Can geoengineering be optimised?

Thomas Hornigold

Linacre College

University of Oxford

Doctor of Philosophy

Trinity Term, 2023

Contents

Acknowledgements

This was a longer and stranger journey than most, which means I have accrued a lot of people to thank. I want to thank my supervisors, Myles Allen and William Ingram, for conversations about physics, road-testing of ideas, support, and encouragement. In particular Dr Ingram deserves credit for teaching me most of what I know about HadCM3 and coming up with the initial project idea, and Professor Allen for seeing the project through to the end with me: I am very grateful for your trust and flexibility and wisdom on all things climate change. I also want to give immense thanks to Sarah Sparrow, David Wallom, Richard Millar, Andy Bowery and Meredith Li for teaching me how to use climateprediction.net and general tireless efforts on that project, and everyone on the CPDN project for volunteering their computers and making this research possible. I thank my examiners, Professor Michael Obersteiner and Dr Pete Irvine. When people told me that the viva could be enjoyable, I didn't really believe them, but much is owed to you both for making it so. I owe thanks to the GeoMIP community, including Ben Kravitz, Dan Visioni, Simone Tilmes, and Doug MacMartin for interesting discussions and ploughing the geoengineering furrow for many years, and Jim Haywood and Andy Jones of the Met Office as well. I also owe thanks to the Environmental Research DTP for my doctoral training, to NERC for funding my project, and to the staff of Linacre College and the AOPP Department in Oxford for putting up with countless administrative requests.

I also owe significant thanks to the staff of the House of Lords Committee Office for allowing and encouraging me to complete my DPhil part-time while working with them, and generally being an extremely supportive, fun and intellectually stimulating environment that makes me happy to

go into work each day. Particular thanks in this regard are owed to the Science and Technology Committee staff: George Webber, Cerise Burnett-Stuart and Matthew Manning, as well as our Chair Baroness Brown of Cambridge. We made it in the end!

The personal thanks I owe to people are far too great in number to ignore and I hope that I will send you all individual letters waxing lyrical about your divine, Incorruptible Virtues rather than embarrassing you in public (more than I already do). But I should say a special thanks to my parents and siblings for support of many different kinds during this process, and especially my Mum who had to put up with me at my most insufferable. Beyond this, my friends, fellow Jacobins, and co-conspirators, in alphabetical order: Mantas Abazorius, Diana Avadanii, Viktorija Blazheska, Mirela Bogasieru, Naomi Crisp, Amy Curran, Peigi Devlin, Adam Hawtin, Lisa Hodgkinson, Valentine Iat, G Jerwood, Dilhan Manawadu, Isabelle Naylor, Lizzie Ormerod, Joe Richardson, Adam Stanway, Lucy Valsamidis, Glenn Wagner, and the most important three or four people I forgot when hastily writing up this list who are now feeling vengeful. Without you, I would be sunk. Instead, we sail on. (Or possibly fly off like a fancy pigeon.)

Many of you kept believing in me long after I had stopped believing in myself, confirming once again that I should have listened to you all along. The ability, in a dark moment, to briefly see yourself through the eyes of a friend (and wish that they could see themselves as you can see them) is a wonderful gift. To borrow once more my favourite closing line from The First Three Minutes, it is one of the very few things that elevates the human condition above the level of farce and gives it some of the grace of a tragedy.

Abstract

Geoengineering is the intentional, large-scale manipulation of Earth's climate. It has been suggested that this could be done to counteract or ameliorate the effects of anthropogenic climate change, to reduce its negative impacts or buy time for global greenhouse gas emissions to be reduced. Stratospheric aerosol injection, where aerosols in the stratosphere are used to reflect sunlight and so cool climate, has been widely studied. Altering the altitude, latitude, or timing of aerosol injections could result in different radiative forcing patterns, which suggests there may be potential to "optimise" stratospheric aerosol injection geoengineering to achieve particular climate goals. The extent to which geoengineering could be optimised, beyond idealised studies that counteract the global mean temperature increase of greenhouse gases, has only relatively recently begun to be explored.

Chapter 1 discusses the background of geoengineering as a concept and includes a literature review and discussion of robust results that have emerged from modelling studies of geoengineering. Chapter 2 uses a combination of analytical techniques and the simple climate model, FaIR, to examine different scenarios for "peak-shaving" – temporarily using geoengineering to hold global mean temperatures below a certain threshold – and examines trade-offs between the amount of warming avoided and the implied duration of commitment to geoengineering.

Chapters 3-5 analyse simulations from the HadCM3 climate model, simulated using climateprediction.net, which uses thousands of volunteer computers to generate large ensembles of simulations with differing distributions of stratospheric aerosol optical depth counteracting an abrupt quadrupling of carbon dioxide to represent different attempts at tailoring

geoengineering. Chapter 3 details initial calibration and characterisation of the response to simple patterns of radiative forcing, and establishes that the temperature response to different patterns of forcing is, to a good approximation, linear and additive. Chapter 4 expands on this and discusses various methods for optimising for temperature and precipitation, including analysis of trade-offs for attempting to optimise temperature in different regions, and analysis of whether there are a limited number of fundamental modes of response in the HadCM3 climate model to a range of imposed radiative forcing patterns. Chapter 5 examines the impact of geoengineering on climate and weather extremes, using metrics that represent heatwaves, flooding, and dry periods, and analysing any differences in the distribution of extreme events between simulations of the preindustrial and geoengineered climates. In the final chapter, results and conclusions are summarised, and possible future work is outlined.

Chapter 1

Introduction and Literature Review

1.1 An introduction to geoengineering

Earth's climate is changing due to anthropogenic activity, predominantly the burning of fossil fuels and land use change. These activities have resulted in the emission of greenhouse gases into the atmosphere, which reduce the Earth's ability to radiate away heat, and thus cause the global mean temperature to rise. As of the IPCC's Sixth Assessment Report in 2023, it was estimated that global surface temperatures had reached 1.1 °C above 1850-1900 levels by 2011-20, and that this was unequivocally due to unsustainable energy use, land use, and land use change. [1]. The same report said that widespread adverse impacts and related losses and damages to nature and people were due to the widespread and rapid changes in the atmosphere, ocean, cryosphere and biosphere that were caused by anthropogenic climate change.

The Paris Agreement, an international treaty on climate change adopted in 2015, set out a long-term temperature goal to limit global warming to well below 2 °C and preferably limit the increase to 1.5 °C. It said that this should be achieved by rapid, near-term emissions reductions, with global greenhouse gas emissions reaching net zero by the second half of this century. [2] It has been estimated that to keep global warming to below 1.5 °C, emissons would need to be cut by 43% relative to

2019 levels by 2030. [1] Mitigation of climate change has proved difficult, despite international agreements and the stated intentions of policymakers. The UN Environment Programme reported in 2023 that global emissions hit a record high in 2022, and estimated that the world was on track for 3 ℃ of warming if current policies were continued. [3] The first Conference of Parties meeting of the United Nations Framework Convention on Climate Change was held in March 1995, a month before this author was born: since then, global CO_2 emissions have increased by more than 50%. [4]

As efforts to mitigate climate change continue, and emissions continue to rise, geoengineering - intentional, large-scale action to modify the climate system - has been suggested as a potential intervention [5], or tool to aid in mitigation [6]. The Earth is in radiative equilibrium because incoming, shortwave, solar radiation is balanced out by outgoing longwave radiation which the Earth radiates back into space. Anthropogenic climate change (to first order) occurs because greenhouse gases make the atmosphere more opaque to longwave radiation, meaning the Earth is less efficient at radiating energy into space. As the longwave energy the Earth radiates increases with temperature, the Earth heats up as a result to maintain radiative equilibrium.

Solar radiation management (SRM) attempts to balance anthropogenic greenhouse gases and reduce anthopogenic global warming by modifying the Earth's radiative budget. This includes techniques to increase Earth's albedo such as marine cloud brightening, stratospheric aerosol injection (SAI), and land albedo modification - reflecting additional shortwave radiation back into space and thus cooling the climate. Carbon dioxide removal (CDR) seeks to increase the amount of CO_2 stored in sinks. [7] It has been argued that, by contrast to mitigation which requires significant investment and time to replace fossil-fuel infrastructure, geoengineering could be deployed rapidly and at a much lower cost to stabilise the Earth's climate or reduce global temperatures. [8] [9]

Although no large-scale field trials of geoengineering have been conducted, stratospheric aerosol injection has a natural analogue in volcanic eruptions, as discussed below, which suggests that it is not implausible and could lead to cooling. It has been suggested that modified military aeroplanes could accomplish stratospheric aerosol injection at scale, on the order of teragrams of sulphate aerosols, for a cost of a few

billion dollars annually. [10] This has raised the possibility that individual nations could attempt to undertake geoengineering unilaterally, and we explore some of the tensions between regions that may arise as a result of this in Chapter 3.

Shepherd's napkin diagram, reproduced as Figure 1.1 illustrates the concept behind a temporary deployment of SRM to "shave the peak" of warming while large-scale carbon dioxide removal can reduce the temperature to below a given threshold. The intention is that this would minimise negative climate impacts while buying time for the world to reduce global greenhouse gas emissions by deploying renewable energy and low-carbon solutions for heating and transport. We explore some of the scenario space for idealised "peak-shaving" geoengineering interventions in Chapter 2. Reducing the planet's temperature through geoengineering could prevent harmful "tipping points" which may occur in the climate system due to sustained global warming, such as the destabilisation of Antarctic ice sheets, some of which may be very difficult to reverse even with emissions reductions. [11]

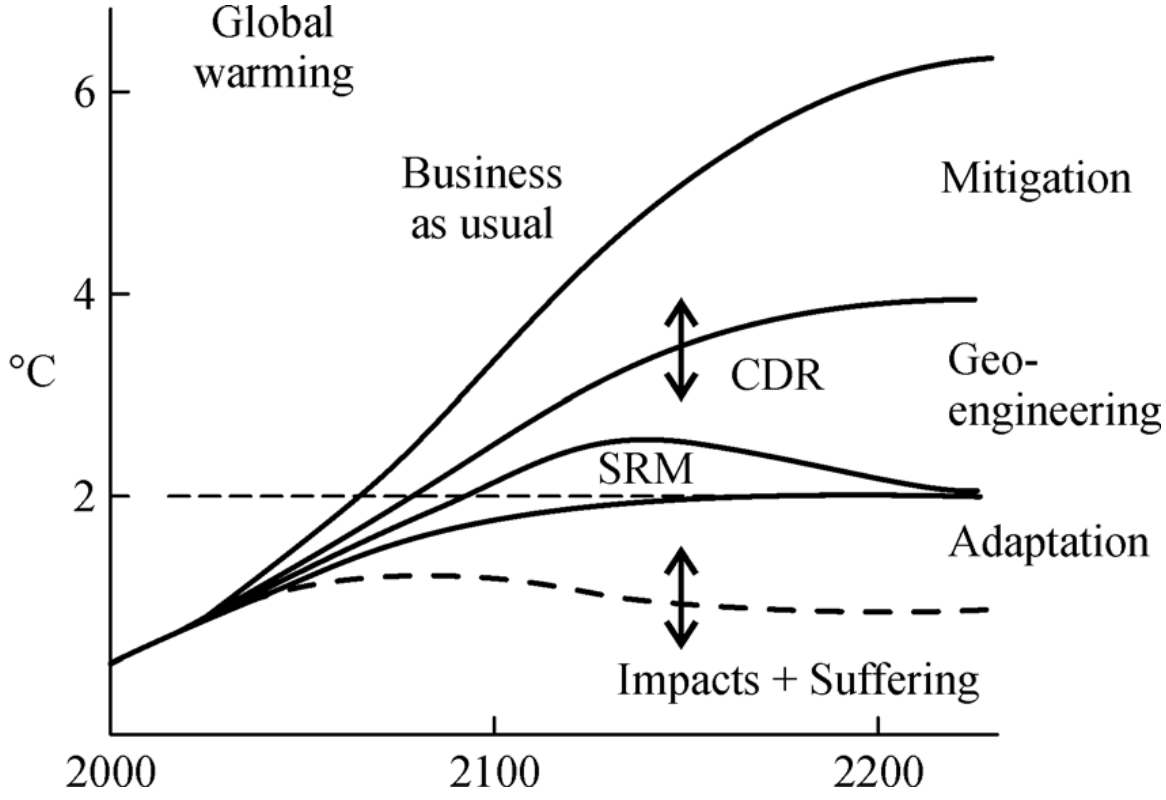

Figure 1.1: **Shepherd's 'napkin diagram', first published in Long and Shepherd (2014).**

When Paul Crutzen "broke the taboo" about geoengineering in his 2006 paper [5], he noted that there are currently significant tropospheric anthropogenic emissions of sulphate aerosols which act to cool the climate. These tropospheric aerosols cause significant harm to human health, and so there is a policy imperative to reduce these emissions. However, this would inevitably result in further global warming. He therefore framed geoengineering as a possible option to resolve the dilemma of simultaneously cleaning up air pollution while keeping global warming under control. [12] Given recent results around the impact of tighter standards on sulphur emissions in the shipping industry, which have been shown to have a warming effect, this framing has been subject to increased discussion. [13]

1.2 Stratospheric aerosol injection

I focus on SAI, first proposed by Budyko (1974, 1977) [14]. SRM studies have mostly focused on SAI, or analogously reducing the solar constant [15]. It has a clear natural analogue: evidence from historical volcanic eruptions, summarised by Robock (2000) [16] and exemplified by the Pinatubo eruption (1992), has demonstrated that aerosols injected into the stratosphere reflect incoming solar radiation and temporarily cool the climate. This effect can be comparable in magnitude to the radiative forcing and temperature changes under anthropogenic climate change. For example, after Pinatubo, radiative forcing of $-2.7 \pm 1.0 Wm^{-2}$ was observed in the global mean, rising to $-6 Wm^{-2}$ in some regions, along with a global mean cooling peaking at 0.5K, and a peak in aerosol optical depth of 0.35 [17] [18].

Volcanic injections can be substantial, injecting on the order of 10Tg of sulphates. [19] Geoengineering modelling experiments that explicitly include injection rates use similar orders of magnitude. Modelling studies suggest SRM can return global mean temperature to preindustrial levels, but the effectiveness of SRM at returning individual regions to preindustrial conditions varies, as does the magnitude and nature of the response to SRM. [20] This, combined with geopolitical concerns over deployment and the risk of a "termination shock" [21] if geoengineering is rapidly halted, are major concerns with any deployment. Models consistently indicate that, when geoengineering is terminated, temperature rapidly rebounds (within a few years)

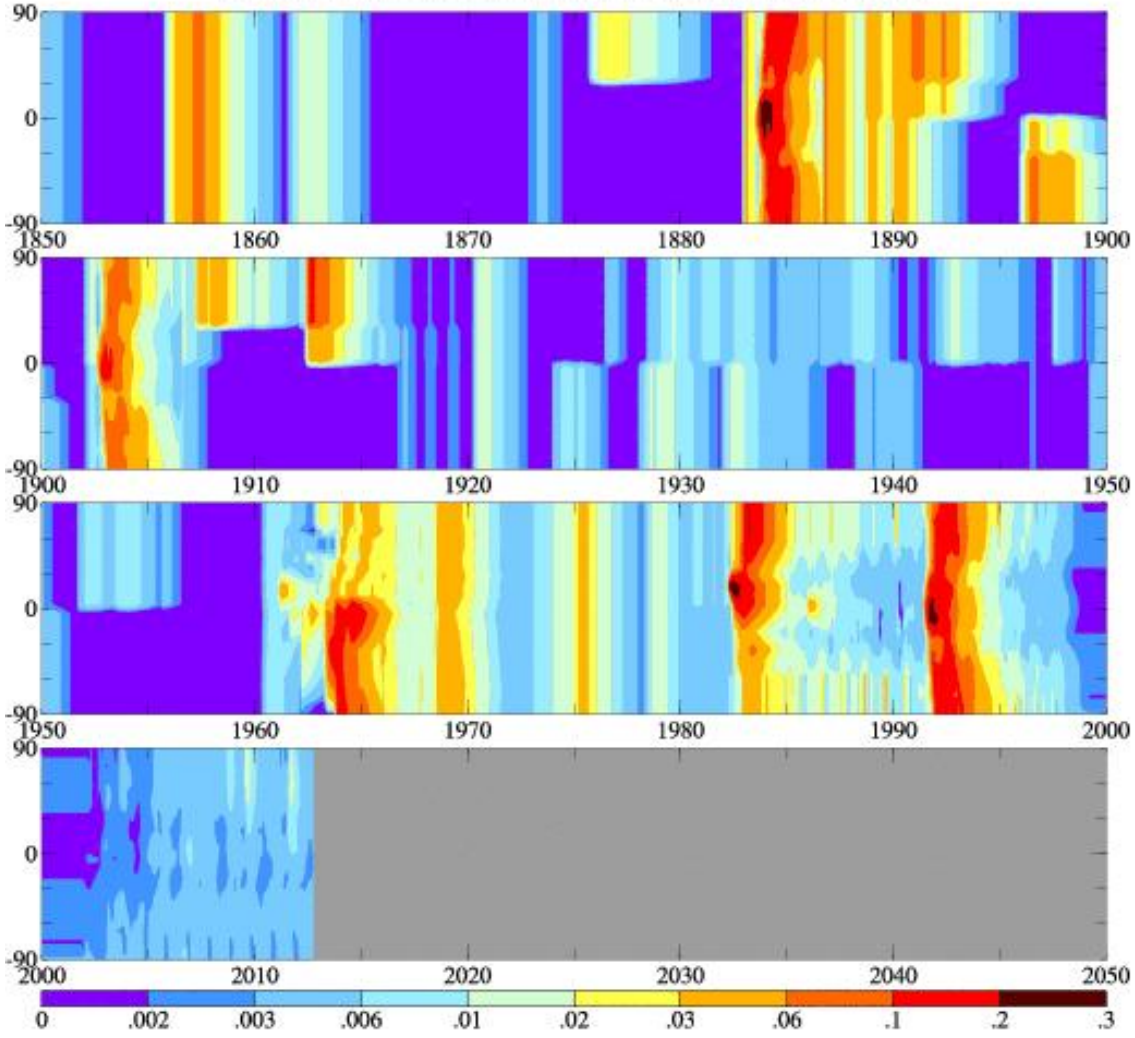

Figure 1.2: Timeseries of zonal mean stratospheric aerosol optical depth at 550nm, reproduced from Sato et al. (1993)

to the level it would have reached without geoengineering; this rapid climate change is referred to as the termination shock. [22]

Stratospheric aerosols affect global precipitation and the hydrological cycle. Geoengineering seeks to offset the effects of decreasing outgoing longwave radiation, due to greenhouse gases, by reflecting incoming solar radiation. As Trenberth and Dai (2007) [23] noted, "In between the incoming solar radiation and the outgoing longwave radiation is the entire weather and climate system and the hydrological cycle." They identified a substantial (6%, 3.17 standard deviations) decrease in precipitation over land in the year after Pinatubo. This is in accordance with Allen and Ingram

(2002) [24]. The forced response of precipitation depends on the energy balance of the troposophere and its ability to radiate away latent heat: radiative forcing that affects the shortwave flux to the surface, like that from stratospheric aerosols, reduces precipitation by affecting this energy balance.

Injections of SO_2 and H_2S into the stratosphere result in the formation of sulphuric acid droplets, which cool the troposphere by reflecting incoming shortwave radiation. The stratosphere is heated as the aerosols absorb longwave radiation, which alters the atmospheric circulation. In the case of Pinatubo, Stenchikov et al. (1998) [18] noted stratospheric heating was approximately zonal, and greater in the tropics: the resulting changes in circulation led to a warmer Northern hemisphere winter.

