cases without and some cases with sea ice (see details on allowing sea water to supercool in the appendix). Limitations of our setup are the use of a single model (though according to Rose et al. (2014), it is representative of at least three other models) and the strong sensitivity of sea ice growth to the local Q-fluxes (Rose, 2015). Note that the Q-fluxes are not symmetric about the equator in most cases, shifting the intertropical convergence zone, which is not our focus here, but itself an area of research (e.g., Kang et al., 2008, 2014; Zhang et al., 2010).

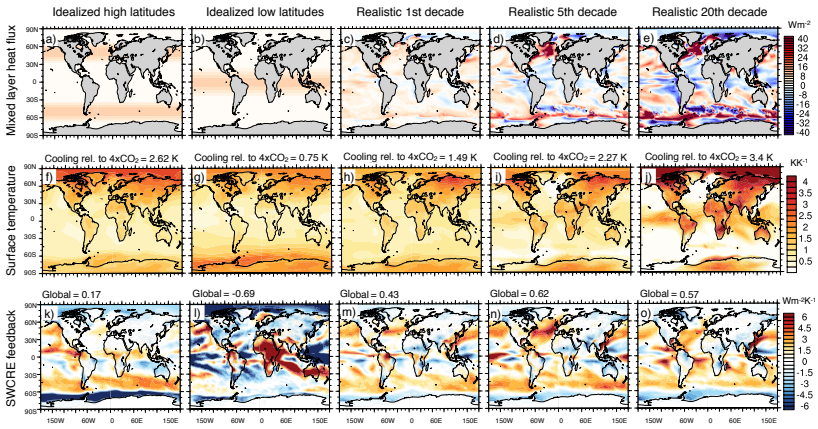

Figure 3.2: *Heat flux imbalances of the same global magnitude but different spatial patterns result in different surface warming and feedback magnitudes. Upper row: Q-flux patterns, all scaled to globally average to 2 Wm^{-2} into the ocean. Idealized high (a) and low (b) latitude heat uptake, realistic decadal averages of the coupled run's 1st (c), 5th (d), and 20th (e) decade. Middle row: Resulting equilibrium warming pattern, after applying both 4xCO_2 (positive forcing, warming) and the Q-flux pattern (negative forcing, cooling, global value indicated above each panel), normalized by the global average equilibrium temperature change. Bottom row: Geographical distribution of the short wave cloud radiative effect (SW CRE). Other feedback components are shown in Appendix Fig. S6.*

3.3 Results

We first discuss the influence of the idealized ocean heat uptake forcing on global and regional temperatures, the global net feedback, its different components, and their geographical distribution. We then show the same metrics for the realistic cases and discuss how they differ from the idealized ones.

3.3.1 Temperature and feedback responses to idealized pattern

Fig. 3.2f and g shows the warming pattern for two of the idealized cases normalized with their global equilibrium temperature and confirm that the overall magnitude and the spatial structure of the warming sensitively depends on the ocean heat uptake pattern. Above each panel, the cooling induced by the ocean heat uptake forcing relative to the warm reference state is indicated. High latitude ocean heat uptake is thus 3.5 times (3.3 without sea ice) more effective in cooling the atmosphere than low latitude ocean heat uptake of the same amount (here globally 2 Wm^{-2}). In other words, the ocean heat uptake efficacy – introduced by Winton et al. (2010) to compare the radiative effect of CO_2 alone (in our case 1.1 $Wm^{-2}K^{-1}$ for the transient response of abrupt4xCO_2 to the control Q-flux pattern) to the radiative effect of ocean heat uptake – is small for the low latitude case ($\epsilon_{\mathrm{low\,lat}} = 0.48$) and large for the high latitude case ($\epsilon_{\mathrm{high\,lat}} = 1.57$). This compares well with Rose et al. (2014), but is substantially smaller than Kang and

Xie (2014) – who find 13 times stronger effects of high versus low latitude heat release, prescribed in a narrower region.

Fig. 3.3a shows the global feedback parameter for the three idealized cases. Every dot is a decadal average of an equilibrated slab ocean simulation with different global imbalances for each ocean heat uptake pattern. Solid lines and filled dots indicate simulations that do include sea ice, dashed lines and open circles those which do not. The regression is a linear least squares fit through all available points (four to six for each ocean heat uptake pattern, including the reference state at (0,0). The horizontal axis shows the cooling relative to the reference state. Overall, for each heat flux pattern the feedbacks are linear for the temperature and heat flux range tested here. The sea ice may or may not, depending on the ocean heat uptake pattern, amplify the response: for the high latitude case, the sea ice response makes up 33% of the total response (similar to Caldeira and Cvijanovic (2014)), while for the homogeneous and low latitude case it is only 11 and 3%.

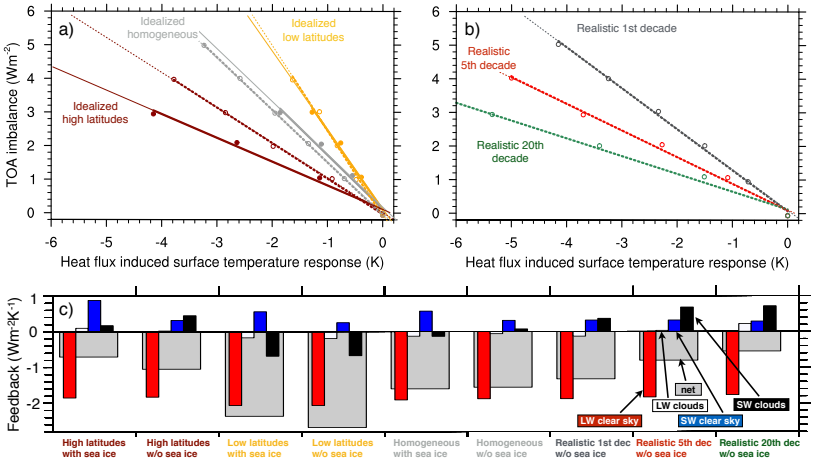

Figure 3.3: *Global feedbacks and their components. a) Idealized cases with sea ice (solid lines) and without sea ice (dashed). The regression line is thin where it is extrapolated beyond the data. Each dot is a decadal average of an equilibrated slab ocean simulation, the horizontal axis shows the coolings relative to a 4xCO_2 climate, see explanations in the text. b) Same as a) for the realistic heat flux cases without sea ice. c) Break down of the feedback components for each case in panel a and b.*

Fig. 3.3c shows the different feedback components for each case. The LW clear sky (red) feedback is more negative in the low latitude case, but otherwise does not change much, since we do not cover a large range of temperatures (Meraner et al., 2013). The SW clear sky (blue), reflecting the surface albedo, differs most strongly between the cases with and without sea ice. The LW cloud feedback (white) is generally small but changes sign between the low and high latitude cases. The largest contribution to the smaller magnitude of the global feedback parameter in the high versus low latitude case arrises from the SW cloud radiative effect (SW CRE, black). The bottom row in Fig. 3.2 shows its geographical distribution. In each grid box four equilibrated end points of the scaled pattern simulations are regressed against the global temperature anomaly.

The feedback pattern is completely different in most regions and sensitively depends on the sea ice formulation for the low latitude case. Compared to the Rose et al. (2014) aquaplanet study the signs of the global feedback are the same, but their magnitudes are substantially smaller in our idealized setup. In the Appendix we discuss other feedback components (Appendix Fig. S5), the simulations without sea ice (Appendix Fig. S6), and cloud masking effects (Appendix Fig. S3 and S4), which do not impact our conclusions.

3.3.2 Comparison to realistic patterns and ocean heat release

In the coupled simulations, the heat uptake patterns differ from the idealized in many aspects. The three representative realistic Q-flux patterns (Fig. 3.2c-e) feature the "Realistic 1st decade" with a relatively homogeneous heat uptake, the average response of the "Realistic 5th decade", in which the reduction of the Atlantic Meridional Overturning Circulation (AMOC) is strong (red patch in the North Atlantic) and the Southern ocean heat uptake becomes important, and finally, the "Realistic 20th decade", in which the AMOC has re-strengthened and the Southern Ocean becomes the dominant heat sink (e.g., Frölicher et al., 2014; Li et al., 2013a). These latter two patterns are representative of CMIP5 models surface heat flux after 100 years (Marshall et al., 2014). We cannot differentiate which part of the pattern is due to the atmospheric or oceanic influence only (e.g., Stouffer and Manabe, 2003; Xie et al., 2010; Exarchou et al., 2014), since we deduce the Q-flux from the coupled model. In parts of the tropics but also high latitudes the heat uptake capacity of the local ocean saturates, and the circulation shifted enough to make some places strong heat *sources* for the mixed layer and the atmosphere. Since the heat fluxes are locally large enough to lead to excessive sea ice growth, we simulate the whole suite of realistic experiments without sea ice for cleaner comparison.

Fig. 3.2h-j show again the normalized temperature response to a $2\,\mathrm{Wm}^{-2}$ ocean heat uptake forcing in a $4\mathrm{xCO}_2$ climate. The 1st decade pattern induces a relatively homogeneous warming, while stronger local heat uptake leads to a small local temperature response (white patches in Fig. 3.2i and j). The polar amplification is reduced (Fig. 3.2h and i) compared to the idealized cases. However, a strong Arctic amplification occurs in the 20th decade, in which the realistic coupled models' sea ice is almost completely melted, causing large upward heat fluxes into the mixed-layer. Note that the ocean heat uptake can influence land surface temperatures on all continents even centuries after the application of the abrupt$4\mathrm{xCO}_2$ forcing and shift the ocean-land warming contrast. Absolute temperatures are not shown here, but land warming is stronger in India and South Africa and weaker in Europe in the 20th than in the 5th decade, even though the global temperature is smaller.

Fig. 3.3b and c show again the global feedback parameter and its components for each case. In contrast to the idealized cases in the realistic cases the LW cloud feedback explains more of the difference between the cases (see SI Table 1) and is the only feedback changing sign. The spatial SW CRE distribution (Fig. 3.2m - o and Appendix Fig. S7) does – in most regions – only change its magnitude. Exceptions are the Eastern topical Pacific, the North Atlantic, and the Southern Ocean. This comparably subtle change in local feedback (compare Fig. 3.2 k versus l with m versus o) – which still aggregates to a strong global signal – might be the reason why Armour et al. (2013) found a description of locally constant feedbacks suitable. In agreement with the idealized setup the SW CRE contributes strongest to the difference between the cases.

We re-iterate the reason for the different temperature response from the heat flux perspective with the help of an illustrative Figure. Fig. 3.4 depicts the heat fluxes at the surface, the TOA, and their difference – the local divergence – at each latitudinal band for the idealized cases of low (yellow) and high (red) latitude ocean heat uptake, relative to the $4\mathrm{xCO}_2$ case. The low latitude ocean heat uptake is balanced mostly by the TOA fluxes within the same latitudinal bands (yellow range between 30°S and 30°N), thus,

meridional heat transport is necessary to compensate the exact regional surface heat flux pattern, but occurs within the latitudes of the prescribed anomaly. Because the atmosphere is unstable, the surface anomalies readily reach tropopause height, and the outgoing LW radiation strongly depends on the temperature (see more details in Kang and Xie (2014)). The result is a weak global cooling (Fig. 3.2g). For other shapes of surface heat flux forcings the meridional heat transport into the tropics can be larger and caused by the suppression of subtropical evaporation and poleward latent heat transport (Rose et al., 2014). The high latitude ocean heat uptake, however, is not balanced locally at the TOA due to the stably stratified atmosphere (red shaded areas and arrows). Large heat transport into the regions of the ocean heat uptake are necessary and lead to a stronger polar amplified warming compared to the low latitude case. Thus, high latitude ocean heat uptake has stronger far field effects and is more efficient in changing global temperatures than low latitude ocean heat uptake. Appendix Fig. S7 confirms this, showing the geographical distribution of each feedback component: The case with the stronger high latitude heat flux ("Realistic 20th decade") has a stronger influence on the low latitude cloud feedbacks than the case with a less amplified high-latitude pattern ("Realistic 5th decade"). However, with our setup we cannot tell whether the cloud response causes or reacts to the meridional heat transport. This highlights the necessity of studying the cause and effect of global ocean surface flux pattern, meridional heat fluxes, and local cloud responses (e.g., Rose and Rencurrel, 2016).

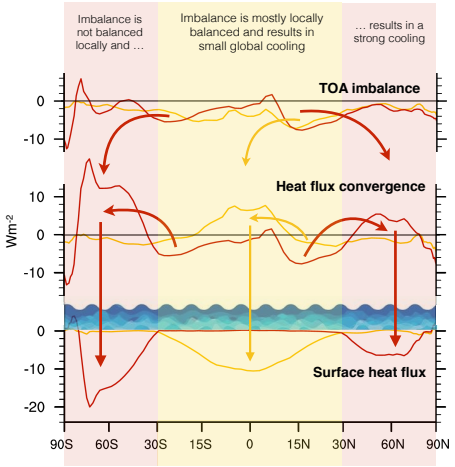

Figure 3.4: *High latitude surface heat flux imbalances result in a far field effect and large meridional heat transport. Zonally averaged top-of-the-atmosphere imbalance, atmospheric heat transport divergence (convergence positive), and heat flux into the interior ocean – relative to the $4xCO_2$ equilibrium for Idealized low (yellow shading and lines) and high (red) latitude pattern each averaging globally to a $3\,Wm^{-2}$ imbalance. Arrows indicate local and far field balancing of the surface perturbation. Appendix Fig. S8 shows realistic cases.*

The heterogeneity of ocean heat uptake, even to the point of local heat release (blue in Fig. 3.1c – e) is little appreciated so far in the literature. The Q-flux patterns differ

little in the regions of heat uptake between the 5th and 20th decade (the Nordic Seas being an exception), but the magnitude of the local surface fluxes differ at smaller spatial scales. For example, the Southern Ocean takes up a lot more heat at later stages during the equilibration, but it also releases large amounts a few degrees farther north and south. In reality and coupled models, the dominant mixing length scale of the surface ocean controls how strong an imprint the heterogeneity of (deep) ocean heat uptake has on the SSTs and surface heat flux patterns. Further studies on heat uptake and release and their relative, local and far field impact are needed to determine whether the effect found here in the slab ocean setting carries over to the real world. In the supplemental information we discuss other spatial homo- and heterogeneous patterns as a first sensitivity test.

Overall our results show for the first time that, the findings from idealized cases do carry over to realistic cases, with three exceptions. First, strong low latitude focused ocean heat uptake or release seems not to occur without changing pattern elsewhere in a decadal-averaged coupled system. Consequently, the difference in feedback components between two idealized cases is more than twice as large as the differences between two realistic cases. Next to the SW CRE, the difference between two cases is set mostly by the SW clear sky feedback for the idealized cases, and by the LW CRE for the realistic cases (Fig. 3.3c and supplemental Table 1 in the appendix). The magnitude of feedbacks is generally smaller in the realistic cases. Second, sea ice may be – depending on the spatial Q-flux forcing pattern warming – an important part of the overall feedback response. Third, the ocean heat release and the exact pattern and distance of release and uptake might play an important role in the noisy real world.

3.4 Implications and outlook

We have shown that different patterns of SST and surface heat fluxes – induced and constrained by heat fluxes from the mixed layer into the deep ocean – can lead to a continuous time evolution of the global climate feedback parameter in a coupled model, which is usually assumed to be constant. For each specified ocean heat flux (Q-flux) pattern global climate feedbacks are linear in the temperature and heat flux range we test, but the magnitude of the feedback parameter and its components depends strongly on the Q-flux pattern.

Fig. 3.1b compares the global feedback parameter of all Q-flux forcing cases (colored arrows) to the one of the fully coupled simulation. Because the slab ocean simulations include some mixed layer averaging, a fixed mixed layer depth, a different reference state, and in some cases no sea ice, the feedback parameters do not match the specific "realistic" decade they represent (Shell, 2013). Overall however, the cases mimic the evolution of homogeneous to high latitude ocean heat uptake. Additionally, we show that the homogeneity of the ocean heat uptake pattern can explain variations in the global feedback parameter of the coupled system. The 150 year regression (blue arrow, calculated by regressing the coupled transient response of the first 150 years following the abrupt4xCO_2 forcing) results in an arbitrary feedback parameter (Andrews et al., 2012, 2015) and the low latitude ocean heat uptake cases result in a overly unsensitive state never reached in the coupled model's reality.

As a sideline, our analysis confirms that slab ocean simulations with a prescribed constant climatological Q-flux might not be the preferred tool to study transient behavior of atmosphere and land (e.g., Jonko et al., 2013; Bitz et al., 2012; Shell, 2013; Donohoe et al., 2014; Deser et al., 2015), even if global mean ocean heat uptake is negligible.

Resulting follow up questions include: Which local features and physical mechanisms of the heat uptake in which geographical combination are most efficient in changing the magnitude of the global feedback parameter and are we sure they will occur in

reality and are not modeling artifacts (e.g., Zhang et al., 2010; Exarchou et al., 2014; Liang et al., 2015)? What is the critical size of a region with a certain ocean heat uptake to influence the magnitude of the global feedbacks (L'Hévéder et al., 2015; Kang and Xie, 2014, see discussion and sensitivity tests in the appendix)? How do decadal variability, the time scales of changing patterns, and the superposition of different forcings (aerosols versus ocean heat uptake) modulate our findings and the transferability to the real world (Hansen et al., 1997; Ban-Weiss and Caldeira, 2010; Hsieh et al., 2013; Long and Collins, 2013; Gregory and Andrews, 2016; Dallafior et al., 2016)? When and where does local ocean heat uptake saturate and anomalous heat release become important relative to the climatological heat fluxes in realistic scenarios (Kravitz et al., 2013)? Subtle changes in wind stress or ocean circulation might have a small direct influence in terms of net heat uptake but a large indirect impact on surface temperature, through modifying feedback magnitudes. How much of polar amplification and its evolution is solely due to ocean heat flux pattern induced atmospheric heat flux convergence? What is the physical mechanism of ocean heat uptake and cloud response and how does it evolve in the coupled system (Rose and Rencurrel, 2016)? How much of the disagreement among models in simulated feedback strengths is due to the difference among the models in ocean heat uptake patterns (Winton et al., 2010) or ocean heat uptake induced SST changes (Ma and Xie, 2013)? The use of a global linear feedback parameter in intermediate complexity models, prediction, and impact studies should be – depending on the purpose of use – carefully considered and may turn out to be problematic. The ability to predict transient and equilibrium behavior on any time scale beyond a decade critically depends on our understanding, regional observations, and ability to correctly simulate SST pattern formation, ocean heat uptake pattern, and the atmospheric response to those changing fluxes – as they evolve through time.

4 Nonlinearities in patterns of long term ocean warming

published by Maria Rugenstein, Jan Sedláček, and Reto Knutti in Geophysical Research Letters (2016) doi: 10.1002/2016GL068041

The ocean dominates the planetary heat budget and takes thousands of years to equilibrate to perturbed surface conditions, yet those long time scales are poorly understood. Here we analyze the ocean response over a range of forcing levels and time scales in a climate model of intermediate complexity and in the CMIP5 model suite. We show that on century to millennia time scales the response time scales, regions of anomalous ocean heat storage, and global thermal expansion depend non-linearly on the forcing level and surface warming. As a consequence, it is problematic to deduce long term from short term heat uptake or scale the heat uptake patterns between scenarios. These results also question simple methods to estimate long term sea level rise from surface temperatures, and the use of deep sea proxies to represent surface temperature changes in past climate.

4.1 Ocean heat uptake in CMIP5 models

Thermal expansion of ocean waters due to heat uptake from the atmosphere is a large contributor to recent and near future sea level rise (Church et al., 2011, 2013; Levermann et al., 2013). General circulation models (GCM) differ in the amplitude of simulated thermal expansion due to different base states, the total amount and vertical extent of the heat uptake, heat redistribution, and differences in the representation of vertical heat transport processes, advection, isopycnal and diapycnal mixing (Gregory, 2000; Kuhlbrodt and Gregory, 2012; Hallberg et al., 2012; Church et al., 2013; Exarchou et al., 2014; Melet and Meyssignac, 2015; Liang et al., 2015). Fig. 4.1a-c shows zonal averaged ocean temperature anomaly patterns at the end of the century for three different scenarios simulated by the Coupled Model Intercomparison Project Phase 5 (CMIP5) models. 95-97% of the ocean heat uptake is confined to the upper kilometer, although locally, deep ocean heat uptake can contribute a large fraction of the total amplitude already decades after the perturbation (Kuhlbrodt and Gregory, 2012; Marshall et al., 2014). The standard deviation between the models increases with the forcing level and in Southern and Northern hemispheric high latitudes, but is generally smaller in magnitude than the mean signal shown in Fig. 4.1, see also Appendix Fig. 1-3 (Yin, 2012; Sallée et al., 2013; Heuzé et al., 2015). Fig. 4.1d-e demonstrate that locally, pattern scaling between the different scenarios is not possible with high accuracy (Bilbao et al., 2015). If the ocean warming pattern would respond linearly to the surface forcing the differences between the scaled scenarios in the upper row should be zero. However, the

lower forced scenario takes up relatively more heat into the Southern Ocean and to low latitude intermediate depth of around 500 – 2000m (panel d) than the intermediate and higher forced scenarios (panel f). In the surface ocean higher forcing lead to relatively more heat uptake (see also Appendix Fig. B.1 for a zoom into the upper ocean).

The near future sea level rise has been studied extensively, but is also known to be a poor indicator of centennial, millennial, or equilibrated conditions (e.g., Stouffer and Manabe, 2003; Li et al., 2013b), but only very few GCMs have been run over millennia due to computational cost. Here we explore the centennial to millennia patterns of ocean warming and their dependence on time and on forcing levels. We show that long term thermal expansion is not proportional to surface warming, and that deduction of one time frame or forcing scenario to another is limited, not only for the transient response as shown for the CMIP5 models, but also for equilibrated conditions.

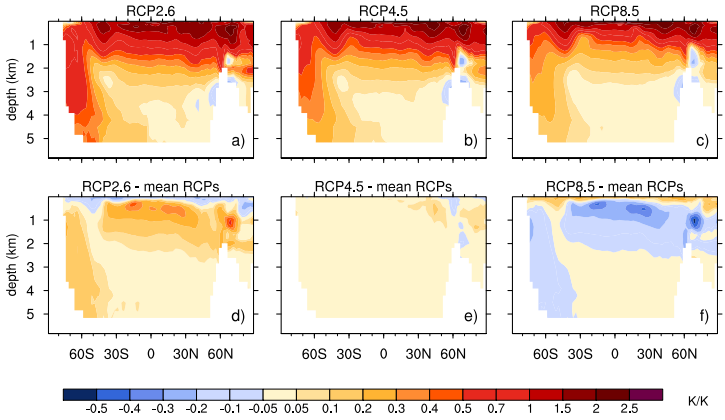

Figure 4.1: *Zonal mean ocean temperature anomaly averaged over year 2081 to 2100 and normalized with the average ocean temperature of Representative Concentration Pathways (RCP) 2.6 (a), RCP4.5 (b), and RCP8.5 (c) of 14 CMIP5 Models, described in Collins et al. (2013), Figure 12.12. Differences between RCP2.6 (d), RCP4.5 (e), RCP8.5 (f) and the mean of the three RCPs. The color scale is not linear (around 0 and above .5). See Appendix Fig. B.1-B.3 for more details on each model.*

4.2 Model and Simulations

To explore the limits of scaling over a wide range of forcing levels and time-scales up to equilibrium, we use the intermediate complexity model (EMIC) ECBILT-CLIO, which consists out of an ocean, sea ice, and atmospheric component. The ocean general circulation model has a free surface, 20 unevenly spaced layers, a 3°x3° horizontal resolution, and a thermodynamic-dynamic sea ice model (Goosse and Fichefet, 1999, and more details on the model formulation in the Appendix). The atmospheric model solves the quasi-geostrophic equations on a spectral grid with three vertical levels and a horizontal resolution of 5.6°x5.6°, parameterizations of diabatic heating and surface heat fluxes, and prescribed seasonal varying cloud cover (Opsteegh et al., 2011). This results

in a low Equilibrium Climate Sensitivity (ECS) of 1.7 K (see also Friedrich et al., 2010; Levermann et al., 2013; Eby et al., 2013; Church et al., 2013, and references therein). Given the limitations of the model, we emphasize that we aim at insight and not prediction. We will interpret the scenarios relative to each other, rather than their absolute values compared to other models with the same forcing. The advantage of the model is that large ensembles of hundreds of members, and simulations of 10,000 years are possible, a range unfeasible with a full general circulation model.

We conduct a range of step forcing experiments, each consisting – roughly according to the ratio of anomaly signal to internal variability noise – of several initial condition ensemble members: 1.07 times preindustrial CO_2 concentrations of 280ppm (90 members), 1.4xCO_2 (90 members), 2xCO_2 (50 members), 4xCO_2 (20 members), 8xCO_2 (20 members), and 16xCO_2 (10 members), each 1000 years long. The forcing levels are chosen to simulate roughly no change in circulation for the very low forced cases to an initially strongly stratified response and large heat uptake for the high forced cases. One member of each forcing level is further integrated to 10,000 years. We show anomalies of these simulations with 1000 or 10,000 year long control simulations (no CO_2 change, 94 initial condition ensemble members), i.e. each ensemble member's anomaly is determined with a slightly different control simulation. The temperature of the control simulations shows no drift, but the salinity shows both a small linear control run drift (globally 0.05 kg m^{-3} per 10,000 yrs) and a smaller forcing level dependent drift, due to the handling of the sea ice. Both drifts are accounted for by scaling the salinity pattern at each time step, location, and forcing level with the drift, so that the global mean salinity is constant at all times, but allowed to change its spatial pattern.

4.3 Equilibration of ocean heat uptake and circulation changes

Fig. 4.2 a-d shows – analogous to Fig. 4.1 – the equilibrium zonally averaged ocean temperature anomalies normalized with the equilibrium global ocean temperature for four representative scenarios in color, and the unscaled chan-ges in salinity in grey contours. The temperature anomaly is not distributed homogeneously even though the global ocean temperature is in equilibrium with the surface. This agrees with models showing gradients of several degrees (Gillett et al. (2011) and Knutti (2002)), but disagrees with the rather homogeneous warming patterns of Stouffer and Manabe (2003) and Li et al. (2013b), in which the vertical gradient of temperature is less than 1.5 K. Contrary to the transient pattern analyses of Kuhlbrodt and Gregory (2012) and Melet and Meyssignac (2015), the equilibrated simulations do not scale with their global temperature especially in the Southern Ocean, which is widely recognized to be important in shaping the transient and equilibrium global heat uptake (e.g., Schneider and Thompson, 1981; Manabe and Stouffer, 1994; Bi et al., 2001; Bryan et al., 2006). The Northern Hemisphere high latitudes become relatively more important in lower forced scenarios, as suggested also in Fig. 4.1 d-f. Salinity and temperature anomalies roughly follow the same pattern, but impact the density in opposite ways through the equation of state. In regions with increased warming, the salinity anomaly is positive, thus compensating the change in density due to the increased heat content (Lowe and Gregory, 2006). This explains why the maximum warming can be sustained in the tropical subsurface ocean without destabilizing the water column (Yin et al., 2011). Panels e-h show the time evolution of global ocean warming with depth. While the higher forced scenarios take longer to equilibrate they are more efficient in transporting the anomaly into the deep ocean. However, taking the average subsurface maximum versus deep ocean temperature as an indicator of homogeneity, we find that the vertical gradient does not evolve linearly with the forcing level: The subsurface maximum warming is a factor of 3.3, 5,

7, 4.3, 1.9, and 1.2 greater than the deep ocean warming for the scenarios of 1.07, 1.4, 2, 4, 8, 16xCO_2. On century time scales, for very low forced scenarios even negative deep ocean temperature trends are possible (panel e blue lines). This is consistent with slightly decreasing trends in recent decadal observations (Wunsch and Heimbach, 2014; Llovel et al., 2014; Liang et al., 2015). Locally, slightly negative and positive trends can occur after the upper 2000m are fully equilibrated.

Fig. 4.3 further explains the time evolution of the different forcing scenarios. All panels share the same color coding (ranging from 1.07 in yellow to 16xCO_2 in blue) and logarithmic time axis. The global average surface air temperature anomaly (panel a) scales only roughly with the forcing level (i.e., the equilibrium climate sensitivity for higher forcings is less than expected from linearly scaling up lower forcing levels). Panel b shows the global average ocean temperature anomaly with the global value as solid and the upper most kilometer evolution as dashed lines. In some scenarios, the deep ocean reaches its equilibrium value at almost the same time scale as the upper ocean (4xCO_2), while for other scenarios it takes several thousand years longer (8xCO_2), dependent on stratification, overturning, and mixing response. Panel c depicts the sea level rise due to ocean thermal expansion and salinity changes, calculated from the global detrended in-situ density anomaly, volume, control run reference density and surface area. The gray dashed lines show the sea level rise due to thermal expansion only, assuming a constant salinity pattern. While locally the salinity changes are important to set the dynamics and sea level change, globally they are negligible. In equilibrium, the deep ocean (below 1km) accounts for 64, 55, 60, 67, and 72 % of the total thermal expansion for the 1.07, 1.4, 2, 4, and 16xCO_2 forcing, respectively. This compares well with the 60% contribution of the deep (below 1.5km) ocean of the general circulation model of Li et al. (2013b).

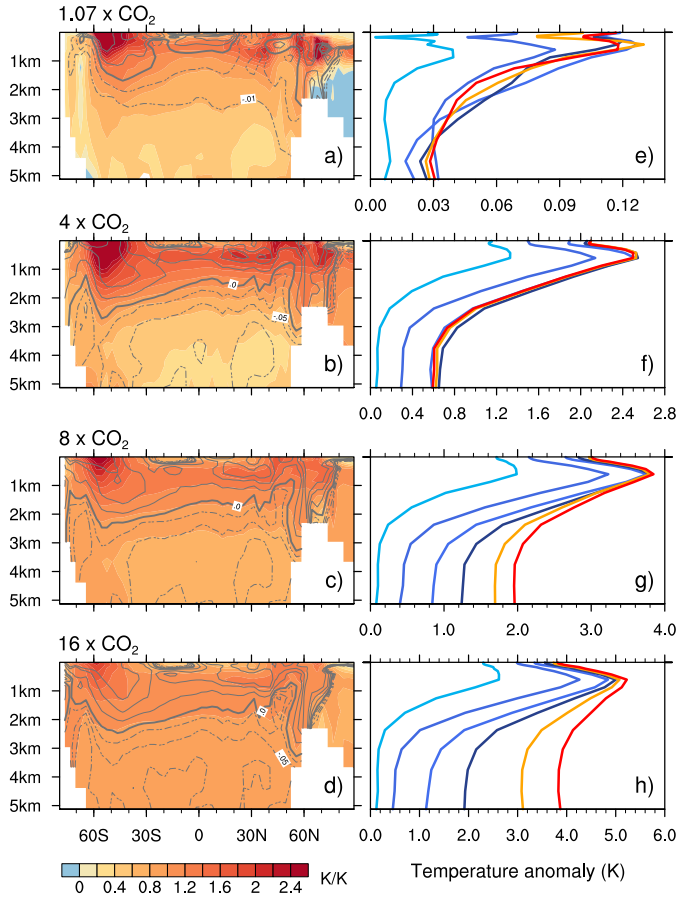

Figure 4.2: *(a-d) Zonal average equilibrium (year 9000-10 000) ocean warming pattern for four forcing scenarios, normalized with the average ocean temperature of that time frame. Gray lines show salinity anomalies contours of -.075, -.05, -.025 (dashed) 0 (thick) and .05, .1, .15 psu unscaled. (e-h) Time evolution of global average warming at all depth for the same scenarios. Lines from light blue to dark blue to red indicate 100-year averages around year 100, 300, 600, 1000, 2000, and 9900 after the perturbation. Only the single long simulations are used for this figure. Scale for panel e-h changes.*

The patterns of ocean heat uptake and redistribution, as well as changes in the patterns of salinity depend on circulation changes, thus on surface temperature anomaly and the forcing level. The Atlantic Meridional Overturning Circulation (AMOC, panel

d), here defined as the maximum in the stream function at 30° N, initially declines, as expected, more with a higher forcing level (Manabe and Stouffer, 1994; Gregory et al., 2005; Zhu et al., 2014). However, both the recovery level and recovery rate vary with the forcing level and some scenarios recover to a greater degree than others. While almost all GCMs and EMICs show a reduction of the AMOC with increased radiative forcing, the time and strength of the recovery differs strongly between them (Stouffer and Manabe, 2003; Li et al., 2013b; Zickfeld et al., 2013). The recovery or increased instability can be caused by a surface salt-advection feedback (Stommel, 1961; Latif et al., 2000; Bryan et al., 2013; Drijfhout et al., 2011) or increased bottom water temperatures due to slower Antarctic Bottom Water Formation (AABWF) (Manabe and Stouffer, 1994; Stouffer and Manabe, 2003). Since the stronger the AMOC declines, the more heat can be taken up by the deep North Atlantic (Rugenstein et al., 2013), the very strong reduction and slow recovery in the higher forced cases likely causes the more vertically homogeneous equilibrium temperature anomaly (Fig. 4.2).

Finally, panel e and f show the time evolution the AABWF strength – defined as maximum overturning at 70° S – and the sea ice volume – depicted as fraction of the control run value. The AABWF reduces in all cases, but does not scale linearly with the radiative forcing. The time of minimum AABWF, the recovery or overshooting amplitude, and rate of recovery vary several thousand years between the different forcing cases. The reason for the recovery is not well understood and attributable to either the long-term warming or the salinization of the deep ocean, destabilizing the water column from below (Manabe and Stouffer, 1994; Bi et al., 2001), or strong enough convective events triggered by changed seasonality (Yamamoto et al., 2015). The higher forced levels, which show relatively more deep ocean warming (Fig. 4.2), indeed recover and overshoot. The Southern Ocean sea ice reduces in all cases roughly proportional to the forcing level until year 20, before the rate of change as well as the equilibrium level becomes forcing level dependent.

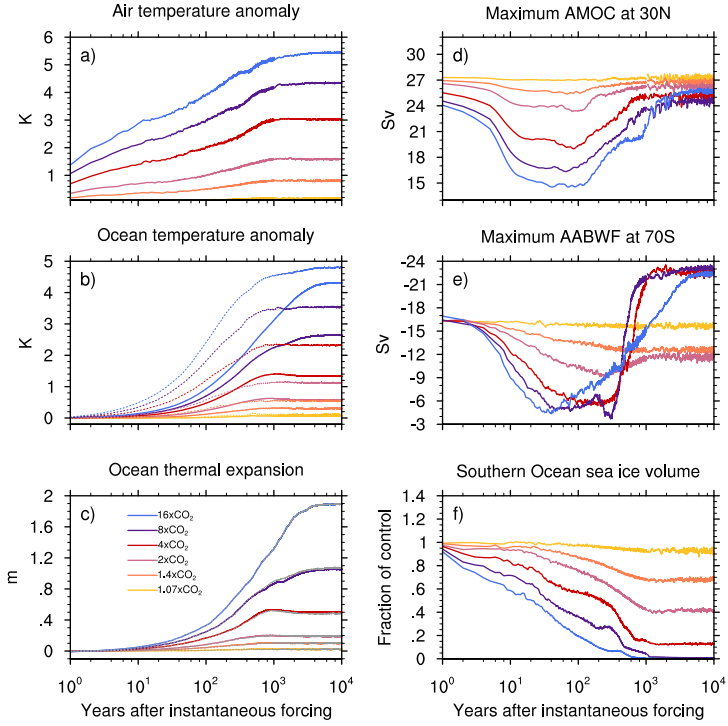

Figure 4.3: *Time evolution of (a) global average surface air temperature anomaly, (b) global average ocean temperature anomaly for the whole (solid) and upper-most kilometer ocean (dashed), (c) global average ocean thermal expansion, (d) maximum Atlantic Meridional Overturning Circulation at 30° N, (e) Antarctic Bottom Water Formation, as maximum of overturning at 70° S, and (f) the sea ice volume, expressed as fraction of the control simulation value. In all panels, year 1-1000 is the annual and ensemble average; year 1000 – 10'000 is a 100 year running mean of one simulation for each forcing level.*

4.4 Thermal expansion

We now explore the consequences of the in-homogeneous and forcing dependent warming pattern on the sea level rise due to thermal expansion. Previous studies used the equilibrium surface temperature anomaly as dependent variable and found an approximately linear relation to the thermal expansion (Knutti and Stocker, 2000; Meehl et al., 2007; Levermann et al., 2013), but the evidence for this is mostly based on intermediate complexity, 2.5D, or single basin models. However, there is no physical reason why this should be the case, and Pardaens et al. (2011) and Körper et al. (2013) find indeed, but

do not explain, non-linearities for transient states at the end of the century for several GCMs.

We discuss three mechanisms which impact the relationship between the thermal expansion and surface temperature anomaly (Fig. 4.4). (1) The non-linearity of the equation of state: The expansion coefficient of sea water increases with warmer temperature and lower pressure (e.g., Palter et al., 2014). (2) The transient effects of taking up more heat with time but moving a larger fraction of heat into the deep ocean, where thermal expansion is less. (3) The forcing level and circulation dependent heat uptake as discussed in Fig. 4.2 and 4.3. Knutti and Stocker (2000) pointed out that thermal expansion is 0.5 m higher with a shutdown AMOC compared to a state of the same forcing and recovered AMOC or North Pacific overturning.

We reproduce this result in ECBILT-CLIO through a fresh water perturbation in the North Atlantic and the sea level rise due to the AMOC collapse – without any radiative forcing – is 0.3 m. To fill up gaps between simulations shown above and to explore higher warming levels, we show 18 additional simulations (depicted by dots in Fig. 4.4). The lowest forcing is $1.15\text{x}CO_2$ (dots farthest left), increasing in steps of $2^{0.2}$ up to $2^{4.4} = 21\text{x}CO_2$ (dots on the far right). The surface temperature sensitivity even becomes lower at high CO_2 values, so $21\text{x}CO_2$ should not be interpreted at face value. The time dimension is depicted in colors, showing five decadal to millennia time slices. Fig. 4.4a and b show that the total ocean heat uptake (which dominates the Top of the Atmosphere (TOA) radiative imbalance), is linearly related to the surface air temperature anomaly at all times, but does not translate into a linear relationship between the ocean and surface air temperature anomalies on centennial to millennial time scales. For one unit of surface warming the ocean warms more through time (equilibration, curves moving up) and higher forcing levels (circulation changes, curves bend up, see also Fig. 4.2 and 4.3). Fig. 4.4c brings the equation of state into play: The sea level rise due to thermal expansion is linearly related to the ocean temperature anomaly on decadal time scales (dark red line). If the equation of state were linear and the circulation were constant, the equilibrium values would lie on the black line, far away from the equilibrated model output (green line). To assess how much of this non-linearity is due to the equation of state, we use the 3D warming and salinity pattern after 50 years and linearly scale it up, mimicking 5K more warming under the assumption of no circulation change. We then calculate the thermal expansion with the non-linear equation of state (McDougall et al. (2003), rho_mwif function, see also Appendix, gray dashed line). The green line is close to that estimate, i.e., the greatest part of the non-linearity of the equilibrated situation is due to the fact that the equation of state is non-linear. The remaining discrepancy between the gray and green line is due to the effect that for longer time scales and higher forcing levels more heat is transported from the surface layers to the deep ocean. Fig. 4.4d brings all effects together, showing that thermal expansion from centennial time scales onwards – increasingly with higher forcing levels – is non-linearly related to the surface temperature anomaly. The dashed gray lines show again two artificially constructed cases to assess the impact of the non-linear equation of state without circulation changes (explained in the Appendix). In summary, thermal expansion is approximately proportional to the atmospheric warming on timescales of a century and for small forcings, but on long timescales or for stronger forcings the linearity assumption is no longer valid. The reasons are that the globally integrated heat uptake itself is not proportional to surface warming and that the distribution of warming changes with changes in circulation and mixing, which in combination with the nonlinearity of the equation of states affects thermal expansion.

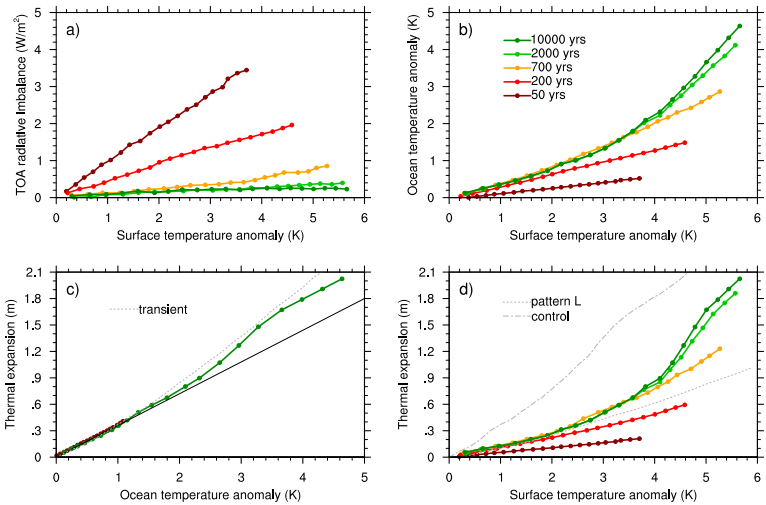

Figure 4.4: *(a) Top of the Atmosphere (TOA) radiative imbalance versus global surface temperature anomalies averages around year 50, 200, 700, 2000, and 10 000. (b) Global ocean temperatures versus global surface temperature anomalies, (c) Global sea level rise due to thermal expansion versus global ocean temperature anomaly. Constructed thermal expansion without circulation change and a linear (black) or non-linear equation of state (gray). (d) Thermal expansion versus global surface temperature anomaly and constructed thermal expansion based on a pattern of equilibrated warming at 1.07xCO_2 (pattern L) and a uniform anomaly added to the control simulation pattern (control, both dashed gray).*

4.5 Implications for paleo and modern studies

A common assumption in paleo oceanography is that the temperature (anomaly) of the intermediate or deep ocean – indicated by e.g. benthic foraminifera – represents a time averaged (sea) surface temperature (SST) record of the regions where deep waters formed (e.g., Savin, 1977; Huber, 1998; Zachos et al., 2001; Voigt et al., 2004; Cramer et al., 2009; Friedrich et al., 2012) or the SST anomalies represent deep ocean temperature anomalies (Jaccard et al., 2014). Modeling studies found the deep water temperature anomaly was the same as the southern high latitude SST anomaly (Manabe and Bryan, 1985) or that it resembles the SST anomaly of low latitude (Stouffer and Manabe, 2003). We find that the anomalous heat uptake by the intermediate and deep ocean does not correspond to the atmospheric temperature increase at low or high latitudes. This implies that, without knowledge of forcing and circulation history, the interpretation of ocean temperature proxies beyond their region may be more problematic than currently appreciated. Another long term implication of our finding concerns climate sensitivity. Fig. 4.3a indicates that the effective climate sensitivity in the model increases with increased temperatures for all simulations (see more elaborate argument in Knutti and Rugenstein (2015)). This implies that potentially, ocean circulation can have a large effect on the TOA radiative imbalance even in the absence of a cloud feedback. A more

general consequence concerns the marine carbon-climate feedback and regions of melting sea ice, which both are likely influenced by the places of anomalous heat storage and thus, dependent on the forcing history (Randerson et al., 2015). Finally and also on shorter timescales, pattern scaling between scenarios might break down already at the end of this century, or more likely in future centuries. We therefore cannot infer future or past equilibria from transient behavior or past transient behavior from equilibria. A detailed knowledge of the dependence of circulation responses to different forcing levels and scenarios is key to understand and compare not only transient, but also equilibrium warming patterns.

In summary, although for sub-centennial time scales and low forcing levels the linear relationship between thermal expansion and surface temperature anomaly seems to hold, our analysis suggests that we do not properly understand the centennial to millennia ocean warming patterns, mainly due to a limited understanding of circulation and mixing changes. Complex enough models – simulating long enough time scales and different ranges of scenarios (as e.g., in Krasting et al. (2016)) – are necessary to explore these effects.

5 Effective and equilibrium climate sensitivity

unpublished, in collaboration with Jonah Bloch-Johnson, Chao Li, Thorsten Mauritsen, Jonathan Gregory, Timothy Andrews, Thomas Frölicher, David Paynter, Shuting Yang, Gavin Schmidt, Ayako Abe-Ouchi, Gokhan Danabasoglu, Alex Jonko, Long Cao, and Jean-Louis Dufresne

Atmosphere-ocean general circulation models, as well as the real world, take thousands of years to equilibrate to CO_2-induced radiative perturbations. Equilibrium climate sensitivity – the global mean surface temperature response to a fully equilibrated $2xCO_2$ perturbation – has been used for decades as a benchmark in model intercomparisons, as a test of our understanding of the climate system and paleo proxies, and to predict or project future climate change. Computational costs and limited time lead to the widespread practice of extrapolating equilibrium conditions from just a few decades of coupled simulations. The most common workaround is the "effective climate sensitivity" – defined here through a linear extrapolation of a 150 year abrupt$2xCO_2$ simulation. We present an ongoing Model Intercomparison Project of millennia scale simulations ("LongRunMIP"), to study century and millennia time scales of AOGCM equilibration and the linearity assumptions around feedback analyses. As a first result, we show that the two most widespread models of extrapolating transient conditions fail to predict equilibrium conditions for most models. I discuss preliminary results on surface temperature patterns, the oceanic Meridional Overturning Circulation, and scaling the response to different forcing levels. The chapter is discursive in character, intending to encourage debate, and reflecting the on-going character of the project. Section 5.6 sketches out the planned publication and options for further research.

5.1 Current state of LongRunMIP

The investigation of changing feedback strengths in the CMIP5 framework is hampered by the fact that only 150 years of the abrupt4xCO_2 simulations are collected by protocol. Methods like the moving bin regression (chapter 1, 2, and 3) are limited when used on just 150 years, and methods to extrapolate transient to equilibrium behavior are rarely if ever verified (e.g., Gregory et al., 2004; Danabasoglu and Gent, 2009; Winton et al., 2010; Bitz, 2012; Andrews et al., 2012; Geoffroy et al., 2013a,b). In addition, the evolution of single models towards the equilibrium indicate that feedbacks do change after the first few decades (chapter 1 and 3, and e.g., Colman and McAvaney, 2009; Senior and Mitchell, 2000; Li et al., 2013b; Jonko et al., 2013). Analyses on the standard CMIP5 contribution of abrupt4xCO_2 and extended RCP8.5 simulations show that the water vapor feedback increases with temperature and that the short wave cloud

response sensitively depends on the evolution of the sea surface temperature pattern (Meraner et al., 2013; Andrews et al., 2015, chapter 6.3). Very recently, studies show that the feedback parameter changes also during the historical period, due to the surface temperature evolution (Gregory et al., 2016; Zhou et al., 2016).

Next to the verification of extrapolation methods and the quest to reduce the uncertainty of equilibrium climate sensitivity estimates, the century to millennial long runs are highly relevant to understand transient and equilibrium processes. Chapter 4 shows that the physical understanding of century to millennial time scales mostly depends on intermediate complexity models with usually very idealized representations of the atmosphere, especially clouds.

Since model comparisons with a protocol in the sense of CMIP6 are time consuming and century to millennial scale GCM scenarios require at least months – usually years – to simulate, we organize a "model comparison of opportunity". We collect existing GCM simulations, some of which are published, others are not and a few are extended or simulated for this comparison. The only condition to participate is a minimum of 1000 years of simulation of constant CO_2 concentration scenario of any kind. Fig. 5.1 indicates the degree of equilibration of the global mean surface temperature and Table **??** shows the submissions and people involved. Table 5.2 lists the requested variables.

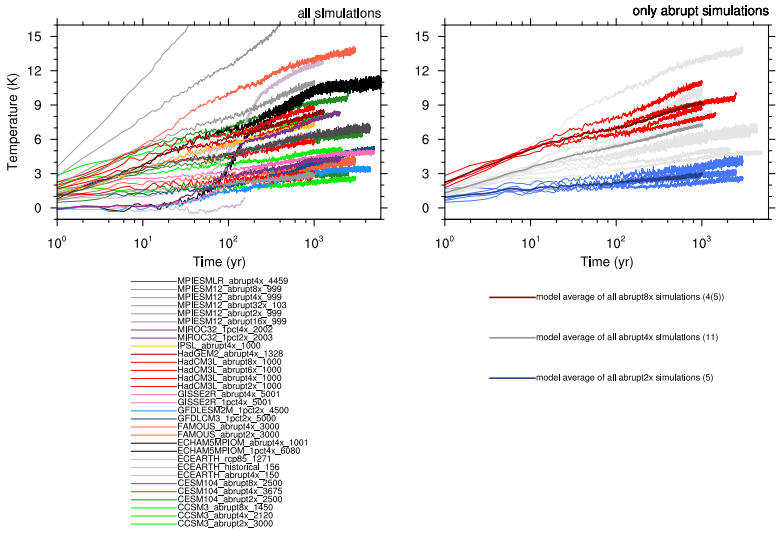

Figure 5.1: *Time evolution of surface temperature anomaly of all model submissions.*

Model	Simulation	Length [yr]	final temp [K]	true ECS [K]	scaled ECS [K]	ECS_{eff} [K]
CCSM3	abrupt2x	3000	2.53	2.44	4.88	2.30 (2.66)
	abrupt4x	2120	5.03	5.41	5.41	4.78 (5.16)
	abrupt8x	1450	8.14	8.94	5.96	7.85 (8.26)
CESM 1.0.4	abrupt2x	2500	3.09	3.14	6.28	2.55 (2.55)
	abrupt4x	3700	6.53	6.57	6.57	5.68 (5.99)
	abrupt8x	2300	9.62	9.92	6.61	8.95 (9.43)
EC-Eearth-PISM	abrupt4x	150	-	-	-	6.61 (7.03)
	RCP8.5+	1270	12.68	12.17	8.67	-
Echam5-MPIOM	abrupt4x	1000	10.18	11.47	11.47	10.43 (9.90)
	1pct4x	6080	10.91	10.91	10.91	-
FAMOUS	abrupt2x	3000	4.15	4.40	8.8	3.71 (4.00)
	abrupt4x	3000	13.71	15.388	15.388	11.20 (12.19)
GFDL-CM3	abrupt4x	150	-	-	-	7.62 (8.25)
	1pct2x	5000	5.15	4.95	9.90	-
GFDL-ESM2M	abrupt4x	150	-	-	-	5.62 (5.92)
	1pct2x	4500	3.37	3.33	6.66	-
GISS E2R	abrupt4x	5000	4.9	4.81	4.81	4.28 (4.48)
	1pct4x	5000	4.82	4.88	4.88	-
HadCM3L	abrupt2x	1000	3.01	3.34	6.68	3.15 (3.34)
	abrupt4x	1000	5.82	7.82	7.82	5.83 (6.30)
	abrupt6x	1000	7.60	11.87	9.22	7.42 (8.10)
	abrupt8x	1000	8.70	11.52	7.685	8.31 (9.25)
HadGEM2	abrupt4x	1299	8.36	9.71	9.71	9.04 (10.82)
IPSL	abrupt4x	1000	7.31	10.87	10.87	8.05 (8.37)
MIROC 3.2	abrupt4x	150	-	-	-	8.27 (8.78)
	1pct2x	2000	4.24	4.41	8.82	-
	1pct4x	2000	8.27	9.12	9.12	-
MPIESM 1.2	abrupt2x	1000	2.82	2.94	5.68	2.79 (2.98)
	abrupt4x	1000	6.25	6.58	6.58	6.04 (6.54)
	abrupt8x	1000	10.88	12.02	8.02	10.23 (10.65)
	abrupt16x	1000	18.15	22.73	11.42	17.00 (18.81)
MPIESM LR	abrupt4x	4459	6.88	6.87	6.87	6.03 (6.33)

Table 5.1: *Overview of all simulations. "final temp" is the average of the final 50 years of each simulation. "true ECS" is the final temperature for equilibrated simulations and the linear regression of the last few hundred years for un-equilibrated simulations."scaled ECS" is the "true ECS" scaled to 4xCO_2 assuming logarithmic forcing dependency. ECS_{eff} is based on the linear regression of first 150 years; two layer model including ocean heat uptake efficacy in brackets.*

Atmospheric variables (monthly)	Oceanic variables (annual)
surface latent heat flux	sea ice area fraction (monthly)
surface sensible heat flux	meridional overturning circulation
precipitation	net ocean heat uptake
sea level pressure	sea surface temperature
surface downwelling LW radiation	sea surface salinity
surface upwelling LW radiation	net water flux into sea water
TOA outgoing LW radiation	water evaporation
TOA outgoing cs LW radiation	precipitation
surface downwelling SW radiation	wind stress in x direction
TOA Incident SW radiation	wind stress in y direction
Surface upwelling SW radiation	salinity
TOA outgoing SW radiation	(3D, annual, Feb, and Sept)
TOA outgoing cs SW radiation	potential temperature
near-surface air temperature	(3D, annual, Feb, and Sept)
Surface Temperature	

Table 5.2: *Collected variables, cs means clear-sky, LW and SW mean long and short wave, respectively, and TOA means Top-of-the-atmosphere. All fields are 2D unless indicated differently.*

5.2 Testing energy balance models

As discussed in chapter 1, the linearization around an equilibrated state following a perturbation is a more or less valid description of the evolution of the global energy imbalance. Although long thought of as a "good enough" description of the coupled models (e.g., Murphy, 1995; Gregory et al., 2004; Andrews et al., 2012; Vial et al., 2013; Flato et al., 2013), this simple energy balance equation is questioned and alternatives are discussed (e.g., chapter 2, Winton et al., 2010; Caldeira and Myhrvold, 2013; Geoffroy et al., 2013b,a; Armour et al., 2013; Andrews et al., 2015; Huybers and Proistosescu, 2016)

We test the two-layer energy balance model without ocean heat uptake efficacy (equivalent to the ordinary least square regression of Gregory et al. (2004)) following the method described in Geoffroy et al. (2013b) and the two-layer energy balance model including ocean heat uptake efficacy following Geoffroy et al. (2013a). This method fits the following equations – representing the mixed layer ocean and atmosphere, and the deep ocean to GCM output.

$$c_{surf}\frac{dT_{surface}}{dt} = F - \lambda T_{surface} - \epsilon\gamma(T_{surface} - T_{deep}) \tag{5.1}$$

$$c_{deep}\frac{dT_{deep}}{dt} = \gamma(T_{surface} - T_{deep}) \tag{5.2}$$

where $c_{surf}\frac{dT_{surface}}{dt} + c_{deep}\frac{dT_{deep}}{dt}$ is equivalent to N in other chapters, i.e., the heat capacities of the upper and lower layers multiplied with their temperature evolutions $T_{surface}$ and T_{deep}. $T_{surface}$ is equivalent to ΔT in other chapters, γ is the heat exchange coefficient between the two layers. ϵ is the ocean heat uptake efficacy – here, as usually – assumed to be a constant. Chapter 6 discusses the option of expressing the ocean heat uptake efficacy as time dependent (as is also implied in chapter 3).

Fig. 5.2 shows that almost all simulations equilibrate warmer than predicted from the effective sensitivity using the first 150 years (up to 38%, left panel). The two-layer energy balance model including efficacy (right panel) points into the right direction, but still equilibrates below the actual equilibration temperature. Fig. 5.4 - **??** show individual models in each panel with the two-layer energy balance model without (yellow) and including (blue) ocean heat uptake efficacy.