The distribution of aerosol after an eruption depends on stratospheric winds, hence the latitude and timing of the eruption. Strong zonal winds in the stratosphere maintain strong meridional concentration gradients: to first order, AOD varies with latitude. The Brewer-Dobson meridional circulation of the stratosphere results in aerosol being lofted upwards in the tropics, advecting polewards, before precipitating back into the troposphere at higher latitudes after a 1-2 year residence time in the stratosphere, depending on particle size. High-latitude volcanoes tend to produce clouds predominantly in their hemisphere, which recent research suggests may lead to a disproportionate cooling effect from extratropical volcanoes [25]. (Figure 1.2) However, the distribution that results from mid-latitude injections is less certain, as demonstrated by the differences in the AOD from the Pinatubo and El Chichon eruptions [26]. (Figure 1.3)

1.3 Tailoring of geoengineering and motivations for the project

The impacts of stratospheric aerosol injection can be altered by changing the latitude, altitude, season [27], amount, or particle type [28] of the aerosol injected. Models have suggested that a 'tailored' approach can achieve specific climate objectives, such as maintaining the Arctic sea ice extent [29], but that hemispherically asymmetric SRM

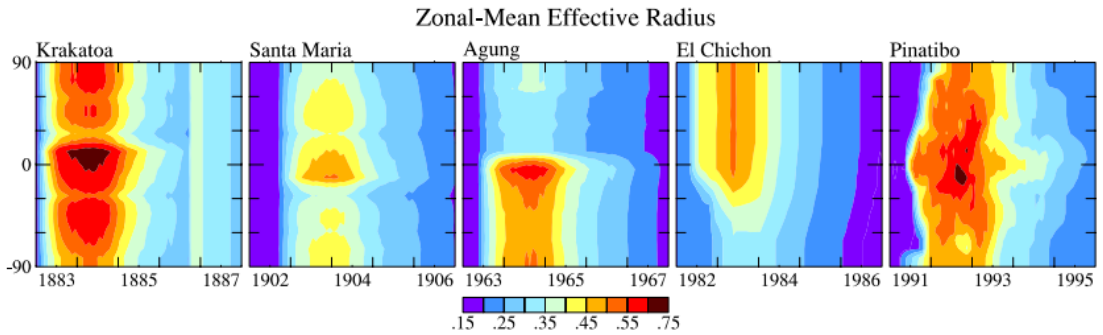

Figure 1.3: Effective radius of stratospheric aerosols after major volcanic eruptions from Krakatoa (6S), Santa Maria (14N), Agung (8S), El Chichon (17N), and Pinatubo (15N). Reproduced from Sato et al. (1993)

can disrupt monsoons and cause drought in the Sahel [30]. Some studies suggest counterbalancing, Antarctic injections can reverse the shift to the ITCZ [31], but the impact of latitudinal variations in AOD on large-scale circulation and precipitation has only just begun to be investigated.

Recent work at Oxford [32] has shown that higher CO_2 levels alone influence extremes in tropical precipitation and wet bulb temperature, suggesting that the direct, local radiative forcing effect of CO_2 is important in its own right. In geoengineered climates, the imperfect compensation between increased reflected shortwave radiation due to aerosols and reduced outgoing longwave radiation due to the greenhouse effect will result in distinct patterns of local radiative forcing that differ from preindustrial conditions, even if temperature is similar. The longwave forcing effect of CO_2 on the troposphere is still in effect under stratospheric aerosol injection, partially compensated by a reduction in incoming shortwave radiation and cooling at the surface, but added to by more longwave absorption by the aerosols in the stratosphere, resulting in a vertical temperature profile that would differ from that in the preindustrial case even if surface temperature is the same. Evaluating and quantifying the effect of these radiative forcing and heating patterns on climate and weather extremes is therefore crucial to understanding whether geoengineering can be optimised effectively. Some of the effects of geoengineering on weather and climate extremes, including 3-day temperature and maximum precipitation, are examined in Chapter 5.

The climate impacts of removing CO_2 also need to be investigated. One such modelling study [33] noted that, when CO_2 is instantaneously returned to preindus-

trial levels from 510ppm, surface warming only decreased by half within the first few years: residual warming remained for centuries. Outgassing of CO_2 from the ocean and land sinks require additional CO_2 to be drawn down and imply a long-term commitment in a range of negative-emissions scenarios [34] [35] [36] [37]. In all cases, reducing CO_2 concentrations cooled the land faster than the oceans, and there was a lag in the climate response due to ocean heat content.

This could affect the interaction between SRM and CDR if used together. For example, in a scenario where a limited deployment of SRM halts temperature rise, while CDR then allows SRM to be ramped down over time. Questions arising from this notion of "peak-shaving" include:

- How long would SRM need to be deployed? How much depends on the emissions pathway, rate of negative emissions available, climate sensitivity, and ocean heat uptake?

- How does a climate stabilised using SRM compare to one where the temperature target is attained by stringent or adaptive mitigation, or the high-emissions alternative: do the benefits outweigh the risks?

- Does limited geoengineering have a smaller impact on the hydrological cycle than attempting to return temperature to preindustrial conditions, as recent work [38] suggests? For example, in a 1.5K peak-shaving scenario, does the warmer world compensate for the reduced shortwave radiation at the surface, resulting in smaller precipitation anomalies?

Answering these questions requires a greater understanding of the climate system's response to SRM, the potential to control that response with latitudinal tailoring of AOD, and the dependence on that response on the state of the climate system when SRM is applied: all of which we can explore with the uniquely large ensembles provided by CPDN.

Advocates of SRM argue that, unlike mitigation or CDR, it could take effect quickly: but the effects likely depend on when it is deployed. Many geoengineering studies [39] ramp up SRM over time to compensate for increasing radiative forcing in climate-change scenarios, or deploy SRM simultaneously with increased greenhouse

gas concentrations. This is true of the Geoengineering Model Intercomparison Project (GeoMIP) [15] experiments, which expand on CMIP experiments and compare model responses to geoengineering. Comparing these simulations to experiments where geoengineering is first deployed later on in a high-emissions trajectory would give insight into the effect of ocean heat uptake on the reversibility of climate change. It is likely to result in a world with greater land-sea temperature contrasts, which are important drivers of, for example, Asian monsoons [40]. I discuss this further in the Conclusions chapter.

1.4 The distribution of aerosols in the stratosphere: stratospheric dynamics and aerosol microphysics

Models disagree about the size and spatial distribution of aerosols resulting from a particular injection strategy, and hence the amount and distribution of radiative forcing that might result from identical injection strategies. [41] Differences in stratospheric dynamics and aerosol microphysics, such as differing coagulation rates, meridional transport barriers, or stratospheric responses to injection, are behind this. [42]

Sustained injection and stratospheric heating alters the stratospheric circulation, which feeds back on the distribution of aerosols, changing the magnitude and pattern of radiative forcing achieved. Models disagree over the magnitude and nature of dynamical shifts in the stratosphere under geoengineering. The combination of stratospheric heating with tropospheric cooling is likely to weaken the Brewer-Dobson circulation in the stratosphere. Stronger meridional temperature gradients under SRM can produce a stronger wintertime westerly stratospheric vortex, and reduce wave-driven mixing of air at high latitudes, slowing the Brewer-Dobson circulation. These changes can result in strengthening of the polar vortex [43] and changes to the quasi-biennial oscillation (QBO) [44] [45] [46]. Strengthening of polar vortices can cool the polar stratosphere and reduce the frequency of sudden stratospheric warming events [47], but this may depend on aerosol particle properties [48].

As injection rates increase, so can coagulation of the aerosols into coarse mode. Larger aerosol particles reflect light less efficiently and have a shorter stratospheric

lifetime, while absorbing more longwave radiation. Niemeier and Timmreck (2015) [49] found that radiative forcing declined exponentially with injection rate, but that a radiative forcing of 8.5Wm^{-2} - comparable to the high-end of climate-change scenarios - was attainable with an injection rate of 5-7 Pinatubos per year. The efficacy for aerosol injection to produce radiative forcing depended on the timing, meridional extent, and altitude of injections. However, Kleinschmitt et al. (2018) [50], using a different model found that 2Wm^{-2} was the maximum attainable radiative forcing - at higher injection rates, stratospheric heating resulted in a greater meridional transport barrier and a strengthened QBO, which increased coagulation, and decreased radiative forcing.

The wide intermodel spread of radiative forcing in the G4sulfur experiment in GeoMIP [51] reflects the range of responses in stratospheric models. It deploys a constant injection rate of 5 Tg SO2 per year at a single point on the equator, but across models, mean RF varies from $-3.6Wm^{-2}$ to $-1.6Wm^{-2}$. Variations arise from different aerosol microphysics and stratospheric dynamics across models, but also from feedbacks and rapid adjustments in water vapour and clouds. [52] In Chapter 2, I explore the range of distributions of AOD that can result from intra-model differences in G4, and their effects on climate.

The effects of geoengineering on the stratosphere are not decoupled from those on the troposphere; for example, Bednarz et al. (2023) [53] found that different injection strategies could achieve the same global mean surface temperature, but owing to differences in the location and timing of aerosol injections, the effects on lower stratospheric warming and stratospheric circulation differed and this in turn drove climate impacts at the surface.

When attempting to understand the effects of a particular injection strategy, there are two major sources of uncertainty. The first surrounds the aerosol distribution that results from a given injection strategy, which depends on the aerosol microphysics and the stratospheric dynamics, including feedbacks on the stratosphere. The second surrounds the response of the tropospheric climate to a different aerosol distributions and imposed forcings: including extreme weather events, regional climate, and patterns of temperature and precipitation change. Existing multimodel comparisons of particular injection strategies, such as the G4 experiment, sample both sources of

uncertainty simultaneously, as they result in different patterns of radiative forcing which are them imposed on climate models which have different responses to that forcing. I show that these different patterns of aerosol optical depth from the G4 experiment can result in different regional climate outcomes when used to drive the same model in Chapter 3.

The nature of climate modelling at present means that different tools are appropriate to investigate each of these uncertanties. The aim of my research is to use the uniquely large ensembles of HadCM3 provided by climateprediction.net (CPDN) to examine uncertainty of the second kind. HadCM3 does not have the ability to explicitly simulate injections of aerosols into the stratosphere, but we can use CPDN's large ensembles to investigate the climate response to a range of different imposed forcings.

Running a fully-coupled, high-resolution stratosphere alongside a state-of-the-art Earth System Model is too computationally expensive at present to provide enough model years to fully characterise the effects of latitudinal variations in AOD, or the effects of SRM on the extreme weather events that are the costliest consequences of climate change. Additionally, the fact that these high-resolution stratospheric models disagree about the AOD distribution that arises from a given injection strategy merely underlines our need to understand how sensitive the climate is to variations in the aerosol distribution and local radiative forcing, such that potential negative consequences of geoengineering can be better anticipated, and research efforts focused.

1.5 Global climate response to geoengineering in idealised studies

Idealised geoengineering experiments since Govindasamy and Caldeira (2000) [54] have consistently shown that global mean temperature increase can be reversed by uniform solar constant reduction or idealised stratospheric aerosol implementations, with undercooling of the poles and overcooling of the tropics, alongside a decrease in precipitation compared to a control scenario.

The GeoMIP project was first described in Kravitz et al. (2011) [15] and expanded in Kravitz et al. (2015) [55] with an experiment, G6, designed to examine differences between solar constant reduction and stratospheric aerosols. A schematic detailing the scenarios is reproduced in Figure 1.4.

The level of geoengineering required to cancel the radiative forcing increase due to quadrupled CO_2 varies from 3.8-5% solar constant reduction across the range of models used, due to cloud and thermodynamic feedbacks [56], with HadCM3 requiring a 4.1% reduction in the solar constant to cancel this forcing. [57] Residual warming is typically small compared to baseline scenarios. Kravitz et al. (2013) [58] found an average residual warming of around 1K in polar regions compared to 10K under quadrupled CO_2. The average overcooling of the tropics was around 0.3K.

Models robustly simulate an overall reduction in precipitation under SRM, beyond that attributable to the cooler climate. [57] In a multimodel comparison of the G1 experiment, SRM aimed to restore preindustrial temperature reduced global mean precipitation to an average of 4.5% below preindustrial levels. When radiative forcing is imposed, the model's response can be decomposed into a fast response - directly due to changes in radiative forcing - and a slow response, incorporating feedbacks as temperature changes. For example, in experiments where CO_2 is abruptly increased, the fast response is an initial strong decrease in precipitation, which is due to the heating of the upper troposphere when carbon dioxide absorbs outgoing longwave radiation. The surface heats more slowly, which results in a weaker temperature gradient across the troposphere, which in turn reduces circulation in the atmosphere and reduces precipitation. The slow response, by contrast, is an increase in precipitation as surface temperature rises, as the atmosphere is able to hold more moisture in accordance with the Clausius-Clapeyron equation [59]. This is explored much further in Chapter 4, where we analyse the separate patterns of fast and slow response and whether there is any mechanism by which the"slow response" of gradual warming can compensate for the "fast response" under 4xCO2 and SRM.

In the G1 experiment, changes to the surface and tropospheric energy budgets due to the initial reduction in solar radiation and abrupt quadrupling of CO_2 results in a "fast response" precipitation decrease. However, as temperature remains roughly constant, there is no slow feedback. Decomposing the response into a fast feedback

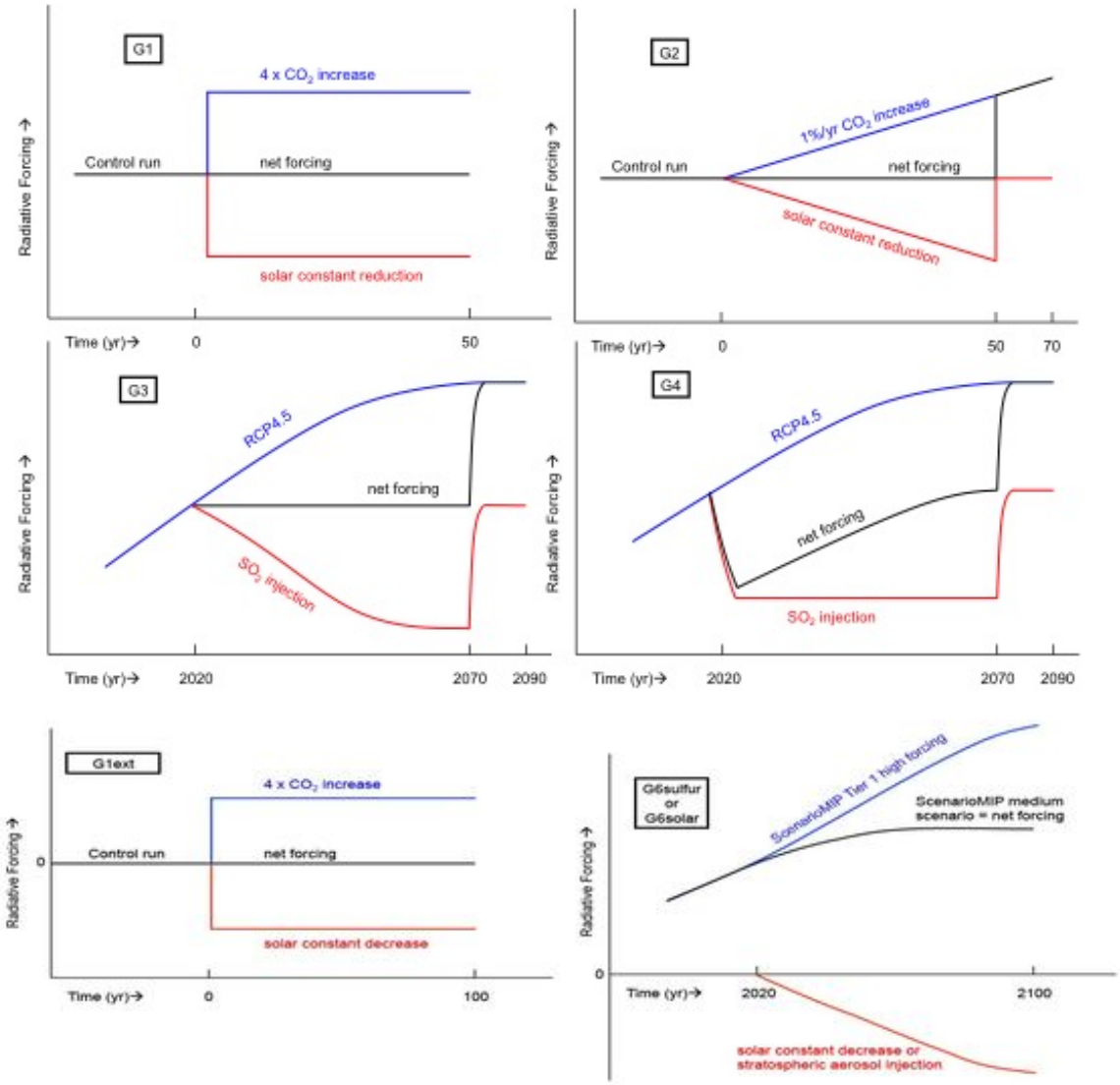

Figure 1.4: Scenarios G1-G6 of the GeoMIP, reproduced from Kravitz et al. (2011) and (2015)

and a slow feedback suggests that geoengineering applied to a warmer world would have different impacts on precipitation [60], before we take into account the hysteresis inherent in other subcomponents of the earth system, such as ice sheets or forest cover.

Niemeier et al. (2013) [61] noted that increased longwave absorption by aerosols means the shortwave flux to Earth must be reduced by a greater amount to balance top-of-atmosphere radiative forcing. This resulted, in their model, in doubled precipitation reduction compared to solar-constant reduction implementations of geoengineering. Physically, changes to downward shortwave flux have a stronger effect on precipitation, which is controlled by the radiative budget in the troposphere, than temperature: precipitation is therefore more sensitive to solar and volcanic forcing than that from greenhouse gases. [24] This then raises further questions we can start to explore: given that using different types of aerosol with different optical properties raises the possibility of changing the ratio of longwave to shortwave forcing impacts for SRM, is it possible to reduce the negative climate impacts and side-effects of SRM? This potential is discussed further in the Conclusions chapter.

Some precipitation change may be attributed to CO_2 concentration changes, which in turn feed back on plant evapotranspiration over land, [57] especially over densely vegetated regions like the Amazon. [62]. Precipitation - evaporation (P-E) changes are arguably more relevant in assessing changing water availability for agriculture and society, and because evaporation is also reduced by the reduction in shortwave radiation reaching the surface under SRM, these changes are often less significant than changes to precipitation, as there is a reduction in the overall intensity of the hydrological cycle which reduces both precipitation and evaporation. Kravitz et al. (2013) [58] found that the changes in P - E due to geoengineering were within 2 standard deviations of natural variability for 91% of the land surface. The largest differences in P - E over land were in the tropics, especially in the East of China and the Amazon basin. More analysis of the effects of SAI on the monsoons, droughts, water availability - but also the effects of CO_2 and diffuse solar radiation - is necessary.

Models suggest SAI could mitigate [63] or even reverse [64] some sea level rise, especially the thermosteric component (directly caused by temperature increases expanding the volume of water, rather than ice sheet melt), but the ice sheet response

is less certain [65]: surface melting is reduced by SAI, but ocean-driven melt is less affected [66] and hysteresis in non-linear ice sheet collapse makes conclusions dependent on deployment time. [67]

1.6 Impacts of geoengineering on climate and weather extremes

Investigation of the idealised G1 experiment suggests that geoengineering may be less effective at restoring preindustrial climate extremes than overall means. [68] Modelling studies suggested SAI was generally as effective at reducing temperature extremes as solar dimming, except in the Arctic, but has a greater impact on precipitation extremes [69]. The coldest night warms by 0.74K in G1 compared to 6.65K in abrupt4xCO_2, while 5-day maximum precipitation decreases over the tropics, especially over the oceans.

Given that SRM studies generally target reducing global mean temperatures to the preindustrial baseline when applied on a global warming scenario, it is perhaps unsurprising that modelling results tend to show that they are more effective at reducing temperature extremes - especially warm temperature extremes - than precipitation extremes. Similarly, as described in the previous subsection, as geoengineering entails a reduction in the overall hydrological cycle relative to a case without geoengineering, it is not surprising that it can also be effective at reducing flooding and extreme rainfall events. Idealised studies have indicated that SRM would decrease heatwave frequency and duration in almost all regions [70], although changes to soil moisture, associated feedbacks, and undercooling of the poles lead to increases in heatwaves in Russia. Wei et al. (2018) [71] found that SAI could decrease flood risk in most regions, especially Southeast Asia, but the pattern of streamflow response was complicated and flooding may increase in Australia, Mexico, and the US Southwest.

Jones et al. (2018) [62] analysed the regional climate impacts of stabilising global warming at 1.5K using solar geoengineering to partially offset the RCP scenarios where they found it could stabilise global mean temperature, Arctic sea ice, and thermosteric sea level rise rate. Only in the geoengineered RCP8.5 scenario did global

mean precipitation decline below 2000 levels, although it decreased in all SAI scenarios. With the exception of P-E in the Amazon basin, there was a "universal reduction in detrimental impacts when compared to RCP scenarios." The Amazon was a specific area of interest due to its dense foliage and the consequent importance of transpiration from plants - under increased CO_2 scenarios, even where SRM is also present, transpiration from plants is reduced and consequently the hydrological impact of CO_2 alone is significant. Tilmes et al. (2020) [72] explored the use of stratospheric aerosol geoengineering to stabilise temperatures at 1.5K or 2K targets, noting that some variables (such as surface air temperature and the impact on the Atlantic Meridional Overturning Circulation) depended primarily on the surface temperature target, while other variables associated with precipitation depended predominantly on the amount of stratospheric aerosol injected. Discussions at recent sessions of GeoMIP have focused on whether scenario definition for CMIP7 and GeoMIP might include a policy-relevant peak-shaving scenario of this kind as a standard model intercomparison, [73] and several such scenarios have been proposed as necessary for modelling focus before the IPCC could consider making statements about SRM. [74]

In specific case studies, they found European heatwaves and Katrina-level storm surge events were decreased in frequency under geoengineering, reflecting results in Wang et al. (2018) [75] which analysed thermodynamic tropical cyclone (TC) statistical indices. Jones et al. (2017) [76] noted that SAI in the Northern Hemisphere alone reduced TC frequency, while Southern Hemisphere SAI increased it.

Studies of some of the patterns of extremes under the Geoengineering Large Ensemble (GLENS) experiment in general show that, while global mean temperature can be maintained near 2020 levels, there are regional variations in the average climate response which also leave an imprint on relevant extremes - for example, undercooled regions have greater temperature extremes (warmer hottest days) while overcooled regions have cooler coldest days. Furthermore, under GLENS and in the Community Earth System Model (CESM) that it used, the distribution of extreme precipitation events is found to change: wetter regions generally have decreases in wet extremes, while drier regions tend to see more extreme precipitation. There is also a general shift to precipitation becoming more consistent, with more days with light rain and fewer very intense rainfall events than the baseline climate. The diurnal difference

in the climate response to aerosols and greenhouse gases (which reduce outgoing longwave radiation at night as well) results in more extreme temperatures at night under GLENS, even where extreme daytime temperatures are reduced compared to the baseline. [77]

Some studies of specific regions under specific SRM scenarios, such as a study of South Asia under GeoMIP scenario G6sulfur, suggest that extreme weather events are generally more common under SRM scenarios than preindustrial scenarios. Compared to unmitigated climate change, some extreme event indices are worse and others are better in different subregions, with a notable simulated increase in warm events under the SRM scenario over the Himalayas compared to the climate change scenario. [78] Other studies illustrate a more mixed picture. For example, a study which used the GLENS experiment (see below) found that meteorological drought characteristics in West Africa initially decreased in most regions compared to the unmitigated climate change scenario, in later decades this was reversed and there was an increase in most drought characteristics driven by reduced precipitation, and a weakened monsoon circulation owing to the reduction in land-sea temperature gradients. [79]

These studies provide evidence that geoengineering may reduce the frequency or severity of some extreme weather events, but geoengineering studies such as GeoMIP involve model intercomparisons from a small number of runs of idealised experiments. Few studies exist that have a sufficient number of model years to make robust statistical, probabilistic statements about changes to extreme weather events. Yet this analysis is crucial to understanding whether SRM avoids the most damaging negative impacts of climate change as intended. A large number of model years will be necessary for studies that implement a latitudinally-varying distribution of aerosols to distinguish the effects of varying the distribution. The effects of tailored geoengineering on the hydrological cycle also require further investigation. We explore the response of some extreme indices in temperature and precipitation to SRM in Chapter 5.

1.7 Regional or non-uniform applications, injection and control strategies

Geoengineering is not equally effective in all regions. Ricke et al. (2010) [20] found that, as geoengineering was ramped up, simulated temperature and precipitation in regions such as China and India drifted from the targeted control scenario in different directions. Multimodel comparisons of the response to the idealised G1 scenario broke down regional analysis into the Giorgi regions used by the IPCC for analysis of regional climates. [80] Assuming the response scales linearly with solar constant reduction, studies have altered the total amount of radiative forcing [81] or constructed simple latitudinal distributions of forcing, with linear superpositions of a few orthogonal basis functions, to attempt to ameliorate this [82].

They found that regions broadly agree on the amount of geoengineering needed to restore preindustrial average global mean temperature, but the differences in precipitation are far less uniform and cannot be reconciled - however, as results in Chapter 2 will indicate, the assumption of linearity may not be valid here. These regional disparities are found in GeoMIP scenarios across models and for various implementations of SRM, and may worsen when aerosols are explicitly simulated. [83] Feedbacks in the climate system which affect local radiative forcing, such as changes in cloud cover, can contribute to regional extremes. [68]

Some alternative methods of geoengineering, such as marine cloud brightening or albedo modification, have been suggested which would apply inhomogenous, localised radiative forcing patterns. When the net radiative forcing is globally significant, as in cases where desert albedo enhancement or significant marine cloud brightening are modelled, there are substantial shifts to climates outside of the targeted region, [84] [85] but there are suggestions that applying albedo modification over very small domains might results in fewer global changes. [86]

Given that it is not possible to exactly restore preindustrial temperature and precipitation using geoengineering, there is a significant parameter space of outcomes that can be explored. It is not at all clear that idealised simulations which aim to restore preindustrial global mean temperatures constitute an "optimal deployment"

of geoengineering - for example, a target of halving warming was found to moderate policy-relevant climate outcomes (such as water availability) relative to an unmitigated climate change scenario. [38] We will explore the concept of optimal levels of residual warming for individual regions further in Chapter 3.

Distributed injections across several degrees of latitude could reduce particle size and hence enhance reflectivity [49]. However, in some models, [50] the shorter lifetimes of aerosols injected at higher latitudes compensated for this and prevented any substantial increase in radiative forcing. Tilmes et al. (2017) [87] found that nonequatorial injections resulted in more efficient transport of aerosols into middle and high latitudes, suggesting unwanted residual latitudinal temperature gradients could be reversed. Injections at higher altitude have also been found to be more efficient at producing radiative forcing. [88]

Overall, although models disagree, studies suggest that, despite limitations imposed by aerosol microphysics and stratospheric dynamics, degrees of freedom exist in the latitudinal aerosol distribution that can be obtained through injections.

For example, MacMartin et al. (2017) [89] demonstrated that, in their model, it was possible to use injections at four different latitudes to adjust three spatial degrees of freedom in AOD: "an approximately spatially uniform AOD distribution, the relative difference in AOD between Northern and Southern Hemispheres, and the relative AOD in high versus low latitudes." (See Equations 1.1-3). At low injection rates, the patterns of aerosol and forcing from different injection locations were additive: for higher injection rates, coagulation, non-linear interactions, and feedbacks on stratospheric dynamics became more important. Future GeoMIP projects are looking into constructing emulators that will allow a wide range of scenarios to be explored and which might rely on basis simulations similar to the ones we have outlined. [90]

Previous studies of "fine-tuning" aerosol optical depth as a function of latitude have looked at three orthogonal spatial patterns of forcing: Legendre polynomials in the sine of latitude. [91] [92] [93] MacMartin et al. (2017) [89] obtained approximations to these patterns of aerosol optical depth with combinations of injections at $\pm 15^\circ$, $\pm 30^\circ$.

$$L_0 = 1 \tag{1.1}$$

$$L_1 = sin(\varphi) \tag{1.2}$$

$$L_2 = \frac{1}{2}\left(3sin^2(\varphi) - 1\right) \tag{1.3}$$

where φ is the latitude of injection.

Both Kravitz (2015) [93] and MacMartin et al. (2017) [89] consider modelling geoengineering as a simplified linear design problem, where degrees of freedom in AOD - represented by these basis functions - are adjusted to achieve specific patterns of temperature as a function of latitude. (Equations 1.4-7). In an optimal strategy, each degree of freedom you control allows you to adjust one variable, using coefficients determined by a simple linear algebra system. The temperature parameters, T_0, T_1, T_2 defined by Kravitz (2015) [93] were Legendre polynomials:

$$T_0 = \frac{1}{A}\int T(\varphi)cos(\varphi)d\varphi \tag{1.4}$$

$$T_1 = \frac{1}{A}\int sin(\varphi)T(\varphi)cos(\varphi)d\varphi \tag{1.5}$$

$$T_2 = \frac{1}{A}\int \left(\frac{3sin^2(\varphi) - 1}{2}\right) T(\varphi)cos(\varphi)d\varphi \tag{1.6}$$

where

$$A = \int cos(\varphi)d\varphi \tag{1.7}$$

and $T(\varphi)$ is the temperature as a function of latitude.

Observations of these temperature parameters are taken from the model, and then the aerosol optical depth is adjusted accordingly to target a reduction in these temperature anomalies, by a process of feedback which aims to control the climate by increasing certain patterns of aerosol optical depth to reduce particular temperature anomalies. In practice, such a system of feedback and control would draw (uncertain) observations from the real world and then use them in a climate model to attempt to determine how to adjust the aerosol injection strategy: this introduces further uncertainties, since the model's physics will differ from reality.

For this reason, Kravitz et al. (2016) [94] used two ESMs. One is used as a "design model", from which observations are drawn and a strategy is designed, while the

other is treated as the "evaluation model", where the feedback strategy is actually implemented and the results evaluated. This aims to account for some of the uncertainties that arise from using an imperfect model of reality. In Chapter 3, I discuss approaches that might achieve a similar assessment of the robustness of conclusions drawn.

Kravitz et al. (2016) [94] found that, in general, they were able to simultaneously balance these three temperature objectives to within 0.2K, by adjusting the coefficients of Legendre polynomials in solar constant reduction. This was maintained for 80 years against a background of 1% increase of CO_2 per year. They also demonstrated that, in a separate design strategy, it was possible to attempt to balance the competing demands of maintaining the Arctic sea ice extent and ITCZ position, by adjusting Arctic and Antarctic SRM levels.

Recently, these researchers have expanded their dynamical feedback approach in the Geoengineering Large Ensemble Experiment (GLENS) [95], which deployed injections of aerosol at the equator, 15°N, 15°S, 30°N, and 30°S to minimise the temperature anomaly objectives (Equations 1.4-7) in the Community Earth System Model (CESM) against a background of RCP8.5. In total, 20 ensemble members were generated. The closed-loop feedback approach successfully reduced regional disparities in temperature and precipitation, but persistent polar warming still resulted in their model, attributed in part to changing ocean circulation. [39] The data from GLENS has been made available for general use in the community, and I will describe in Chapters 3 and 4 my analysis of some of that data, alongside aspects of how we have used the distributions generated in this way in our modelling.

In addition, there are scenarios in the GeoMIP testbed that investigate using SRM to "peak-shave", i.e. limit temperature increases to below a specific threshold [72] - we will explore temporally varying and idealised scenario deployment of SRM of this nature in the next chapter.

1.8 Introduction to the rest of the thesis

Chapter 2 will use a combination of analytical techniques and the simple climate model, FaIR, to examine different scenarios for "peak-shaving" – temporarily using

geoengineering to hold global mean temperatures below a certain threshold, as in the Shepherd napkin diagram - to investigate scenarios for geoengineering deployment and the length of time we might be committed to peak-shaving. Chapters 3-5 analyse simulations from the HadCM3 climate model, simulated using climateprediction.net, to generate large ensembles of simulations with differing distributions of stratospheric aerosol optical depth counteracting an abrupt quadrupling of carbon dioxide to represent different attempts at tailoring geoengineering - in other words, trying to reduce the uncertainty of how the climate response is affected by different patterns of imposed aerosol optical depth. Chapter 3 will detail the initial calibration and characterisation of the response to simple patterns of radiative forcing. Chapter 4 will expand on this by considering various methods for optimising for temperature and precipitation. Chapter 5 will examine the impact of geoengineering on climate and weather extremes. In the final chapter, results and conclusions from previous chapters are summarised, and possible future work is outlined.

Chapter 2

Tradeoffs inherent in solar geoengineering peak-shaving strategies

2.1 Introduction

Note: an earlier version of this chapter has been available as a preprint on arXiv since 2021. [96]

Under what circumstances could we consider solar geoengineering? Many authors have observed that solar geoengineering cannot be a substitute for mitigation and the reduction of emissions [97] [98]. Solar radiation management (SRM) imperfectly compensates for the effects of greenhouse gases, leading to residual climatic changes. These residual climate changes are likely to worsen as the amount of anthropogenic radiative forcing that is offset by SRM is increased. This is the case because SRM fundamentally seeks to compensate for the effect of tropospheric warming from an increased greenhouse effect and reduced outgoing longwave radiation with a reflection of incoming shortwave radiation in the stratosphere, and these patterns of forcing compensate for each other imperfectly. [54] One direct consequence of this is that SRM reduces precipitation more quickly than temperature, chiefly as the net result of cooling the surface with SRM by reducing the incoming shortwave flux, while the

troposphere is warmed by greenhouse gases, is an atmosphere that is more stable to convection and a consequent reduction in precipitation.

Consequently, trade-offs exist between regions and between returning temperature and precipitation to preindustrial baselines using SRM. [99] (This is explored in further detail in our GCM ensembles in Chapter 4.) Applied alone, SRM does little to address ocean acidification [100], and may result in residual polar ocean warming [101]. The greater the radiative forcing that is offset by SRM, the more dangerous it would be to suddenly cease its use due to the termination effect: a rapid rebound of temperatures [102]. In addition, the radiative forcing, and latitudinal distribution of forcing, that can be obtained with the most viable method, stratospheric aerosol injection, is not yet clear, with model studies in disagreement [103] [104]. A scenario where humanity is committed to ever-increasing deployment of SRM while emissions continue to climb is evidently fraught with risk.

"Peak-shaving" scenarios, where SRM is used temporarily to keep temperatures below a threshold – for example, the 1.5C or 2C targets of the Paris Agreement – have been suggested as a more reasonable form of deployment. The first prominent heuristic example is the so-called "Shepherd Napkin Diagram", reproduced as Figure 2.1 from Long and Shepherd (2014). [105] This illustrates the concept behind a temporary deployment of SRM to "shave the peak" of warming while large-scale carbon dioxide removal can reduce the temperature to below a given threshold. The original napkin diagram was not based on physical modelling, so the estimates for the duration and the peak overshoot are approximate and not related to emissions trajectories.

This concept has been explored elsewhere, such as in the IPCC SR1.5 report (Section 4.3.8 and Cross-Chapter Box 10) [106] [6], as a temporary use of SRM to reduce global mean temperatures, which would buy time for mitigation to take effect. Modelling suggests that limited deployment of SRM - for example, to cancel half of the warming due to greenhouse gases - could reduce the negative residual climatic changes in most regions. [107] Peak-shaving scenarios which combine rapid mitigation with temperature overshoots have begun to receive more focus from the modelling community [72] and recent discussions at GeoMIP have posited a greater focus on

policy-relevant scenarios such as peak-shaving alongside a core of scientifically relevant scenarios. [73]

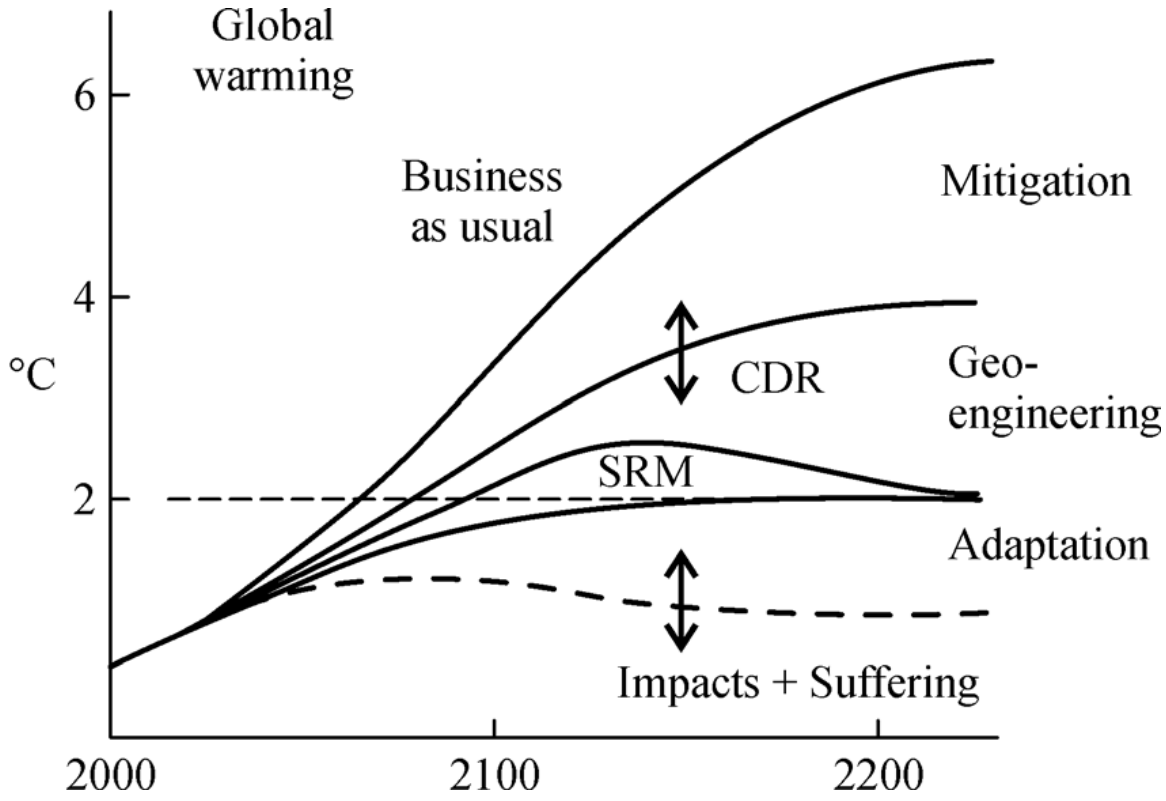

Figure 2.1: **Shepherd's 'napkin diagram', first published in Long and Shepherd (2014).**

Here, I do not seek to evaluate whether the climate benefits of such peak-shaving outweigh the risks, but instead to explore the parameter space of peak-shaving scenarios. In particular, this study aims to explore the possible relationships between avoided temperature increase, the duration of committment to SRM, and the necessary cumulative negative emissions that are required to end SRM. First, this is done using a combination of simple approximate relationships between these quantities (like the TCRE) and idealised emissions scenarios to derive analytical formulae that approximately relate these quantities together, in the "Peak-shaving in the abstract" section. This is followed by idealised simple climate modelling which relates radiative forcing and cumulative emissions to global mean temperature and allows us to estimate the trade-offs between avoided temperature peaks and cumulative emissions in a range of different scenarios.

2.2 Peak-shaving in the abstract

2.2.1 A toy example: parabolic peak

First, I note that if SRM is to be deployed temporarily, CO_2 emissions must peak and decline to net zero. This follows from the approximate proportionality of temperature changes to cumulative carbon emissions [108]. Once emissions reach net zero, carbon dioxide concentrations in the atmosphere will gradually decline due to ocean uptake and weathering, but substantial net-negative CO_2 emissions are likely to be necessary to avoid long-term (i.e. multi-centennial) commitment to SRM in a peak-shaving scenario. [109]

As a simple example, consider only CO_2 emissions and a very simple emissions trajectory. Assume that no mitigation occurs until a particular threshold temperature - say, the 1.5C threshold in the Paris Agreement - is crossed. Thereafter, CO_2 emissions begin falling at a constant rate, becoming net-negative emissions at this same rate. Under these circumstances, the temperature peak will be approximately parabolic, as the emissions in a given year are approximately proportional to the rate of temperature change in that year; the second derivative of temperature is therefore the first derivative of emissions. [108] I then assume that geoengineering is applied to "shave" this peak and prevent temperature from increasing above the threshold.

Let D be the duration of continued geoengineering above a given temperature threshold, T be the temperature excess above threshold, and R be the rate of warming when the threshold is first exceeded (noting that this is approximately the Transient Climate Response to Cumulative Emissions, TCRE, multiplied by the emissions when the threshold is exceeded.)

This quadratic curve can then be expressed as $T(t) = at(t–D)$, where D is the length of time for which the threshold is exceeded. Differentiating gives $dT/dt = a(2t–D)$, and $R = -aD$ by definition. The maximum temperature overshoot above the threshold, T_{os}, is equal to $–aD^2/4$. I then express the maximum temperature above the threshold in terms of the duration and the warming rate when the threshold is first crossed: $T_{os} = RD/4$ or, equivalently, $4T_{os}/R = D$.

In such an idealised scenario, the duration of the SRM commitment is twice the time taken to reach net zero. The attainable mitigation rate therefore determines the cumulative negative emissions required, and the duration of commitment to SRM. When this is combined with the climate sensitivity - for example, through the rate of warming - the temperature overshoot can be inferred.

This allows us to see the tradeoff between the implied duration of deployment and the height of the peak that has been avoided. For example, at the current rate of warming of 0.25K per decade, shaving a peak of 2C to 1.5C implies a commitment to deployment of around 80 years. (In reality, this is likely to be an underestimate, as I will explain.)

Considered in the context of the Paris Agreement, current scenarios that limit global warming to 1.5K (or 2K) include "overshoot" scenarios where temperature temporarily rises above the target, before reaching it at the end of the century. SRM could be deployed to prevent such overshoot, as in SR1.5 Section 4.3.8, Cross-Chapter Box 10, Figure 1. [106] In the example considered in the IPCC report, overshoot of around 0.1K is avoided (T_{os}).

Assuming rates of warming and cooling (R) are similar to today's magnitude of warming at 0.25K/decade, the rule-of-thumb $4T_{os}/R = D$ implies a commitment (D) of 32 years, which is similar to the indicative time-scale shown on the graph in SR1.5 Section 4.3.8, which is approximately 35 years. This exemplifies the trade-off between the duration of the implied commitment to SRM and the magnitude of the temperature peak avoided.