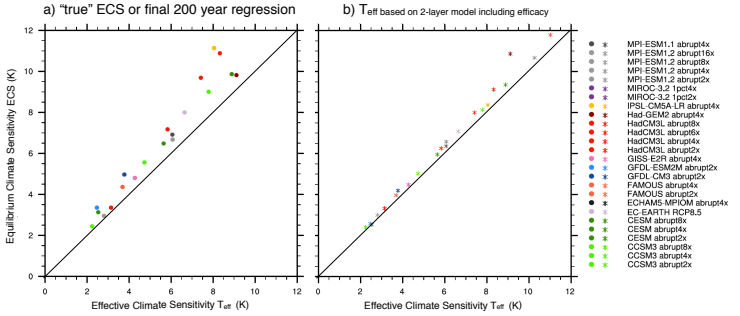

Figure 5.2: *Relation of effective climate sensitivity, calculated based on the linear regression of the first 150 years after the application of the forcing, and the x-axis intersect of the equilibrated simulation, or the linear regression of its last 200 years, if the simulation did not equilibrate in surface temperature. Thus, for these simulations the true value will be likely higher. MPIESM1.2 abrupt8x and abrupt16x and FAMOUS abrupt4x are outside the plotted range but all have a 30-40% higher ECS than T_{eff}. Climate sensitivity is calculated for the forcing level of the simulations (thus, 2x-16xCO_2). To compare it with the standard definition for doubling CO_2 one would have to scale the number down.*

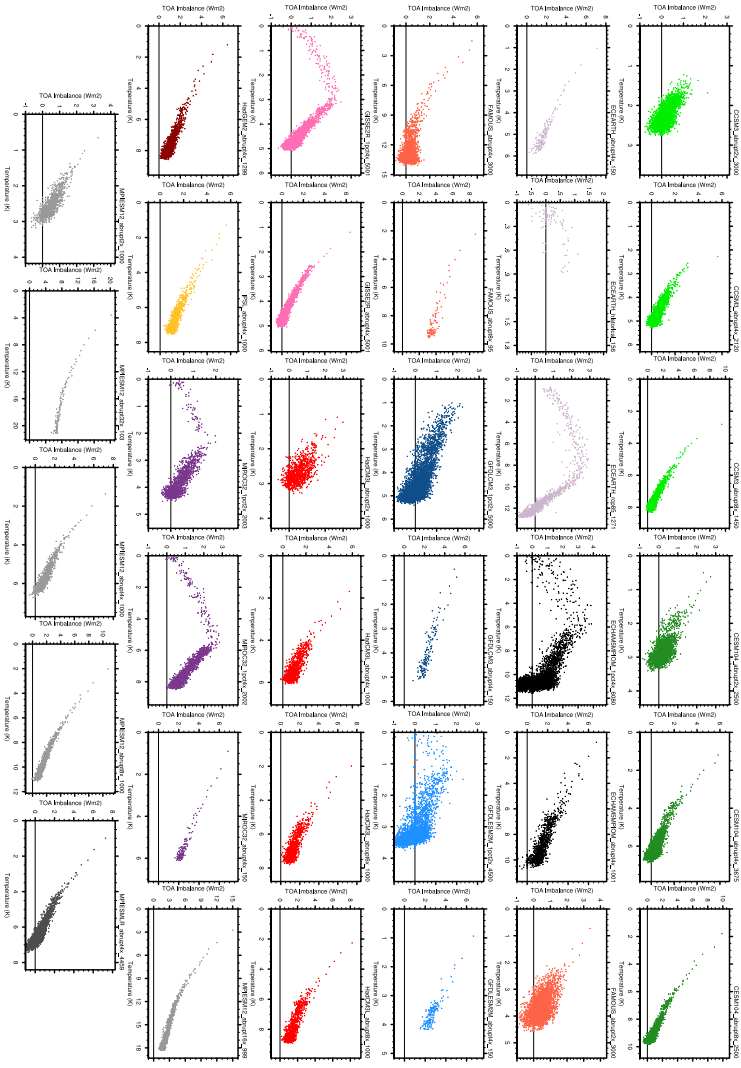

Figure 5.3: *Evolution of temperature and TOA imbalance for all simulations taking part in LongRunMIP. Note the different ranges of x and y-axes.*

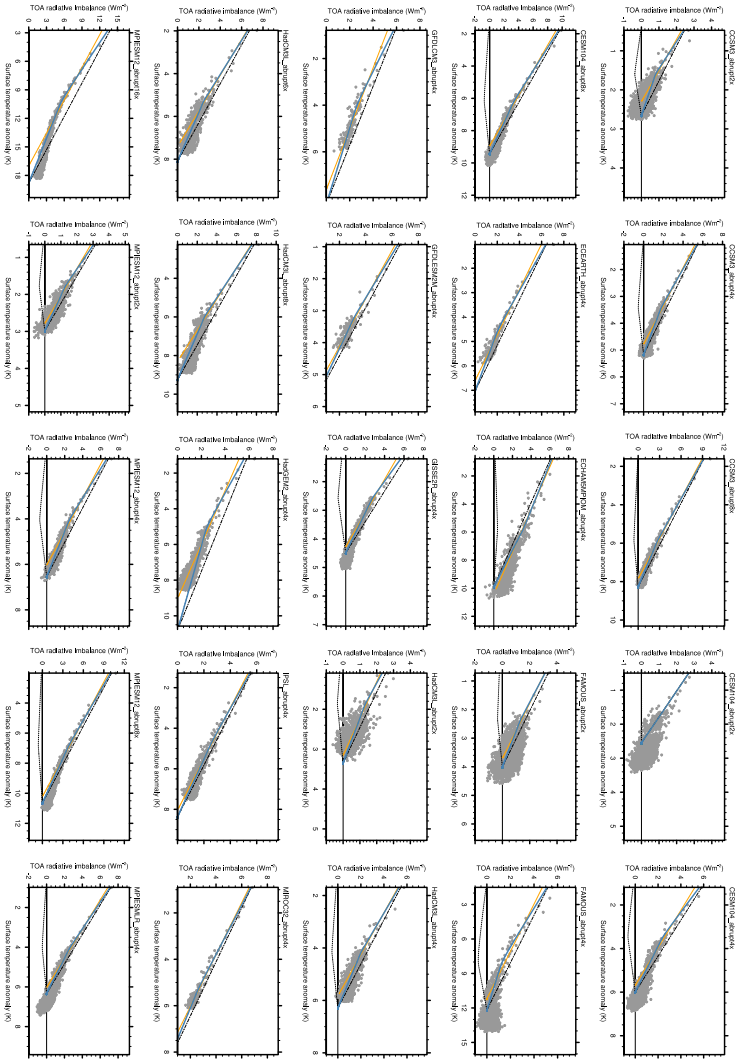

Figure 5.4: *Evolution of temperature and TOA imbalance, annual average (gray dots). Linear regression of the first 150 years (yellow line, following Gregory et al. (2004)), and the two layer model including ocean heat uptake efficacy (blue, following Geoffroy et al. (2013a), with the linear and nonlinear components in dashed black). Note the different ranges of x and y-axes.*

5.3 Estimating the feedback evolution

Chapter 1-3 uses the moving bin regression to estimate the continuously changing feedback magnitude during the equilibration. With a large ensemble this method works best, since otherwise the number of annual means to regress differs vastly in each bin. Fig. 5.5 shows one realization of this method on the LongRunMIP models. The bin width (in temperature space) is adjusted for each model. The exact time evolution critically depends on the bin width. What was apparent in the temperature–TOA evolution figures, becomes clearer in this way of depicting the same data, but against time instead of temperature.

The global feedback parameter does not only change during the very first decades, as suggested by the two-layer energy balance model including ocean heat uptake efficacy, but rather continuously for at least 200 years and, in many models, beyond 1000 years (e.g., IPSL in yellow, GISS in pink, FAMOUS in orange-red, or one HadCM3L realization in red). This implies that when using the conventional method, one would have to regress the last years of for example a 200 year simulation.

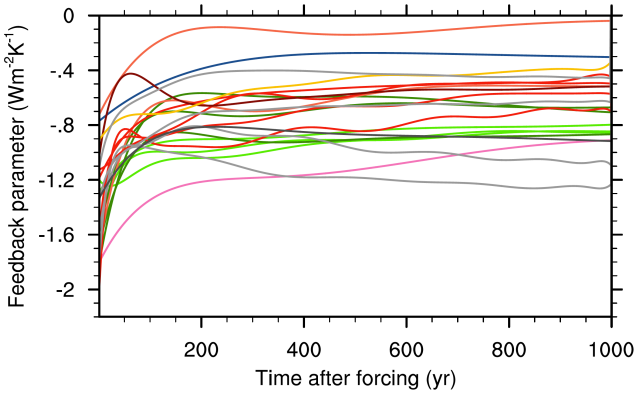

Figure 5.5: *Evolution of the feedback parameter estimated by the moving window regression. Same color coding as in Fig. 5.2.*

The leading hypothesis to explain the non-constancy of the global feedback parameter is the short wave cloud response to changing surface temperature and heat flux patterns (chapter 3, chapter 6.3, Winton et al., 2010; Rose et al., 2014; Kang and Xie, 2014; Andrews et al., 2015).

Fig. 5.6 shows the pattern evolution of the models in a highly condensed way:

Panel a shows the model mean surface temperature response of the first decade after the application of the step forcing. Each model grid point is first divided by the model's global mean warming to account for the different scenarios (of abrupt2x to abrupt16xCO_2 forcing). The right column shows the standard deviation of the models. Appendix Fig. C.1 - C.9 show each individual model. Common to all models is an Arctic Amplification in the first decade, to various degrees over the land masses (standard deviation locally about half of the signal). Southern high latitudes either warm almost as strongly as the Northern (FAMOUS), barely warm (GFDL ESM2M), or even cool compared to the control simulation (ECHAM, and CESM abrupt2x), see Ap-

pendix Figures. Notably, the patterns differ a lot between simulations of one model (CCSM, CESM, and MPIESM1.2), with a tendency to warm the Southern Hemisphere more with higher forcing levels. In the Equatorial Pacific (see discussion in chapter 2) the response differs vastly as well: from La Niña like cooling (CESM) to El Niño like warming (HadCM3L, MPIESM1.2, CCSM, and GISS). This indicates that abrupt step forcing simulations might not easily be comparable to the historic time frame (Gregory et al., 2016; Zhou et al., 2016).

Panels c and d show the same fields for the 7th decade after the step forcing, while panels e and f show the difference between the 7th and the 1st decade (i.e. panel e = c – a and panel f = d – b). In this decadal timescale, in most models the Southern Hemispheric high latitudes catch up with the warming of the Northern Hemispheric high latitudes, which is cooler than the global mean. Again, the degree of Antarctic Amplification differs vastly between the models and between simulations of the same model (notably CESM and MPIESM1.2; models higher forced scenarios show a smaller warming lag in the Southern Hemisphere). In the Equatorial Pacific the response is opposite to the first decade. The decadal time scale could be described as "equalizing" the temperature response of the first decade (HadCM3L and CCSM). However, some models show barely a structure in the warming (both GFDL models and FAMOUS), only obvious in the Appendix Figures.)

Finally panels g and h show the difference between the 100th decade (pattern not shown) and the 7th decade, i.e. the centennial time scales. The models do still show a strong evolution of temperature patterns (roughly equally strong in the years 70-140 than in year 140 to 1000, not shown). Especially the lower CO_2 forced models (abrupt2x of CCSM, CESM, HadCM3L, MPIESM) still "fulfill" a large part of their fraction of equilibration (see discussion around 1.3). Although most models show Antarctic amplification and more Southern than Northern Hemispheric warming, the models show a range of responses in e.g., the Equatorial Pacific, North Atlantic, or South Pacific (subtle in panel h). In the outlook (chapter 6) I further discuss the potential influence of time varying sea surface temperature and heat flux patterns on clouds and possible ways to study their relationship.

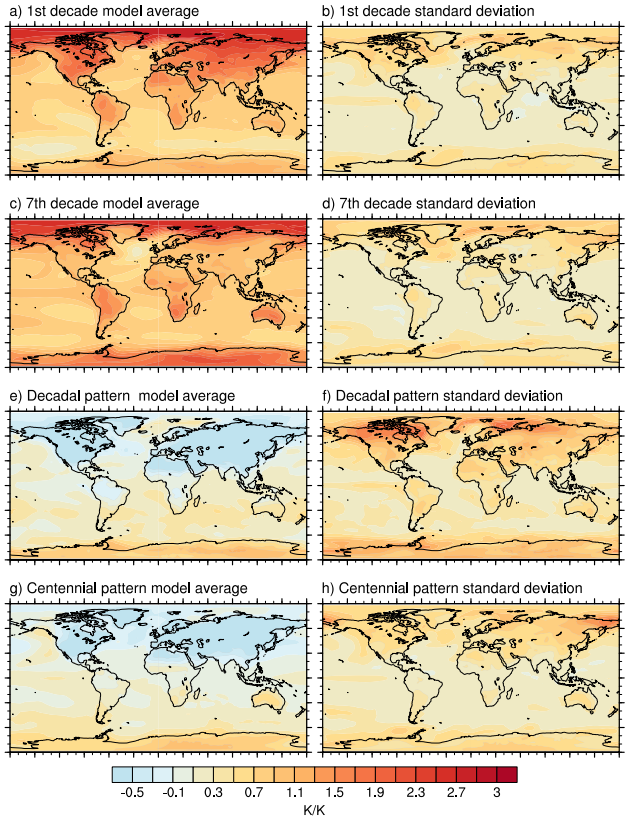

Figure 5.6: *Agreement and discrepancy between the model's surface temperature responses at different time scales. See text for explanation.*

5.4 Scaling the response to different forcing levels

It is very common to simply divide an abrupt4xCO_2 scenario or its equilibrium temperature anomaly by two to get the nominal climate sensitivity, defined as the equilibrium temperature anomaly for doubling CO_2. This approach assumes that the temperature dependence of feedbacks is negligible, i.e., feedbacks acting in a warm world of for example 5° global warming are the same as in a 2.5° warmer world. However, it has been shown that this scaling does not hold for all models (Boer et al., 2005; Jonko et al., 2013) and regional responses in many models (Good et al., 2015). Fig. 5.7 shows for all models, which contributed more than one forcing scenario, the scaled up abrupt2x

simulation and scaled down abrupt8x or abrupt16x simulation. If the forcing of CO_2 were logarithmic, curves for each model would align. The scaling assumption works reasonably well for some models (MIROC and HadCM3L on decadal time scales), but fails for higher forcing scenarios and warmer states (closer to the equilibration). As shown in the previous section, this might be due to the surface warming pattern, which differs between scenarios for one model. Temperature and state dependent feedbacks might be other reasons for the discrepancy (see discussion in chapter 6.3).

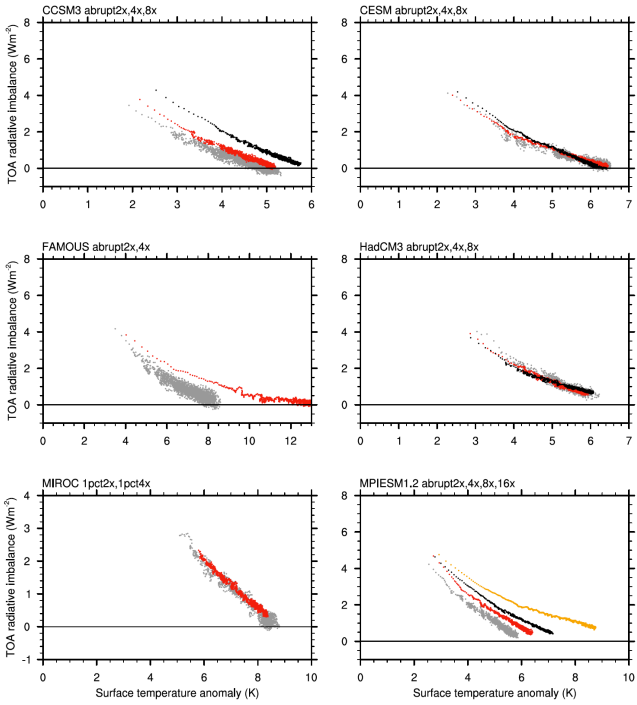

Figure 5.7: *Linearity with forcing levels: Model submissions with several forcing levels scaled according to their estimated forcing (intercept of linear regression of the first three years) relative to the abrupt4xCO_2 simulation.*

5.5 The Meridional Overturning Circulation hypothesis

Fig. 5.8 depicts the global ocean Meridional Overturning Circulation (MOC) anomaly through time. As in Fig. 4.3 shown for a model of intermediate complexity, the response can depend strongly on the forcing level (MIROC and FAMOUS do permanently reduce or re-strengthen their MOC). The time scale of the MOC minimum lays consistently between 70 and 100 years after the forcing is applied. In CESM the time scale of Atlantic MOC recovery coincides exactly with the time of slowdown of the global

feedback parameter reduction (not shown). In other models it has been shown that the initial reduction of the Atlantic MOC can have a strong effect on local cloud feedbacks, reflecting non-negligibly in the global feedback parameter (Zhang et al., 2010; Winton et al., 2013; Trossman et al., 2016; Rose and Rencurrel, 2016). Since the MOC strength also sets the effective heat capacity of part of the ocean, this process could potentially link ocean heat uptake and surface heat flux induced feedback response, as discussed in chapter 6.2 and 6.3.

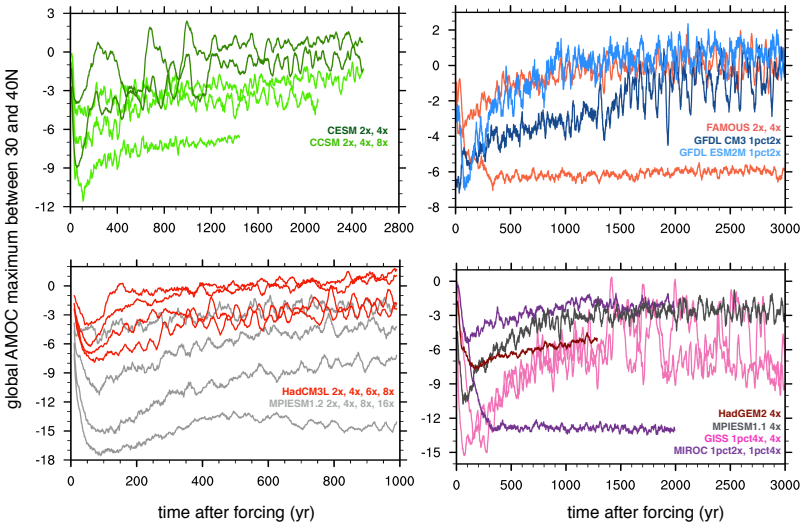

Figure 5.8: *Meridional overturning circulation anomaly (global, maximum value between 30 °and 40 °N, 20 year running average) of models available in January 2017. Note that x and y axes differ for each plot.*

5.6 Future prospects of LongRunMIP

For the publication we plan to analyze which feedback components contribute to the the nonlinear feedbacks at which time scales (see also discussion in the outlook chapter 6.2) as a first step to differentiate how much of the changing feedback parameter is due to sea ice and snow albedo feedback, water vapor, and short versus long wave cloud response. We will then test the hypotheses of (1) the influence of the surface temperature patterns, (2) temperature dependent feedbacks, (3) the time scales of MOC recovery (both in the Northern Hemisphere, dominated by the Atlantic MOC, and the Southern Hemisphere (not shown)), (4) the hemispheric temperature gradient (Senior and Mitchell, 2000), and (5) the influence of the fraction of the ocean taking up heat on the cloud response (see discussion in chapter 3, 6.2, and 6.3). Another idea is the normalize the models according to their fraction of total heat uptake and potentially be able to predict the feedback evolution better. Ultimately, the question is, how many years or which process knowledge of the transient simulations would be sufficient to predict the equilibrium response.

Many other questions could be answered with this dataset. Examples include: (1) When does the Southern Hemispheric warming fully catch up with the Northern Hemispheric warming and how are these processes linked to inter-hemispheric energy transport? (2) When and how does the Arctic become sea-ice-free and how sensitive is the sea ice and snow albedo feedback towards its extinction? (3) When do high latitude warming patterns and feedback activation become more important than tropical zonally heterogeneous feedbacks? (4) How does the ocean heat uptake saturate and are surface heat flux patterns substantially different then, compared to the control simulation? (5) What is the difference in committed sea level rise between forcing scenarios and where is the heat ultimately stored in complex ocean models? (6) Do intermediate or deep water temperature anomalies reflect surface condition anomalies? (7) Does Southern Ocean convection re-strengthen on centennial time scales and if so why? (8) Do continental landmasses get dryer or wetter in a stable warm climate and does the response scale with CO_2 forcing levels? (9) Do dominant modes of internal variability differ detectably between transient and equilibrium conditions? (10) How much do thousand year estimates of Southern Hemisphere warmth diverge from paleoclimatic estimates that presumably include the effect of a changing ice-sheet?

6 Outlook

A few conclusions unify all chapters: (1) Process understanding is key to predictability. (2) Concepts like ocean heat uptake efficacy help to phrase a problem, even if they turn out to be not the best description. They point to the open questions, force one to think about process understanding in a broader context, and question basic assumptions. (3) Simplified models require ongoing scrutiny by complex models and observations – and vice versa. The back and forth between energy balance models and GCMs shows that one without the other is incomplete.

In the following, I discuss issues which go beyond the conclusions made in the different chapters.

6.1 Usefulness of the concept of virtual forcing

Thus far, I have yet to convince others (and myself) that the concept of virtual forcing is a useful alternative to the effective radiative forcing or the fixed-SST forcing. However, there are solid reasons why this might be the case: (1) Ringer et al. (2014) shows that effective radiative forcing and short wave cloud feedback are anti-correlated, which renders the current separation invalid. Folding part of the cloud response into the forcing would solve this problem. (2) Fig. 2.3 shows that the short wave cloud radiative effect is truly non-linear, changing sign after three years and then leveling off around year 12. We show that this can be accurately described as a time dependent forcing adjustment. (3) The large ensemble allows us to use the method of the "moving bin regression" in very high resolution (Fig. 2.10) and shows that the TOA – temperature evolution is substantially different within the first five years compared to later states. We argue that there is evidence that oceanic adjustments induce changing SST patterns which influence the global cloud response. This is fundamentally different from SST pattern formation later in the equilibration time, and thus, could be argued is part of the response to the forcing and not the global mean temperature increase. (4) Along the same lines, not only surface heat fluxes, wind pattern, and land heat uptake, but also precipitation and evaporation require several years before they increase with global mean temperature. (5) Models show different degrees of adjustments, which would be normalized by folding adjustments into the forcing term, based on process understanding opposed to methodological brute force definitions, such as fixed-SST forcing. In other words, using effective radiative forcing and fixed-SST forcing leads to an apparent time dependence of the feedbacks, which has to be accounted for. It is unclear whether the multi-annual time scale we found is only a modeling problem: In transient simulations of more realistic scenarios, although tropospheric adjustment occurs, the internal variability might be so large that for example a reduction in the Atlantic meridional overturning circulation might not be detectable.

The problem remains to quantify the "virtual forcing", but trying to understand the

processes behind it might help to more carefully or cautiously interpret feedbacks and forcings.

6.2 Usefulness of the concept of ocean heat uptake efficacy

Chapter 3 and 5 demonstrated that ocean heat uptake efficacy is very likely not a constant, but time dependent (see also Paynter and Frölicher, 2015). A constant ocean heat uptake efficacy increases the fit of temperature and TOA evolution (Winton et al., 2010; Geoffroy et al., 2013a,b), but not necessarily beyond the time used for the fit (chapter 5). Estimating the ocean heat uptake efficacy based on for example 100 versus 200 years of abrupt4xCO_2 simulations yields a different value (not shown). Other work has stressed the sensitivity of the global feedback parameter on sea surface temperature pattern for the historical period (Gregory et al., 2016; Zhou et al., 2016) or idealized equilibration time scales (Andrews et al., 2015) without referring to the ocean heat uptake efficacy concept, but implying the same mechanism: patterns of surface temperature or heat fluxes change radiative feedbacks, dominated by the cloud response.

The central question remains: is the best description of general temperature and TOA evolution achieved with a time or process dependent feedback parameter (λ(t)) and no ocean heat uptake efficacy ($\epsilon = 1$), a constant feedback parameter (λ) and time dependent ocean heat uptake efficacy (ϵ(t)), or both a time dependent feedback parameter (λ(t)) and a time dependent ocean heat uptake efficacy (ϵ(t)). In the notation of the energy balance equation (with the same notation as in chapter 1)

$$\epsilon(\mathrm{t})\mathrm{N}(\mathrm{t}) = \mathrm{F}(\mathrm{t}) - \lambda(t)\Delta\mathrm{T}(\mathrm{t}) \tag{6.1}$$

Note that "it is simply a matter of convenience to attach ϵ as a factor to N" (Winton et al., 2010), but the two layer model formulation attaches ϵ more intuitively to the heat exchange between the surface ($\mathrm{T_{surf}}$) and deep ($\mathrm{T_{deep}}$) box

$$\mathrm{N_{total}} = \mathrm{F} - \lambda(\mathrm{t})\mathrm{T_{surf}} - (\epsilon(\mathrm{t}) - 1)\gamma(\mathrm{T_{surf}} - \mathrm{T_{deep}})$$

Next to the time dependence, there are other arguments against using the concept. The wording is somewhat unfortunate that it suggest that the ocean heat uptake itself can be more or less effective, while physically, all atmospheric processes react to the surface conditions, but are locally indifferent about whether a local heat flux results in a net global ocean heat uptake, or whether the ocean releases the same amount of heat elsewhere. This is why in chapter 3 we use the wording "ocean heat flux induced surface temperature or heat flux pattern change", i.e., surface heat flux does not imply a net ocean heat uptake. The ocean heat uptake efficacy can be larger than one while the ocean does not take up any heat, which is counter-intuitive. Andrews et al. (2015) shows in their Fig.10a surface temperature response (in this case an adjustment to the application of the forcing) without any global warming response, i.e., with zero ocean heat uptake. A future study could answer the question: What are the SST patterns that result in zero ocean heat uptake but differ strongly in atmospheric radiative feedbacks?

In reality, the case of zero heat uptake but strongly changing SST pattern is an exception, i.e., the ocean and its heat uptake *do* influence the SST patterns. The idea behind ocean heat uptake efficacy is – in my view – a valid one. The concept emphasizes that the ocean's influence on the atmosphere comprises not only the amount of heat taken out of the atmosphere, but also the indirect effect of influencing the radiative feedbacks.[1]

[1] "The implication of our simple model interpretation is that one would be more effective reducing AOGCM uncertainties in transient climate sensitivity by reducing uncertainty in the radiative response to ocean heat uptake than in the relationship of the uptake magnitude to the surface climate perturbation." (Winton et al., 2010)

If the concept of ocean heat uptake efficacy was to be dropped, it would be useful to differentiate how much of $\lambda(t)$ is due to the ocean base state, the changing circulation, and the net heat uptake, versus solely due to the atmospheric influence (e.g., Ma and Xie, 2013; Long et al., 2014, and see discussion below).

6.3 Understanding the pattern effect

The influence of SST patterns on radiative feedbacks – in the absence of global warming – has recently been termed "pattern effect" (Stevens et al., 2016), although the concept itself has been discussed for some years (e.g., chapter 3, Senior and Mitchell, 2000; Armour et al., 2013; Andrews et al., 2015). It could be described as a generalized version of ocean heat uptake efficacy according to the interpretation discussed above. Zhou et al. (2016) introduces the following notation:

$$\Delta R_{cloud} = \lambda_c \Delta T + \Delta R_{PSST} + \Delta R_{cf} + \epsilon \tag{6.2}$$

Where ΔR_{cloud} is the total cloud response, which is the sum of the cloud feedback under uniform warming ($\lambda_c \Delta T$), the cloud induced radiation anomaly in response to the changing SST pattern only (ΔR_{PSST}), the rapid cloud adjustment (ΔR_{cf}), and an error term (ϵ, not the ocean heat uptake efficacy).

As a sideline, we discuss in chapter 2 why ΔR_{cf} should be quantified based on several years. If one fails to account for the multi-annual adjustment effect, one might mistake the time evolution of the cloud feedback as pattern effect.

Eq. 6.2 only covers the cloud response, while the pattern effect could also act on other feedbacks. For example sea ice might be a dominating part of the response to different heat flux patterns (chapter 3). For the following discussion I concentrate on the cloud response, but come back to other feedbacks below.

Spatial features causing the pattern effect

The tropical Pacific has probably received the most attention – for the historical period (Zhou et al., 2016; Gregory et al., 2016) and related to the aim of constraining climate sensitivity by observations (e.g., Sherwood et al., 2014b; Su et al., 2014; Qu et al., 2014; Tian, 2015; Tan et al., 2016). Others argue that the forced response will be dominated by the high latitudes (e.g., Winton et al., 2010; Armour et al., 2013; Rose et al., 2014). Chapter 3 shows that more equilibrated surface heat flux patterns are globally heterogeneous and different regions could potentially play a role next to the high latitudes (see also heat flux patterns in Kravitz et al. (2013) or Marshall et al. (2014)). The interaction of the North Atlantic and its cloud layer has been shown to be strong enough to influence global feedback parameter (Zhang et al., 2010; Trossman et al., 2016; Rose and Rencurrel, 2016), while the interaction of the North Pacific and its cloud layer has been mostly overlooked even though its spatial extent would allow for a large impact (Praetorius et al., in prep). The Southern Ocean – cloud interaction are potentially the most important contributor to forced pattern effect (Senior and Mitchell, 2000; Armour et al., 2013; Grise and Polvani, 2014a), but its relative impact compared to the North Atlantic is unclear. A pressing research focus would be to understand which features contribute (efficiently) to the pattern effect (see a first sensitivity discussion around Fig. A.9 and A.10). What is the critical size of a region to influence the magnitude of the global feedbacks? Are we sure that the efficient features would also occur in reality and are not modeling artifacts? Vice versa, do the coupled models reproduce the real world patterns (Zhou et al., 2016; Huber and Knutti, 2014; Marotzke and Forster, 2015)? Is the pattern effect strong enough to hamper comparisons between different climate states, for example with the Last Glacial Maximum or the Miocene (e.g., Rohling et al., 2012;

Kutzbach et al., 2013)? How much of the spread in cloud responses in the CMIP5 model suite is due to the difference in surface temperature patterns – in constrast to the cloud schemes themselves? As far as I know this is not known for the historical period or for idealized and RCP-type scenarios. Fig. 6.1 shows that in idealized scenarios the spatial pattern of surface heat fluxes is large.

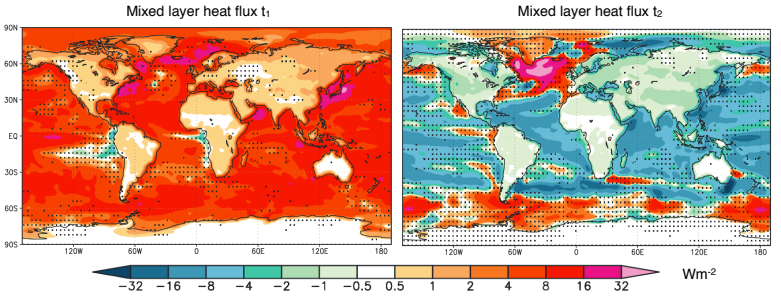

Figure 6.1: *Surface heat fluxes from CMIP5 models at two time slices of the abrupt4xCO_2 simulations (first year (left), difference between year 50 and year 11 (right)). Stippling indicates where fewer than 8 of 11 models agree on the sign of change. Adapted from Kravitz et al. (2013).*

Internal variability, forced response, and frequency dependence

The natural follow up question is: When and on which time scale do the efficient features of surface conditions appear in coupled GCMs and the real world? It is unclear whether the historical pattern is dominated by internal variability (Mauritsen, 2016), which effect aerosol and volcanic forcing play in setting the pattern effect, and what the time of emergence of forced patterns would be (Serreze et al., 2009; Hawkins and Sutton, 2012; Hawkins et al., 2015). For the influence of the ocean on SST patterns, it is unclear how important the base state circulation (e.g., Armour et al., 2016), the forced response (e.g., Zhang et al., 2010; Rugenstein et al., 2013), local air-sea interaction (e.g., Ma and Xie, 2013; Long et al., 2014), or model biases (such as the upwelling zones, open ocean convection, ENSO frequencies) are.

The pattern dependence could also be expressed in a temporal frequency dependence (Hasselmann, 1976; MacMynowski et al., 2011), and first tests show that this is indeed the case for unforced control simulations (Fig. 6.2).