By integrating, one can determine the area under the curve and hence the number of degree–years avoided, as a rough metric for avoided damages [110], although due to SRM's imperfect compensation for increased GHG concentrations, this should not be considered fungible with the equivalent temperature increase avoided by mitigation. This gives $RD^2/6$ and consequently would equate to around 267 degree–years avoided in the above example.

This illustrates the trade-off between duration of the commitment and the temperature avoided for a given rate of warming. The more warming we use SRM to avoid – and hence, the more damaging any potential termination, which would entail

a more rapid climate change of greater magnitude as the suppressed warming rapidly rebounds [21] – the longer our commitment must be.

2.2.2 Different rates of mitigation and net-negative emissions deployment

This analysis can be simply extended to consider the more general case for overshoot trajectories where the rate of warming when the target is first crossed and the rate of cooling are not the same, but parabolic trajectories are assumed either side of the peak - in other words, mitigation and net-negative emissions deployment occur at different rates. This could occur, for example, as net-negative emissions are expensive to deploy and the urgency to deploy them is reduced once global warming is stabilised, so mitigation occurs relatively rapidly compared to negative emissions.

In this case, one can add up the components of the duration of SRM on either side of the parabola's peak (where warming and cooling rates are positive-definite magnitudes):

$$D_{warm} = \frac{2T_{os}}{R_{warm}} \tag{2.1}$$

$$D_{cool} = \frac{2T_{os}}{R_{cool}} \tag{2.2}$$

$$D_{overshoot} = D_{warm} + D_{cool} = \frac{2T_{os}}{R_{warm}} \left(1 + \frac{R_{warm}}{R_{cool}}\right) \tag{2.3}$$

If the rate of warming (absent SRM) on the way up exceeds the rate of cooling on the way down, then overshoot is longer. Given that rates of warming and cooling are approximately proportional to the emissions of CO_2, the ratio of R_{warm} to R_{cool} is roughly the ratio of emissions for the year when the carbon budget is exceeded and the negative emissions when the "carbon debt" has been repaid.

The generalised formula for degree-years of exposure avoided by the SRM is similarly obtained:

$$DY_{\text{overshoot}} = \frac{4T_{\text{peak}}^2}{3R_{\text{warm}}}\left(1 + \frac{R_{\text{warm}}}{R_{\text{cool}}}\right) \tag{2.4}$$

Denoting the carbon budget overshoot by ΔC_{ex}, then we have, with λ as the TCRE:

$$D \approx 2 \cdot \lambda \cdot \Delta C_{ex}\left(\frac{1}{\lambda \cdot em\left(t_{\text{warm}}\right)} + \frac{1}{\lambda \cdot \left|em\left(t_{\text{cool}}\right)\right|}\right) \tag{2.5}$$

$$D \approx 2 \cdot \Delta C_{ex}\left(\frac{1}{em\left(t_{\text{warm}}\right)} + \frac{1}{\left|em\left(t_{\text{cool}}\right)\right|}\right) \tag{2.6}$$

It can be seen from this that the duration of the overshoot – and hence the implied SRM commitment – depends on the amount the carbon budget is exceeded, the timescale over which you continue to emit carbon dioxide and warm, and the timescale to remove the full anthropogenic excursion of CO_2. In this approximation, the dependence on climate sensitivity arises from its implications for the carbon budget for 1.5K or 2K. Higher climate sensitivity reduces the cumulative emissions required for a given overshoot, which shortens both the timescale of continued emissions and the timescale of negative emissions required to remove them.

2.2.3 The implications of a negative emissions floor

Suppose that global net-negative emissions can only be increased up to a certain "floor" (n). This could be due to biophysical limitations in the case of large-scale BECCS, afforestation, or enhanced weathering, or societal willingness to pay in the case of Direct Air Capture. One might also anticipate that the average rate at which large-scale net-negative emissions are deployed (m_2) will differ from the average rate of mitigation (m_1). If mitigation and geoengineering start when the temperature threshold is crossed, and carbon emissions are at C_{max}, then our carbon budget formulation gives us:

$$\Delta C_{ex} = \frac{1}{2}\frac{C_{max}^2}{m_1} \tag{2.7}$$

With t_{neg} as the total time committment to any negative emissions, and t_{floor} the time commitment to negative emissions at their lowest value, we have:

$$t_{neg} = \frac{n}{m_2} + t_{floor} \tag{2.8}$$

Requiring that the total negative emissions cancel out the exceedance of the carbon budget, we have:

$$\Delta C_{ex} = \frac{1}{2}\frac{n^2}{m_2} + t_{floor} n \tag{2.9}$$

Solving for t_{floor} and with the total duration of the overshoot as D, we obtain:

$$D = \frac{C_{max}^2}{2m_1 n} + \frac{C_{max}}{m_1} + \frac{n}{2m_2} \tag{2.10}$$

Which simplifies in the special case that $m_1 = m_2 = m$ to:

$$D = \frac{1}{2nm}(C_{max} + n)^2 \tag{2.11}$$

We also have that

$$T_{peak} = \Delta C_{ex} \cdot \lambda = \frac{C_{max}^2 \cdot \lambda}{2 \cdot m} \tag{2.12}$$

Which allows us to re-express the duration in terms of the TCRE, the peak shaved, the negative emissions floor, and the mitigation rate, if desired.

2.2.4 The double exponential emissions trajectory

An alternative model of emissions trajectories uses a double exponential function. This combines emissions which fall exponentially, or by a given percentage each year, with negative emissions which rise exponentially. This trajectory is idealised, but it reflects certain likely realities - reducing emissions to zero gets more difficult over time, as relatively easy-to-decarbonise sectors such as electricity and waste see rapid reductions in emissions, while harder-to-decarbonise sectors and recalcitrant nations lag further behind. For peak-shaving emissions trajectories, subject to a negative emissions floor, n, you can model the emissions after mitigation begins as:

$$em(t) = C_{max}exp(-m_1 t) - n(1 - exp(-m_2 t)) \tag{2.13}$$

where

$$m_{1,2} = ln(2)/t_{1,2} \tag{2.14}$$

and $t_{1,2}$ are the timescales for halving (gross) emissions and scaling up the deployment of negative emissions to half of its maximum value, respectively. To represent emissions which change by x% per year, one could consider $m_i = \ln(1+0.01x)$. In this idealised emissions scenario, rapid mitigation is combined with expanding deployment of negative emissions at scale to reach net zero, while absolute mitigation rate slows down over time while the rate of negative emissions deployment saturates. The first term then has emissions decaying from a maximum value of C_{max} to zero, while the second term the scale-up of negative emissions.

This models well the trajectory of the Low Energy Demand scenario [111] and the rapid decarbonisation pathway referred to as the "Carbon Law", which proposes halving emissions each decade while scaling up negative emissions to a multi-gigatonne scale. [112]

This gives a cumulative exceedance of the carbon budget as:

$$\Delta C_{ex} = \frac{C_{max}}{m_1}(1 - exp(-m_1 t)) + \frac{n}{m_2}(1 - exp(-m_2 t)) - nt \tag{2.15}$$

This is a transcendental equation which can be solved numerically: the duration of the implied overshoot occurs for the (non-trivial) root of $\Delta C_{ex} = 0$, while the maximum carbon budget overshoot - and, approximately, the maximum temperature overshoot - occurs when emissions are zero. Therefore, in a peak-shaving scenario, maximum SRM deployment is only reached when net zero emissions are also reached.

As an example, one could consider a best-case scenario of emissions peaking immediately close to their present value of around $40GtCO_2$/year, then declining (through conventional mitigation) at a given annual average percentage while negative emissions are scaled up to their maximum value. Different examples illustrate the trade-off between the implied duration of and height of the peak across a range of assumed

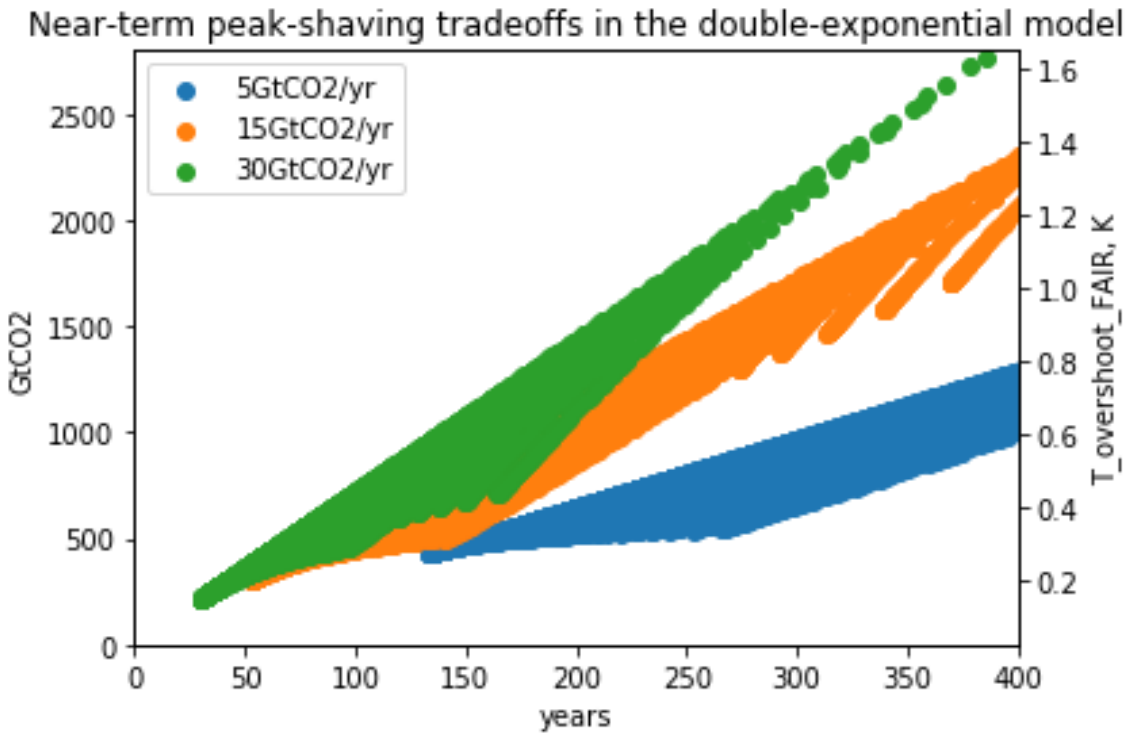

Figure 2.2: **The tradeoff between duration and height of overshoot in the double-exponential emissions model.** The TCRE is taken to be 1.4K/GtC as in the FaIR model. Mitigation and negative emissions deployment rates are both chosen in the interval [1, 7]% and three negative emissions floors are sampled.

mitigation rates. For this example, I choose mitigation rates between a maximum of 7% per annum, which corresponds to halving emissions in just under a decade, and 0.5%, which corresponds to halving emissions in around 140 years. This is illustrated in Figure 2.2, with each data point representing the solution to the transcendental equation for a different combination of mitigation and negative emissions deployment rates. This allows you to compare the relative effects of the negative emissions floor and changing mitigation rates on the possible peak-shaving trajectories across a wide range of idealised scenarios. Note the tradeoff between the duration of geoengineering and the height of the peak shaved in all scenarios. Having access to a greater maximum feasible rate of negative emissions (I illustrate the cases of 5, 15, 30 $GtCO_2$/yr) reduces the severity of this tradeoff. For example, a negative emissions floor of $30GtCO_2$ allows you to erase an overshoot of 0.5K in between 70 and 170 years, depending on the mitigation rate. With a negative emissions floor of $5GtCO_2$, this same 0.5K overshoot requires between 220 and 350 years of SRM and negative

emissions.

It has been argued that the best case for peak-shaving is a scenario where conventional mitigation is slow or expensive at first, but delaying near-term emissions cuts allow mitigation efforts and negative emissions to be deployed more rapidly and cheaply later in the century. [113] Scenarios like this are sampled here, but even in this extreme case shaving a peak of 0.5K involves 70+ years of geoengineering.

The maximum feasible "negative emissions" that could feasibly be obtained by techniques like afforestation, nature-based solutions, soil carbon sequestration, or enhanced weathering are on the order of 1-3GtCO_2/yr each. [114] While the overall contribution that bioenergy with carbon capture and storage can make is disputed [115], deployment at the scale of tens of gigatonnes will put substantial pressure on global land use, and even the median assumption in IPCC scenarios of 12GtCO_2/yr is controversial, with some researchers suggesting a lower feasible ceiling. [116] [117] It is also unclear for whether these techniques could maintain these maximum negative emissions rates in the multi-decadal or centennial scenarios considered here, or act as stable stores of carbon during that time This implies that avoiding a centennial commitment to SRM for peak-shaving of 0.5K will require the development and large-scale deployment of negative emissions technologies like direct air capture, which can be scaled up to tens of gigatonnes of CO_2 (constrained only by societal willingness to pay) in this century.

The preceding idealised analysis has assumed that the TCRE is constant, and takes the same value, for both negative and positive emissions. This is not necessarily the case, owing to carbon-climate cycle feedbacks, which illustrate that the TCRE will appear to exhibit hysteresis and depend on the historical path of emissions and temperatures, owing to the lagged thermal response of the ocean. This has implications for the results above, as it may mean that negative emissions appear to be less effective at cooling than emissions are at warming, due to feedbacks in the carbon cycle which could result in feedbacks, such as reduced efficacy of carbon sinks, or greater emissions from the biosphere; and feedbacks in the climate cycle, such as the lagged thermal response of the ocean, which reduce the efficacy of cooling. I will discuss this in greater detail in the discussion section 2.5.2.

2.3 Methods

2.3.1 The SSP-Overshoot Scenario

One can validate these approximations by considering concrete peak-shaving scenarios. Tilmes et al. (2016) [118] used the Community Earth System Model (CESM) to simulate the SSP5-3.4-OS scenario [119]. In this scenario, emissions follow the RCP8.5 pathway until 2040, the year emissions peak. Then, emissions follow an idealised pathway, declining to become net-negative by the end of the century, before reaching a floor of negative emissions.

The mitigation rate of -1.2Gt CO_2/yr/yr, as well as the floor of negative emissions at -18.5GtCO_2/yr, are both drawn from the Integrated Assessment Modelling literature: specifically, the Shared Socioeconomic Pathway database [120], representing the maximum decarbonisation and negative emissions rates respectively. This pathway therefore represents a peak-shaving scenario at the upper end of what is considered feasible in the current scenario literature: there is a significant exceedance of temperature targets due to following RCP8.5 to 2040, then rapid decarbonisation and large-scale negative emissions. Tilmes et al. (2016) [118] simulate SRM by using stratospheric aerosol injections to reduce temperature anomalies to 2K from this pathway.

In CESM, which has an equilibrium climate sensitivity of around 4K, [121] this pathway results in temperature anomalies which peak at 3K above the preindustrial. The warming rate in CESM approaches 0.05K/year at RCP8.5 2040 when its peak-shaving begins (compare to 0.045K/year in the CMIP5 ensemble for RCP8.5 in 2040) [122]. The cooling rate during the sustained period of negative emissions is just over 0.01K/year.

The analytical approximation above therefore suggests a deployment time for SRM to "peak-shave" from 3K to 2K as around 40 years while temperatures rise and around 100 years as they decline for a total duration of 140 years. CESM's explicit simulation of this scenario found that geoengineering is deployed for around 130 years to shave the peak below 2C.

2.3.2 The FaIR simple climate model

In subsequent exercises, I will determine the global mean temperature response to a range of forcing scenarios using the FaIR simple climate model. [123]. The FaIR model emulates the carbon cycle and climate response to a range of different forcings, including short-lived greenhouse gases and methane, but for simplicity I consider only additional forcings due to CO_2 emissions and aerosol optical depths using the modelling for volcanic forcing. FaIR converts emissions of greenhouse gases and other climate forcers to a concentration and radiative forcing time series using its carbon cycle component and radiative forcing calculations, and then uses these radiative forcings to drive the simple climate model.

The carbon cycle component of FaIR is based on a modified four time-constant impulse response function that has been shown to be a good emulator of more complex Earth System Models. [124] The advantage of using this simple model is that I can explore a large parameter space very rapidly while its fidelity of emulation for carbon cycle means that I can test the key assumption in the simple, analytically derived model - that temperature change is proportional to cumulative emissions at all times. It also allows us to investigate the implications of path-dependence for the TCRE, which I cover in the discussion section.

For a given emissions trajectory, uncertainties around the climate sensitivity affect the overshoot trajectory. I use the FaIR simple climate model to sample climate sensitivities from across the IPCC AR5 "likely" range for climate sensitivity [125]. Non-CO_2 emissions are assumed to follow a 1.5C-compatible mitigation pathway from 2020, stabilising at a constant radiative forcing of $0.45 W/m^2$ in the long term. In these versions of the scenario, the negative emissions are switched off when temperatures reach 1.5C.

Note that in the simulations that use FaIR, for simplicity and to avoid needing to implement a computationally costly feedback loop in FaIR, I only simulate the emissions trajectories rather than explicitly simulating SRM that holds temperatures to below a given threshold. This means that the "peak-shaving" and duration of commitment to SRM is estimated, rather than simulated. I discuss the justification

for this simplification, and the implications of this for our results, in the discussion section. I will now characterise the simulations used to drive FaIR.

2.3.3 Scenarios used to drive FaIR

2.3.3.1 The SSP-OS scenario in 2020 and 2040

In the following sections, I drive FaIR with a range of idealised peak-shaving emissions scenarios, based on the SSP-OS scenario described above, and scenarios taken from the Integrated Assessment Modelling literature.

I begin by looking at a simple case in detail - two variants of the SSP-OS scenario described above, where mitigation begins in 2020 or in 2040, illustrating the impact of delayed mitigation and varying one parameter, the year that mitigation begins. The detailed model output from this is illustrated in Figure 2.3 including plotting the emissions trajectory, the cumulative emissions, the temperature trajectory, and the radiative forcing that would be required to suppress the temperature overshoot, subject to a confidence interval for the climate sensitivity.

2.3.3.2 Overshoot scenarios in the Integrated Assessment Modelling Literature

The Integrated Assessment Modelling (IAM) literature includes many examples of emissions scenarios that are compatible with 1.5C of warming by 2100. These scenarios vary in terms of their assumptions about the rates of mitigation and negative emissions that are attainable. The emissions in these scenarios that are deemed 1.5C compatible also vary, as the different integrated assessment models use differing climate sensitivites, which renders different emissions pathways compatible with 1.5C depending on the model. In addition, some of the scenarios have a more pronounced peak-and-decline - those which depend more on late-in-the-day negative emissions - while others rely on rapid mitigation to avoid peak-and-decline altogether. These scenarios are illustrated in 2.4 and listed in the Appendix, along with the cumulative carbon emissions associated with each scenario.

The 1.5C compatible scenarios from the IIASA database comprise 43 scenarios from six different intergrated assessment models - GCAM, IMAGE, MERGE, MESSAGE, REMIND-MAGPIE, and WITCH-EMF. [126] To illustrate the nature of some of the temperature peaks and declines that are modelled by integrated assessment models, I select only the subset of scenarios which have an explicit temperature peak that begins to decline before 2100, which leaves 22 scenarios. In order to be compatible with the simple CO_2-only analysis in the earlier idealised treatment, I take only the CO_2 emissions associated with these pathways, and run them through FaIR with both default and high climate sensitivities to illustrate the typical range of peak-and-decline scenarios that are plotted by integrated assessment models.

I can extend these scenarios so that they resemble the idealised analysis elsewhere in this chapter. I assume that mitigation (and SRM to shave the peak) begins in 2020 when the scenarios diverge. The "peak" is then defined as the duration for which the temperature is above its value in 2020. I can extend the scenarios past 2100 by assuming that the negative emissions hit a "floor" in 2100 and can continue at that rate indefinitely thereafter.

2.3.3.3 Idealised variants on the SSP-OS overshoot scenario

Finally, I examined a much wider range of variants of the SSP-OS scenario. Variants can be characterised by varying three different parameters - the year that emissions peak (i.e., the year until which they follow the SSP8.5 trajectory, with no mitigation), the mitigation rate (in $GtCO_2$/yr/yr), and the negative emissions floor. The trade-offs between the temperature overshoot avoided and the duration of geoengineering for each of these scenarios, and for two different values of the climate sensitivity, are plotted in Figures 2.6 and 2.7.

Here, I consider variants of the idealised SSP5-34-OS case where emissions follow RCP8.5 until a given year (year_peak), then decline and go negative with an average mitigation rate (mit_rate), eventually reaching a negative emissions floor (neg_floor). I randomly sample 5000 sets of year_peak, mit_rate, neg_floor parameters from "feasible" ranges, discussed below. These parameters are used to calculate CO_2 emissions trajectories, which in turn are used as scenarios to drive FaIR in CO_2-only mode. The peak of the temperature trajectory that would be avoided by SRM which holds

anomalies to 1.5K, and the duration of the overshoot, are calculated from the temperature trajectory simulated by FaIR.

The aim of this is to examine a much wider range of possible peak-shaving scenarios than are considered in the Integrated Assessment Modelling literature, to test the robustness and outer limits of the trade-offs that I identify in FaIR. Note that FaIR includes a sufficiently complex carbon cycle that its TCRE is path-dependent, as discussed in the Discussions section.

The year that mitigation starts is randomly sampled between [2020, 2060] – CO_2 emissions follow RCP8.5 prior to this year. Scenarios that do not overshoot 1.5K, or which do not decline below 1.5K during the 750-year integration time of the model, are filtered out. The mitigation rates are sampled from the interval [0.1, 2.3] $GtCO_2/yr/yr$. $2.3GtCO_2/yr/yr$ is the maximum mitigation rate achieved at any point in the ambitious Low Energy Demand scenario [111].

The negative emissions rate is sampled from the range [3.5, 30.0] $GtCO_2/yr$. While the lower end of this range assumes that sustained large-scale BECCS and DAC are unavailable, and is determined by the land-use change negative emissions from the Low Energy Demand (LED) scenario of Grubler et al. (2018) [111], the high end of this range is well beyond the maximum considered feasible for BECCS by any 1.5C-compliant scenario, and substantially higher than the largest negative emissions of $18.5GtCO_2/yr$ in the scenario database.

Realmonte et al. (2019) [127] compared the use of DAC across two integrated assessment models, considering a maximum deployment by 2100 of $30GtCO_2/yr$ and a maximum scale-up rate of 1.5 $GtCO_2/yr/yr$. This does not imply that the most extreme scenarios sampled here are feasible - for example, the rapid mitigation under the LED scenario may not be compatible with the large-scale expansion in energy supply required for DAC on a $30GtCO_2/yr$ scale. If large-scale direct air capture is the predominant form of negative emissions, then constraints are likely to be set by societal willingness-to-pay, energy availability, and the speed that the technology can be scaled up and deployed; the implications of this large-scale deployment are explored in Realmonte et al. (2019) [127]

2.4 Results

2.4.1 The SSP-OS scenario in 2020 and 2040

The results, for the standard overshoot parameters with mitigation beginning in 2020 and 2040, are illustrated in Figure 2.3. The uncertainty in the emissions and cumulative emissions pathways illustrated corresponds to the 67% confidence interval for climate sensitivity sampled in FaIR, as the time for which negative emissions are switched off and 1.5C is reached depends on the TCRE.

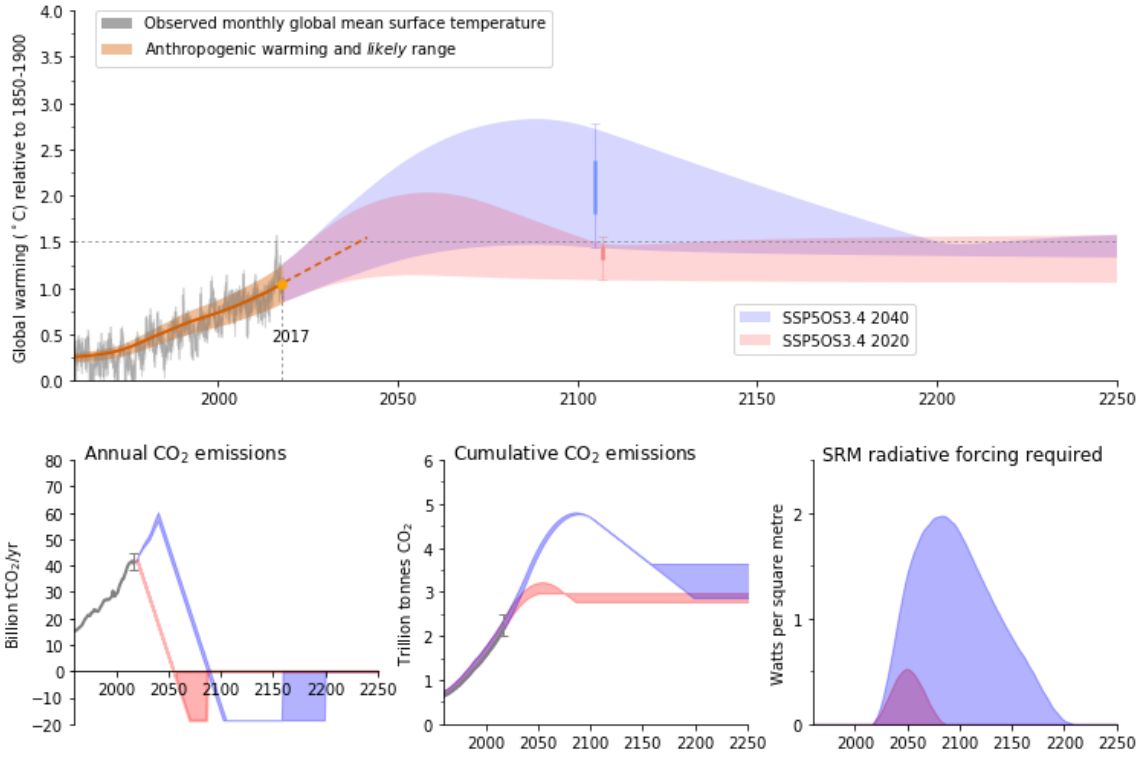

Figure 2.3: **Variants of the SSP5-34-OS scenario, which follows the RCP8.5 scenario with rapid mitigation beginning in 2020 and 2040 respectively.** The grey line corresponds to observed temperatures. The necessary radiative forcing from SRM to keep temperatures to below 1.5C in these overshoot scenarios is calculated and plotted in the lower right-hand graph. Uncertainty in the temperature response and radiative forcing required, as well as the emissions pathways, corresponds to the 66% confidence interval for the TCRE from IPCC's AR5 and is plotted with shading.

Note that there are regions on the graph where the 66% confidence ("uncertainty") interval intersects the axis. This illustrates scenarios where no SRM or negative

emissions are required to stay below 1.5C - for example, in the case where mitigation begins in 2020, and the TCRE is at the lower end of estimates. For the high-end climate sensitivity, when mitigation begins in 2020, 1.5K is not exceeded if a period of low-level SRM is maintained out from around 50 years. In the case where mitigation does not begin until 2040, negative emissions and peak-shaving are a much longer-term commitment, with SRM for peak-shaving required for up to 170 years. When climate sensitivity is high, this corresponds to avoiding a peak of almost 3C of warming if mitigation is delayed to 2040, while the peak is closer to 0.5C if climate sensitivity is high and mitigation begins in 2020.

Notably, the difference in the length of commitment to negative emissions (and hence SRM) depending on the range of climate sensitivity is only around 50 years in both scenarios, compared to the difference between scenarios of around a century of implied commitment to negative emissions and SRM due to delaying the onset of mitigation by 20 years. This illustrates that faster near-term mitigation is more important than uncertainties in the climate sensitivity for limiting the length of a commitment to SRM for peak-shaving. In case with highest climate sensitivity, at its peak, SRM is needed to offset around $2W/m^2$, similar in scale to the entire anthropogenic contribution to radiative forcing in 2019 of $2.7W/m^2$.

2.4.2 Overshoot scenarios in the Integrated Assessment Modelling Literature

I simulate the explicit overshoot scenarios in the IIASA database in FaIR. In extending these scenarios, the inherent trade-offs between the duration of peak-shaving and the suppressed temperature overshoot can be determined.

Figure 2.4 illustrates the trade-offs in peak-shaving deployment under the range of mitigation and negative emissions rates considered feasible by the integrated assessment modelling literature, under indicative high-end and low-end climate sensitivities. Note that when climate sensitivity is higher, peaks are also higher due to the greater sensitivity of temperature to cumulative emissions. We can see that for most of the scenarios, in the lower-end of climate sensitivity, the temperature anomalies do not exceed 1.5K - this is in accordance with these scenarios being 1.5C compatible even

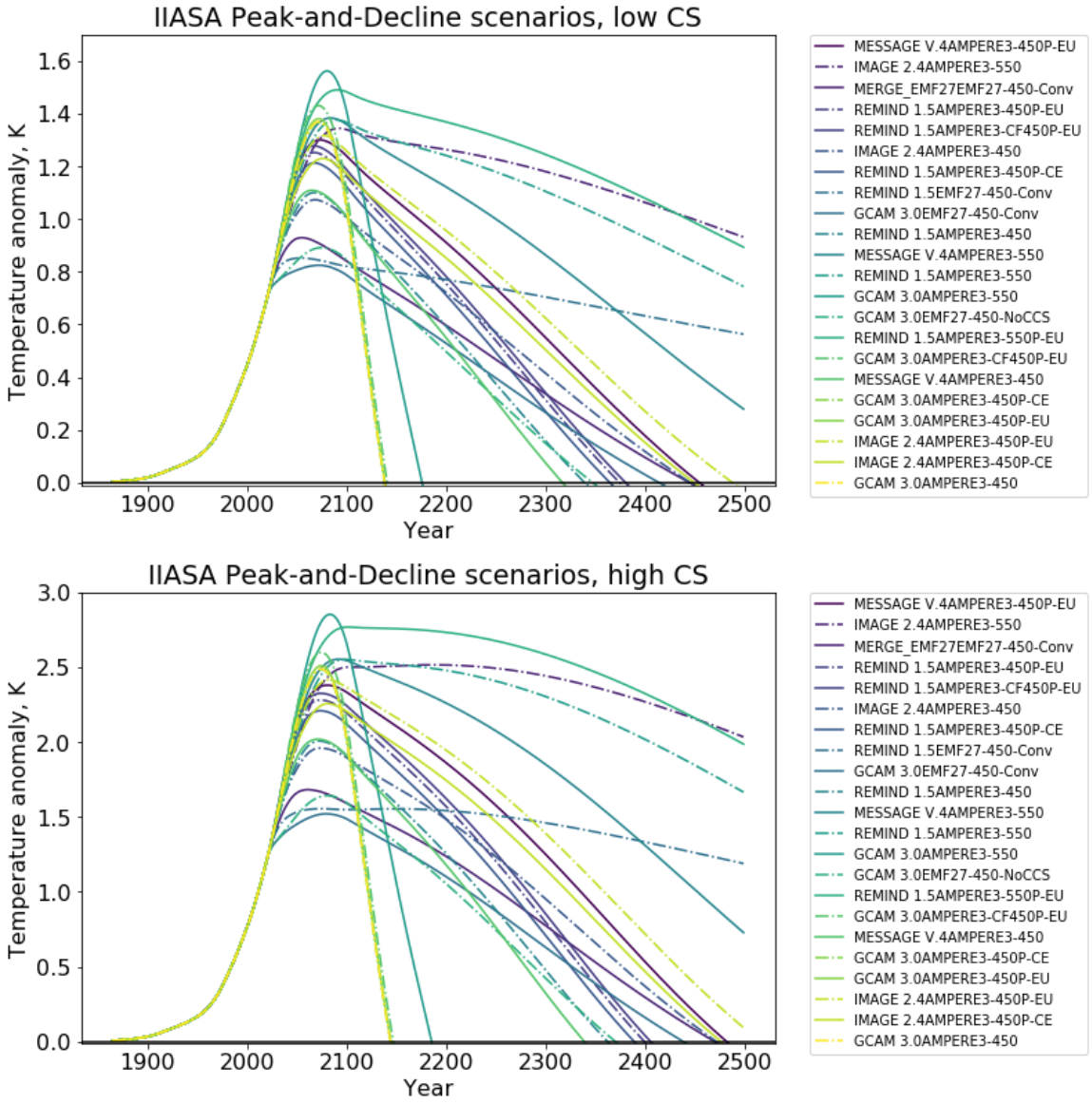

Figure 2.4: **Illustration of the IIASA database climate scenarios described.** The upper plot uses the FaIR default climate sensitivities (which are towards the low end of the CMIP6 range), while the lower plot illustrates the same CO2 emissions pathways but with the higher UKESM1 ECS and TCR of 5.4K and 2.7K respectively.

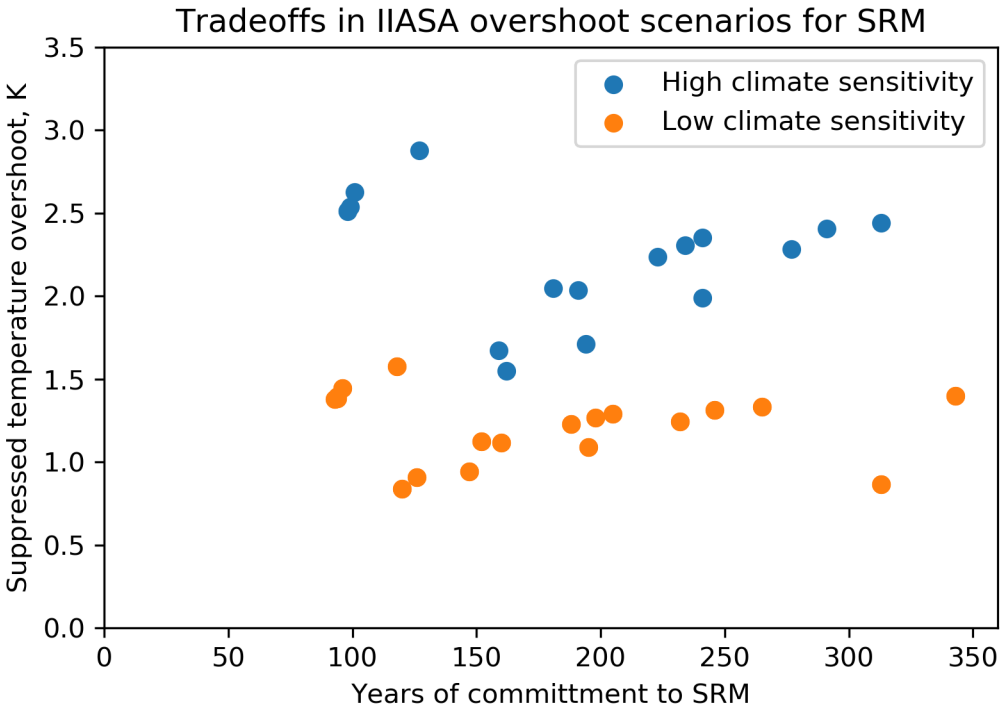

Figure 2.5: **Illustration of the IIASA database climate scenarios that are 1.5C compatible but have an explicit peak-and-decline, using CO2 emissions only.**

when non-CO2 emissions are taken into account and under median climate sensitivity, where they will reach 1.5K by the end of the century. In the high climate sensitivity case, the same set of scenarios peaks with a temperature anomaly at more than 2.7K as opposed to just over 1.5K in the low sensitivity case, illustrating the increased efficacy of peak-shaving when climate sensitivity is high.

Figure 2.4 plots the temperature trajectories for each of these scenarios, while Figure 2.5 plots the duration of peak-shaving and the height of the peak shaved for both sets of climate sensitivities across these scenarios. In the most extreme scenario of the IIASA database, and with a high climate sensitivity, we can see that temperature anomalies of a degree over 1.5K can be suppressed with geoengineering that takes place over around a century, but for less extreme cases where mitigation and negative emissions scale up less rapidly, suppressing a similar temperature overshoot could entail 300 years of geoengineering, illustrating the strong dependence of the commitment to geoengineering on the underlying emissions trajectory even when climate sensitivity is high.

The simple rules of thumb attainable from the toy model analysis apply to these scenarios: there is an inherent trade-off where suppressing higher peaks of temperature requires commitment to SRM for more than a century. This arises due to the requirement to remove all of the excess cumulative CO_2 emitted over the 1.5C limit before SRM can be discontinued. For this reason, the trade-off between the excess temperature avoided by peak-shaving and the years of commitment to SRM is less dramatic under higher climate sensitivities. Under higher climate sensitivity, smaller carbon budget excursions are responsible for higher temperature overshoots, so they can be more rapidly corrected by net-negative emissions.

2.4.3 Idealised variants on the SSP-OS overshoot scenario

The idealised scenarios which are variants on the SSP-OS overshoot scenario are simulated in FaIR, and are generated by randomly selecting mitigation rate, maximum negative emissions rate, and the year that mitigation begins from ranges of feasible values as described in the methods section.

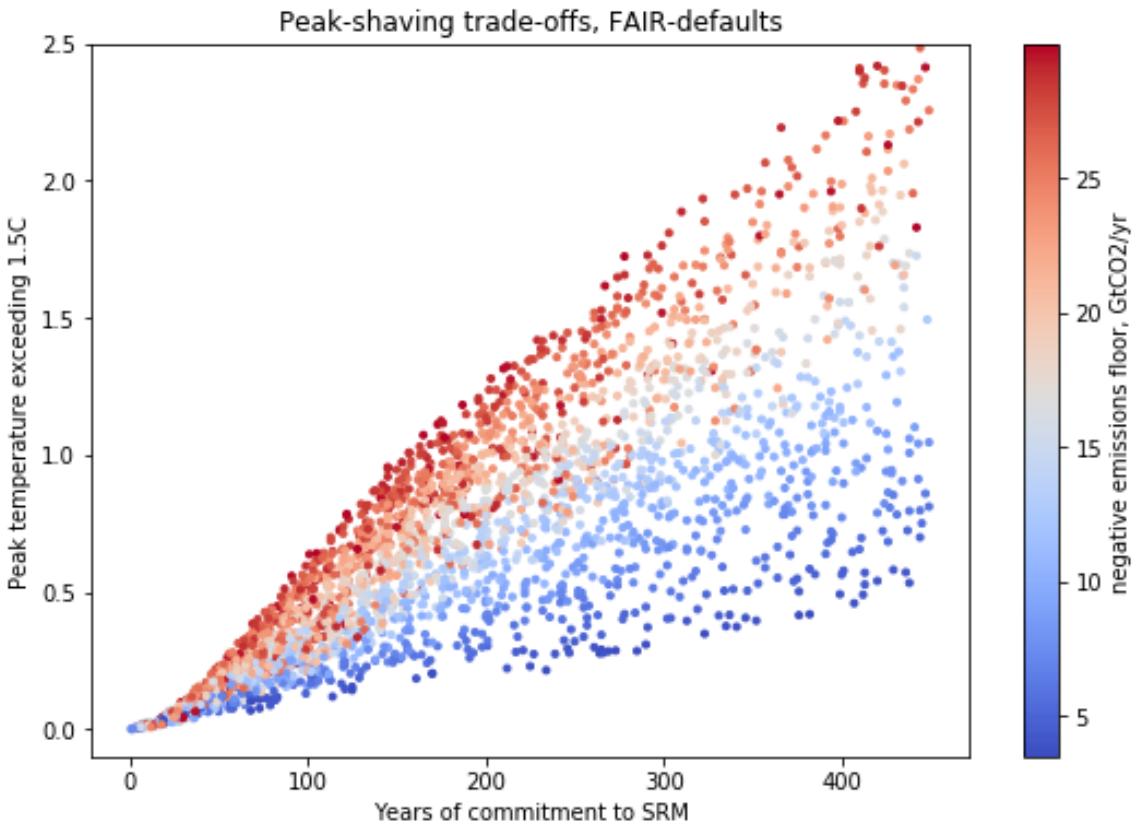

Figure 2.6: **Trade-off between years of committment to solar radiation management and the peak in temperature avoided by a range of idealised overshoot and negative emissions scenarios, under the FaIR default ECS and TCR of 2.86K and 1.53K respectively.**

Figure 2.6 illustrates the trade-off between the duration of the commitment and the magnitude of the temperature overshoot under a range of different scenarios. We can see that even if very high levels of negative emissions are permitted, there is a sharp trade-off between the number of years of committment to SRM and the suppressed temperature overshoot. Even if the most rapid mitigation and large-scale negative emissions are assumed, shaving a peak of 2C to 1.5C requires a deployment in excess of 80 years. If negative emissions are not available at sufficient scale ($> 2 - 3\text{GtCO}_2/\text{yr}$), shaving a substantial (>0.5K peak) becomes impossible on the approximately 450-year timescale considered here.

This is dependent on climate sensitivity, as illustrated by Figure 2.7, which illustrates that the trade-off between geoengineering duration and the temperature reduction it achieves is arguably more favourable when climate sensitivity is high.