Connected to the forced response is also how – towards equilibrium – the forced response levels off and internal variability becomes dominant again. This might be important to climate sensitivity estimates. Subtle changes in wind stress or ocean circulation might have a small direct influence in terms of net heat uptake but a large indirect impact on surface temperature, through modifying feedback magnitudes.

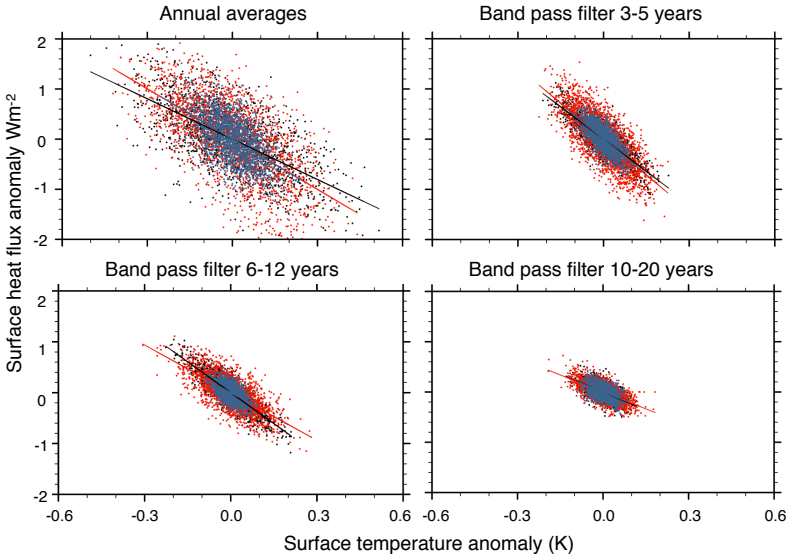

Figure 6.2: *Control simulations of three different model versions of CESM (blue, red, black, each simulating more than 2000 years) are bandwidth selected for different frequencies. The lower the frequency the smaller the global feedbacks (decreasing more than 50% for lower frequencies compared with the annual values).*

Connection between heat storage and pattern effect

Although the atmosphere reacts to the sea surface temperature and heat fluxes, there might be a connection between global net ocean heat uptake and radiative feedback evolution, thus $\Delta R_{PSST}(\lambda_c \Delta T)$ in Eq. 6.2, or in other words $\lambda_c \Delta T$ and ΔR_{PSST} can not diverge endlessly as they do since 1980 (Zhou et al., 2016, Fig.2a). Potential connections are

- The Pacific, where clouds are sensitive to SST pattern but it is also capable of taking up a lot of heat (Trenberth and Fasullo, 2013; Yan et al., 2016).
- The North Atlantic and the reduction and re-strengthening of the AMOC (Zhang et al., 2010; Rugenstein et al., 2013; Kostov et al., 2014; Trossman et al., 2016).
- The Southern Ocean heat uptake – cloud interaction is not well understood (Senior and Mitchell, 2000; Armour et al., 2013, 2016).
- The aforementioned saturation of heat uptake could relate to the pattern effect, as for example a more heterogeneous surface heat flux pattern might influence radiative feedbacks less efficiently (Fig. A.9, or Rose and Rencurrel, 2016).

Alternatives[2] to bringing the ocean heat uptake efficacy back into the discussion could be

$$N = F - (\lambda_c \Delta T + \lambda_1 \Delta T_1 + \lambda_2 \Delta T_2 + ...),$$

[2] I acknowledge a discussion with Dan Murphy on this topic.

where $\lambda_1 \Delta T_1$ and $\lambda_2 \Delta T_2$ are dominating patterns of forced SST responses such as a reduced equator to pole temperature gradient or

$$N = F - \lambda_c \Delta T - \lambda' \frac{\delta T}{\delta t},$$

with $\lambda' \frac{\delta T}{\delta t}$ being a term that expresses the degree of SST pattern change through equilibration time.

Limitations of the pattern effect

The overall role of the pattern effect in changing the feedback magnitudes through equilibration time is unclear. Chapter 3 suggests it dominates; however, several other processes are known to influence different feedbacks.

The cancellation of water vapor and lapse rate feedback can fail in warmer climates, where the water vapor feedback is thought to become dominant, implying the overall feedback magnitude becomes more positive (Meraner et al., 2013; Caballero and Huber, 2013). A decrease in feedback magnitude is expected from diminishing sea ice and snow albedo feedback (Boer and Yu, 2003a; Manabe and Bryan, 1985; Colman and McAvaney, 2009; Kutzbach et al., 2013). The role of short and long wave cloud forcing and the Plank feedback is unclear but likely positive (Boer and Yu, 2003a; Jonko et al., 2013; Caballero and Huber, 2013).

The North Atlantic and Southern Ocean show a minimum of the ratio of transient to equilibrium response (e.g., Manabe et al., 1991; Armour et al., 2013), and thus become dominant later in the equilibration process, while likely activating more positive feedbacks. While this could be described as a pattern effect, it is intrinsically connected to the overall warming itself and the distance to equilibrium conditions, thus $\Delta R_{PSST}(\lambda_c \Delta T)$.

Similarly, the SST pattern itself can be influenced by clouds and it is hard to differentiate a state dependence from the pattern effect. For example, a narrowing of the Hadley circulation cell and a poleward shift of the mid latitude westerlies due to overall warming can influence the cloud response as much as the surface temperature pattern (Caballero and Huber, 2013; Grise and Polvani, 2014a).

In summary, a research priority should be to understand how GCMs produce surface temperature and heat flux patterns, which features influence radiative feedback efficiently, and – critically – how ΔR_{PSST} and $\lambda_c \Delta T$ depend on each other.

Appendices

A Appendix for Chapter 3

Generation of the Q-flux forcing files

We use the routine `pop_frc.csh` available in the standard CESM download, which uses monthly oceanic fields of temperature, salinity, surface velocities, boundary layer depth, surface heat flux, melting and freezing fluxes to generate the Q-flux for that specific mixed layer depth. The mixed layer is an annual average and does not change monthly to easier balance the annual cycle. The Q-fluxes aim to be consistent with a specific state in the coupled simulation, not observations. The annual net Q-flux is a small residual of a strong seasonal cycle, which follows the surface heat flux with a smaller amplitude.

The mixed layer depth is kept at the annual average spatial variable climatological control simulation state for all experiments. This is strictly inconsistent with a changing Q-flux, but the depth of the mixed layer sensitively influences the atmospheric response (Donohoe et al. 2014). Since we are mostly concerned with the heat flux pattern, we keep the setup simple, i.e. fix the mixed layer depth and only vary the Q-fluxes, at the expense of realistically reproducing the exact transient state. Fig. A.1 shows the control mixed layer depth and small change after 1000 years.

The anomalous idealized Q-flux patterns are constructed in the following way: For the cases which globally averaged to 1 Wm^{-2} radiative imbalance a local maximum of 5 Wm^{-2} at 50°N and S (for the high latitude pattern) and 4.5 Wm^{-2} at the equator (for the low latitude pattern) was set and declines to 0 Wm^{-2} at 30 and 70°N and S (for the high latitude case) versus 30°N and S (for the low latitude case). Manuscript Figure 4 shows this pattern – due to the topography it leads to a non-symmetrical pattern around the equator. The maximum is tuned to lead to the global desired heat flux.

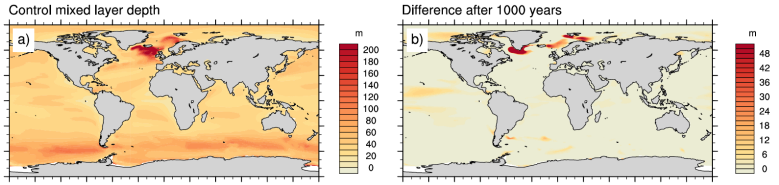

Figure A.1: *a) Control state annual mean mixed layer depth from the coupled control run used to generate the climatological Q-flux, b) change in mixed layer depth after 1000 years of a coupled abrupt 4xCO_2 simulation.*

Reference states

We chose a warm reference state four times pre-industrial CO2 concentration since we want to analyze the influence of OHU in a warming world. Rose et al. 2014 show that the positive CO_2 forcing and the negative OHU forcing add linearly. We chose to apply the OHU forcing to a warm state since in the slab ocean setup the sea ice is very sensitive to the local Q-flux. If the mixed layer depth is very shallow and the the Q-flux large, the atmosphere cannot sustain warm temperatures and sea ice grows, even in the tropics. Once a few grid cells are filled with sea ice, clouds form above and lead to more cooling and quickly to either a full snowball Earth or sea ice up to the mid latitudes, depending on the geographical pattern of the Q-flux. Though choosing a warm world as reference state we circumvent this problem (for the idealized cases). For the realistic cases we had to inhibit the sea ice growth anyways, since it was locally too high, leading to excessive sea ice growth even under warm conditions. To test the linearity assumption and whether our argument would hold for other reference states, Fig. A.2 shows the homogeneous case run under 2xCO_2 conditions (without sea ice). The feedback is still linear (each dot is again one equilibrated slab ocean simulation) and very close to the 4xCO_2 case.

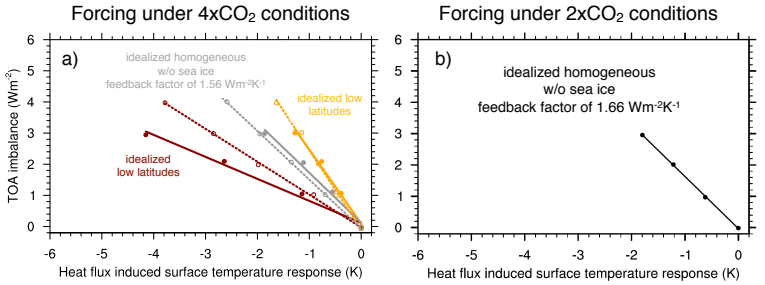

Figure A.2: *This figure indicates that our qualitative results would hold in a colder world. In the main paper the reference state is four times pre-industrial CO_2 concentration (left panel). Here we test for a reference state of two times pre-industrial CO_2 concentration, and apply again the Q-fluxes as negative forcing (right panel). The magnitude of the feedback parameter is approximately the same.*

Inhibiting sea ice growth

As described above, the sea ice is very sensitive to the Q-flux forcing. The slab ocean model is usually used with preindustrial control Q-flux averaging globally to zero. We perturb the Q-flux locally more than $50\,\mathrm{Wm}^{-2}$, especially in the realistic case. To get a reasonable sea ice distribution not covering more area than today, one can heavily smooth the Q-flux field deduced from the coupled runs to make it less patchy, mask out the sea ice areas (no change in grid boxes which can be covered by sea ice), or run under extremely high CO_2 forcings. We decided to artificially inhibit sea ice growth to get around this technical problem. To do so in CESM, we set the freezing temperature of salt water (SHR_CONST_TKFRZSW) in `shr_const_mod.F90` to some very low number. Fig. A.3 and Fig. A.4 below show the difference between cases with and without sea ice. In general, an ice free world is warmer, i.e. a positive perturbation increasing the temperature, will be reached faster. A dense cloud cover is forming above the supercooled water, and changes less with warming than the cases including sea ice, which have large changes over the melting sea ice. Midlatidudes and tropics are not impacted by the sea ice growth inhibition (except for the overall warming).

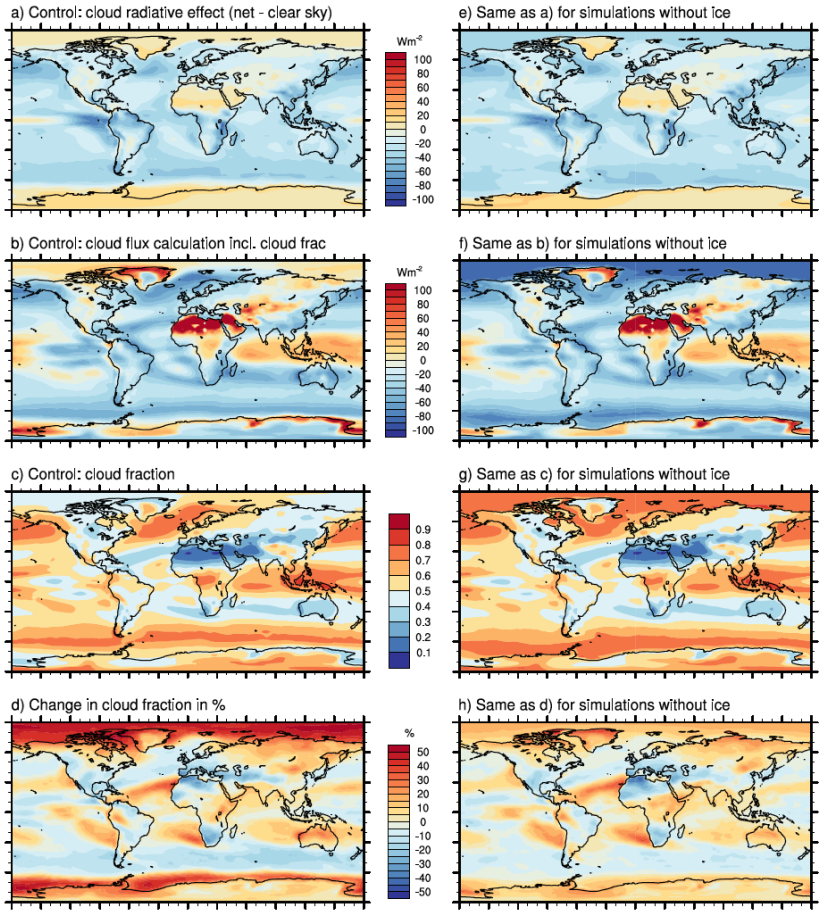

Figure A.3: *a) Control simulation CRE as used in the main body of the paper. b) Control simulation* F_{cloud} *as defined in Eq. A.1, and thus capturing some of the cloud masking issues. c) Control simulation's TOA cloud cover fraction and d) its change in a four times pre-industrial* CO_2 *world. The method in Eq. A.1 and A.2 has of course limitations where the cloud fraction is close to zero or one. e)-h) are the same quantities for the simulations inhibiting sea ice growth. Above the cold water inhibiting sea ice growth clouds are forming, but in the midlatitudes and tropics the pattern and magnitude of the cloud and cloud cover are similar to the case including the sea ice (the global temperature is about 2K warmer for the no-ice world).*

Cloud masking

We show with a first order estimation of the error of the Cloud Radiative Effect as defined in the main paper: CRE is the netTOA - clear sky TOA, for both the short and the long wave components. With this definition, processes which do not include clouds are falsely attributed the cloud forcing. A reduction in sea ice – through changing the clear sky SW component – will appear as SW CRE. Increasing CO_2 causes a stronger reduction in upwelling LW in clear sky compared to cloudy conditions, thus decreasing LW CRE, and humidity profiles do look different between the two cases (Soden et al. 2004, Soden et al. 2008, Zelinka et al. 2013). To properly quantify cloud feedbacks, a kernel analysis would be helpful, but comes with its own problems. Since the focus of our paper is not in the exact number of the cloud or clear sky feedback, we show a first order approach to the cloud masking effect here, using the information of the fraction of a grid cell covered with clouds. The whole analysis is done at the top of the atmosphere, frac means total cloud fraction, not differentiated by cloud types or height.

$$F_{total} = F_{clearssky}(1 - frac) + F_{cloud}(frac) \tag{A.1}$$

In Fig. A.3a CRE, as used in the main body of the paper, is compared to F_{cloud} (panel b). It is broadly the same pattern, except in regions where frac is very small or large. Panel c shows the control simulation's cloud fraction and panel d) the change in % for a 4xCO_2 case. The same approach gives

$$\Delta F_{total} \approx (1 - \Delta frac)\, F_{clearsky} + \Delta frac\, F_{cloud} + (1 - frac)\, \Delta F_{clearsky} + frac\, \Delta F_{cloud} \tag{A.2}$$

On basin scales the magnitudes of the two methods agree, except in the Indian Ocean and Western Tropical Pacific. We conclude that the cloud masking effect is large, but that the basin and continental scale pattern, sign, and anomalies get captured *well enough* with the simplified definition of CRE in the paper.

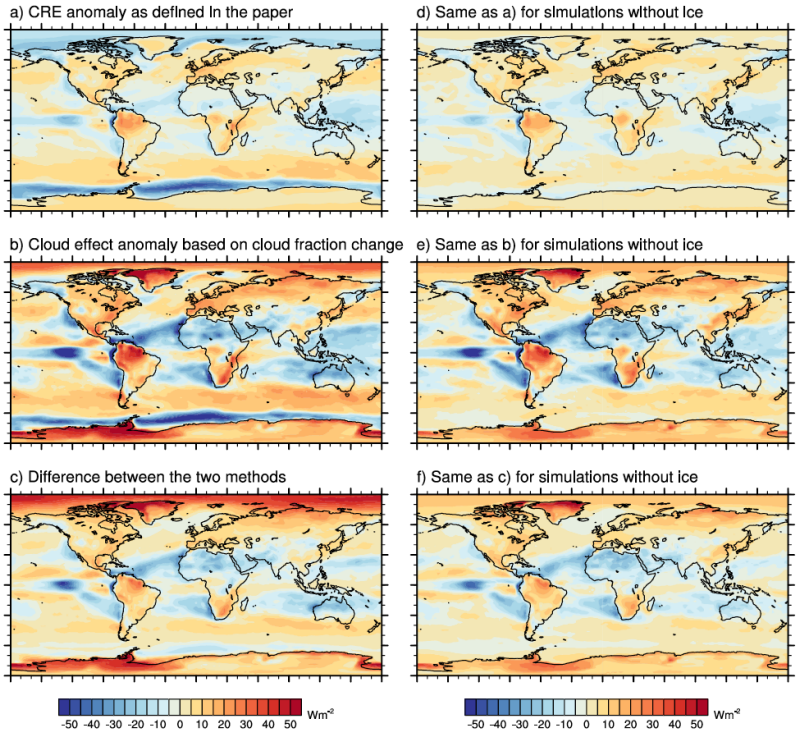

Figure A.4: *Same as Fig. A.3 a and b, but now for the change in cloud radiative effect as defined in the paper (net TOA SW minus net TOA LW) (a) and in Eq. A.2 using the cloud fraction (b) and the difference of the methods (c). In the mid latitudes and tropics the sign and order of magnitude is captured, but the strength of the feedback differs for regions with both negative and positive feedbacks. In the high latitudes above the sea ice the cloud masking effect becomes large. Panel d-f show the same quantities for the simulations without sea ice and the difference between the methods is reduced.*

Feedback components

Fig. A.6 and A.7 show that the low latitude forcing, the local TOA balance of surface heat flux anomalies occurs along all longitudes and is dominated by the mostly negative SW cloud feedback and LW clear sky feedback (and compensated by the LW cloud feedback). For the high latitude forcing, the TOA response is globally spread out and spatially variable, e.g. even positive above the melting sea ice. The tropical as well as the midlatitude SW cloud feedback becomes positive for that case.

As discussed in the paper, in most places, the difference between the cases comes about through changing the magnitude of a feedback or slightly changing its position, not through drastically changing the geographical patterns.

Table A.1: *Global feedback components (like the boxes in manuscript Fig. 3.3c) and their standard errors. All in* $Wm^{-2}K^{-1}$.

case	net	error	SW CRE	error	LW CRE	error
high lat	-0.7	0.05	0.16	0.02	0.09	0.03
high lat no ice	-1.05	0.04	0.44	0.03	0.02	0.01
low lat	-2.37	0.2	-0.68	0.17	-0.17	0.02
low lat no ice	- 2.68	0.16	-0.67	0.1	-0.19	0.08
homo	-1.6	0.13	-0.13	0.07	-0.13	0.02
homo no ice	-1.55	0.01	0.07	0.07	-0.06	0.03
1st decade	-1.32	0.03	0.36	0.07	-0.13	0.036
5th decade	-0.8	0.06	0.68	0.04	0.03	0.02
20th decade	-0.55	0.04	0.71	0.11	0.22	0.01
high – low lat	1.67		0.84		0.26	
1st – 20th dec	0.77		0.35		0.35	

case	SW cs	error	LW cs	error
high lat	0.88	0.04	-1.85	0.05
high lat no ice	0.3	0.02	-1.8	0.01
low lat	0.55	0.03	-2.07	0.02
low lat no ice	0.25	0.04	-2.07	0.07
homo	0.57	0.03	-1.91	0.01
homo no ice	0.31	0.01	-1.89	0.02
1st decade	0.32	0.005	-1.88	0.02
5th decade	0.32	0.01	-1.8	0.02
20th decade	0.29	0.01	-1.77	0.03
high – low lat	0.33		0.22	
1st – 20th dec	0.03		0.11	

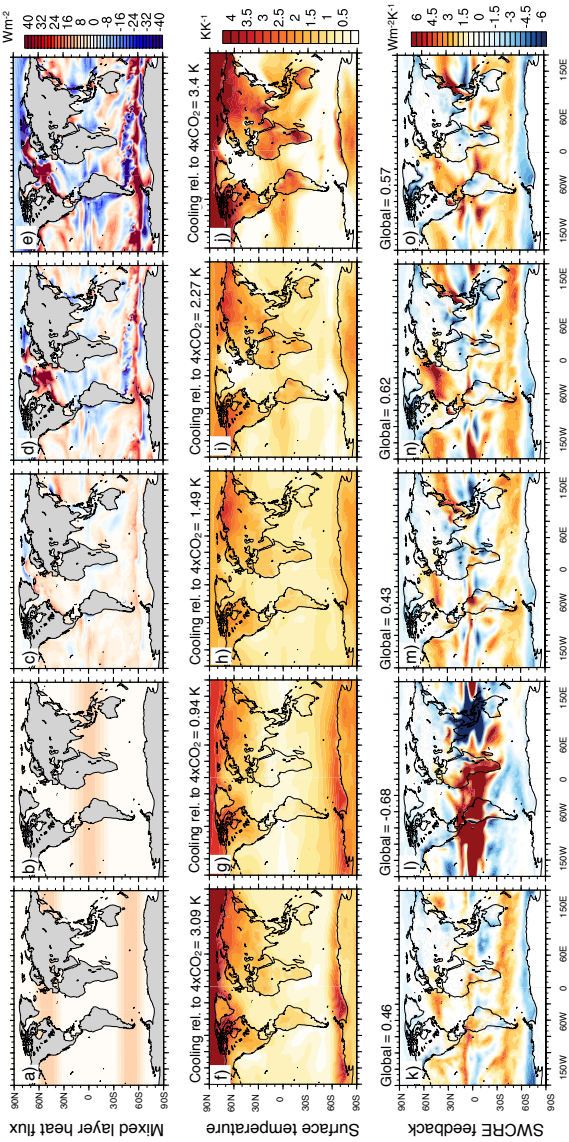

Figure A.5: *As Fig. 2 in the main paper, but cases without sea ice are shown in the first two columns. The polar amplification is higher without sea ice, probably due to their bases state in low clouds, see Fig. A.3.*

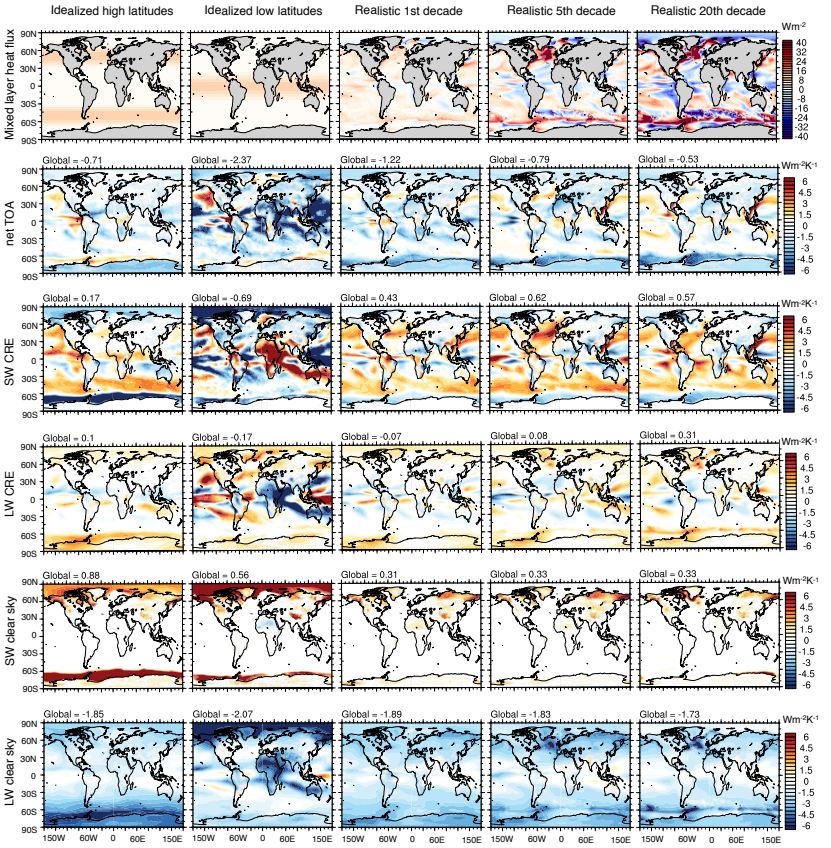

Figure A.6: *Geographically distribution of feedback parameters for all TOA feedback components, as in manuscript Fig. 2. The idealized cases (first two columns) include sea ice, which is evident in the SW components.*

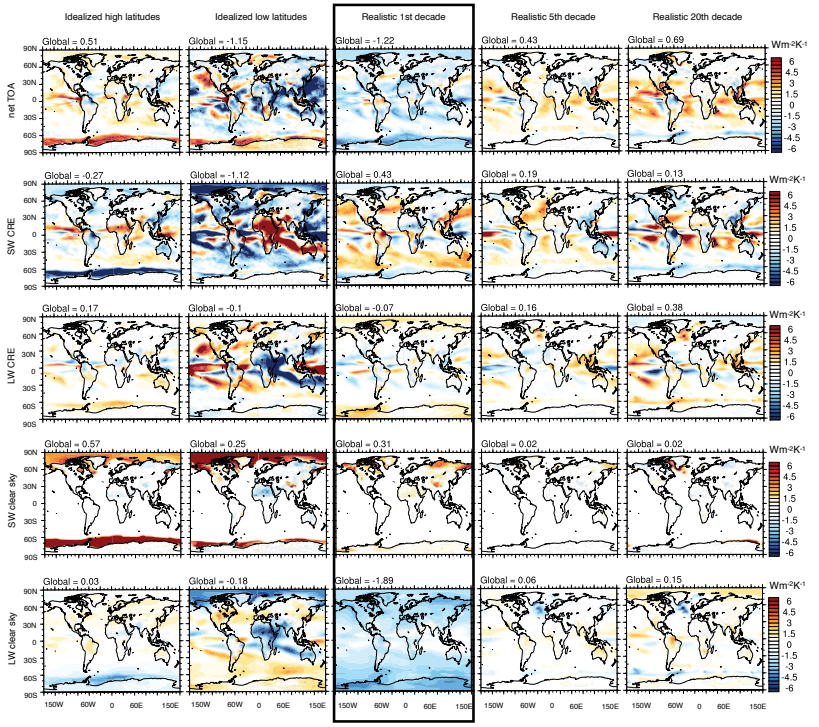

Figure A.7: *Same as in Fig. A.6, but depicted as difference to the middle column (framed in black, "Realistic 1st decade"). Same color magnitude as in Fig. A.6 for direct comparison of the feedback magnitude.*

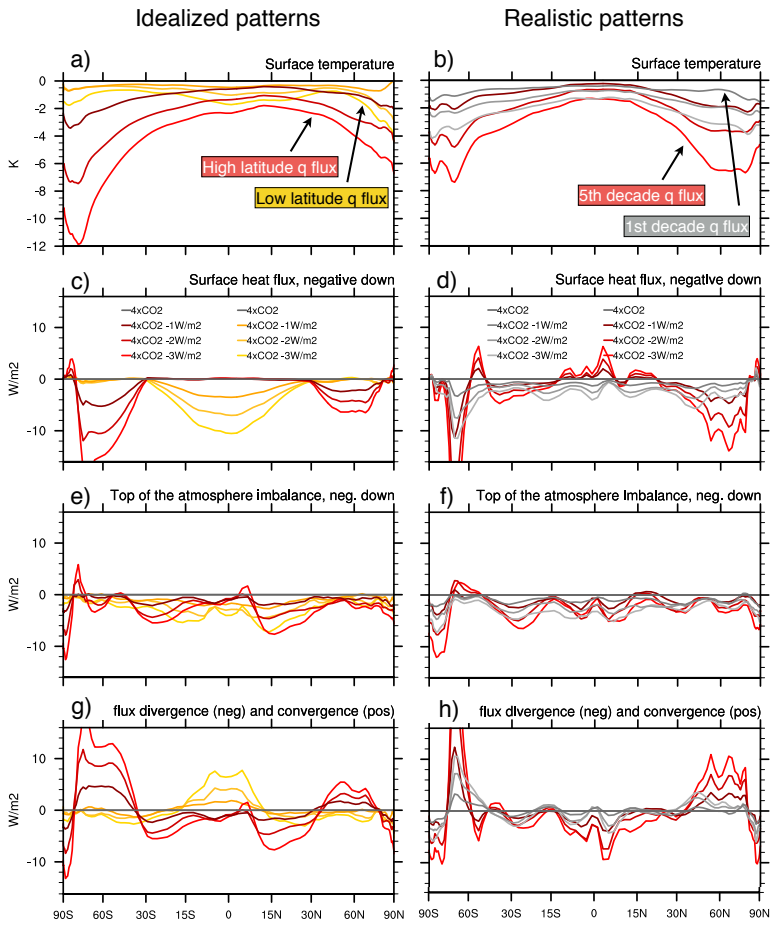

Figure A.8: *Equivalent to manuscript Fig. 4, for idealized (left) and realistic (right) cases. The "Realistic 1st decade" already shows polar amplification.*

Smooth versus heterogeneous Q-flux patterns

In Fig. A.9 and A.10 we show a short and non-exhaustive sensitivity analysis of the importance of the homo- versus heterogeneity of the Q-flux pattern. We compare four versions of two hetero- versus homogeneous patterns:
(A) "Realistic 5th decade" versus "Realistic 20th decade", shown in the main manuscript Fig. 3.2d versus e. The resulting feedbacks differ by $\Delta\lambda = 0.25\,\mathrm{Wm^{-2}K^{-1}}$ (main manuscript Fig. 3.3b red versus green). As discussed before, the two heat-flux patterns differ little at large spatial scale but differ substantially at local scale, resulting in different global feedback magnitude. This leads to the conclusion that small scale pattern and ocean heat release could be an important influence on the global feedback parameter.
(B) Again, already shown in the main manuscript, "Realistic 1st decade" versus "Idealized homogeneous" (both without sea ice) are shown in Fig. A.9 a versus b, resulting in different global feedback parameters: $\Delta\lambda = 0.23\,\mathrm{Wm^{-2}K^{-1}}$ (Fig. A.10a). Thus, in this case a subtle difference of localized maxima and minima in a rather homogeneous pattern is sufficient to change the global feedback parameter somewhat.
(C) Further, we compare "Realistic 5th decade" discussed in the main manuscript with a smoothed version of the same pattern (additional simulations not discussed in the main manuscript), shown in Fig. A.9d. The smoother version was created as sensitivity test. The global feedback parameter difference, shown in Fig. A.10b, is only $\Delta\lambda = 0.12\,\mathrm{Wm^{-2}K^{-1}}$. Thus, in this case, the small scale pattern smoothed out seem to be less important on the global scale.
(D) Finally, we test a case including active sea ice: The "Idealized high latitudes", as shown in the manuscript, and a heterogeneous version of the same pattern, shown in Fig. A.9e, lead to essential the same global feedback factor ($\Delta\lambda = 0.003\,\mathrm{Wm^{-2}K^{-1}}$, Fig. A.10c). Thus, for this pattern the heterogeneity does not significantly impact the global scale.
We conclude that the small scales and heterogeneity of the Q-flux pattern may or may not affect global feedback strength, depending on the sign, magnitude, and location of the forcing. More exhaustive simulations for single features, sea ice, positive versus negative patterns, and time dependence of the patterns are necessary to determine in which cases regional ocean heat flux heterogeneity impacts global scales (see discussion in manuscript Section 4).

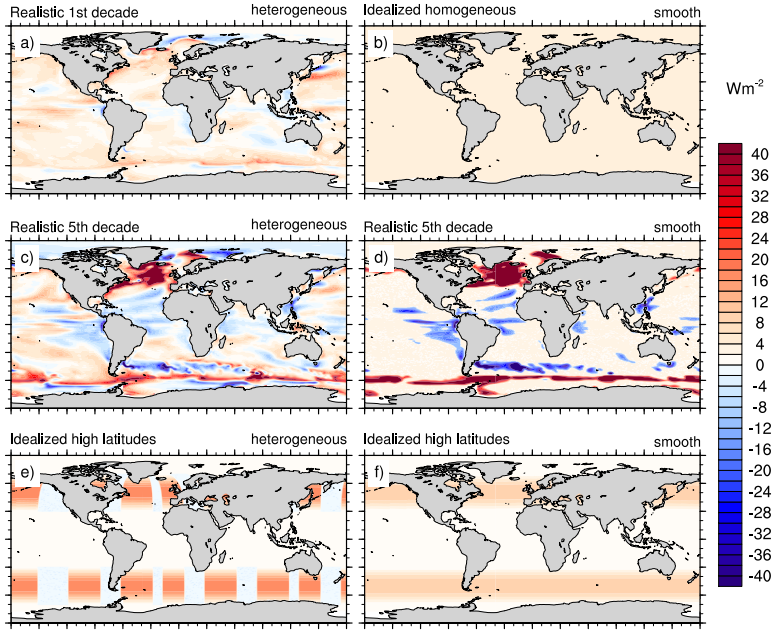

Figure A.9: *Test of whether small scales and heterogeneous patterns matter for the global climate feedback parameter. The left column shows results for heterogeneous patterns, the right column shows results for the smoother, "more homogeneous" versions of the same pattern. Panel a) and c) are identical to Fig. 2c and d in the main manuscript, whereas the heat flux patterns in panel e) and d) are created as sensitivity test. All patterns show a globally averaged ocean heat uptake of 2 Wm^{-2}.*

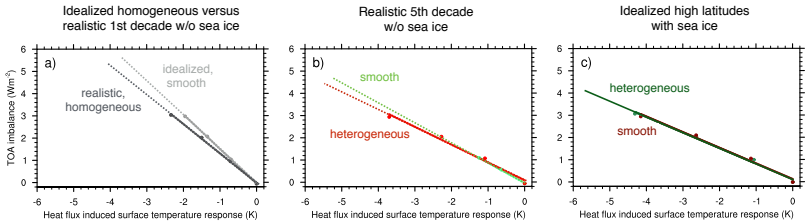

Figure A.10: *The effect of the smoothness of a pattern can, but does not have to be important. Same color coding as in the main manuscript Fig. 3.3a and b. Corresponding Q-flux anomalies are depicted in Fig. A.9. Panel a) shows the patterns shown in Fig. A.9a (dark gray, $\lambda = 1.32\,Wm^{-2}K^{-1}$) and b (light gray, $\lambda = 1.55\,Wm^{-2}K^{-1}$). Panel b) shows the patterns shown in Fig. A.9c (red, $\lambda = 0.8\,Wm^{-2}K^{-1}$) and d (light green, $\lambda = 0.92\,Wm^{-2}K^{-1}$). Panel c) shows the patterns shown in Fig. A.9e (dark green, $\lambda = 0.703\,Wm^{-2}K^{-1}$) and f (dark red, $\lambda = 0.706\,Wm^{-2}K^{-1}$).*

B | Appendix for Chapter 4

ECBilt characteristics

The ocean model has a partly rotated grid in the North Atlantic. Mixing along isopycnals, vertical mixing (based on the Mellow and Yamada model), effects of mesoscales eddies, and down-sloping currents at the bottom of the continental shelfs are parameterizied. The horizontal eddy diffusivity is $150 m^2/s$ and the horizontal eddy viscosity is $10^5 m^2/s$. The vertical diffusivity is set to $10 m^2/s$ when the vertical density profile becomes unstable. Momentum, heat and freshwater fluxes do couple. Tropical trade winds are very weak, and not enough moisture from Atlantic to Pacfic requires a redirection of snow and rainfall over the Atlantic to the North Pacific. In the atmosphere the vertical motion field is used to calculated the a geostrophic forcing terms, which are then added to the prognostic vorticity and thermodynamic equations. The atmospheric grid results in a horizontal resolution of 5.6°x5.6° . Diabatic heating due to radiative fluxes and the sensible and latent heat exchange at the surface are parameterize. ECBilt-CLIO (often used with ice sheets, vegetation, carbon and biogeochemical cycles as LOVECLIM) is routinely used in EMIC intercomparisons, sea ice, and ocean overturning studies (references in the main text).

Construct patterns of thermal expansion

The gray lines in Fig.4.4 are constructed by taking a 3D field of temperature at three different instances in time. Then this temperature pattern, together with a constant salinity pattern, is scaled up or down, assuming warming or cooling without a changing temperature pattern, i.e. no circulation change. From that temperature and salinity field the density is calculated with the rho_mwif function in NCL in each grid cell and the thermal expansion is calculated through the grid volume and control simulation density. This shows the influence of the non-linear equation of state for any giving pattern. Pattern L in Fig.4 is the equilibrium pattern of a low forced run (top panels in Fig. 2), with most of the anomalous heat concentrated in the surface ocean. This line shows the thermal expansion under the assumption that the circulation changes only a little bit (yellow and orange lines in Fig.3) for a wide range of forcing levels. The other case shows thermal expansion under the assumption that the circulation doesn't change at all and one unit of surface warming would translate into the the same amount of ocean warming everywhere.

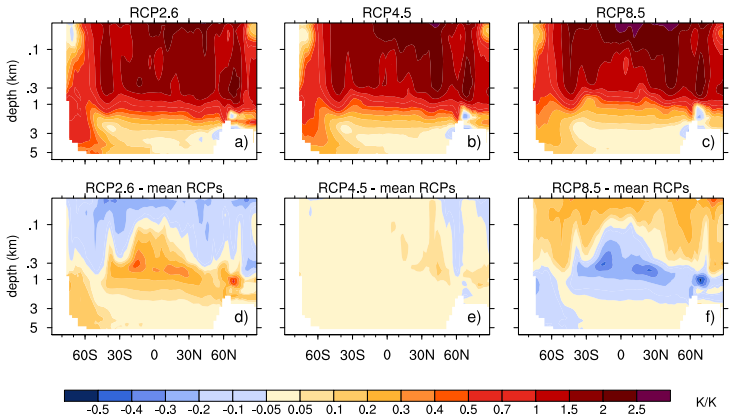

Figure B.1: *Same as Fig. 4.1 in the main text, with zoom into the surface ocean.*

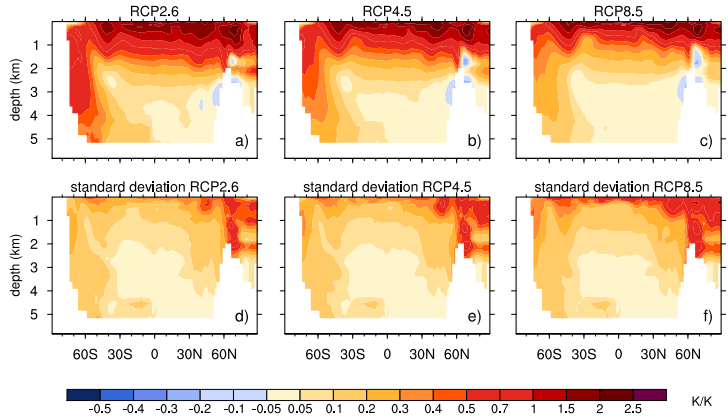

Figure B.2: *Same as Fig. 4.1 in the main text with standard deviation of the models in the lower row. The low latitude shape of the zonal heat penetration with depth is fairly similar in all models, except in two with a very shallow warming. The cooling of the North Atlantic results from seven models, while the other show moderate warming, with the exception of one model with strong warming. The deep Southern Ocean displays a large variety of responses as well.*

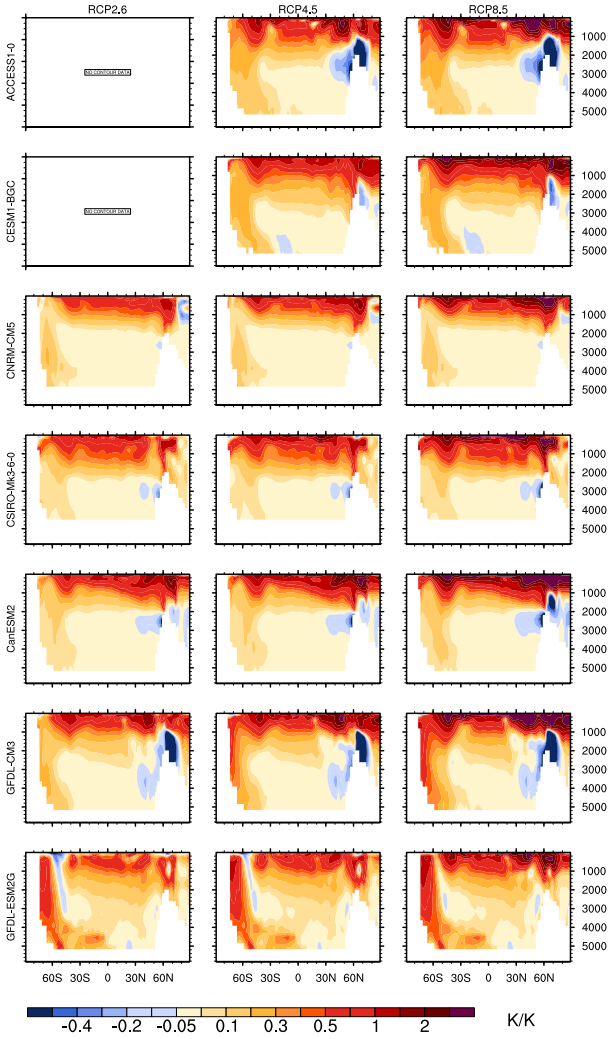

Figure B.3: *As Fig. 4.1 upper row in the main text, each model.*

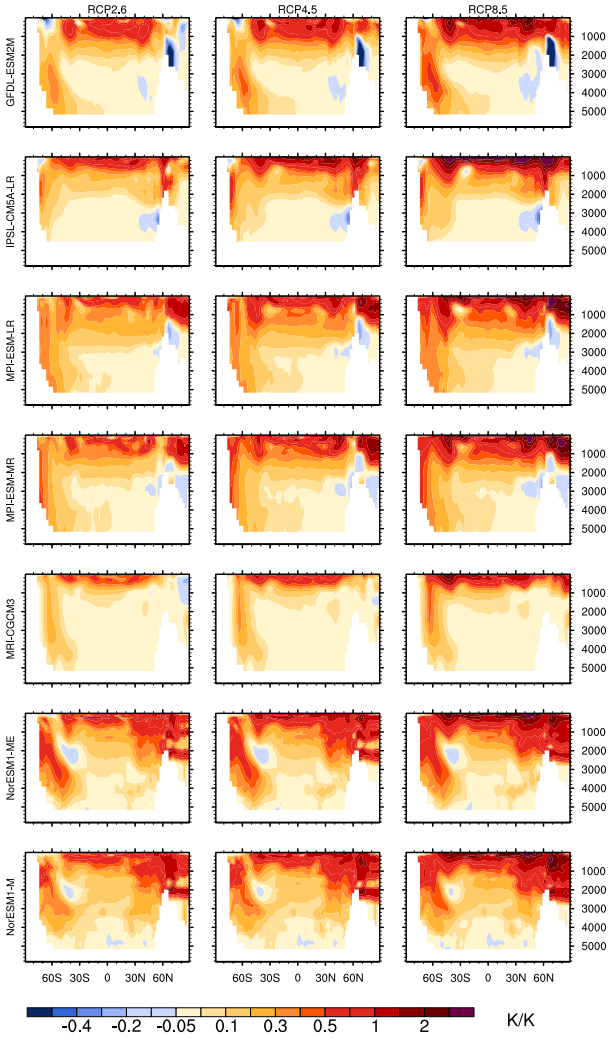

Figure B.4: *As Fig. 1 upper row in the main text, each model, continued.*

C | Appendix for Chapter 5

Surface warming patterns of individual models, discussed around Figure 5.6.

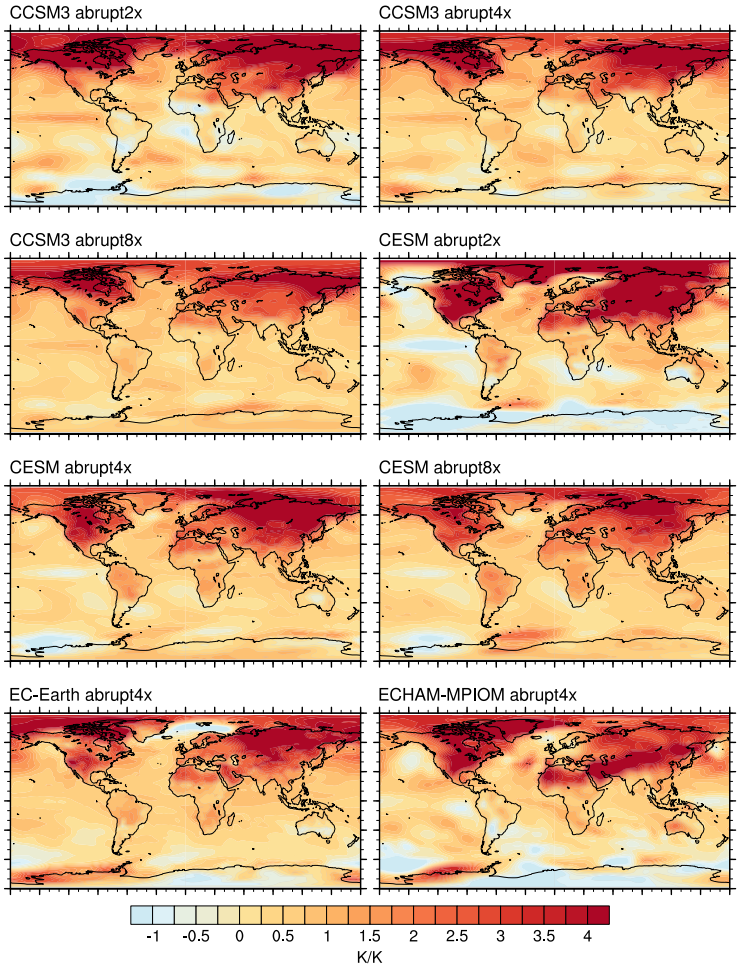

Figure C.1: *Surface temperature anomaly average of the first decade after the step forcing, normalized with the global surface temperature response (thus, unitless). Note that even for simulations which submitted 1pct scenarios, we do use abrupt4x scenarios, from either CMIP5 or additional model runs. However, they are not necessarily matching up with the LongRunMIP scenarios, e.g., abrupt4x with 1pct2x.*

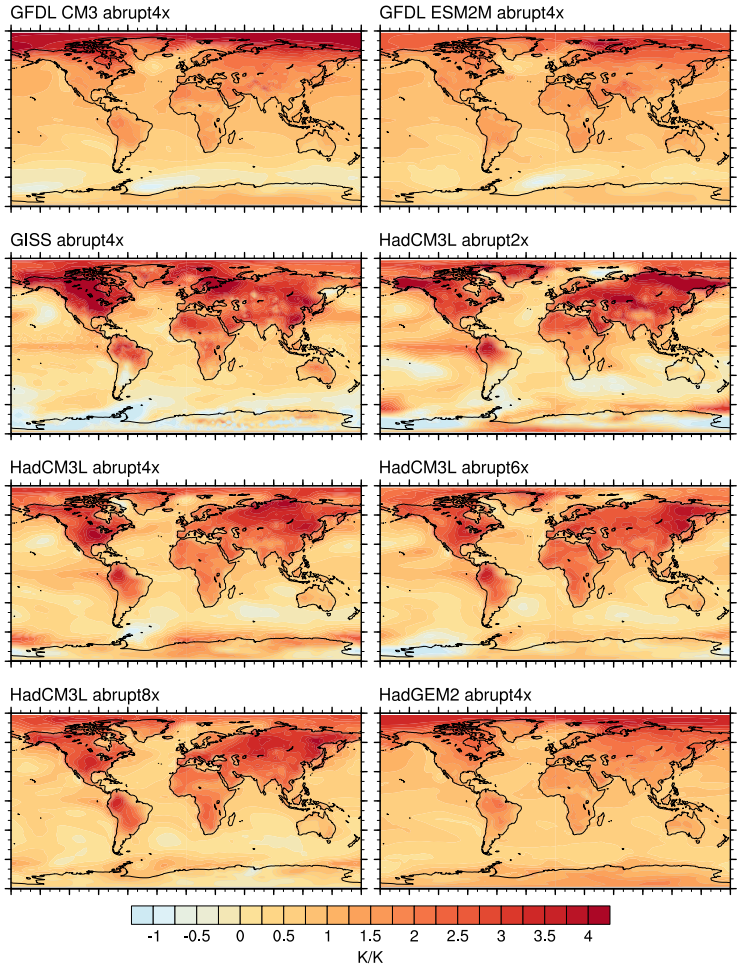

Figure C.2: *As C.1.*

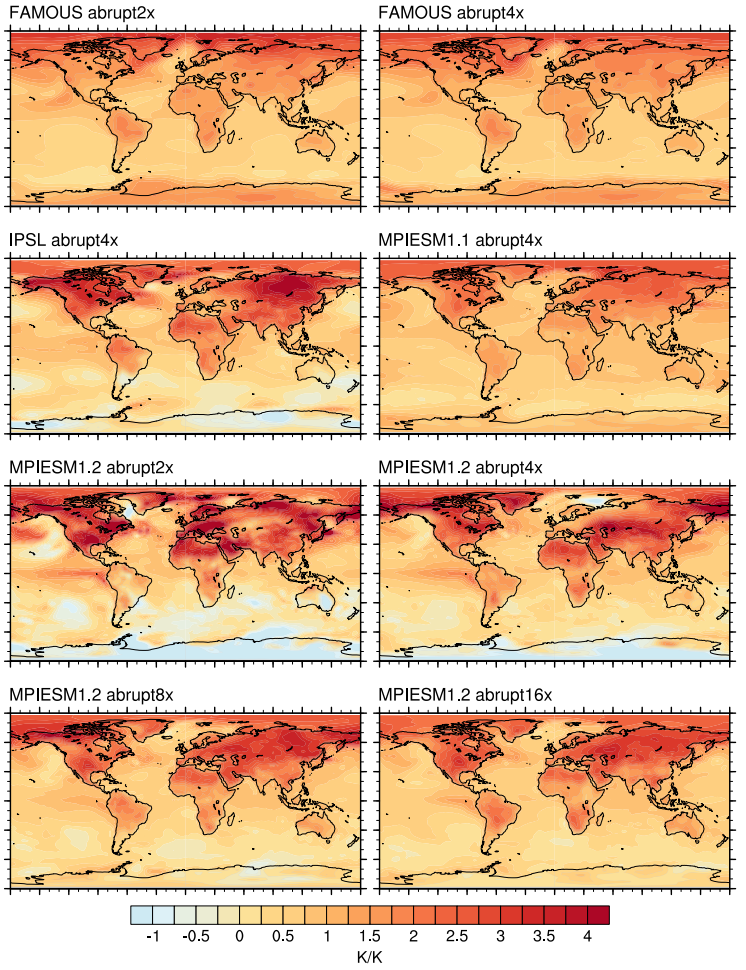

Figure C.3: *As C.1.*

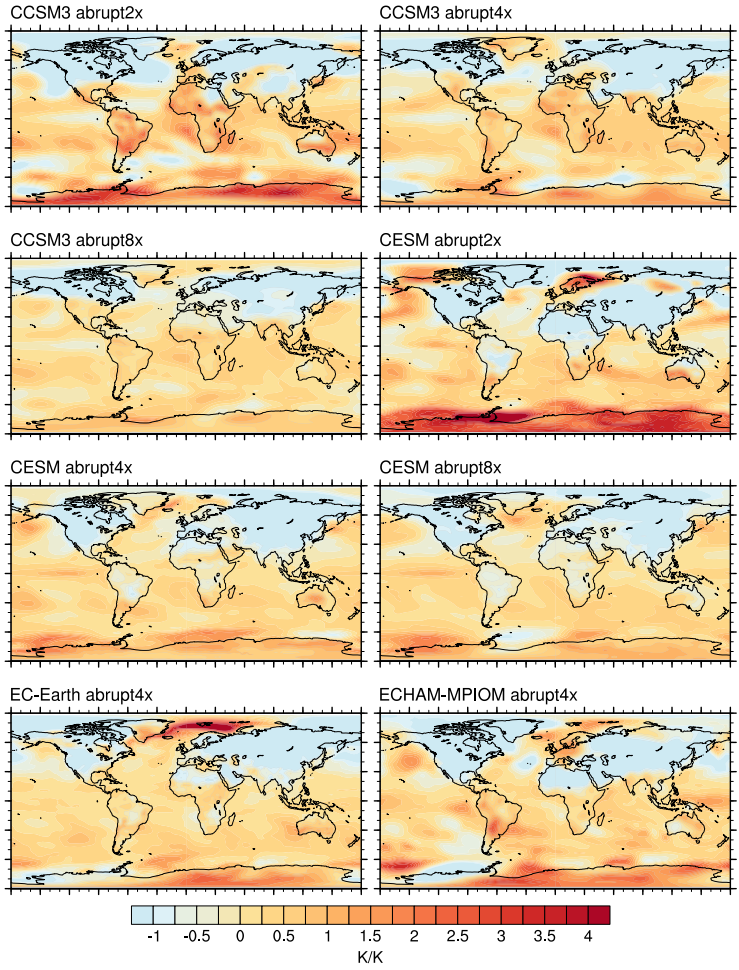

Figure C.4: *Surface temperature anomaly average of year 70-80 minus year 1-10 (as shown in Fig. C.1-C.3). Both periods are normalized with their global surface temperature response.*

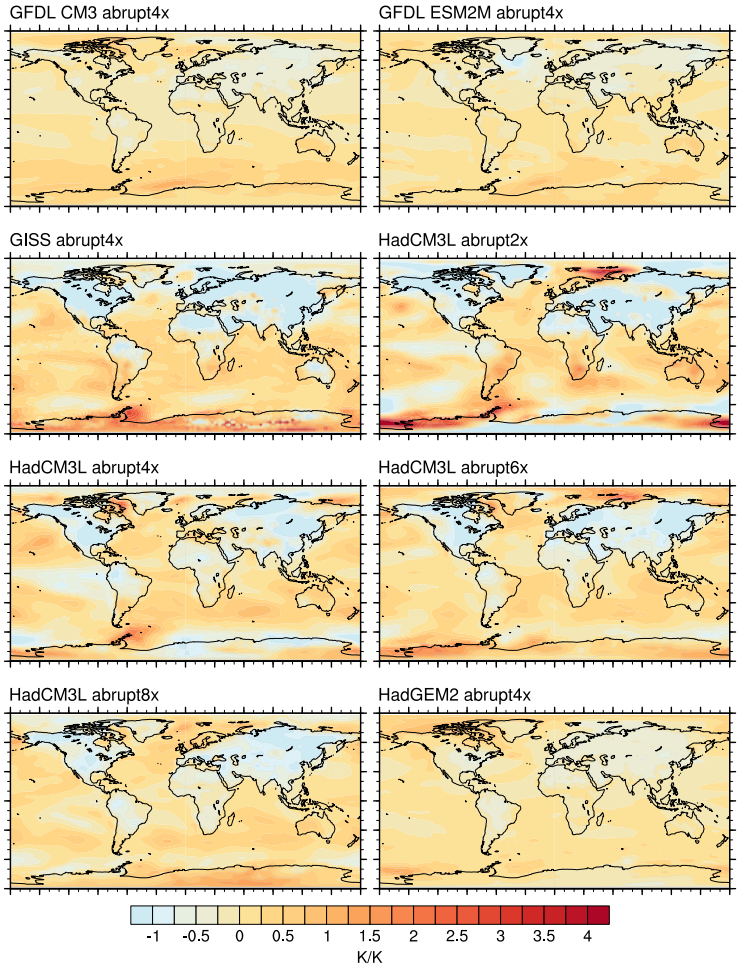

Figure C.5: *As C.4.*

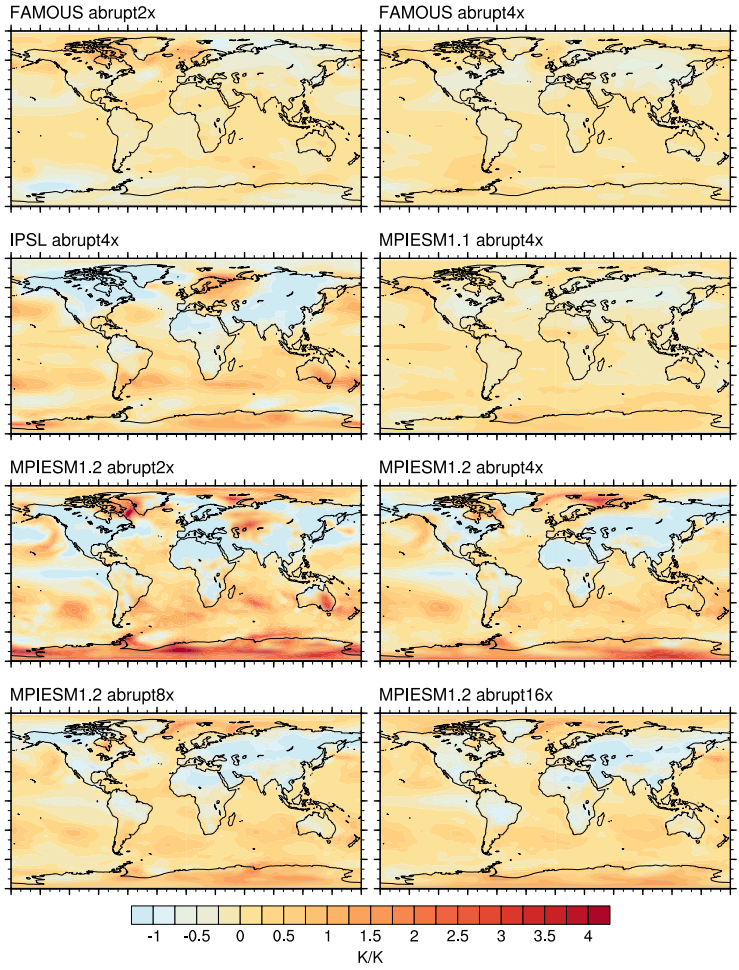

Figure C.6: *As C.4.*

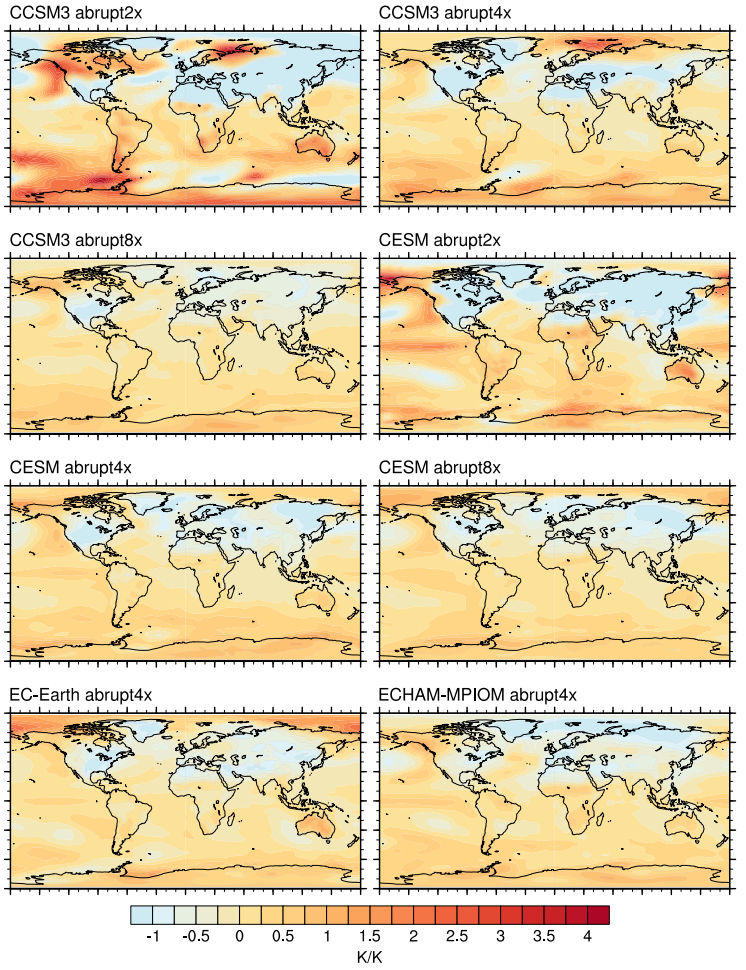

Figure C.7: *Surface temperature anomaly average of year 970-990 minus year 70-80 (as shown in Fig. C.4-C.6). Both periods are normalized with their global surface temperature response.*