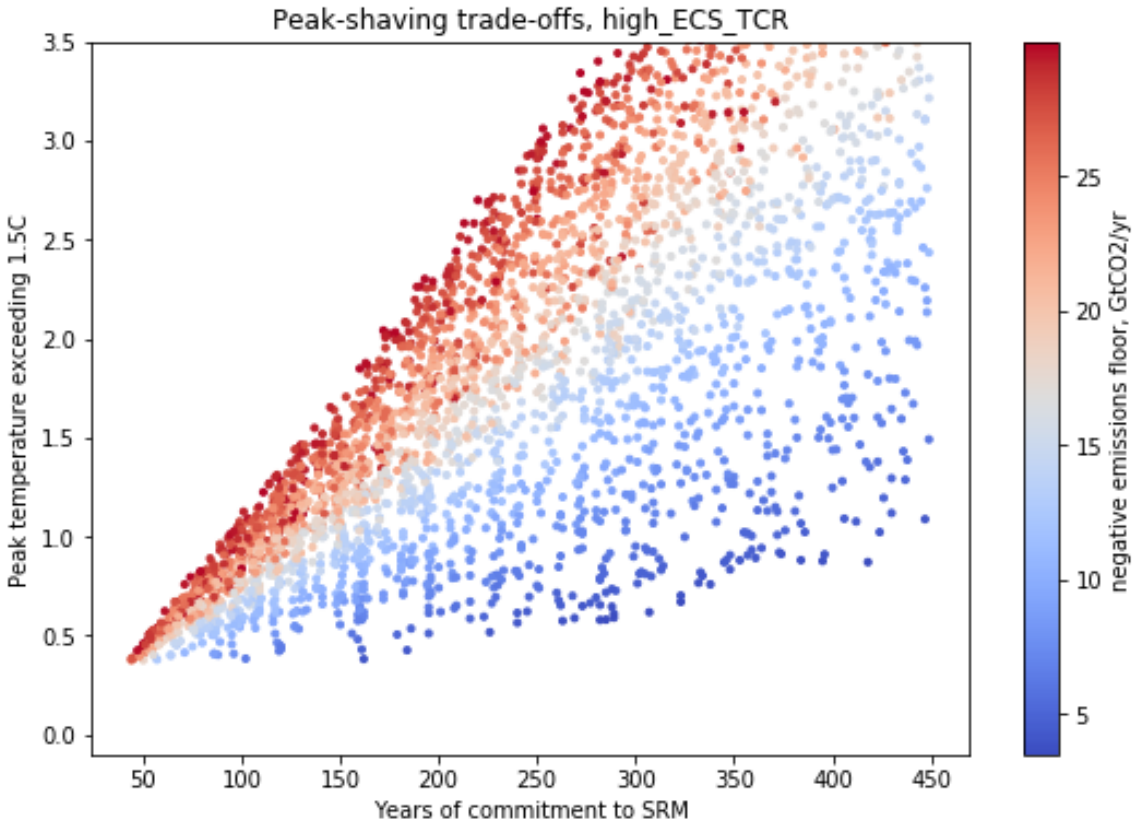

Figure 2.7: **As in the previous figure, but with the UKESM1 ECS and TCR of 5.4K and 2.7K respectively, at the higher end of any estimate for climate sensitivity.**

Because the TCRE is high, small CO_2 excursions lead to significant rises in temperature. This means that the CO_2 corresponding to a temperature rise of 0.5K could be removed in decades, rather than around a century as in the low climate sensitivity case.

With the UKESM climate sensitivity [128], even the fastest mitigation rate considered (where emissions start declining in 2020, hit net zero by 2038, and become negative thereafter) overshoots the 1.5K target by 0.4K, entailing around 40 years of commitment to SRM. As in the case of modelling IIASA scenarios with different climate sensitivites, higher climate sensitivites mean smaller carbon budget overshoots correspond to larger temperature overshoots. Given that the duration of SRM deployment is equivalent to the time spent overshooting the carbon budget, added to the time required to remove the carbon budget overshoot, higher climate sensitivity implies shorter overshoot commitment times to avoid the same level of warming.

The requirement to remove the entire overshoot of the carbon budget before temperatures will decline to below the threshold results in a high sensitivity of the duration of the commitment to any "floor" in achievable negative emissions, especially for large overshoots. Table 2.1 shows how the various carbon budget overshoots depend on the TCRE and the temperature overshoot, illustrating the commitment to global net-negative emissions implied by various temperature overshoots with different TCREs. Note that, as most overshoot scenarios require negative emissions to cancel out residual anthropogenic emissions as well as to provide global net-negative emissions, the actual realised negative emissions will be higher in these scenarios. [129]

TCRE/Overshoot	0.5K	1K	1.5K	2K
0.92K (17%)	543	1086	1629	2172
1.26K (33%)	396	793	1189	1586
1.5K (50%)	333	666	999	1332
1.71K (67%)	292	584	876	1168
2.05K (83%)	244	487	731	975

Table 2.1: **Implied carbon budget overshoots ($GtCO_2$) for various levels of temperature overshoot and TCREs from different percentiles of the IPCC AR5 WG1 distribution of likely TCREs.** These correspond to the cumulative net-negative emissions required to return to 1.5K in each overshoot case.

2.5 Discussion of results and limitations, and conclusions

Our results illustrate that, in a hierarchy of models from simple analytical rules of thumb, through simple climate model simulations of overshoot trajectories in the Integrated Assessment Modelling Literature, and through to a simulation of a very wide range of idealised overshoot scenarios, there are always significant trade-offs between the length of time that geoengineering must be deployed and the actual temperature overshoot that is suppressed by geoengineering. These trade-offs are such that avoiding a significant overshoot of (say) 0.5K generally entails a commitment to geoengineering and negative emissions for at least a century under most assumptions

around emissions trajectories and climate sensitivity. Analysis of scenarios from integrated assessment models extended into the 22nd century, as well as the very wide range of variants on the overshoot scenario, suggests that even this relatively short duration of implied commitment is only likely to be realised where mitigation rates and negative emissions potential are at the higher end: for example, the IIASA scenarios illustrate that peak-shaving to shave a peak from 2.5C to 1.5C could last for between 100 and 300 years depending on the underlying emissions trajectory. However, there are some caveats to the above modelling which may affect these trade-offs, which I turn to discuss now.

2.5.1 Explicit simulation of SRM in FaIR

One of the limitations of the analysis above is that I did not explicitly simulate SRM taking place in FaIR and calculate the temperature trajectory with both SRM forcing and CO_2 forcing imposed together. Instead, I simulated only the response to emissions trajectories and inferred the implied temperature overshoots that geoengineering would avoid, and duration of geoengineering that would be required to accomplish this, from the temperature trajectories themselves. However, the climate stabilised at 1.5K would be different to that which is allowed to overshoot past 1.5K and then return to 1.5K with negative emissions, in terms of climate and carbon-cycle feedbacks and the heat absorbed by the ocean during the years of overshoot. For this reason, one might expect overshoot calculations as conducted above to slightly overestimate the actual length of time that SRM would need to be deployed to sustain temperatures with this emissions trajectory, because the carbon-cycle feedbacks and ocean heat absorption in a world where the SRM is actually deployed would be smaller.

To test the size of this effect in FaIR, I ran a simulation of the standard SSP-OS scenario, with the model's default climate sensitivity, with mitigation beginning in 2040 such that there is an overshoot (of 1.5K) that has a temperature peak at around 2K. I then manually diagnosed the required volcanic radiative forcing in each year to keep temperature anomalies below 1.5K by running FaIR once for each year of the simulation's duration, and adjusting the radiative forcing accordingly for each individual year until it was large enough to ensure that the temperature remained

below 1.5K. The resulting diagnosed radiative forcing was then imposed on top of the emissions trajectory to assess its effect.

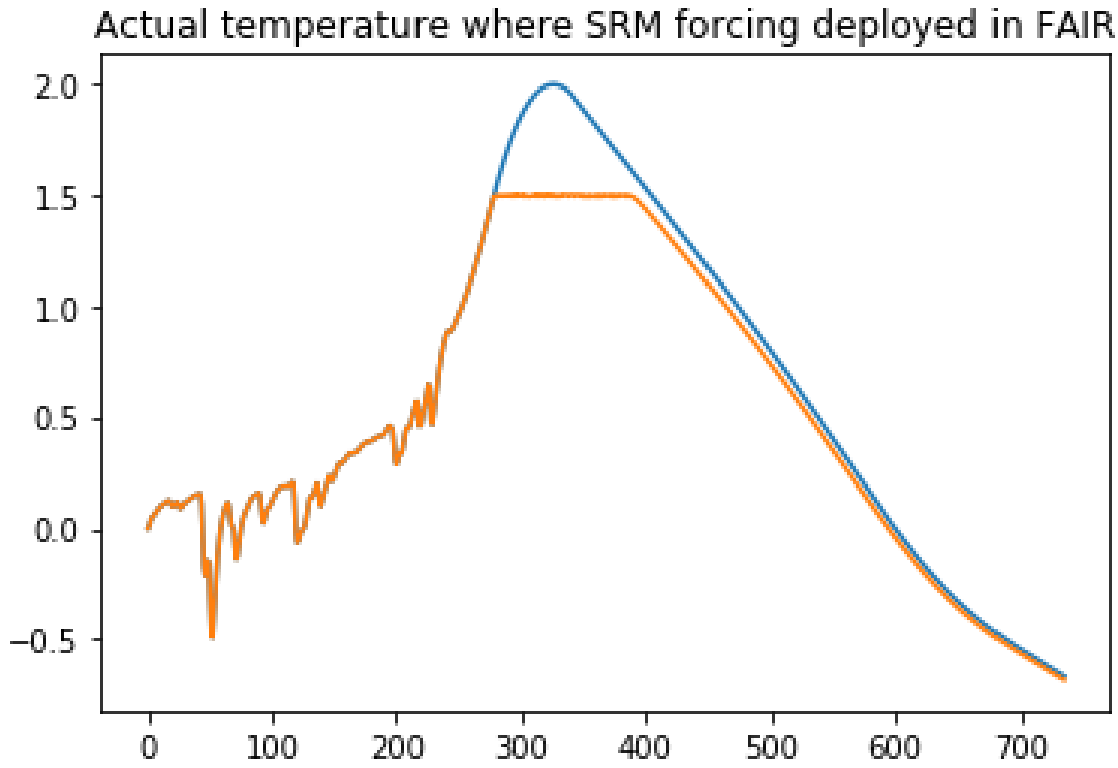

Figure 2.8: **Temperature trajectories in the scenario where SRM is explicitly deployed (orange) and emissions-only (blue) SSP-OS scenarios.** The x-axis is years since 1750 and the y-axis is temperature above the preindustrial mean.

The results are plotted in Figures 2.8 and 2.9. We can see that the impact on the carbon-cycle is relatively small for this level of peak-shaving, corresponding to 3.1ppm at the peak difference between the scenarios, which corresponds to around a year of 2020 SSP-OS 8.5 emissions, or in other words delaying the onset of mitigation in this scenario by one year. The corresponding difference between the estimated duration of geoengineering from considering the emissions-only temperature trajectory and the actual duration that SRM needs to be deployed before temperature reduces to below 1.5K is approximately 5 years. Given that, in this scenario, geoengineering needs to be deployed for around 140 years to "shave the peak" to 1.5K, this discrepancy is relatively small in this case, suggesting that for these climate sensitivity parameters and given the temperature-interactivity of the carbon cycle in FaIR, approximating

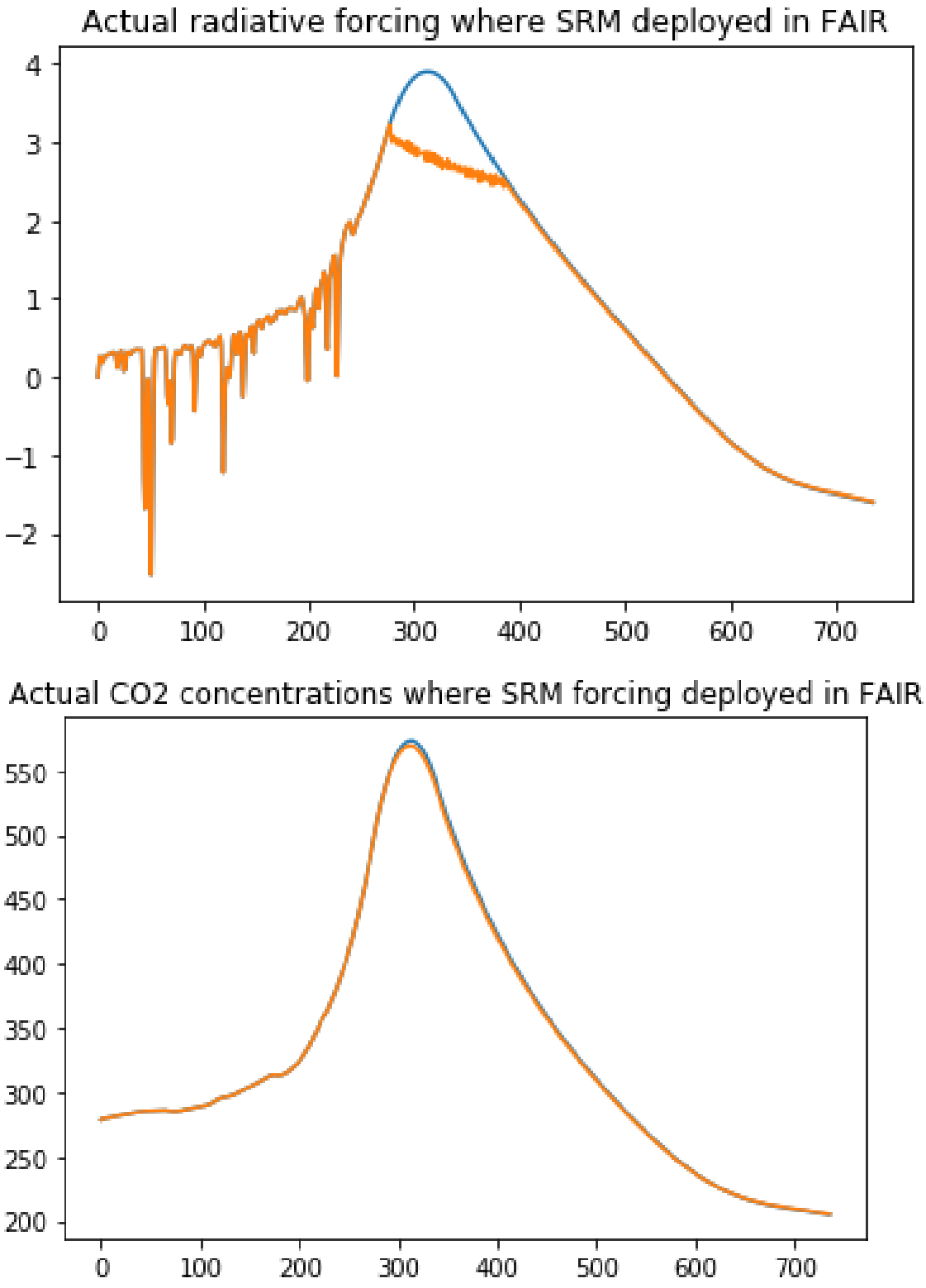

Figure 2.9: **Radiative forcing and carbon dioxide concentration trajectories in the scenario where SRM is explicitly deployed (orange) and emissions-only (blue) SSP-OS scenarios.** The x-axis is years since 1750 and the y-axis for Radiative Forcing is the net radiative forcing for anthropogenic greenhouse gases plus SRM in Watts per square metre, while the y-axis for CO2 is the concentration in parts per million in the atmosphere.

the implied duration of geoengineering using the emissions-driven temperature trajectory rather than explicitly simulating SRM is a reasonably good approximation. However, this error is likely to increase for trajectories with a longer overshoot, or if the climate sensitivity of the model differs, as I discuss in the next section.

2.5.2 Path-dependence of the TCRE

The preceding idealised analysis in Section 2.2 assumed that the TCRE is constant, and takes the same value, for both negative and positive emissions. However, more detailed modelling of the carbon-climate cycle suggests that the TCRE exhibits hysteresis, depending on the historical path of emissions, owing to the lagged thermal response of the ocean [130], which means that negative emissions are less effective at cooling than emissions are at warming, at least in the short term. This has implications for overshoot trajectories which worsen as the overshoot increases, although analysis of ESMs suggests that the concept of a carbon budget and the proportionality of temperature to emissions remains robust up to excursions of around 300 gigatonnes of carbon, which with a median TCRE (of around 1.6K per trillion tonnes of carbon, or 0.44K per trillion tonnes of CO_2) [131] would correspond to temperature overshoots of around half a degree of warming. [132]

This non-linearity is most prominent when the lagged ocean response is large, which occurs when the carbon budget overshoot is large. For example, Zickfeld et al. (2016) [130] modelled an emissions trajectory where CO_2 concentrations were increased by 1% per annum to 4x preindustrial levels and then decreased in the same way. They found that the TCRE was approximately 1.59K in the ascending branch and 1.48K in the descending branch, i.e. 7% lower. In scenarios with smaller total cumulative emissions, the difference between the ascending and descending branch was smaller (or even of the opposite sign, implying negative emissions were slightly more effective at cooling than positive emissions were at warming.) This implies that the analytical estimates above are likely to be underestimates for the real-world commitment to peak-shaving.

Depending on the emissions scenario and the model used, the approximation that global mean temperature is proportional to cumulative emissions - and hence that

TCRE is constant and well-defined - can be less accurate. Physically, the concept of a constant TCRE depends on the approximate cancellation of two factors: residual surface warming as the ocean comes into equilibrium with the high-CO_2 atmosphere, and uptake of carbon dioxide by the ocean sink and biosphere [108]. Consequently, factors that vary between models - such as the rate of uptake of CO_2 by the oceans and the biosphere, as well as the timescale and magnitude of long-term warming as the deep ocean comes into equilibrium with the atmosphere - can introduce a hysteresis effect in temperature as cumulative emissions are reduced, and hence an apparent time-dependence of the TCRE.

This behaviour can be modelled with FaIR by using idealised emissions trajectories and varying the TCR/ECS ratio, known as the "Realised Warming Fraction" (RWF). RWFs for the CMIP5 models lie in the range 0.45-0.75, with an average of 0.55: observationally constrained estimates tend to fall in the upper half of this range. [133] In these scenarios, emissions linearly decrease from an initial maximum value E to -E, resulting in net cumulative emissions of zero. The TCRE is evaluated for the ascending and descending branches by calculating the gradient of the temperature-cumulative emissions curve, then averaging across the periods of positive and negative emissions respectively. The ratio of the TCRE in the ascending and descending branches is plotted against the RWF for several emissions scenarios in Figure 2.12.

If the RWF is low, or the total carbon perturbation is high, the assumption of linearity is less accurate. In these cases, global mean temperature can continue to rise even as emissions go negative, as the large "unrealised" warming dominates the effect of declining CO_2 concentrations, reducing the average TCRE across this descending branch. This will result in the analytical formulae of Section 2 underestimating the commitment time to peak-shaving geoengineering. By contrast, if the RWF is high, and the carbon perturbation is small enough, the TCRE ratio can be less than 1: in other words, natural removal of CO_2 by sinks exceeds the effect of long-term warming, which results in faster cooling during the negative emissions branch, as concentrations decline more quickly.

The non-constant nature of the TCRE is an important caveat for the early, analytical calculations of the trade-off between duration and the height of the peak shaved - especially if the RWF is at the extreme ends of current estimates, or the carbon per-

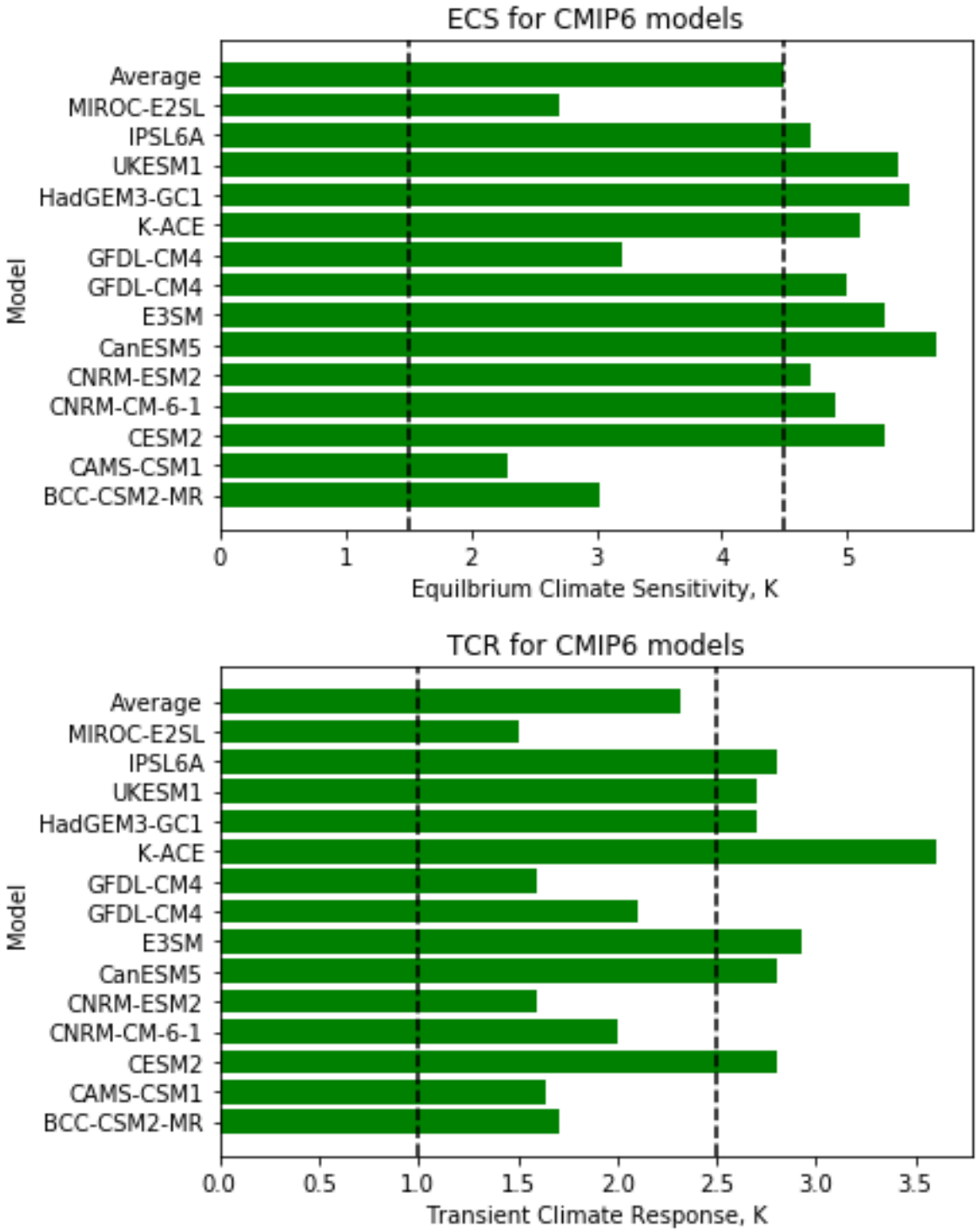

Figure 2.10: **Climate sensitivities (top to bottom: ECS, TCR) for the CMIP6 ensemble.** The ECS range is between 2.2-5.7K and the TCR range between 1.6K and 3.6K. The IPCC AR5 "likely" ranges are illustrated for AR5 with dashed lines in both cases. Results earlier in the chapter sampled the (relatively low) standard FaIR sensitivites (1.53K, 2.86K), as well as high-end climate sensitivites (2.7K, 5.4K).

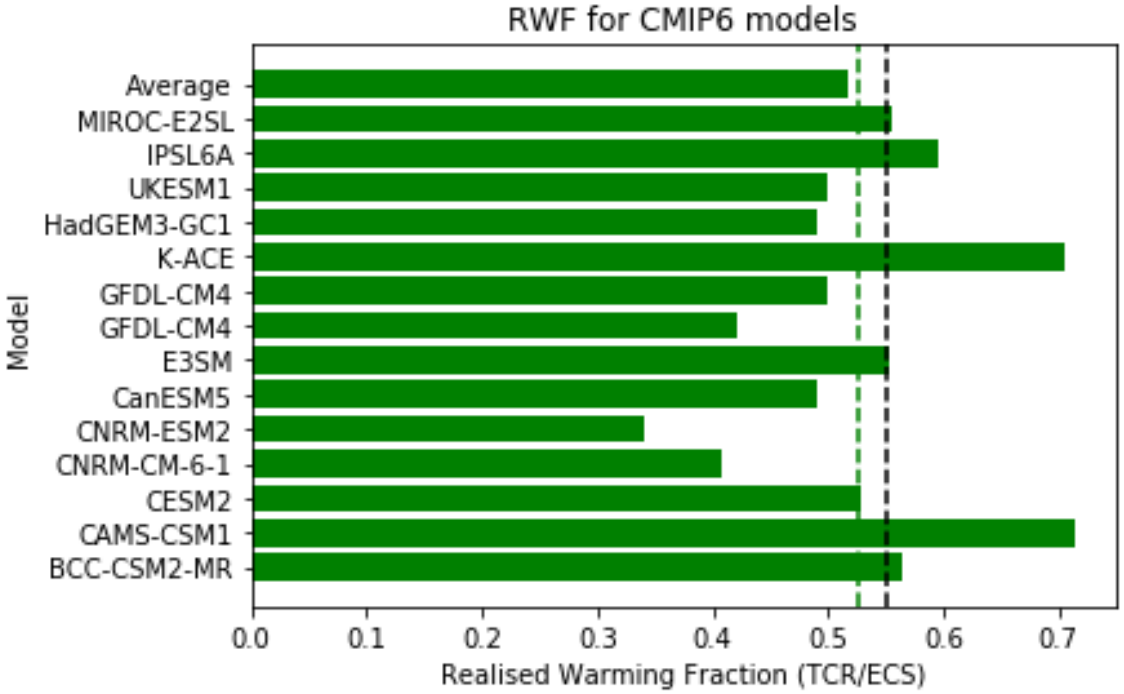

Figure 2.11: **Realised warming fractions (RWFs) for the CMIP6 ensemble.** Note the FaIR default of 0.53. The CMIP6 range (0.34-0.72) is wider than the CMIP5 (0.45-0.75) and has a slightly lower average (0.52 vs 0.55). Note that the results earlier in the chapter sampled the standard FaIR sensitivites (1.53K, 2.86K), as well as high-end climate sensitivites (2.7K, 5.4K), but with an RWF close to the multimodel average in both cases (0.53 and 0.5 respectively.)

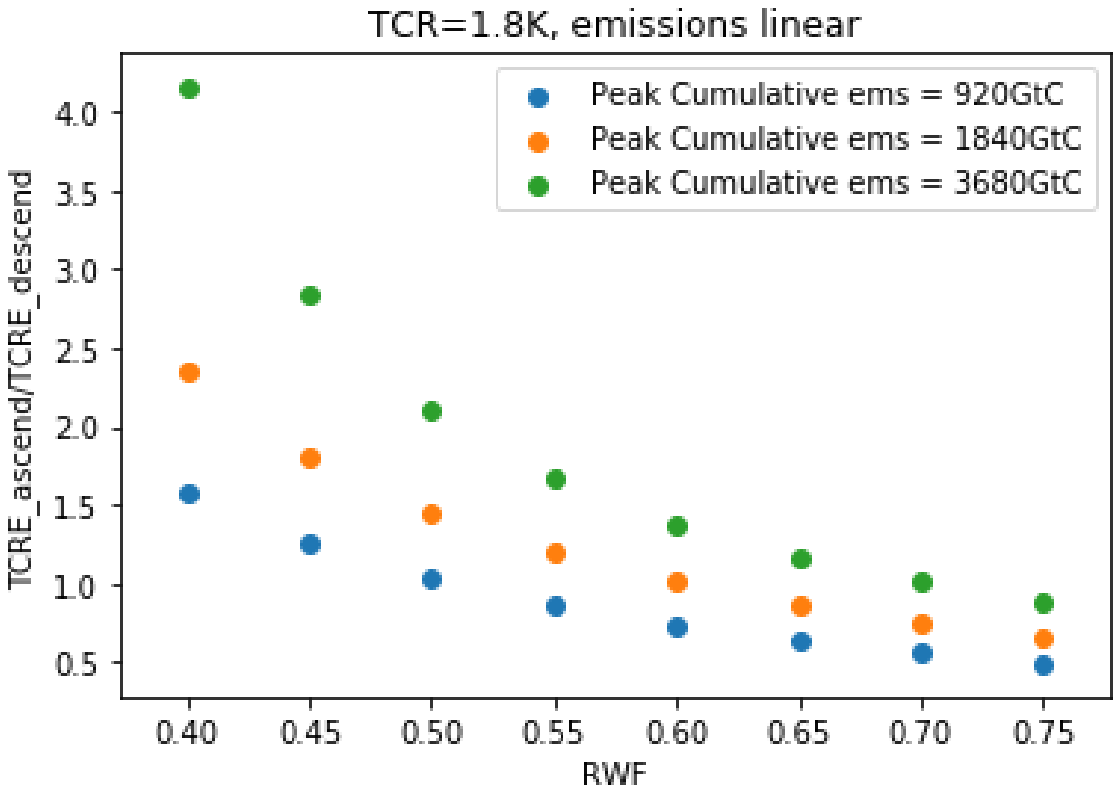

Figure 2.12: **Ratio of TCRE during positive and negative emissions periods for various Realised Warming Fractions (TCR/ECS ratios).** This was calculated by simulating simple emissions trajectory where emissions linearly decrease to zero and then become negative at the same rate, until cumulative emissions are zero. Differences both in the total anthropogenic CO2 excursion and the RWF give rise to different levels of hysteresis. The approximation that the TCRE is constant for removals and additions of CO2 - which is relied on for the simple algebraic models, but not the FaIR analysis - breaks down for RWF far from the FaIR default of 0.53, or for large anthropogenic CO2 excursions. Lower RWF or greater CO2 excursions would result in a greater implied duration of geoengineering to maintain temperatures below a given threshold.

turbation is high. However, this effect is relatively small for RWF of around 0.5-0.6, similar to the multi-model CMIP5 average of 0.55 and the FaIR default of 0.53.

2.5.3 Sensitivity of the results for the idealised overshoot scenarios to different realised warming fractions

To illustrate the sensitivity of the conclusions to the climate sensitivity, I have tested two sets of parameters. The first, using the default FAIR parameters, represents a TCR and ECS towards the centre of the AR5 likely ranges. The second uses UKESM1 as a representative of the high end of climate sensitivity from the new generation of CMIP6 models. Both of these have a similar ratio to TCR and ECS, or Realised Warming Fraction, which represents the commitment to long-term vs immediate warming when a given tonne of carbon is emitted (0.53 for FAIR, 0.52 for UKESM1), and captures some of the effect of long-term climate-carbon cycle feedbacks and the lagged thermal ocean response to surface warming. Models with a low RWF correspond to cases where these effects are high.

Given that many of these scenarios evolve over centuries, it is reasonable to ask whether sampling TCR and ECS values with a significantly different Realised Warming Fraction affects our conclusions. For example, consider the case where the RWF is low, and the negative emissions floor is also low. Here, we might anticipate longer durations for peak-shaving, as long-term warming due to excess CO2 emissions cancels the effect of removing CO2 from the atmosphere.

To investigate the effect of different RWF on our conclusions, I simulate two additional parameter sets. One is [TCR, ECS] = [1.53, 5.1], corresponding to an RWF = 0.3. The other is [1.53, 1.9]. corresponding to an RWF = 0.8 – both with the same TCR associated with the original FAIR model. I am not suggesting that either of these parameter combinations is especially likely: they simply correspond to extreme cases to illustrate any dependence of our results on the RWF (and, indeed, few models are this far from the mean RWF). [134]

One can see from Figure 2.13 that low RWF models give you a more "exponential" curve, representing a greater dependence of the peak temperature overshoot on the duration of the overshoot itself; this is because, for simulations with much longer durations, there is more time for carbon-cycle and other feedbacks to kick in and generate additional warming beyond that directly due to the transient climate

response to emissions. Under these circumstances, temperature overshoots are generally higher. It is also worth noting that in this low-RWF case, the effect of not explicitly simulating SRM, as discussed in the previous discussion section, becomes more important. The more "s-shaped" curve in Figure 2.14 shows that temperature overshoots and their duration are much more dependent on the emissions trajectories that I simulated rather than properties of the climate system in the case where long-term feedbacks are less important, as we would expect.

Figure 2.13: **Trade-off between years of committment to solar radiation management and the peak in temperature avoided by a range of idealised overshoot and negative emissions scenarios, under the FaIR default TCR of 1.53K, but with .**

However, it is worth noting that to more fully analyse these effects, we would need to use a model with a more robust representation of the carbon cycle, and carbon-climate feedbacks, than the impulse-response model possesses. In particular, we would expect these non-linearities to become more important for greater carbon budget

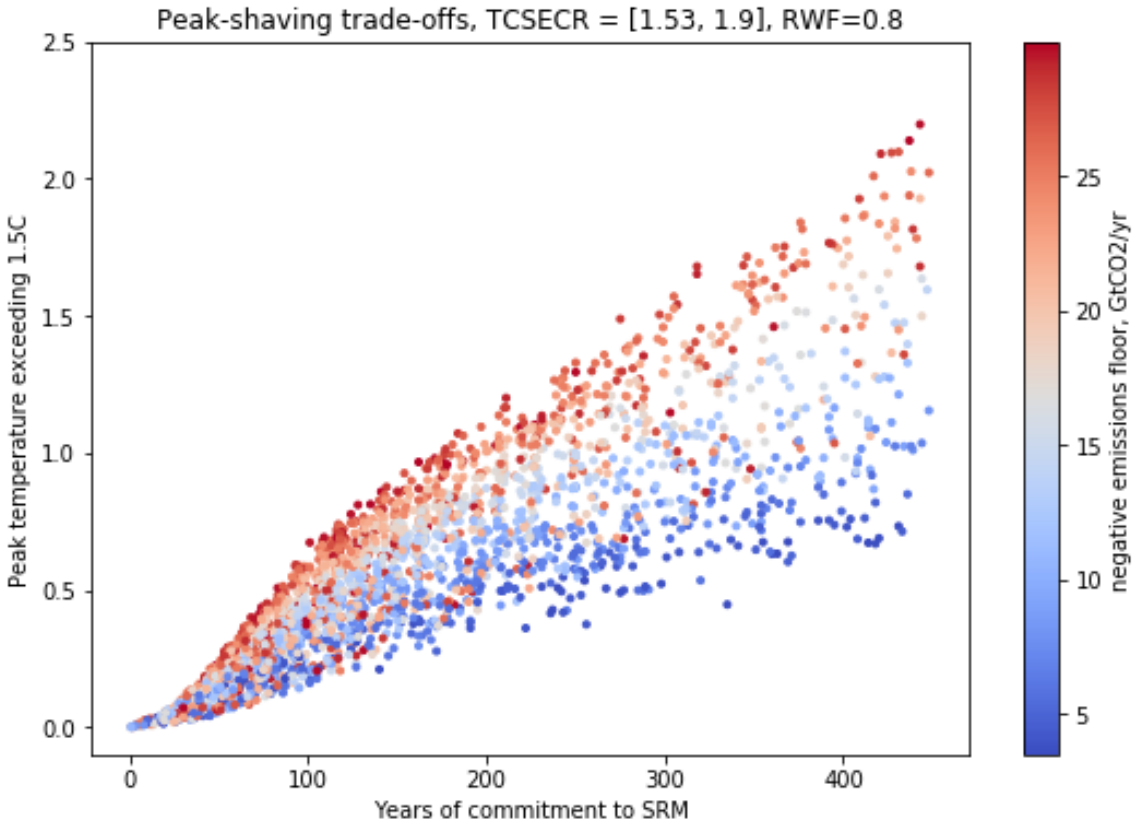

Figure 2.14: **Trade-off between years of committment to solar radiation management and the peak in temperature avoided by a range of idealised overshoot and negative emissions scenarios, under the FaIR default ECS and TCR of 2.86K and 1.53K respectively.**

excursions, and so we might expect them to have a greater impact on our estimates for implied overshoot durations and magnitudes for given emissions trajectories.

2.5.4 Conclusions

This exercise is not intended to span the entire conceivable scenario space of SRM deployment, nor to evaluate the circumstances under which SRM deployment might result in improved outcomes, at least for some. Instead, I aim to illustrate tradeoffs inherent in "peak-shaving" scenarios: between the implied commitment time of SRM deployment and the size of the peak, as well as to outline implications for the mitigation rate and scale of negative emissions that are required to ensure the deployment is limited to a particular timescale.

This analysis is arguably conservative, both for the timescales and negative emissions commitment concerned. It assumes that the carbon budget is exceeded rapidly, with high emissions, and that mitigation begins immediately after emissions peak. When discussing negative emissions, I consider only the scale of global net-negative emissions. However, additional negative emissions will likely be required even to reach net-zero emissions, with integrated assessment modelling often indicating large-scale negative emissions will be required to offset residual emissions from "hard-to-mitigate" sectors [129], so the actual requirement for negative emissions will be higher.

This analysis also allows for the largest mitigation and negative emissions rates found in the ambitious 1.5C-compliant literature. We have considered an extremely large range of mitigation and negative emissions rates in the idealised variants of the SSP-OS scenario, many of which are unlikely to be plausible in reality. In a world where SRM is deployed, assuming it is effective in small doses at reducing climate risks [107], the perceived urgency of reducing emissions and deploying negative emissions at scale may be reduced; this consideration of the political economy of geoengineering and the possibility of reduced urgency may lengthen the timescale of actual deployment for geoengineering beyond the estimates here, which would assume the same level of urgency for mitigation is maintained even if climate impacts are less damaging and geoengineering is effetively reducing climate risks, such as tipping points. We have also assumed, for simplicity, that "peak-shaving" sticks to a certain global mean temperature target and is always committed to keeping temperatures below that threshold. However, political considerations may differ - if geoengineering proves politically difficult or expensive to maintain, it may be the case that would-be geoengineers opt to slowly ramp it down and slow the rate of warming that would otherwise be experienced, particularly if it was only suppressing a relatively negligble amount of warming (say, 0.1-0.2K, or within interdecadal variability.)

Finally, as discussed, the analytic section assumes that the TCRE is the same for both positive and negative emissions, while more detailed carbon-cycle and climate modelling suggests that the TCRE is path-dependent and that negative emissions are less effective at reducing temperature than positive emissions are at increasing temperature, especially when the cumulative emissions are large. Where this is important, it is likely to increase the timescales considered, unless the realised warming

fraction is high. In the discussions section, we showed modelling results in FaIR that suggest this effect is likely to be relatively small for trajectories like the SSP-OS scenario which might be considered more feasible overshoot scenarios.

Unless climate sensitivity is at the high end of estimates, deploying SRM to shave a substantial peak over 1.5C – say, stabilising temperatures at 1.5C when they might have peaked at 2C – is likely to require a commitment to SRM on the order of a century, and to large-scale negative emissions, cumulatively at least 350GtCO_2. We find that, under median estimates of climate sensitivity, deployments of SRM that are complete within a few decades are unlikely to shift global mean temperatures a great deal beyond multidecadal variability. Cost-benefit analyses that assume SRM deployment is not indefinite must consider the cost and feasibility of cumulative negative emissions equivalent to the exceedance of the carbon budget, as well as the timescales required to draw these emissions down. Unless it is maintained on a centennial timescale, using SRM for peakshaving does not free the world of its obligation to rapidly mitigate CO_2 emissions and scale up negative emissions technologies substantially within a few decades.

During the COVID-19 pandemic, to avoid loss of life, many of the world's governments have been compelled to implement severe social distancing measures in order to "flatten the curve" of infection. We see an analogy between these two situations, where the R-rate during the uncontrolled initial phase of the epidemic is analagous to our emissions before net-zero is reached, and the rate of negative emissions we can attain is analagous to the R-rate in lockdown. The curves of new case numbers in countries that implemented lockdowns resembled the asymmetric parabolas of section 2.2, as in general the R during the uncontrolled epidemic phase was larger than the reciprocal of the R during lockdown. Therefore, each week prior to lockdown required several weeks of commitment to lockdown to return case levels to normal. Countries that were able to implement early lockdowns, prior to substantial spread of the disease, were generally able to adopt less stringent, less expensive, and shorter-duration social distancing measures, while experiencing a lower rate of mortality.

In both cases, the inescapable conclusion is that the optimal scenario involves early, swift and decisive action to tackle the root of the problem, reducing emissions,

which keeps options open and can even obviate the need for a long-term commitment to a costly suppression policy. When this becomes impossible, the trade-offs illustrated here must be confronted.

2.6 Appendix

2.6.1 List of scenarios from the IIASA database with overshoots

The full list of 22 IAM scenarios used in the modelling is:

Scenario	Peak Cumulative Emissions
MESSAGE V.4AMPERE3-450P-EU	1246
IMAGE 2.4AMPERE3-550	2132
MERGE EMF27EMF27-450-Conv	874
REMIND 1.5AMPERE3-450P-EU	1014
REMIND 1.5AMPERE3-CF450P-EU	1047
IMAGE 2.4AMPERE3-450	1017
REMIND 1.5AMPERE3-450P-CE	963
REMIND 1.5EMF27-450-Conv	1364
GCAM 3.0EMF27-450-Conv	771
REMIND 1.5AMPERE3-450	832
MESSAGE V.4AMPERE3-550	1678
REMIND 1.5AMPERE3-550	2026
GCAM 3.0AMPERE3-550	665
GCAM 3.0EMF27-450-NoCCS	718
REMIND 1.5AMPERE3-550P-EU	2250
GCAM 3.0AMPERE3-CF450P-EU	532
MESSAGE V.4AMPERE3-450	794
GCAM 3.0AMPERE3-450P-CE	507
GCAM 3.0AMPERE3-450P-EU	513
IMAGE 2.4AMPERE3-450P-EU	1336
IMAGE 2.4AMPERE3-450P-CE	1170
GCAM 3.0AMPERE3-450	508

Table 2.2: **Peak cumulative emissions for each of the 1.5C overshoot/peak-and-decline scenarios in the IIASA database, plotted in Figure 2.4.** With a given TCRE and hence carbon budget, this can be used to approximate temperature overshoots.