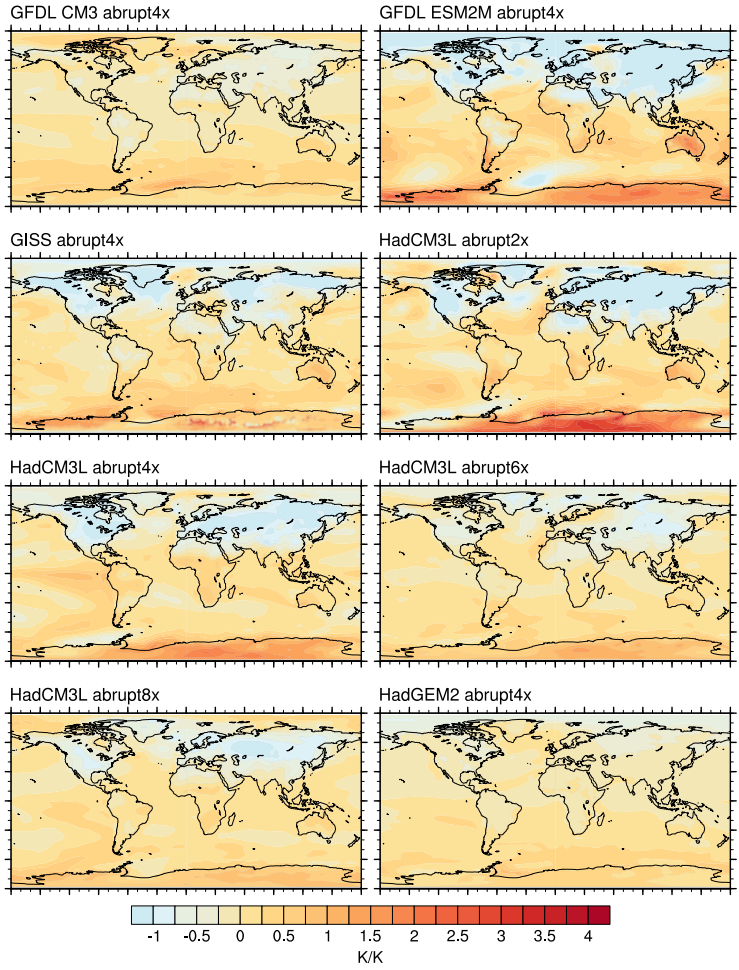

Figure C.8: *As C.7.*

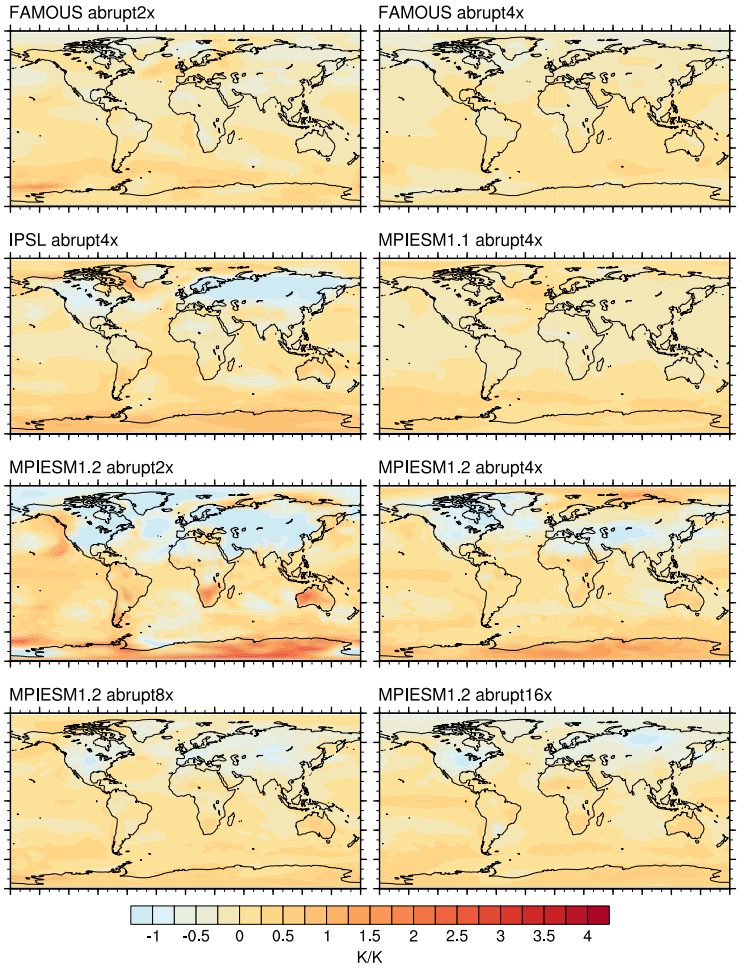

Figure C.9: *As C.7.*

Bibliography

Aldrin, M., M. Holden, P. Guttorp, R. B. Skeie, G. Myhre, and T. K. Berntsen, 2012: Bayesian estimation of climate sensitivity based on a simple climate model fitted to observations of hemispheric temperatures and global ocean heat content. *Environmetrics*, **23 (3)**, 253–271, URL http://dx.doi.org/10.1002/env.2140.

Allen, M. R., 2006: Observational constraints on climate sensitivity. *In Avoiding dangerous climate change*, H. J. Schellnhuber, W. Cramer, N. Nakicenovic, T. M. L. Wigley, and G. Yohe, Eds., Cambridge University Press, Cambridge, 281–289.

Allen, M. R., D. J. Frame, C. Huntingford, C. D. Jones, J. A. Lowe, M. Meinshausen, and N. Meinshausen, 2009: Warming caused by cumulative carbon emissions towards the trillionth tonne. *Nature*, **458 (7242)**, 1163–1166, URL http://dx.doi.org/10.1038/nature08019.

Andrews, T. and P. M. Forster, 2008: CO_2 forcing induces semi-direct effects with consequences for climate feedback interpretations. *Geophysical Research Letters*, **35 (4)**, URL http://dx.doi.org/10.1029/2007GL032273.

Andrews, T., P. M. Forster, O. Boucher, N. Bellouin, and A. Jones, 2010: Precipitation, radiative forcing and global temperature change. *Geophysical Research Letters*, **37 (14)**, URL http://dx.doi.org/10.1029/2010GL043991.

Andrews, T., J. M. Gregory, and M. J. Webb, 2015: The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *Journal of Climate*, **28 (4)**, 1630–1648, URL http://dx.doi.org/10.1175/JCLI-D-14-00545.1.

Andrews, T., J. M. Gregory, M. J. Webb, and K. E. Taylor, 2012: Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophysical Research Letters*, **39 (9)**, URL http://dx.doi.org/10.1029/2012GL051607.

Andronova, N. G. and M. E. Schlesinger, 2001: Objective estimation of the probability density function for climate sensitivity. *Journal of Geophysical Research: Atmospheres*, **106 (D19)**, 22 605–22 611, URL http://dx.doi.org/10.1029/2000JD000259.

Annan, J. and J. Hargreaves, 2015: A perspective on model-data surface temperature comparison at the Last Glacial Maximum. *Quaternary Science Reviews*, **107**, URL http://www.sciencedirect.com/science/article/pii/S0277379114003679.

Armour, K. C., C. M. Bitz, and G. H. Roe, 2013: Time-Varying Climate Sensitivity from Regional Feedbacks. *Journal of Climate*, **26 (13)**, 4518–4534, URL http://dx.doi.org/10.1175/JCLI-D-12-00544.1.

Armour, K. C., J. Marshall, J. R. Scott, A. Donohoe, and E. R. Newsom, 2016: Southern ocean warming delayed by circumpolar upwelling and equatorward transport. *Nature Geosci*, **9 (7)**, 549–554, URL http://dx.doi.org/10.1038/ngeo2731.

Baker, M. B. and G. H. Roe, 2009: The Shape of Things to Come: Why Is Climate Change So Predictable? *Journal of Climate*, **22 (17)**, 4574–4589, doi:10.1175/2009JCLI2647.1, URL http://dx.doi.org/10.1175/2009JCLI2647.1.

Bala, G., K. Caldeira, and R. Nemani, 2010: Fast versus slow response in climate change: implications for the global hydrological cycle. *Climate Dynamics*, **35 (2-3)**, URL http://dx.doi.org/10.1007/s00382-009-0583-y.

Ban-Weiss, G. A. and K. Caldeira, 2010: Geoengineering as an optimization problem. *Environmental Research Letters*, **5 (3)**, 034009, URL http://stacks.iop.org/1748-9326/5/i=3/a=034009.

Bi, D., W. F. Budd, A. C. Hirst, and X. Wu, 2001: Collapse and reorganisation of the Southern Ocean overturning under global warming in a coupled model. *Geophysical Research Letters*, **28 (20)**, 3927–3930, URL http://dx.doi.org/10.1029/2001GL013705.

Bilbao, R. A., J. M. Gregory, and N. Bouttes, 2015: Analysis of the regional pattern of sea level change due to ocean dynamics and density change for 1993–2099 in observations and CMIP5 AOGCMs. *Climate Dynamics*, **45 (9-10)**, 2647–2666, URL http://dx.doi.org/10.1007/s00382-015-2499-z.

Bitz, C., 2012: Antarctic climate response to stratospheric ozone depletion in a fine resolution ocean climate model. *Geophys. Res. Lett.*, **39 (20)**, L20705, doi:10.1029/2012GL053393, URL http://dx.doi.org/10.1029/2012GL053393.

Bitz, C. M., K. M. Shell, P. R. Gent, D. A. Bailey, G. Danabasoglu, K. C. Armour, M. M. Holland, and J. T. Kiehl, 2012: Climate Sensitivity of the Community Climate System Model, Version 4. *Journal of Climate*, **25 (9)**, URL http://dx.doi.org/10.1175/JCLI-D-11-00290.1.

Block, K. and T. Mauritsen, 2013: Forcing and feedback in the MPI-ESM-LR coupled model under abruptly quadrupled CO_2. *Journal of Advances in Modeling Earth Systems*, **5 (4)**, 676–691, URL http://dx.doi.org/10.1002/jame.20041.

Bloomfield, S. D. and D. A. Updegrove, 1981: Modeling for insight, not numbers. *New Directions for Higher Education*, **(35)**, 93–104, URL http://dx.doi.org/10.1002/he.36919813509.

Boer, G. and B. Yu, 2003a: Climate sensitivity and climate state. *Climate Dynamics*, **21 (2)**, 167–176, URL http://dx.doi.org/10.1007/s00382-003-0323-7.

Boer, G. J., K. Hamilton, and W. Zhu, 2005: Climate sensitivity and climate change under strong forcing. *Climate Dynamics*, **24 (7)**, 685–700, URL http://dx.doi.org/10.1007/s00382-004-0500-3.

Boer, G. J., M. Stowasser, and K. Hamilton, 2007: Inferring climate sensitivity from volcanic events. *Climate Dynamics*, **28 (5)**, 481–502, URL http://dx.doi.org/10.1007/s00382-006-0193-x.

Boer, G. J. and B. Yu, 2003b: Dynamical aspects of climate sensitivity. *Geophysical Research Letters*, **30 (3)**, n/a–n/a, doi:10.1029/2002GL016549, URL http://dx.doi.org/10.1029/2002GL016549.

Bony, S., G. Bellon, D. Klocke, S. Sherwood, S. Fermepin, and S. Denvil, 2013: Robust direct effect of carbon dioxide on tropical circulation and regional precipitation. *Nature Geosci*, **6 (6)**, 447–451, URL http://dx.doi.org/10.1038/ngeo1799.

Bony, S. and J.-L. Dufresne, 2005: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophysical Research Letters*, **32 (20)**, URL http://dx.doi.org/10.1029/2005GL023851.

Bony, S., et al., 2006: How Well Do We Understand and Evaluate Climate Change Feedback Processes? *Journal of Climate*, **19 (15)**, 3445–3482.

Bony, S., et al., 2015: Clouds, circulation and climate sensitivity. *Nature Geosci*, **8 (4)**, 261–268, URL http://dx.doi.org/10.1038/ngeo2398.

Boucher, O., et al., 2013: Clouds and Aerosols. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Bouttes, N., J. M. Gregory, and J. A. Lowe, 2013: The Reversibility of Sea Level Rise. *Journal of Climate*, **26 (8)**, 2502–2513, URL http://dx.doi.org/10.1175/JCLI-D-12-00285.1.

Braconnot, P., S. P. Harrison, M. Kageyama, P. J. Bartlein, V. Masson-Delmotte, A. Abe-Ouchi, B. Otto-Bliesner, and Y. Zhao, 2012: Evaluation of climate models using palaeoclimatic data. *Nature Clim. Change*, **2 (6)**, 417–424, URL http://dx.doi.org/10.1038/nclimate1456.

Bryan, F., N. Nakashiki, Y. Yoshida, and K. Maruyama, 2013: Response of the Meridional Overturning Circulation During Differing Pathways Toward Greenhouse Gas Stabilization. *Ocean Circulation: Mechanisms and Impacts—Past and Future Changes of Meridional Overturning*, A. Schmittner, J. C. H. Chiang, and S. R. Hemming, Eds., American Geophysical Union, 351–363, URL http://dx.doi.org/10.1029/173GM22.

Bryan, F. O., G. Danabasoglu, N. Nakashiki, Y. Yoshida, D.-H. Kim, J. Tsutsui, and S. C. Doney, 2006: Response of the North Atlantic Thermohaline Circulation and Ventilation to Increasing Carbon Dioxide in CCSM3. *Journal of Climate*, **19 (11)**, 2382–2397, doi:10.1175/JCLI3757.1.

Budyko, M. I., 1969: The effect of solar radiation variations on the climate of the earth. *Tellus*, **21 (5)**, 611–619, URL http://dx.doi.org/10.1111/j.2153-3490.1969.tb00466.x.

Caballero, R. and M. Huber, 2013: State-dependent climate sensitivity in past warm climates and its implications for future climate projections. *Proceedings of the National Academy of Sciences of the United States of America*, **110 (35)**, 14162–14167, URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3761583/.

Cai, W., et al., 2015: Increased frequency of extreme la nina events under greenhouse warming. *Nature Clim. Change*, **5 (2)**, 132–137, URL http://dx.doi.org/10.1038/nclimate2492.

Caldeira, K. and I. Cvijanovic, 2014: Estimating the Contribution of Sea Ice Response to Climate Sensitivity in a Climate Model. *Journal of Climate*, **27 (22)**, 8597–8607, URL http://dx.doi.org/10.1175/JCLI-D-14-00042.1.

Caldeira, K. and N. P. Myhrvold, 2013: Projections of the pace of warming following an abrupt increase in atmospheric carbon dioxide concentration. *Environmental Research Letters*, **8 (3)**, 034 039, URL http://stacks.iop.org/1748-9326/8/i=3/a=034039.

Cao, L., G. Bala, and K. Caldeira, 2011: Why is there a short-term increase in global precipitation in response to diminished CO_2 forcing? *Geophysical Research Letters*, **38 (6)**, URL http://dx.doi.org/10.1029/2011GL046713.

Cao, L., G. Bala, and K. Caldeira, 2012: Climate response to changes in atmospheric carbon dioxide and solar irradiance on the time scale of days to weeks. *Environmental Research Letters*, **7 (3)**, 034 015, URL http://stacks.iop.org/1748-9326/7/i=3/a=034015.

Cess, R. D., et al., 1989: Interpretation of Cloud-Climate Feedback as Produced by 14 Atmospheric General Circulation Models. *Science*, **245 (4917)**, 513–516, URL http://science.sciencemag.org/content/245/4917/513.

Chadwick, R., P. Good, T. Andrews, and G. Martin, 2014: Surface warming patterns drive tropical rainfall pattern responses to CO_2 forcing on all timescales. *Geophysical Research Letters*, **41 (2)**, 610–615, URL http://dx.doi.org/10.1002/2013GL058504.

Charney, J., 1979: Carbon Dioxide and Climate: A Scientific Assessment. Tech. rep., Washington, DC.

Chung, E.-S. and B. J. Soden, 2015a: An assessment of direct radiative forcing, radiative adjustments, and radiative feedbacks in coupled ocean–atmosphere models. *Journal of Climate*, URL http://dx.doi.org/10.1175/JCLI-D-14-00436.1.

Chung, E.-S. and B. J. Soden, 2015b: An assessment of methods for computing radiative forcing in climate models. *Environmental Research Letters*, **10 (7)**, 074 004, URL http://stacks.iop.org/1748-9326/10/i=7/a=074004.

Church, J. A., et al., 2011: Revisiting the Earth's sea-level and energy budgets from 1961 to 2008. *Geophysical Research Letters*, **38 (18)**, URL http://dx.doi.org/10.1029/2011GL048794.

Church, J. A., et al., 2013: Sea Level Change. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press.

Clement, A. C., R. Seager, M. A. Cane, and S. E. Zebiak, 1996: An Ocean Dynamical Thermostat. *Journal of Climate*, **9 (9)**, 2190–2196, URL http://dx.doi.org/10.1175/1520-0442(1996)009<2190:AODT>2.0.CO;2.

Collins, M., et al., 2013: Long-term Climate Change: Projections, Commitments and Irreversibility. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press.

Colman, R. and B. McAvaney, 2009: Climate feedbacks under a very broad range of forcing. *Geophysical Research Letters*, **36 (1)**, n/a–n/a, doi:10.1029/2008GL036268, URL http://dx.doi.org/10.1029/2008GL036268.

Colman, R. and B. McAvaney, 2011: On tropospheric adjustment to forcing and climate feedbacks. *Climate Dynamics*, **36 (9-10)**, 1649–1658, doi:10.1007/s00382-011-1067-4, URL http://dx.doi.org/10.1007/s00382-011-1067-4.

Cramer, B. S., J. R. Toggweiler, J. D. Wright, M. E. Katz, and K. G. Miller, 2009: Ocean overturning since the Late Cretaceous: Inferences from a new benthic foraminiferal isotope compilation. *Paleoceanography*, **24 (4)**, URL http://dx.doi.org/10.1029/2008PA001683.

Dallafior, T. N., D. Folini, R. Knutti, and M. Wild, 2016: Mixed-layer ocean responses to anthropogenic aerosol dimming from 1870 to 2000. *Journal of Geophysical Research: Atmospheres*, **121 (1)**, 49–66, URL http://dx.doi.org/10.1002/2015JD024070.

Danabasoglu, G., S. C. Bates, B. P. Briegleb, S. R. Jayne, M. Jochum, W. G. Large, S. Peacock, and S. G. Yeager, 2011: The CCSM4 Ocean Component. *Journal of Climate*, **25 (5)**, 1361–1389, URL http://dx.doi.org/10.1175/JCLI-D-11-00091.1.

Danabasoglu, G. and P. R. Gent, 2009: Equilibrium Climate Sensitivity: Is It Accurate to Use a Slab Ocean Model? *Journal of Climate*, **22 (9)**, 2494–2499, URL http://dx.doi.org/10.1175/2008JCLI2596.1.

Danabasoglu, G., S. G. Yeager, Y.-O. Kwon, J. J. Tribbia, A. S. Phillips, and J. W. Hurrell, 2012: Variability of the Atlantic Meridional Overturning Circulation in CCSM4. *Journal of Climate*, URL http://dx.doi.org/10.1175/JCLI-D-11-00463.1.

Deser, C., R. A. Tomas, and L. Sun, 2015: The Role of Ocean–Atmosphere Coupling in the Zonal-Mean Atmospheric Response to Arctic Sea Ice Loss. *Journal of Climate*, **28 (6)**, 2168–2186, URL http://dx.doi.org/10.1175/JCLI-D-14-00325.1.

Dessler, A. E., 2010: A Determination of the Cloud Feedback from Climate Variations over the Past Decade. *Science*, **330 (6010)**, 1523–1527, URL http://science.sciencemag.org/content/330/6010/1523.

Dong, B., J. M. Gregory, and R. T. Sutton, 2009: Understanding Land–Sea Warming Contrast in Response to Increasing Greenhouse Gases. Part I: Transient Adjustment. *Journal of Climate*, **22 (11)**, 3079–3097, URL http://dx.doi.org/10.1175/2009JCLI2652.1.

Donohoe, A., D. Frierson, and D. Battisti, 2014: The effect of ocean mixed layer depth on climate in slab ocean aquaplanet experiments. *Climate Dynamics*, **43 (3-4)**, 1041–1055, URL http://dx.doi.org/10.1007/s00382-013-1843-4.

Doutriaux-Boucher, M., M. J. Webb, J. M. Gregory, and O. Boucher, 2009: Carbon dioxide induced stomatal closure increases radiative forcing via a rapid reduction in low cloud. *Geophysical Research Letters*, **36 (2)**, URL http://dx.doi.org/10.1029/2008GL036273.

Drijfhout, S. S., S. L. Weber, and E. van der Swaluw, 2011: The stability of the moc as diagnosed from model projections for pre-industrial, present and future climates. *Climate Dynamics*, **37 (7)**, 1575–1586, doi:10.1007/s00382-010-0930-z, URL http://dx.doi.org/10.1007/s00382-010-0930-z.

Eby, M., et al., 2013: Historical and idealized climate model experiments: an intercomparison of Earth system models of intermediate complexity. *Climate of the Past*, **9 (3)**, 1111–1140, URL http://www.clim-past.net/9/1111/2013/.

Exarchou, E., T. Kuhlbrodt, J. M. Gregory, and R. S. Smith, 2014: Ocean Heat Uptake Processes: A Model Intercomparison. *Journal of Climate*, **28 (2)**, 887–908, URL http://dx.doi.org/10.1175/JCLI-D-14-00235.1.

Fasullo, J. T. and K. E. Trenberth, 2012: A Less Cloudy Future: The Role of Subtropical Subsidence in Climate Sensitivity. *Science*, **338 (6108)**, 792–794, URL http://science.sciencemag.org/content/338/6108/792.

Feldl, N. and G. H. Roe, 2013: The Nonlinear and Nonlocal Nature of Climate Feedbacks. *Journal of Climate*, **26 (21)**, 8289–8304, URL http://dx.doi.org/10.1175/JCLI-D-12-00631.1.

Flato, G., et al., 2013: Evaluation of Climate Models. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Forest, C. E., P. H. Stone, A. P. Sokolov, M. R. Allen, and M. D. Webster, 2002: Quantifying Uncertainties in Climate System Properties with the Use of Recent Climate Observations. *Science*, **295 (5552)**, 113–117, URL http://science.sciencemag.org/content/295/5552/113.

Forster, P. M., T. Andrews, P. Good, J. M. Gregory, L. S. Jackson, and M. Zelinka, 2013: Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *Journal of Geophysical Research: Atmospheres*, **118 (3)**, 1139–1150, URL http://dx.doi.org/10.1002/jgrd.50174.

Forster, P. M. F. and J. M. Gregory, 2006: The Climate Sensitivity and Its Components Diagnosed from Earth Radiation Budget Data. *Journal of Climate*, **19 (1)**, 39–52, URL http://dx.doi.org/10.1175/JCLI3611.1.

Frame, D. J., B. B. B. Booth, J. A. Kettleborough, D. A. Stainforth, J. M. Gregory, M. Collins, and M. R. Allen, 2005: Constraining climate forecasts: The role of prior assumptions. *Geophysical Research Letters*, **32 (9)**, URL http://dx.doi.org/10.1029/2004GL022241.

Friedrich, O., R. D. Norris, and J. Erbacher, 2012: Evolution of middle to Late Cretaceous oceans-A 55 m.y. record of Earth's temperature and carbon cycle. *Geology*, **40 (2)**, 107–110, doi:10.1130/g32701.1.

Friedrich, T., A. Timmermann, L. Menviel, O. Elison Timm, A. Mouchet, and D. M. Roche, 2010: The mechanism behind internally generated centennial-to-millennial scale climate variability in an earth system model of intermediate complexity. *Geoscientific Model Development*, **3 (2)**, 377–389, doi:10.5194/gmd-3-377-2010, URL http://www.geosci-model-dev.net/3/377/2010/.

Frölicher, T. L., M. Winton, and J. L. Sarmiento, 2014: Continued global warming after CO_2 emissions stoppage. *Nature Clim. Change*, **4 (1)**, 40–44, URL http://dx.doi.org/10.1038/nclimate2060.

Gent, P. R., et al., 2011: The Community Climate System Model Version 4. *Journal of Climate*, **24 (19)**, 4973–4991, URL http://dx.doi.org/10.1175/2011JCLI4083.1.

Geoffroy, O., D. Saint-Martin, G. Bellon, A. Voldoire, D. J. L. Olivié, and S. Tytéca, 2013a: Transient Climate Response in a Two-Layer Energy-Balance Model. Part II: Representation of the Efficacy of Deep-Ocean Heat Uptake and Validation for CMIP5 AOGCMs. *Journal of Climate*, **26 (6)**, 1859–1876, URL http://dx.doi.org/10.1175/JCLI-D-12-00196.1.

Geoffroy, O., D. Saint-Martin, D. J. L. Olivié, A. Voldoire, G. Bellon, and S. Tytéca, 2013b: Transient Climate Response in a Two-Layer Energy-Balance Model. Part I: Analytical Solution and Parameter Calibration Using CMIP5 AOGCM Experiments. *Journal of Climate*, **26 (6)**, 1841–1857, URL http://dx.doi.org/10.1175/JCLI-D-12-00195.1.

Gillett, N. P., V. K. Arora, K. Zickfeld, S. J. Marshall, and W. J. Merryfield, 2011: Ongoing climate change following a complete cessation of carbon dioxide emissions. *Nature Geosci*, **4 (2)**, 83–87, URL http://dx.doi.org/10.1038/ngeo1047.

Good, P., J. Gregory, J. Lowe, and T. Andrews, 2013: Abrupt CO_2 experiments as tools for predicting and understanding CMIP5 representative concentration pathway projections. *Climate Dynamics*, **40 (3-4)**, 1041–1053, URL http://dx.doi.org/10.1007/s00382-012-1410-4.

Good, P., J. M. Gregory, and J. A. Lowe, 2011: A step-response simple climate model to reconstruct and interpret AOGCM projections. *Geophysical Research Letters*, **38 (1)**, URL http://dx.doi.org/10.1029/2010GL045208.

Good, P., et al., 2015: Nonlinear regional warming with increasing co2 concentrations. *Nature Clim. Change*, **5 (2)**, 138–142, URL http://dx.doi.org/10.1038/nclimate2498.

Goosse, H. and T. Fichefet, 1999: Importance of ice-ocean interactions for the global ocean circulation: A model study. *Journal of Geophysical Research: Oceans*, **104 (C10)**, 23 337–23 355, URL http://dx.doi.org/10.1029/1999JC900215.

Gregory, J. and M. Webb, 2008: Tropospheric Adjustment Induces a Cloud Component in CO_2 Forcing. *Journal of Climate*, **21 (1)**, 58–71, URL http://dx.doi.org/10.1175/2007JCLI1834.1.

Gregory, J. M., 2000: Vertical heat transports in the ocean and their effect on time-dependent climate change. *Climate Dynamics*, **16**, 501–515, URL http://dx.doi.org/10.1007/s003820000059, 10.1007/s003820000059.

Gregory, J. M. and T. Andrews, 2016: Variation in climate sensitivity and feedback parameters during the historical period. *Geophysical Research Letters*, URL http://dx.doi.org/10.1002/2016GL068406.

Gregory, J. M., T. Andrews, and P. Good, 2015: The inconstancy of the transient climate response parameter under increasing CO_2. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **373 (2054)**, URL http://rsta.royalsocietypublishing.org/content/373/2054/20140417.

Gregory, J. M., T. Andrews, P. Good, T. Mauritsen, and P. M. Forster, 2016: Small global-mean cooling due to volcanic radiative forcing. *Climate Dynamics*, 1–13, URL http://dx.doi.org/10.1007/s00382-016-3055-1.

Gregory, J. M., R. J. Stouffer, S. C. B. Raper, P. A. Stott, and N. A. Rayner, 2002: An Observationally Based Estimate of the Climate Sensitivity. *Journal of Climate*, **15 (22)**, 3117–3121, URL http://dx.doi.org/10.1175/1520-0442(2002)015<3117:AOBEOT>2.0.CO;2.

Gregory, J. M., et al., 2004: A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, **31 (3)**, URL http://dx.doi.org/10.1029/2003GL018747.

Gregory, J. M., et al., 2005: A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing atmospheric CO_2 concentration. *Geophys. Res. Lett.*, **32 (L12703)**.

Grise, K. M. and L. M. Polvani, 2014a: Southern Hemisphere Cloud–Dynamics Biases in CMIP5 Models and Their Implications for Climate Projections. *Journal of Climate*, **27 (15)**, 6074–6092, URL http://dx.doi.org/10.1175/JCLI-D-14-00113.1.

Grise, K. M. and L. M. Polvani, 2014b: The response of mid-latitude jets to increased CO_2: Distinguishing the roles of sea surface temperature and direct radiative forcing. *Geophysical Research Letters*, URL http://dx.doi.org/10.1002/2014GL061638.

Hallberg, R., A. Adcroft, J. P. Dunne, J. P. Krasting, and R. J. Stouffer, 2012: Sensitivity of Twenty-First-Century Global-Mean Steric Sea Level Rise to Ocean Model Formulation. *Journal of Climate*, **26 (9)**, 2947–2956, URL http://dx.doi.org/10.1175/JCLI-D-12-00506.1.

Hamming, R. W., 1962: *Numerical Methods for Scientists and Engineers*. McGraw-Hill.

Hansen, J., A. Lacis, D. Rind, G. Russel, P. Stone, I. Fung, R. Ruedy, and J. Lerner, 1984: Analysis of feedback mechanisms. In: Climate processes and climate sensitivity. *Climate sensitivity: Analysis of feedback mechanisms*, J. Hansen and T. Takahashi, Eds., AGU Geophysical Monograph 29, Maurice Ewing, American Geophysical Union, Washington, DC, Vol. 5, 130–163.

Hansen, J., G. Russell, A. Lacis, I. Fung, D. Rind, and P. Stone, 1985: Climate Response Times: Dependence on Climate Sensitivity and Ocean Mixing. *Science*, **229 (4716)**, 857–859, doi:10.1126/science.229.4716.857, URL http://www.sciencemag.org/content/229/4716/857.abstract.

Hansen, J., M. Sato, and R. Ruedy, 1997: Radiative forcing and climate response. *Journal of Geophysical Research: Atmospheres*, **102 (D6)**, 6831–6864, doi:10.1029/96JD03436, URL http://dx.doi.org/10.1029/96JD03436.

Hansen, J., et al., 2005a: Earth's Energy Imbalance: Confirmation and Implications. *Science*, **308 (5727)**, 1431–1435, URL http://science.sciencemag.org/content/308/5727/1431.

Hansen, J., et al., 2005b: Efficacy of climate forcings. *Journal of Geophysical Research: Atmospheres*, **110 (D18)**, URL http://dx.doi.org/10.1029/2005JD005776.

Hargreaves, J. C., A. Abe-Ouchi, and J. D. Annan, 2007: Linking glacial and future climates through an ensemble of GCM simulations. *Climate of the Past*, **3 (1)**, 77–87, doi:10.5194/cp-3-77-2007, URL http://www.clim-past.net/3/77/2007/.

Hargreaves, J. C., J. D. Annan, M. Yoshimori, and A. Abe-Ouchi, 2012: Can the Last Glacial Maximum constrain climate sensitivity? *Geophysical Research Letters*, **39 (24)**, URL http://dx.doi.org/10.1029/2012GL053872.

Hasselmann, K., 1976: Stochastic climate models part i. theory. *Tellus*, **28 (6)**, 473–485, URL http://dx.doi.org/10.1111/j.2153-3490.1976.tb00696.x.

Hasselmann, K., R. Sausen, E. Maier-Reimer, and R. Voss, 1993: On the cold start problem in transient simulations with coupled atmosphere-ocean models. *Climate Dynamics*, **9 (2)**, 53–61, URL http://dx.doi.org/10.1007/BF00210008.

Hawkins, E., R. Smith, J. Gregory, and D. Stainforth, 2015: Irreducible uncertainty in near-term climate projections. *Climate Dynamics*, 1–13, doi:10.1007/s00382-015-2806-8, URL http://dx.doi.org/10.1007/s00382-015-2806-8.

Hawkins, E. and R. Sutton, 2012: Time of emergence of climate signals. *Geophysical Research Letters*, **39 (1)**, URL http://dx.doi.org/10.1029/2011GL050087.

Hegerl, G. C., T. J. Crowley, W. T. Hyde, and D. J. Frame, 2006: Climate sensitivity constrained by temperature reconstructions over the past seven centuries. *Nature*, **440 (7087)**, 1029–1032, URL http://dx.doi.org/10.1038/nature04679.

Held, I., M. Winton, K. Takahashi, T. L. Delworth, F. Zeng, and G. Vallis, 2010: Probing the Fast and Slow Components of Global Warming by Returning Abruptly to Preindustrial Forcing. *Journal of Climate*, **23**, 2418 – 2427.

Held, I. M., 2005: The Gap between Simulation and Understanding in Climate Modeling. *Bulletin of the American Meteorological Society*, **86 (11)**, 1609–1614, URL http://dx.doi.org/10.1175/BAMS-86-11-1609.

Herger, N., B. M. Sanderson, and R. Knutti, 2015: Improved pattern scaling approaches for the use in climate impact studies. *Geophysical Research Letters*, **42 (9)**, 3486–3494, URL http://dx.doi.org/10.1002/2015GL063569.

Heuzé, C., K. J. Heywood, D. P. Stevens, and J. K. Ridley, 2015: Changes in Global Ocean Bottom Properties and Volume Transports in CMIP5 Models under Climate Change Scenarios. *Journal of Climate*, **28 (8)**, 2917–2944, URL http://dx.doi.org/10.1175/JCLI-D-14-00381.1.

Hsieh, W.-C., W. D. Collins, Y. Liu, J. C. H. Chiang, C.-L. Shie, K. Caldeira, and L. Cao, 2013: Climate response due to carbonaceous aerosols and aerosol-induced SST effects in NCAR community atmospheric model CAM3.5. *Atmospheric Chemistry and Physics*, **13 (15)**, 7489–7510, doi:10.5194/acp-13-7489-2013, URL http://www.atmos-chem-phys.net/13/7489/2013/.

Huber, B. T., 1998: Tropical Paradise at the Cretaceous Poles? *Science*, **282 (5397)**, 2199–2200, doi:10.1126/science.282.5397.2199, URL http://www.sciencemag.org/content/282/5397/2199.short.

Huber, M. and R. Knutti, 2012: Anthropogenic and natural warming inferred from changes in Earth's energy balance. *Nature Geosci*, **5 (1)**, 31–36, URL http://dx.doi.org/10.1038/ngeo1327.

Huber, M. and R. Knutti, 2014: Natural variability, radiative forcing and climate response in the recent hiatus reconciled. *Nature Geosci*, **7 (9)**, 651–656, URL http://dx.doi.org/10.1038/ngeo2228.

Hurrell, J. W., et al., 2013: The Community Earth System Model: A Framework for Collaborative Research. *Bulletin of the American Meteorological Society*, **94 (9)**, 1339–1360, URL http://dx.doi.org/10.1175/BAMS-D-12-00121.1.

Huybers, P. and C. Proistosescu, 2016: Paleoclimate records needed for observationally constraining equilibrium climate sensitivity. *AGU Fall Meeting Abstracts*, 130956, URL https://agu.confex.com/agu/fm16/meetingapp.cgi/Paper/130956.

Jaccard, S., E. D. Galbraith, T. L. Frölicher, and N. Gruber, 2014: Ocean (de)oxygenation across the last deglaciation: Insights for the future. *Oceanography*, **27 (1)**, 26–35, URL http://dx.doi.org/10.5670/oceanog.2014.05.

Johansson, D. J. A., B. C. O/'Neill, C. Tebaldi, and O. Haggstrom, 2015: Equilibrium climate sensitivity in light of observations over the warming hiatus. *Nature Clim. Change*, **5 (5)**, 449–453, URL http://dx.doi.org/10.1038/nclimate2573.

Jonko, A. K., K. M. Shell, B. M. Sanderson, and G. Danabasoglu, 2012: Climate Feedbacks in CCSM3 under Changing CO2 Forcing. Part I: Adapting the Linear Radiative Kernel Technique to Feedback Calculations for a Broad Range of Forcings. *Journal of Climate*, **25 (15)**, 5260–5272, URL http://dx.doi.org/10.1175/JCLI-D-11-00524.1.

Jonko, A. K., K. M. Shell, B. M. Sanderson, and G. Danabasoglu, 2013: Climate Feedbacks in CCSM3 under Changing CO2 Forcing. Part II: Variation of Climate Feedbacks and Sensitivity with Forcing. *Journal of Climate*, **26 (9)**, 2784–2795, URL http://dx.doi.org/10.1175/JCLI-D-12-00479.1.

Kamae, Y. and M. Watanabe, 2013: Tropospheric adjustment to increasing CO_2: its timescale and the role of land–sea contrast. *Climate Dynamics*, **41 (11-12)**, 3007–3024, URL http://dx.doi.org/10.1007/s00382-012-1555-1.

Kamae, Y., M. Watanabe, T. Ogura, M. Yoshimori, and H. Shiogama, 2015: Rapid Adjustments of Cloud and Hydrological Cycle to Increasing CO_2: a Review. *Current Climate Change Reports*, 1–11, URL http://dx.doi.org/10.1007/s40641-015-0007-5.

Kang, S. M., I. M. Held, D. M. W. Frierson, and M. Zhao, 2008: The Response of the ITCZ to Extratropical Thermal Forcing: Idealized Slab-Ocean Experiments with a GCM. *Journal of Climate*, **21 (14)**, 3521–3532, URL http://dx.doi.org/10.1175/2007JCLI2146.1.

Kang, S. M., I. M. Held, and S.-P. Xie, 2014: Contrasting the tropical responses to zonally asymmetric extratropical and tropical thermal forcing. *Climate Dynamics*, **42 (7-8)**, 2033–2043, URL http://dx.doi.org/10.1007/s00382-013-1863-0.

Kang, S. M. and S.-P. Xie, 2014: Dependence of Climate Response on Meridional Structure of External Thermal Forcing. *Journal of Climate*, **27 (14)**, 5593–5600, URL http://dx.doi.org/10.1175/JCLI-D-13-00622.1.

Klocke, D., J. Quaas, and B. Stevens, 2013: Assessment of different metrics for physical climate feedbacks. *Climate Dynamics*, **41 (5-6)**, 1173–1185, URL http://dx.doi.org/10.1007/s00382-013-1757-1.

Knutti, R., 2002: Modelling studies on the probability and predictability of future climate change. Ph.D. thesis, Physics Institute, University of Bern, 138pp.

Knutti, R. and G. C. Hegerl, 2008: The equilibrium sensitivity of the Earth's temperature to radiation changes. *Nature Geosci*, **1 (11)**, 735–743, URL http://dx.doi.org/10.1038/ngeo337.

Knutti, R., F. Joos, S. A. Müller, G.-K. Plattner, and T. F. Stocker, 2005: Probabilistic climate change projections for co_2 stabilization profiles. *Geophysical Research Letters*, **32 (20)**, URL http://dx.doi.org/10.1029/2005GL023294.

Knutti, R., G. A. Meehl, M. R. Allen, and D. A. Stainforth, 2006: Constraining Climate Sensitivity from the Seasonal Cycle in Surface Temperature. *Journal of Climate*, **19 (17)**, 4224–4233, URL http://dx.doi.org/10.1175/JCLI3865.1.

Knutti, R. and J. Rogelj, 2015: The legacy of our co_2 emissions: a clash of scientific facts, politics and ethics. *Climatic Change*, **133 (3)**, 361–373, URL http://dx.doi.org/10.1007/s10584-015-1340-3.

Knutti, R. and M. A. A. Rugenstein, 2015: Feedbacks, climate sensitivity and the limits of linear models. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **373 (2054)**, doi:10.1098/rsta.2015.0146.

Knutti, R. and J. Sedlacek, 2013: Robustness and uncertainties in the new cmip5 climate model projections. *Nature Clim. Change*, **3 (4)**, 369–373, URL http://dx.doi.org/10.1038/nclimate1716.

Knutti, R. and T. F. Stocker, 2000: Influence of the Thermohaline Circulation on Projected Sea Level Rise. *Journal of Climate*, **13 (12)**, 1997–2001, URL http://dx.doi.org/10.1175/1520-0442(2000)013<1997:IOTTCO>2.0.CO;2.

Knutti, R., T. F. Stocker, F. Joos, and G.-K. Plattner, 2002: Constraints on radiative forcing and future climate change from observations and climate model ensembles. *Nature*, **416 (6882)**, 719–723, URL http://dx.doi.org/10.1038/416719a.

Knutti, R., T. F. Stocker, F. Joos, and G.-K. Plattner, 2003: Probabilistic climate change projections using neural networks. *Climate Dynamics*, **21 (3)**, 257–272, URL http://dx.doi.org/10.1007/s00382-003-0345-1.

Knutti, R. and L. Tomassini, 2008: Constraints on the transient climate response from observed global temperature and ocean heat uptake. *Geophys. Res. Lett.*, **35**, L09701.

Köhler, P., R. Bintanja, H. Fischer, F. Joos, R. Knutti, G. Lohmann, and V. Masson-Delmotte, 2010: What caused Earth's temperature variations during the last 800,000 years? Data-based evidence on radiative forcing and constraints on climate sensitivity. *Quaternary Science Reviews*, **29 (1–2)**, 129 – 145, URL http://www.sciencedirect.com/science/article/pii/S0277379109003291.

Körper, J., et al., 2013: The effects of aggressive mitigation on steric sea level rise and sea ice changes. *Climate Dynamics*, **40 (3-4)**, 531–550, URL http://dx.doi.org/10.1007/s00382-012-1612-9.

Kostov, Y., K. C. Armour, and J. Marshall, 2014: Impact of the atlantic meridional overturning circulation on ocean heat storage and transient climate change. *Geophysical Research Letters*, **41 (6)**, 2108–2116, doi:10.1002/2013GL058998, URL http://dx.doi.org/10.1002/2013GL058998.

Krasting, J. P., J. P. Dunne, R. J. Stouffer, and R. W. Hallberg, 2016: Enhanced Atlantic sea-level rise relative to the Pacific under high carbon emission rates. *Nature Geosci*, **9 (3)**, 210–214, URL http://dx.doi.org/10.1038/ngeo2641.

Kravitz, B., et al., 2013: An energetic perspective on hydrological cycle changes in the Geoengineering Model Intercomparison Project. *Journal of Geophysical Research: Atmospheres*, **118 (23)**, 13 087–13 102, URL http://dx.doi.org/10.1002/2013JD020502.

Kuhlbrodt, T. and J. Gregory, 2012: Ocean heat uptake and its consequences for the magnitude of sea level rise and climate change. *Geophys. Res. Lett.*, URL http://dx.doi.org/10.1029/2012GL052952.

Kutzbach, J. E., F. He, S. J. Vavrus, and W. F. Ruddiman, 2013: The dependence of equilibrium climate sensitivity on climate state: Applications to studies of climates colder than present. *Geophysical Research Letters*, **40 (14)**, 3721–3726, URL http://dx.doi.org/10.1002/grl.50724.

Lahellec, A. and J.-L. Dufresne, 2014: A Formal Analysis of the Feedback Concept in Climate Models, Part II: Tangent Linear Systems in GCMs. *Journal of the Atmospheric Sciences*, URL http://dx.doi.org/10.1175/JAS-D-13-0334.1.

Lambert, F. H. and N. E. Faull, 2007: Tropospheric adjustment: The response of two general circulation models to a change in insolation. *Geophysical Research Letters*, **34 (3)**, URL http://dx.doi.org/10.1029/2006GL028124.

Larson, E. J. L. and R. W. Portmann, 2015: A temporal kernel method to compute effective radiative forcing in CMIP5 transient simulations. *Journal of Climate*, URL http://dx.doi.org/10.1175/JCLI-D-15-0577.1.

Latif, M., E. Roeckner, U. Mikolajewicz, and R. Voss, 2000: Tropical Stabilization of the Thermohaline Circulation in a Greenhouse Warming Simulation. *Journal of Climate*, **13 (11)**, 1809–1813, URL http://dx.doi.org/10.1175/1520-0442(2000)013<1809:L>2.0.CO;2.

Levermann, A., P. U. Clark, B. Marzeion, G. A. Milne, D. Pollard, V. Radic, and A. Robinson, 2013: The multimillennial sea-level commitment of global warming. *Proceedings of the National Academy of Sciences*, **110 (34)**, 13 745–13 750, URL http://www.pnas.org/content/110/34/13745.abstract.

Lewis, N., 2013: An Objective Bayesian Improved Approach for Applying Optimal Fingerprint Techniques to Estimate Climate Sensitivity. *Journal of Climate*, **26 (19)**, 7414–7429, URL http://dx.doi.org/10.1175/JCLI-D-12-00473.1.

Lewis, N. and J. A. Curry, 2015: The implications for climate sensitivity of AR5 forcing and heat uptake estimates. *Climate Dynamics*, **45 (3)**, 1009–1023, URL http://dx.doi.org/10.1007/s00382-014-2342-y.

L'Hévéder, B., F. Codron, and M. Ghil, 2015: Impact of anomalous northward oceanic heat transport on global climate in a slab ocean setting. *Journal of Climate*, **28 (7)**, 2650–2664, URL http://dx.doi.org/10.1175/JCLI-D-14-00377.1.

Li, C., D. Notz, S. Tietsche, and J. Marotzke, 2013a: The Transient versus the Equilibrium Response of Sea Ice to Global Warming. *Journal of Climate*, **26 (15)**, 5624–5636, URL http://dx.doi.org/10.1175/JCLI-D-12-00492.1.

Li, C., J.-S. Storch, and J. Marotzke, 2013b: Deep-ocean heat uptake and equilibrium climate response. *Climate Dynamics*, **40 (5-6)**, 1071–1086, URL http://dx.doi.org/10.1007/s00382-012-1350-z.

Liang, X., C. Wunsch, P. Heimbach, and G. Forget, 2015: Vertical Redistribution of Oceanic Heat Content. *Journal of Climate*, **28 (9)**, 3821–3833, URL http://dx.doi.org/10.1175/JCLI-D-14-00550.1.

Llovel, W., J. K. Willis, F. W. Landerer, and I. Fukumori, 2014: Deep-ocean contribution to sea level and energy budget not detectable over the past decade. *Nature Clim. Change*, **4 (11)**, 1031–1035, URL http://dx.doi.org/10.1038/nclimate2387.

Long, D. and M. Collins, 2013: Quantifying global climate feedbacks, responses and forcing under abrupt and gradual CO_2 forcing. *Climate Dynamics*, **41 (9-10)**, 2471–2479, URL http://dx.doi.org/10.1007/s00382-013-1677-0.

Long, S.-M., S.-P. Xie, X.-T. Zheng, and Q. Liu, 2014: Fast and Slow Responses to Global Warming: Sea Surface Temperature and Precipitation Patterns. *Journal of Climate*, **27 (1)**, 285–299, URL http://dx.doi.org/10.1175/JCLI-D-13-00297.1.

Lowe, J. A. and J. M. Gregory, 2006: Understanding projections of sea level rise in a Hadley Centre coupled climate model. *Journal of Geophysical Research: Oceans*, **111 (C11)**, URL http://dx.doi.org/10.1029/2005JC003421.

Ma, J. and S.-P. Xie, 2013: Regional Patterns of Sea Surface Temperature Change: A Source of Uncertainty in Future Projections of Precipitation and Atmospheric Circulation. *Journal of Climate*, **26 (8)**, 2482–2501, URL http://dx.doi.org/10.1175/JCLI-D-12-00283.1.

MacKay, R. M. and M. K. W. Ko, 1997: Normal modes and the transient response of the climate system. *Geophysical Research Letters*, **24 (5)**, 559–562, URL http://dx.doi.org/10.1029/97GL00286.

MacMynowski, D. G., H.-J. Shin, and K. Caldeira, 2011: The frequency response of temperature and precipitation in a climate model. *Geophysical Research Letters*, **38 (16)**, URL http://dx.doi.org/10.1029/2011GL048623.

Manabe, S. and K. Bryan, 1985: CO_2-induced change in a coupled ocean-atmosphere model and its paleoclimatic implications. *Journal of Geophysical Research: Oceans*, **90 (C6)**, 11 689–11 707, URL http://dx.doi.org/10.1029/JC090iC06p11689.

Manabe, S. and R. J. Stouffer, 1994: Multiple-Century Response of a Coupled Ocean-Atmosphere Model to an Increase of Atmospheric Carbon Dioxide. *Journal of Climate*, **7 (1)**, 5–23, URL http://dx.doi.org/10.1175/1520-0442(1994)007<0005:MCROAC>2.0.CO;2.

Manabe, S., R. J. Stouffer, M. J. Spelman, and K. Bryan, 1991: Transient Responses of a Coupled Ocean Atmosphere Model to Gradual Changes of Atmospheric CO_2. Part I. Annual Mean Response. *Journal of Climate*, **4 (8)**, 785–818.

Manabe, S. and R. T. Wetherald, 1967: Thermal Equilibrium of the Atmosphere with a Given Distribution of Relative Humidity. *Journal of the Atmospheric Sciences*, **24 (3)**, 241–259, URL http://dx.doi.org/10.1175/1520-0469(1967)024<0241:TEOTAW>2.0.CO;2.

Manabe, S. and R. T. Wetherald, 1975: The Effects of Doubling the CO_2 Concentration on the climate of a General Circulation Model. *Journal of the Atmospheric Sciences*, **32 (1)**, 3–15, URL http://dx.doi.org/10.1175/1520-0469(1975)032<0003:TEODTC>2.0.CO;2.

Marotzke, J. and P. M. Forster, 2015: Forcing, feedback and internal variability in global temperature trends. *Nature*, **517 (7536)**, 565–570, URL http://dx.doi.org/10.1038/nature14117.

Marshall, J., J. R. Scott, K. C. Armour, J.-M. Campin, M. Kelley, and A. Romanou, 2014: The ocean's role in the transient response of climate to abrupt greenhouse gas forcing. *Climate Dynamics*, **44 (7)**, 2287–2299, URL http://dx.doi.org/10.1007/s00382-014-2308-0.

Martinez-Boti, M. A., et al., 2015: Plio-pleistocene climate sensitivity evaluated using high-resolution co2 records. *Nature*, **518 (7537)**, 49–54, URL http://dx.doi.org/10.1038/nature14145.

Masson-Delmotte, V., et al., 2006: Past and future polar amplification of climate change: climate model intercomparisons and ice-core constraints. *Climate Dynamics*, **26 (5)**, 513–529, URL http://dx.doi.org/10.1007/s00382-005-0081-9.

Mauritsen, T., 2016: Global warming: Clouds cooled the Earth. *Nature Geosci*, **9 (12)**, 865–867, URL http://dx.doi.org/10.1038/ngeo2838.

McDougall, T. J., D. R. Jackett, D. G. Wright, and R. Feistel, 2003: Accurate and Computationally Efficient Algorithms for Potential Temperature and Density of Seawater. *Journal of Atmospheric and Oceanic Technology*, **20 (5)**, 730–741, URL http://dx.doi.org/10.1175/1520-0426(2003)20<730:AACEAF>2.0.CO;2.

Meehl, G. A., W. M. Washington, W. D. Collins, J. M. Arblaster, A. Hu, L. E. Buja, W. G. Strand, and H. Teng, 2005: How Much More Global Warming and Sea Level Rise? *Science*, **307 (5716)**, 1769–1772, URL http://science.sciencemag.org/content/307/5716/1769.

Meehl, G. A., et al., 2007: Global climate projections. *In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Meinshausen, M., N. Meinshausen, W. Hare, S. C. B. Raper, K. Frieler, R. Knutti, D. J. Frame, and M. R. Allen, 2009: Greenhouse-gas emission targets for limiting global warming to 2°C. *Nature*, **458 (7242)**, 1158–1162, URL http://dx.doi.org/10.1038/nature08017.

Melet, A. and B. Meyssignac, 2015: Explaining the spread in global mean thermosteric sea level rise in CMIP5 climate models. *Journal of Climate*, URL http://dx.doi.org/10.1175/JCLI-D-15-0200.1.

Meraner, K., T. Mauritsen, and A. Voigt, 2013: Robust increase in equilibrium climate sensitivity under global warming. *Geophysical Research Letters*, **40 (22)**, 5944–5948, URL http://dx.doi.org/10.1002/2013GL058118.

Merlis, T. M., 2015: Direct weakening of tropical circulations from masked co_2 radiative forcing. *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1508268112, URL http://www.pnas.org/content/early/2015/10/06/1508268112.abstract.

Moss, R. H., et al., 2010: The next generation of scenarios for climate change research and assessment. *Nature*, **463 (7282)**, 747–756, URL http://dx.doi.org/10.1038/nature08823.

Murphy, J. M., 1995: Transient response of the hadley centre coupled ocean-atmosphere model to increasing carbon dioxide. part 1: Control climate and flux adjustment. *Journal of Climate*, **8 (1)**, 36–56, URL http://dx.doi.org/10.1175/1520-0442(1995)008<0036:TROTHC>2.0.CO;2.

North, G. R., R. F. Cahalan, and J. A. Coakley, 1981: Energy balance climate models. *Reviews of Geophysics*, **19 (1)**, 91–121, URL http://dx.doi.org/10.1029/RG019i001p00091.

Opsteegh, J., R. Haarsma, F. Selten, and A. Kattenberg, 2011: ECBILT: a dynamic alternative to mixed boundary conditions in ocean models. *Tellus A*, **50 (3)**, URL http://www.tellusa.net/index.php/tellusa/article/view/14524.

Otto, A., et al., 2013: Energy budget constraints on climate response. *Nature Geosci*, **6 (6)**, 415–416, URL http://dx.doi.org/10.1038/ngeo1836.

Palmer, T. N., F. J. Doblas-Reyes, A. Weisheimer, and M. J. Rodwell, 2008: Toward Seamless Prediction: Calibration of Climate Change Projections Using Seasonal Forecasts. *Bulletin of the American Meteorological Society*, **89 (4)**, 459–470, URL http://dx.doi.org/10.1175/BAMS-89-4-459.

Palter, J. B., S. M. Griffies, B. L. Samuels, E. D. Galbraith, A. Gnanadesikan, and A. Klocker, 2014: The Deep Ocean Buoyancy Budget and Its Temporal Variability. *Journal of Climate*, **27 (2)**, 551–573, URL http://dx.doi.org/10.1175/JCLI-D-13-00016.1.

Pardaens, A. K., J. A. Lowe, S. Brown, R. J. Nicholls, and D. de Gusmão, 2011: Sea-level rise and impacts projections under a future scenario with large greenhouse gas emission reductions. *Geophysical Research Letters*, **38 (12)**, URL http://dx.doi.org/10.1029/2011GL047678.

Paynter, D. and T. L. Frölicher, 2015: Sensitivity of radiative forcing, ocean heat uptake, and climate feedback to changes in anthropogenic greenhouse gases and aerosols. *Journal of Geophysical Research: Atmospheres*, **120 (19)**, 9837–9854, URL http://dx.doi.org/10.1002/2015JD023364.

Praetorius, S. K., M. A. A. Rugenstein, and K. Caldeira, in prep: North Pacific Ocean warming enhances Arctic amplification.

Previdi, M., et al., 2013: Climate sensitivity in the Anthropocene. *Quarterly Journal of the Royal Meteorological Society*, **139 (674)**, 1121–1131, URL http://dx.doi.org/10.1002/qj.2165.

Qu, X., A. Hall, S. A. Klein, and P. M. Caldwell, 2014: On the spread of changes in marine low cloud cover in climate model simulations of the 21st century. *Climate Dynamics*, **42 (9)**, 2603–2626, doi:10.1007/s00382-013-1945-z, URL http://dx.doi.org/10.1007/s00382-013-1945-z.

Randerson, J. T., K. Lindsay, E. Munoz, W. Fu, J. K. Moore, F. M. Hoffman, N. M. Mahowald, and S. C. Doney, 2015: Multicentury changes in ocean and land contributions to the climate-carbon feedback. *Global Biogeochemical Cycles*, **29 (6)**, 744–759, URL http://dx.doi.org/10.1002/2014GB005079.

Ringer, M. A., T. Andrews, and M. J. Webb, 2014: Global-mean radiative feedbacks and forcing in atmosphere-only and coupled atmosphere-ocean climate change experiments. *Geophysical Research Letters*, **41 (11)**, 4035–4042, URL http://dx.doi.org/10.1002/2014GL060347.

Rodwell, M. J. and T. N. Palmer, 2007: Using numerical weather prediction to assess climate models. *Quarterly Journal of the Royal Meteorological Society*, **133 (622)**, 129–146, URL http://dx.doi.org/10.1002/qj.23.

Roe, G., 2009: Feedbacks, timescales, and seeing red. *Annual Review of Earth and Planetary Sciences*, **37 (1)**, 93–115, URL http://dx.doi.org/10.1146/annurev.earth.061008.134734.

Roe, G. H. and M. B. Baker, 2007: Why Is Climate Sensitivity So Unpredictable? *Science*, **318 (5850)**, 629–632, doi:10.1126/science.1144735, URL http://science.sciencemag.org/content/318/5850/629.

Roe, G. H., N. Feldl, K. C. Armour, Y.-T. Hwang, and D. M. W. Frierson, 2015: The remote impacts of climate feedbacks on regional climate predictability. *Nature Geosci*, **8 (2)**, 135–139, URL http://dx.doi.org/10.1038/ngeo2346.

Rogelj, J., M. Meinshausen, and R. Knutti, 2012: Global warming under old and new scenarios using IPCC climate sensitivity range estimates. *Nature Clim. Change*, **2 (4)**, 248–253, URL http://dx.doi.org/10.1038/nclimate1385.

Rogelj, J., M. Meinshausen, J. Sedláček, and R. Knutti, 2014: Implications of potentially lower climate sensitivity on climate projections and policy. *Environmental Research Letters*, **9 (3)**, 031 003, URL http://stacks.iop.org/1748-9326/9/i=3/a=031003.

Rogelj, J., et al., 2011: Emission pathways consistent with a 2C global temperature limit. *Nature Clim. Change*, **1 (8)**, 413–418, URL http://dx.doi.org/10.1038/nclimate1258.

Rohling, E. J., et al., 2012: Making sense of palaeoclimate sensitivity. *Nature*, **491 (7426)**, 683–691, URL http://dx.doi.org/10.1038/nature11574.

Rose, B. E. J., 2015: Stable "Waterbelt" climates controlled by tropical ocean heat transport: A nonlinear coupled climate mechanism of relevance to Snowball Earth. *Journal of Geophysical Research: Atmospheres*, **120 (4)**, 1404–1423, URL http://dx.doi.org/10.1002/2014JD022659.

Rose, B. E. J., K. C. Armour, D. S. Battisti, N. Feldl, and D. D. B. Koll, 2014: The dependence of transient climate sensitivity and radiative feedbacks on the spatial pattern of ocean heat uptake. *Geophysical Research Letters*, **41 (3)**, 1071–1078, URL http://dx.doi.org/10.1002/2013GL058955.

Rose, B. E. J. and M. C. Rencurrel, 2016: The Vertical Structure of Tropospheric Water Vapor: Comparing Radiative and Ocean-Driven Climate Changes. *Journal of Climate*, **29 (11)**, 4251–4268, URL http://dx.doi.org/10.1175/JCLI-D-15-0482.1.

Rowlands, D. J., et al., 2012: Broad range of 2050 warming from an observationally constrained large climate model ensemble. *Nature Geosci*, **5 (4)**, 256–260, URL http://dx.doi.org/10.1038/ngeo1430.

Rugenstein, M. A. A., J. M. Gregory, N. Schaller, J. Sedláček, and R. Knutti, 2016: Multiannual Ocean–Atmosphere Adjustments to Radiative Forcing. *Journal of Climate*, **29 (15)**, 5643–5659, URL http://dx.doi.org/10.1175/JCLI-D-16-0312.1.

Rugenstein, M. A. A., M. Winton, R. J. Stouffer, S. M. Griffies, and R. Hallberg, 2013: Northern High-Latitude Heat Budget Decomposition and Transient Warming. *Journal of Climate*, **26 (2)**, 609–621, URL http://dx.doi.org/10.1175/JCLI-D-11-00695.1.

Sallée, J. B., E. Shuckburgh, N. Bruneau, A. J. S. Meijers, T. J. Bracegirdle, Z. Wang, and T. Roy, 2013: Assessment of Southern Ocean water mass circulation and characteristics in CMIP5 models: Historical bias and forcing response. *Journal of Geophysical Research: Oceans*, **118 (4)**, 1830–1844, URL http://dx.doi.org/10.1002/jgrc.20135.

Sanderson, B. M., R. Knutti, and P. Caldwell, 2015: A Representative Democracy to reduce interdependency in a multi-model ensemble. *Journal of Climate*, URL http://dx.doi.org/10.1175/JCLI-D-14-00362.1.

Savin, S. M., 1977: The History of the Earth's Surface Temperature During the Past 100 Million Years. *Annual Review of Earth and Planetary Sciences*, **5 (1)**, 319–355, URL http://dx.doi.org/10.1146/annurev.ea.05.050177.001535.

Schaller, N., J. Cermak, M. Wild, and R. Knutti, 2013: The sensitivity of the modeled energy budget and hydrological cycle to CO_2 and solar forcing. *Earth System Dynamics*, **4 (2)**, 253–266, doi:10.5194/esd-4-253-2013, URL http://www.earth-syst-dynam.net/4/253/2013/.

Schlesinger, M. E., 1986: Equilibrium and transient climatic warming induced by increased atmospheric CO_2. *Climate Dynamics*, **1 (1)**, 35–51, URL http://dx.doi.org/10.1007/BF01277045.

Schmidt, G. A., D. T. Shindell, and K. Tsigaridis, 2014: Reconciling warming trends. *Nature Geosci*, **7 (3)**, 158–160, URL http://dx.doi.org/10.1038/ngeo2105.

Schneider, S. H. and S. L. Thompson, 1981: Atmospheric CO_2 and climate: Importance of the transient response. *Journal of Geophysical Research: Oceans*, **86 (C4)**, 3135–3147, URL http://dx.doi.org/10.1029/JC086iC04p03135.

Schneider von Deimling, T., H. Held, A. Ganopolski, and S. Rahmstorf, 2006: Climate sensitivity estimated from ensemble simulations of glacial climate. *Climate Dynamics*, **27 (2)**, 149–163, URL http://dx.doi.org/10.1007/s00382-006-0126-8.

Sellers, W. D., 1969: A Global Climatic Model Based on the Energy Balance of the Earth-Atmosphere System. *Journal of Applied Meteorology*, **8 (3)**, 392–400, URL http://dx.doi.org/10.1175/1520-0450(1969)008<0392:AGCMBO>2.0.CO;2.

Senior, C. A. and J. F. B. Mitchell, 2000: The time-dependence of climate sensitivity. *Geophysical Research Letters*, **27 (17)**, 2685–2688, URL http://dx.doi.org/10.1029/2000GL011373.

Serreze, M. C., A. P. Barrett, J. C. Stroeve, D. N. Kindig, and M. M. Holland, 2009: The emergence of surface-based Arctic amplification. *The Cryosphere*, **3 (1)**, 11–19, doi:10.5194/tc-3-11-2009, URL http://www.the-cryosphere.net/3/11/2009/.

Shell, K. M., 2013: Consistent Differences in Climate Feedbacks between Atmosphere–Ocean GCMs and Atmospheric GCMs with Slab-Ocean Models. *Journal of Climate*, **26 (12)**, 4264–4281, URL http://dx.doi.org/10.1175/JCLI-D-12-00519.1.

Sherwood, S. and Q. Fu, 2014: A Drier Future? *Science*, **343 (6172)**, 737–739, URL http://www.sciencemag.org/content/343/6172/737.short.

Sherwood, S. C., S. Bony, O. Boucher, C. Bretherton, P. M. Forster, J. M. Gregory, and B. Stevens, 2014a: Adjustments in the forcing-feedback framework for understanding climate change. *Bulletin of the American Meteorological Society*, URL http://dx.doi.org/10.1175/BAMS-D-13-00167.1.

Sherwood, S. C., S. Bony, and J.-L. Dufresne, 2014b: Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, **505 (7481)**, 37–42, URL http://dx.doi.org/10.1038/nature12829.

Shine, K. P., J. Cook, E. J. Highwood, and M. M. Joshi, 2003: An alternative to radiative forcing for estimating the relative importance of climate change mechanisms. *Geophysical Research Letters*, **30 (20)**, URL http://dx.doi.org/10.1029/2003GL018141.

Shine, K. P., R. Derwent, D. Wuebbles, and J.-J. Morcrette, 1990: Radiative Forcing of Climate. *Climate Change: The IPCC Scientific Assessment*, J. Houghton, G. J. Jenkins, and J. J. Ephraums, Eds., Cambridge University Press.

Siegenthaler, U. and H. Oeschger, 1984: Transient temperature changes due to increasing CO_2 using simple models. *Annals of Glaciology*, **4**, 153–159.

Skeie, R. B., T. Berntsen, M. Aldrin, M. Holden, and G. Myhre, 2014: A lower and more constrained estimate of climate sensitivity using updated observations and detailed radiative forcing time series. *Earth System Dynamics*, **5 (1)**, 139–175, doi:10.5194/esd-5-139-2014, URL http://www.earth-syst-dynam.net/5/139/2014/.

Smith, R. S., R. Sutton, and J. M. Gregory, 2014: The impact of salinity perturbations on the future uptake of heat by the Atlantic Ocean. *Geophysical Research Letters*, URL http://dx.doi.org/10.1002/2014GL062169.

Soden, B. J. and I. M. Held, 2006: An Assessment of Climate Feedbacks in Coupled Ocean-Atmosphere Models. *Journal of Climate*, **19 (14)**, 3354–3360, URL http://dx.doi.org/10.1175/JCLI3799.1.

Soden, B. J., R. T. Wetherald, G. L. Stenchikov, and A. Robock, 2002: Global Cooling After the Eruption of Mount Pinatubo: A Test of Climate Feedback by Water Vapor. *Science*, **296 (5568)**, 727–730, URL http://science.sciencemag.org/content/296/5568/727.

Sokolov, A. P., et al., 2009: Probabilistic Forecast for Twenty-First-Century Climate Based on Uncertainties in Emissions (Without Policy) and Climate Parameters. *Journal of Climate*, **22 (19)**, 5175–5204, URL http://dx.doi.org/10.1175/2009JCLI2863.1.

Staten, P. W., T. Reichler, and J. Lu, 2014: The transient circulation response to radiative forcings and sea surface warming. *Journal of Climate*, URL http://dx.doi.org/10.1175/JCLI-D-14-00035.1.

Stevens, B. and S. Bony, 2013: What Are Climate Models Missing? *Science*, **340 (6136)**, 1053–1054, URL http://science.sciencemag.org/content/340/6136/1053.

Stevens, B., S. C. Sherwood, S. Bony, and M. J. Webb, 2016: Prospects for Narrowing Bounds on Earth's Equilibrium Climate Sensitivity. *Earth's Future*, URL http://dx.doi.org/10.1002/2016EF000376.

Stommel, H., 1961: Thermohaline Convection with Two Stable Regimes of Flow. *Tellus*, **13 (2)**, 224–230, URL http://dx.doi.org/10.1111/j.2153-3490.1961.tb00079.x.

Stott, P. A. and J. A. Kettleborough, 2002: Origins and estimates of uncertainty in predictions of twenty-first century temperature rise. *Nature*, **416 (6882)**, 723–726, URL http://dx.doi.org/10.1038/416723a.

Stouffer, R. and S. Manabe, 2003: Equilibrium response of thermohaline circulation to large changes in atmospheric CO_2 concentration. *Climate Dynamics*, **20 (7-8)**, 759–773, URL http://dx.doi.org/10.1007/s00382-002-0302-4.

Stouffer, R. J., 2004: Time Scales of Climate Response. *Journal of Climate*, **17 (1)**, 209–217, URL http://dx.doi.org/10.1175/1520-0442(2004)017<0209:TSOCR>2.0.CO;2.

Stouffer, R. J., J. Russel, and M. J. Spelman, 2006: Importance of oceanic heat uptake in transient climate change. *Geophys. Res. Lett.*, **33**, L17 704.

Su, H., J. H. Jiang, C. Zhai, T. J. Shen, J. D. Neelin, G. L. Stephens, and Y. L. Yung, 2014: Weakening and strengthening structures in the hadley circulation change under global warming and implications for cloud response and climate sensitivity. *Journal of Geophysical Research: Atmospheres*, **119 (10)**, 5787–5805, doi:10.1002/2014JD021642, URL http://dx.doi.org/10.1002/2014JD021642.

Tan, I., T. Storelvmo, and M. D. Zelinka, 2016: Observational constraints on mixed-phase clouds imply higher climate sensitivity. *Science*, **352 (6282)**, 224–227, doi:10.1126/science.aad5300, URL http://science.sciencemag.org/content/352/6282/224, http://science.sciencemag.org/content/352/6282/224.full.pdf.

Tebaldi, C. and J. M. Arblaster, 2014: Pattern scaling: Its strengths and limitations, and an update on the latest model simulations. *Climatic Change*, **122 (3)**, 459–471, URL http://dx.doi.org/10.1007/s10584-013-1032-9.

Tian, B., 2015: Spread of model climate sensitivity linked to double-intertropical convergence zone bias. *Geophysical Research Letters*, **42 (10)**, 4133–4141, doi:10.1002/2015GL064119, URL http://dx.doi.org/10.1002/2015GL064119.

Tomassini, L., R. Knutti, G.-K. Plattner, D. P. van Vuuren, T. F. Stocker, R. B. Howarth, and M. E. Borsuk, 2010: Uncertainty and risk in climate projections for the 21st century: comparing mitigation to non-intervention scenarios. *Climatic Change*, **103 (3)**, 399–422, URL http://dx.doi.org/10.1007/s10584-009-9763-3.

Tomassini, L., P. Reichert, R. Knutti, T. F. Stocker, and M. E. Borsuk, 2007: Robust Bayesian Uncertainty Analysis of Climate System Properties Using Markov Chain Monte Carlo Methods. *Journal of Climate*, **20 (7)**, 1239–1254, URL http://dx.doi.org/10.1175/JCLI4064.1.

Tomassini, L., et al., 2013: The respective roles of surface temperature driven feedbacks and tropospheric adjustment to co_2 in CMIP5 transient climate simulations. *Climate Dynamics*, **41 (11-12)**, 3103–3126, URL http://dx.doi.org/10.1007/s00382-013-1682-3.

Trenberth, K. E. and J. T. Fasullo, 2013: An apparent hiatus in global warming? *Earth's Future*, **1 (1)**, 19–32, URL http://dx.doi.org/10.1002/2013EF000165.

Trossman, D. S., J. B. Palter, T. M. Merlis, Y. Huang, and Y. Xia, 2016: Large-scale ocean circulation-cloud interactions reduce the pace of transient climate change. *Geophysical Research Letters*, URL http://dx.doi.org/10.1002/2016GL067931.

Vial, J., J.-L. Dufresne, and S. Bony, 2013: On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates. *Climate Dynamics*, **41 (11-12)**, 3339–3362, URL http://dx.doi.org/10.1007/s00382-013-1725-9.

Voigt, S., A. S. Gale, and S. Flögel, 2004: Midlatitude shelf seas in the Cenomanian-Turonian greenhouse world: Temperature evolution and North Atlantic circulation. *Paleoceanography*, **19 (4)**, URL http://dx.doi.org/10.1029/2004PA001015.

von der Heydt, A. S., P. Köhler, R. S. W. van de Wal, and H. A. Dijkstra, 2014: On the state dependency of fast feedback processes in (paleo) climate sensitivity. *Geophysical Research Letters*, **41 (18)**, 6484–6492, URL http://dx.doi.org/10.1002/2014GL061121.

Watanabe, M., H. Shiogama, M. Yoshimori, T. Ogura, T. Yokohata, H. Okamoto, S. Emori, and M. Kimoto, 2012: Fast and slow timescales in the tropical low-cloud response to increasing CO_2 in two climate models. *Climate Dynamics*, **39 (7-8)**, 1627–1641, URL http://dx.doi.org/10.1007/s00382-011-1178-y.

Webb, M., F. Lambert, and J. Gregory, 2013: Origins of differences in climate sensitivity, forcing and feedback in climate models. *Climate Dynamics*, **40 (3-4)**, 677–707, URL http://dx.doi.org/10.1007/s00382-012-1336-x.

Wigley, T. M. L., 2005: The Climate Change Commitment. *Science*, **307 (5716)**, 1766–1769, doi:10.1126/science.1103934, URL http://www.sciencemag.org/content/307/5716/1766.abstract.

Wigley, T. M. L. and S. C. B. Raper, 1990: Natural variability of the climate system and detection of the greenhouse effect. *Nature*, **344 (6264)**, 324–327, URL http://dx.doi.org/10.1038/344324a0.

Wigley, T. M. L. and M. E. Schlesinger, 1985: Analytical solution for the effect of increasing co2 on global mean temperature. *Nature*, **315 (6021)**, 649–652, URL http://dx.doi.org/10.1038/315649a0.

Williams, K. D., W. J. Ingram, and J. M. Gregory, 2008: Time Variation of Effective Climate Sensitivity in GCMs. *Journal of Climate*, **21 (19)**, 5076–5090, URL http://dx.doi.org/10.1175/2008JCLI2371.1.

Winton, M., W. G. Anderson, T. L. Delworth, S. M. Griffies, W. J. Hurlin, and A. Rosati, 2014: Has Coarse Ocean Resolution Biased Simulations of Transient Climate Sensitivity? *Geophysical Research Letters*, URL http://dx.doi.org/10.1002/2014GL061523.

Winton, M., S. M. Griffies, B. L. Samuels, J. L. Sarmiento, and T. L. Frölicher, 2013: Connecting changing ocean circulation with changing climate. *Journal of Climate*, **26 (7)**, 2268–2278, doi:10.1175/JCLI-D-12-00296.1, URL http://dx.doi.org/10.1175/JCLI-D-12-00296.1.

Winton, M., K. Takahashi, and I. M. Held, 2010: Importance of ocean heat uptake efficacy to transient climate change. *Journal of Climate*, **23 (9)**, 2333–2344, doi:10.1175/2009JCLI3139.1, URL http://dx.doi.org/10.1175/2009JCLI3139.1, http://dx.doi.org/10.1175/2009JCLI3139.1.

Wu, Y., R. Seager, M. Ting, N. Naik, and T. A. Shaw, 2011: Atmospheric Circulation Response to an Instantaneous Doubling of Carbon Dioxide. Part I: Model Experiments and Transient Thermal Response in the Troposphere*. *Journal of Climate*, **25 (8)**, 2862–2879, doi:10.1175/JCLI-D-11-00284.1, URL http://dx.doi.org/10.1175/JCLI-D-11-00284.1.

Wunsch, C. and P. Heimbach, 2014: Bidecadal Thermal Changes in the Abyssal Ocean. *Journal of Physical Oceanography*, **44 (8)**, 2013–2030, URL http://dx.doi.org/10.1175/JPO-D-13-096.1.

Wyant, M. C., C. S. Bretherton, P. N. Blossey, and M. Khairoutdinov, 2012: Fast cloud adjustment to increasing CO_2 in a superparameterized climate model. *Journal of Advances in Modeling Earth Systems*, **4 (2)**, URL http://dx.doi.org/10.1029/2011MS000092.

Xie, S.-P., C. Deser, G. A. Vecchi, J. Ma, H. Teng, and A. T. Wittenberg, 2010: Global Warming Pattern Formation: Sea Surface Temperature and Rainfall. *Journal of Climate*, **23 (4)**, 966–986, URL http://dx.doi.org/10.1175/2009JCLI3329.1.

Yamamoto, A., A. Abe-Ouchi, M. Shigemitsu, A. Oka, K. Takahashi, R. Ohgaito, and Y. Yamanaka, 2015: Global deep ocean oxygenation by enhanced ventilation in the Southern Ocean under long-term global warming. *Global Biogeochemical Cycles*, **29 (10)**, 1801–1815, URL http://dx.doi.org/10.1002/2015GB005181.

Yan, X.-H., T. Boyer, K. Trenberth, T. R. Karl, S.-P. Xie, V. Nieves, K.-K. Tung, and D. Roemmich, 2016: The global warming hiatus: Slowdown or redistribution? *Earth's Future*, **4 (11)**, 472–482, URL http://dx.doi.org/10.1002/2016EF000417.

Yin, J., 2012: Century to multi-century sea level rise projections from CMIP5 models. *Geophysical Research Letters*, **39 (17)**, URL http://dx.doi.org/10.1029/2012GL052947.

Yin, J., J. T. Overpeck, S. M. Griffies, A. Hu, J. L. Russell, and R. J. Stouffer, 2011: Different magnitudes of projected subsurface ocean warming around Greenland and Antarctica. *Nature Geosci*, **4 (8)**, 524–528, URL http://dx.doi.org/10.1038/ngeo1189.

Yokohata, T., S. Emori, T. Nozawa, Y. Tsushima, T. Ogura, and M. Kimoto, 2005: Climate response to volcanic forcing: Validation of climate sensitivity of a coupled atmosphere-ocean general circulation model. *Geophysical Research Letters*, **32 (21)**, URL http://dx.doi.org/10.1029/2005GL023542.

Zachos, J. C., M. Pagani, L. C. Sloan, E. Thomas, and K. Billups, 2001: Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. *Science*, **292 (5517)**, 686–693, doi:10.1126/science.1059412, URL http://www.sciencemag.org/content/292/5517/686.abstract.

Zaliapin, I. and M. Ghil, 2010: Another look at climate sensitivity. *Nonlinear Processes in Geophysics*, **17 (2)**, 113–122, doi:10.5194/npg-17-113-2010, URL http://www.nonlin-processes-geophys.net/17/113/2010/.

Zeebe, R. E., 2013: Time-dependent climate sensitivity and the legacy of anthropogenic greenhouse gas emissions. *Proceedings of the National Academy of Sciences*, **110 (34)**, 13 739–13 744, URL http://www.pnas.org/content/110/34/13739.abstract.

Zelinka, M. D., S. A. Klein, K. E. Taylor, T. Andrews, M. J. Webb, J. M. Gregory, and P. M. Forster, 2013: Contributions of Different Cloud Types to Feedbacks and Rapid Adjustments in CMIP5. *Journal of Climate*, **26 (14)**, 5007–5027, URL http://dx.doi.org/10.1175/JCLI-D-12-00555.1.

Zhang, R., S. M. Kang, and I. M. Held, 2010: Sensitivity of Climate Change Induced by the Weakening of the Atlantic Meridional Overturning Circulation to Cloud Feedback. *Journal of Climate*, **23 (2)**, 378–389, URL http://dx.doi.org/10.1175/2009JCLI3118.1.

Zhou, C., M. D. Zelinka, and S. A. Klein, 2016: Impact of decadal cloud variations on the Earth/'s energy budget. *Nature Geosci*, **9 (12)**, 871–874, URL http://dx.doi.org/10.1038/ngeo2828.

Zhu, J., Z. Liu, J. Zhang, and W. Liu, 2014: AMOC response to global warming: dependence on the background climate and response timescale. *Climate Dynamics*, **44 (11)**, 3449–3468, URL http://dx.doi.org/10.1007/s00382-014-2165-x.

Zickfeld, K., et al., 2013: Long-Term Climate Change Commitment and Reversibility: An EMIC Intercomparison. *Journal of Climate*, **26 (16)**, 5782–5809, URL http://dx.doi.org/10.1175/JCLI-D-12-00584.1.

Curriculum Vitae

maria.rugenstein@env.ethz.ch | website | Google scholar

Education

PhD Candidate | ETH Zürich	2014 – 2017
Visiting researcher | Stanford Carnegie Department for Global Ecology, Stanford	2015 – 2016
Master of Science	
joint at ETH Zürich and Princeton University	2009 – 11
Study abroad | UK British Antarctic Survey, Cambridge	2011
Bachelor in Environmental Sciences | ETH Zürich	2006 – 09
Study abroad | Norway NTNU, Trondheim	2009

Professional and teaching experience

Research assistant

- Utrecht University, Netherlands 2011 – 2013
- British Antarctic Survey, Cambridge 2011
- Institute for Environmental decisions, ETH Zürich 2008 – 09

Teaching assistant

- *Climate Change Uncertainty and Risk* ETH Zürich 14 & 15 & 16
- *Earth- and Production Systems* ETH Zürich 2014 & 2015
- *Ocean and Climate* Utrecht University 2012 & 13
- *Introduction to Environmental Systems* ETH Zürich 2007 – 08
- *Introduction to Political Economics* ETH Zürich 2006 – 07

Fellowships and Awards

SNF PostDoc Mobility fellowhip (2017) | **Alexander von Humboldt PostDoc fellowship** (2017) | **AGU OSPA poster prize** (2014) | **Audience and panel prize** for best presentation at Buys Ballot symposium (2012) | **ETH medal for Master's thesis** (2012) | **German National Academic Foundation** student and PhD fellowships (2006–2013) | **NOAA Geophysical Fluid Dynamics Laboratory** sponsored visiting student fees of

Princeton University (2010 – 11) | **British Antarctic Survey, ResClim/Norway** and **European Consortium for Ocean Drilling Research**, each funded participation and travel to summer schools (2011 & 12)

Publications and Presentations

- Knutti, **Rugenstein**, Hegerl, 2017: *Beyond Climate Sensitivity*, Nature Geoscience, accepted
- He, Winton, Vecchi, Jia, **Rugenstein**, 2016: *Transient climate sensitivity depends on base climate ocean circulation*, J. Climate, 30, 1493-1504
- **Rugenstein**, Caldeira, Knutti, 2016: *Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes*, GRL, 43,18
- **Rugenstein**, Gregory, Schaller, Sedláček, Knutti, 2016: *Multi-annual ocean-atmosphere adjustments to radiative forcing* J. Climate, 29, 5643-5659
- **Rugenstein**, Sedláček, Knutti, 2016: *Nonlinearities in patterns of long-term ocean warming*, GRL, 43, 7
- Knutti and **Rugenstein**, 2015: *Feedbacks, climate sensitivity and the limits of linear models* Phil. Trans. R. Soc. A 373: 20150146
- **Rugenstein**, Stocchi, von der Heydt, Dijkstra, Brinkhuis, 2014: *Emplacement of Antarctic ice sheet mass affects circumpolar ocean flow*, Gbl.& Plan. Change 118, 16-24
- **Rugenstein**, Winton, Stouffer, Griffies, Hallberg, 2013: *Northern High-Latitude Heat Budget Decomposition and Transient Warming*, J. Climate, 26, 609 – 621

Scientific talks
AGU Fall Meeting San Francisco (2016 inv., 2015 & 2012) | Lawrence Livermore Seminar (2017 inv.) | MPI Joint Seminar Hamburg (2017 inv.) | Carnegie Internal Seminar Stanford (2016) | EGU meeting Vienna (2015) | Internal Seminar Noah Diffenbaugh Stanford (2014) | Ocean Heat Uptake Workshop Southampton (2014) | Ocean Gateways conference Jerusalem (2013) | Buys Ballot Research School for Fundamental Processes in the Climate System Nijmegen (2012) | Summer School Sea ice in the climate system Svalbard (2011)

Talks on science policy
British Antarctic Service Board Meeting Cambridge (2011) | Institute for Marine and Atmospheric Science Utrecht University (2013)

Scientific posters
AGU Ocean Science meeting New Orleans (2016 (inv.)) | International Paleoceanographic Conference Utrecht (2016) | AGU fall meeting San Francisco (2016, 2015, 2014) | Latsis Conference Zürich (2015) | EGU meeting Vienna (2014) | Swiss Global Change Day Bern (2016, 2015, 2014) | Graduate Climate Conference Seattle (2014) | various summer schools (2011-2015)

Summer schools and Workshops

Extreme Events and Climate (SCSS Switzerland 2015) | Science Meets Practice (CCES Switzerland 2015) | Ocean heat uptake (Southampton 2014) | Geophysical Fluiddynamics (FDSE Paris 2013) | Paleoclimatology (ECORD Urbino 2012) | Climate Modelling (NCAS Cambridge 2011) | Antarctic Funding Initiative (BAS Cambridge 2011) | Sea ice in the climate system (ResClim Svalbard 2011) | Earth's cryosphere and sea level change (ISSI Bern 2010) | German National Academic Foundation: Environmental Physics and Remote Sensing (Greifswald 2010) | Countryside, Landscape, and Scenery (Rot 2009)

| Molecular Biophysics (Bonn & Berlin 2009) | Pseudo Science around 1900 (Marburg 2005)

Miscellaneous

Reviewer for | US National Science Foundation; Journal of Climate; Geophysical Research Letters; Palaeogeography, Palaeoclimatology, Palaeoecology; Current Climate Change Reports; Climatic Change; Science Advances

Workshops organized | Millennia scale model intercomparison MPI Hamburg (2016); Climate sensitivity summer seminar ETH Zürich (2014)

Service | PhD representative IAC ETH Zürich (2014-2015) | student representative departmental committee of teaching (2010, 2013) | Mittelbau representative at the departmental conference (2014-2016) | mentoring undergrads at ETH (2010 - 2011)

Languages | German, English, and Dutch, some French, some Norwegian