Chapter 3

Methods, calibration, and initial characterisation for SRM experiments in HadCM3

3.1 Introduction

The principal aim of this part of the project (in Chapters 3-5) is to explore the question: to what extent can geoengineering be optimised? We have discussed in the literature review in Chapter 1 the shortcomings of uniform geoengineering, such as an undercooling of the poles and overcooling of the tropics. We have also discussed the modelling and volcanic analogue evidence which suggests that it is possible to achieve different temporal and latitudinal distributions of stratospheric aerosol through different injection strategies, resulting in a different climate response. In this and the following chapters, we seek to use climate models to characterise the response to different patterns of applied aerosol optical depth, in order to determine the extent to which geoengineering could be optimised, in the idealised case that we have control over the distribution of aerosol optical depth. We can also use the large number of model years that this process will generate to investigate climate and weather extremes. There are several steps to this process:

- Explore and characterise the broad climate impacts of latitudinally varying SRM using large HadCM3 ensembles. (Chapter 3)
- Investigate the linearity of the climate response with respect to the imposed aerosol patterns in radiative fields, temperature, precipitation, etc. (Chapter 3)
- Use this framework to construct patterns of aerosol optical depth and predicted response that would optimise for attaining particular climate goals, and analyse their effects. (Chapter 4)
- Investigate whether SRM reduces the frequency of extreme weather events compared to preindustrial baseline. (Chapter 5).

The parameter space of potential SRM interventions is extremely large - even with the large ensembles available to climateprediction.net (CPDN), we do not propose to explore the entire parameter space. Instead, we seek to determine whether the variability in the climate response to inhomogenous forcing can be projected on a few main modes of variability. These would correspond to the internal modes of variability of the climate system. This would allow us to broadly illustrate which trade-offs exist and the limits of the potential to optimise the climate with SRM, even if you assume it is possible to obtain a high degree of control of the latitudinal distribution of aerosol. We return to this idea in Chapter 4, but initially our task is to implement SRM in our model and characterise the response.

3.2 Methods

3.2.1 Climateprediction.net and HadCM3

The climateprediction.net project [135] [136], along with its sister project weather@home [137], are distributed computing projects that utilise the world's largest ensemble of climate models, run on volunteer computers through the Berkley Open Infrastructure for Network Computing (BOINC) system. The CPDN project uses the HadCM3 global circulation model (GCM) [138] and was originally designed for quantifying uncertainties in the climate sensitivity when physics is perturbed [139], while

weather@home uses a regional climate model embedded in HadCM3 – which can be driven with sea-surface temperature patterns – to carry out attribution studies [140] for extreme weather events.

CPDN's capacity to generate PPEs (Perturbed Physics Ensembles), which can sample a wide range of responses, feedbacks, and forcings [141], has been used in a number of studies to constrain climate sensitivity [142] and represent uncertainties in the climate system as a multi-model ensemble would. [143] When run in this mode, climateprediction.net allows you to vary different parameters of model's representation of the atmosphere and ocean in the climate system, which parameterize more complicated physical processes. For example, some of these parameters correspond to the sea surface roughness, the diffusion coefficients in the ocean or atmosphere, or the threshold liquid water content for clouds. There are a range of different combinations of parameters in the model which can yield plausible climates; CPDN studies in the past have investigated the impact of varying these parameters on the model's ability to simulate climate and the climate sensitivity.

For our purposes, we will run HadCM3 with the standard physics parameters, but instead drive the model with a range of different patterns of aerosol optical depth, mostly superimposed on an abrupt quadrupling of CO_2 in order to ensure a good signal-to-noise ratio and because this idealised experiment has been widely studied. Ensembles of model runs driven by the same forcing pattern can be generated with initial conditions perturbations, discussed below. The large ensembles of HadCM3 models that we can run enable us to distinguish the forced response to latitudinally dependent forcings from natural climate variability, and distinguish between the response to forcings in general and the specific dependence on the variations of the forcings with latitude and time in our model.

The latest version of HadCM3 used by CPDN builds on the published version. It allows the stratospheric aerosol optical depth (AOD) to be specified in each of 73 latitude bands, which allows inhomogeneous SRM to be implemented. (To first approximation, stratospheric AOD is a function of latitude, as described in Section 1.2). Note that as the AOD can be varied in each month of the annual cycle, and the polar latitudinal bands are not read in by the radiative scheme, this effectively gives us 12 x 71 degrees of freedom to vary the AOD in each year. The model has been

modified to include multiple calls to the radiation code, allowing the user to diagnose radiative forcings with and without aerosols and greenhouse gases and output these as diagnostic variables, which allows us to carry out calibration as discussed later.

Standard physics parametrizations and initial conditions perturbations can provide a large number of model years for analysis. Initial conditions perturbations are implemented by adding a random potential temperature perturbation between 0.1K and 1K to one of the grid cells in the "start file" specifying the initial condition of the model; this is sufficient to lead to divergent model evolution of temperature and precipitation across all grid cells within a few model weeks. The "start file" that is perturbed to generate ensembles is always the same preindustrial start file, so initial conditions differ only by this perturbation, except where runs were explicitly continuations of earlier runs to generate more model years (which also all ultimately originated from the same perturbed start file.)

3.2.2 Ensembles from climateprediction.net

Climateprediction.net relies on using volunteer computers, which have downloaded a version of HadCM3. Simulations are released to volunteers in "batches", which specify the ancillary files to be read into the model to drive it, such as radiative forcing from greenhouse gases and the volcanic forcing files which are used to implement our latitudinally-varying stratospheric aerosol optical depth. These ancillary files are specified by the experimenter and uploaded to a server, where they can then be downloaded by volunteers and used to run individual experiments. For our purposes, greenhouse gases, specified as concentrations, and specifically CO_2, as well as the aerosol optical depths prescribed, were the only ancillary files we modified - the rest were kept at preindustrial levels. The batches provide a unique ID for each run, and also specify how long the simulation should be run for. In this, there are many trade-offs for the experimenter to consider. Volunteer computers are often slow, and have varying speed depending on the hardware that the volunteers have. Volunteers can specify how much computing power CPDN will utilise, and the hours of operation when it will run. This, as well as their internet speed for uploads and downloads, influences how long simulations take to run. The longer a simulation is run for, the

greater the risk that the results will not be returned, due to hardware or software errors from the volunteer, or failure to upload large data files. There were also several other research projects which relied on CPDN when I was doing model runs, so the number of volunteer computers I was able to use was sometimes limited by other experimenters.

In practice, we found that simulations of 20 years offered a reasonable compromise between obtaining a sufficient amount of data and minimising the number of model runs that crashed or did not return sufficient data. In the analysis that follows, I typically treated the first ten years as spin-up years, and analysed data only from the second decade, when the simulations had stabilised. Sometimes, simulations of ten years' duration were run, but the start file they used had already been spun up for ten years, allowing comparable data from shorter simulations. (This standard spin-up period is used for all simulations except where I am explicitly analysing the dynamics of that first decade; for example, in showing how global mean temperatures would increase under 4xCO_2 or regressing precipitation in grid cells onto global mean temperature over time to characterise the "fast" and "slow" response in Chapter 4.) Even so, it was generally necessary to filter out model runs with incomplete data due to upload errors, or that had become unstable due to calculation errors on user equipment. Typically, this would manifest itself as missing data in the output or a run that "failed" and did not return any data; these runs were filtered out before analysis.

The result of this is that the ensembles used are all of varying sizes, depending on how many results were returned by CPDN. Typically, although not exclusively, I would send out batches of 50 ensemble runs and receive 20-30 back within a 3-month timeframe, which was the cutoff point at which I started to analyse the data from that set of model runs. (More runs, from slower user machines, would ultimately be downloaded and some of these were used in subsequent chapters.) Consequently, the ensemble averages of simulations that we analyse in Chapters 3 and 4, and the number of model years of preindustrial simulations and SRM that are analysed in Chapter 5, are of slightly different sizes. I will list the numbers of each simulation analysed in each chapter in a dedicated table. With the exception of the calibration

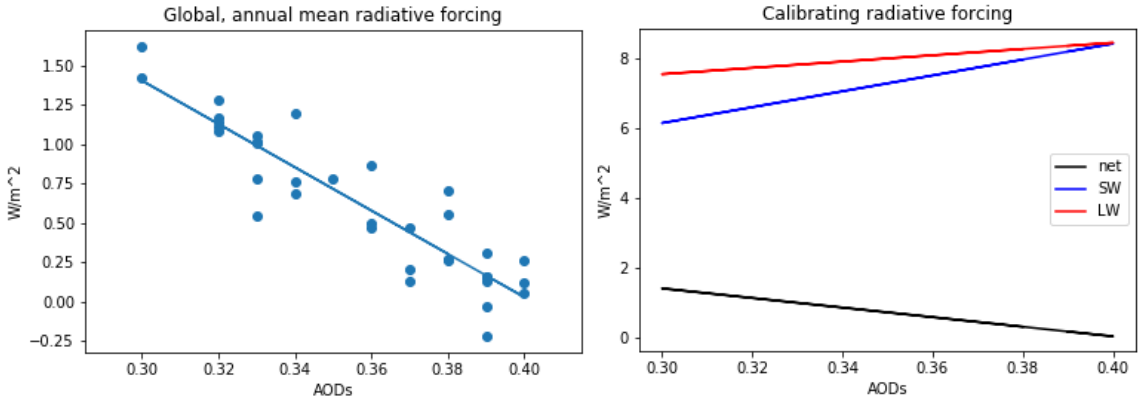

Figure 3.1: **L: Determining the AOD for net zero radiative forcing under 4xCO2; R: global mean shortwave and longwave components of radiative forcing under different aerosol optical depths.** L: Calibration of the model's response to CO_2 and SRM R: Breakdown of longwave and shortwave forcing components and their dependence on AOD in 4xCO2 + SRM simulations.

section below, and the temperatures plotted in Figure 3.6, the rest of the analysis in this chapter is from the ensembles set out in Section 3.2.4.

CPDN implements code which converts model output into netCDF format. Analysis and averaging is then undertaken using a combination of the CDO command line tools and Python, and plots are either in matplotlib or Panoply.

In the next section, we will discuss initial simulations of geoengineering and the necessary calibration of geoengineering in this model. We will then move on to discuss the simulations that are analysed in the results section of this chapter.

3.2.3 Calibration of SRM in HadCM3

Initial calibration of the model and its response to statospheric aerosols and greenhouse gases required running idealised simulations with a large signal-to-noise ratio, aiming to compensate for an abrupt quadrupling of CO_2 levels with a uniform layer of stratospheric AOD. I chose the AOD to attempt to balance the IPCC's standard for radiative forcing from CO_2:

$$RF = \beta \ln\left(\frac{CO_2}{CO_{2,ref}}\right) \tag{3.1}$$

where $\beta = 5.35 Wm^{-2}$ from Myhre et al. (1998) [144] leading to $7.41 Wm^{-2}$ as the radiative forcing from quadrupling CO2 to cancel. Hansen et al. (2005) [145] calculated that, for the stratospheric aerosol lofted by Pinatubo, which the aerosol implementation in HadCM3 attempts to simulate, the shortwave radiative forcing can be approximated by $RF = -24 \cdot \tau$, where τ is AOD.

This gives an estimated aerosol optical depth of $\tau = 0.31$. However, this ansatz for the required AOD to suppress warming from quadrupled CO_2 is too small: when simulations were run, there was 1K of residual warming across the first decade. This is because it does not account for either the longwave radiative absorption of the stratospheric aerosols, which acts to warm the stratosphere, or the differing effects on surface temperature of reflecting incoming shortwave radiation (SRM) and decreasing outgoing longwave radiation (GHGs). These effects mean that the initial guess is an underestimate for the AOD required to prevent surface warming. This has been observed in previous multimodel simulations of geoengineering, where each individual model requires calibration (Figure 3.1) to determine the AOD required to cancel out the increase in global mean surface temperature, depending on its response to the different forcings applied. [58] Note in Figure 3.1 the weak dependence of LW forcing on AOD showing longwave absorption by aerosols, which is why calibration is needed.

A small ensemble was run, with five 'runs' (each with different initial conditions perturbations), of 11 different aerosol optical depths evenly spaced at intervals of 0.01 between $\tau = 0.30$ and $\tau = 0.40$. This was sufficient to establish approximate linearity in the global mean radiative forcing with AOD in this range. Calibration gave us an AOD of $\tau = 0.39$ which is henceforth used as a scaling factor for aerosol distributions imposed on quadrupled CO_2. (See Figure 3.1.) In the uniform case with this AOD, the residual radiative forcing was $0.19 Wm^{-2}$, and the uniform AOD stabilised temperatures at around 0.3K above the preindustrial control for a decade. (Figure 3.5)

Gregory [146] plots provide a faster means of estimating the equilibrium climate sensitivity (ECS) and radiative forcing due to an imposed change. Rather than waiting for the model to come to equilibrium, which can take centuries of model time, the approximate linear relationship between changes in net radiative forcing and temperature is used. A Gregory plot from an ensemble average of 4xCO_2 experiments

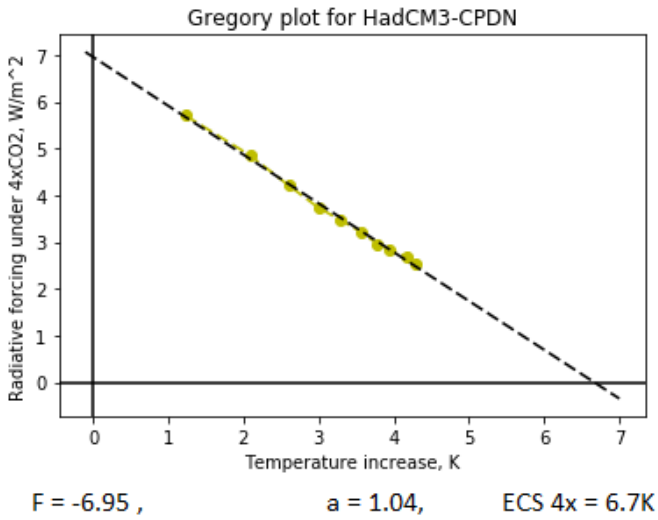

Figure 3.2: **Net radiative forcing plotted against temperature anomaly in the abrupt4xCO2 scenario** - so-called Gregory plot determining the model's climate sensitivity to abrupt 4xCO_2

suggested that the effective radiative forcing of 4xCO_2 in CPDN's HadCM3 was close to $7Wm^{-2}$, (note the difference from Myhre et al. (1998) [144]) with an eventual expected warming of 6.7K. (Figure 3.2)

3.2.4 Simulations and ensemble sizes

3.2.4.1 Characterising the preindustrial, 4xCO2, and SRM scenarios

The baseline set of our scenarios is the preindustrial case, with no external forcings applied at all. The greenhouse gas concentrations are set to their default preindustrial levels, and the volcanic forcing is set to zero. These simulations are referred to as piCO_2 or the preindustrial baseline simulations. The second set analysed here are the abrupt4xCO_2 experiments, where the greenhouse gas concentrations are abruptly quadrupled from their preindustrial levels, but no volcanic forcing is applied. The third main set of simulations analysed are the "SRM" simulations. These simulations abruptly quadruple CO_2 and they also apply a uniform layer of aerosol optical depth, scaled by the scaling factor, $\tau = 0.39$, derived in the calibration section, which - when applied as a global average uniform aerosol optical depth - stabilises global mean temperature at the preindustrial levels. This is the baseline SRM scenario and

scaling factor which others build on. When we describe different distributions - such as the NH or SH SRM scenarios, or the multimodel G4 scenarios - these patterns of aerosol optical depth have been normalised to have the same global mean aerosol optical depth as the uniform SRM simulation, i.e. that which would cancel out the temperature increase from abrupt 4xCO_2 when applied as a uniform layer.

As a first test of inhomogenous radiative forcing, simulations were conducted where the aerosol optical depth was doubled in one hemisphere (relative to the uniform SRM in earlier tests) and set to zero in the other. These simulations are referred to as NH SRM and SH SRM. This is an idealised representation of the effects of stratospheric aerosol injection concentrated in either hemisphere, and has an analogue in historic volcanic eruptions which resulted in AOD predominantly in one hemisphere, such as El Chichon, as discussed in Chapter 1. HadCM3 has 71 effective latitude bands in which we can set an aerosol optical depth, given that the polar latitude bands are not read into the radiative model. This is an odd number, so the central one of these bands spans the equator. For the purposes of these hemispheric simulations, the aerosol optical depth in this band was set to zero in both cases. This means that the NH and SH simulations do not exactly correspond to AOD across the entire Northern or Southern hemisphere, but this was as close as could be achieved with this limitation of our model.

3.2.4.2 Characterising the distributions from the G4 multimodel experiment

To what extent is the climate sensitive to different latitudinal distributions of aerosol? Models with different aerosol microphysics and stratospheric dynamics result in different distributions of aerosol optical depth for the same injection strategy, as described in Chapter 1's literature review and illustrated in Figure 3.4.

As the presence of aerosol in the stratosphere affects the stratospheric temperature, chemistry, and dynamics, even field-trials that fall short of full deployment are unlikely to eliminate these uncertainties. Observational data from volcanic eruptions is unlikely to cover the full range of possible injection scenarios and strategies. For this reason, understanding the effects of variations in the distribution of aerosol -

and the sensitivity of any geoengineering strategy to unintended variations - is important. The GeoMIP G4 experiment [15] aimed to investigate the effects of using the same injection strategy - daily equatorial injections amounting to 5Tg SO2 per year in the stratosphere - in different models. Differences in response between the models [52] arise due to the different distributions of aerosol, which in turn arise from different model physics - for example, in terms of stratospheric dynamics and aerosol microphysics.

To isolate and investigate the impacts of differing aerosol distributions alone, I scaled up aerosol distributions from models that participated in the GeoMIP G4 experiment (for those cases data could be obtained) to the global mean AOD required to cancel $4xCO_2$ in HadCM3. While this is not the same as simulating injections of the required magnitude in each of the stratospheric models used, it benefits from the high signal-to-noise ratio and ability to compare with existing simulations.

The data used were from models HadGEM2, ULAQ, MIROC-CHEM-AMP, and GEOCSSM, as well as the G4 experiment's prescribed stratospheric AOD dataset for models that cannot explicitly simulate aerosol injections, which was based on observational data from the Pinatubo eruption. The annual cycle for AOD in the modelled cases was calulcated by averaging over thirty years in the simulation, starting ten years after injections began. Data was regridded and zonal averages taken to provide inputs for HadCM3; these are plotted in Figure 3.4.

There is a notable bimodality: GEOSSCM, ULAQ and MIROC experience meridional transport barriers that result in tropical AOD confinement: this was not seen in the Pinatubo eruption, the HadGEM2 simulation, or the CESM-WACCM model used by GLENS (see Chapter 4). This confinement owes to greater aerosol coagulation in these models, which leads to a faster sedimentation rate for aerosol at high latitudes, and a strong isolation of aerosols in the so-called 'tropical pipe' [42]. The level of confinement differs: 24% of the total AOD is between 20N and 20S in the MIROC simulation compared to 42% in ULAQ. This is characterised more precisely by examining the kurtosis of the annual average meridional distributions (Figure 3.3): positive kurtosis corresponds to strong confinement, while negative kurtosis corresponds to distributions concentrated at high latitudes.

Seasonal dependence of meridional transport depends on the model: for example, in ULAQ, MIROC and GEOSSCM aerosol transport to higher latitudes is most effective in summer.

Finally, we model an idealised distribution defined such that:

$$AOD_\beta \propto (1 + \sin^2(\varphi))$$

where φ is latitude.

This distribution was normalised to the same AOD average scaling as the other distributions: it represents an idealised distribution where AOD is concentrated at the poles, and therefore contrasts with models with high tropical confinement. Constrasting responses like this have been observed in some models: while AOD data from the GISS-E2-R model could not be obtained for this study, its simulation of G4 provides an example of a stratospheric model where simulations of equatorial injections, with more vigorous horizontal mixing of air in the lower stratosphere, results in a minimum in the tropics and higher concentrations at high latitudes. [42]

3.2.4.3 Ensemble sizes in this chapter

The table below lists the ensemble sizes for completed simulations that are analysed in this chapter, with the exception of the "stripe" simulations.

To investigate linearity a series of geoengineering simulations where the aerosol was confined to ten adjacent model latitude bands (22.5 degrees of latitude). In each stripe, the AOD was set to 0.39, and zero elsewhere, against a background of 4xCO_2. The AOD was fixed at 0.39 for three latitude bands at the poles (two in the NH and one in the SH) to allow for seven "stripe" simulations to cover the entire globe.

The "stripe" simulations were run in smaller batches, of 20 each, and therefore the variability in how many were analysed for the ensembles were smaller. Moving down across the latitude bands from the Northern hemisphere, the latitude bands covered are 2-12, 12-22, 22-32, etc. Denoting these as st1, st2, st3..., the batch sizes were as follows. st1: 10 st2: 5 st3: 7 st4: 5 st5: 8 st6: 7 st7:5 .

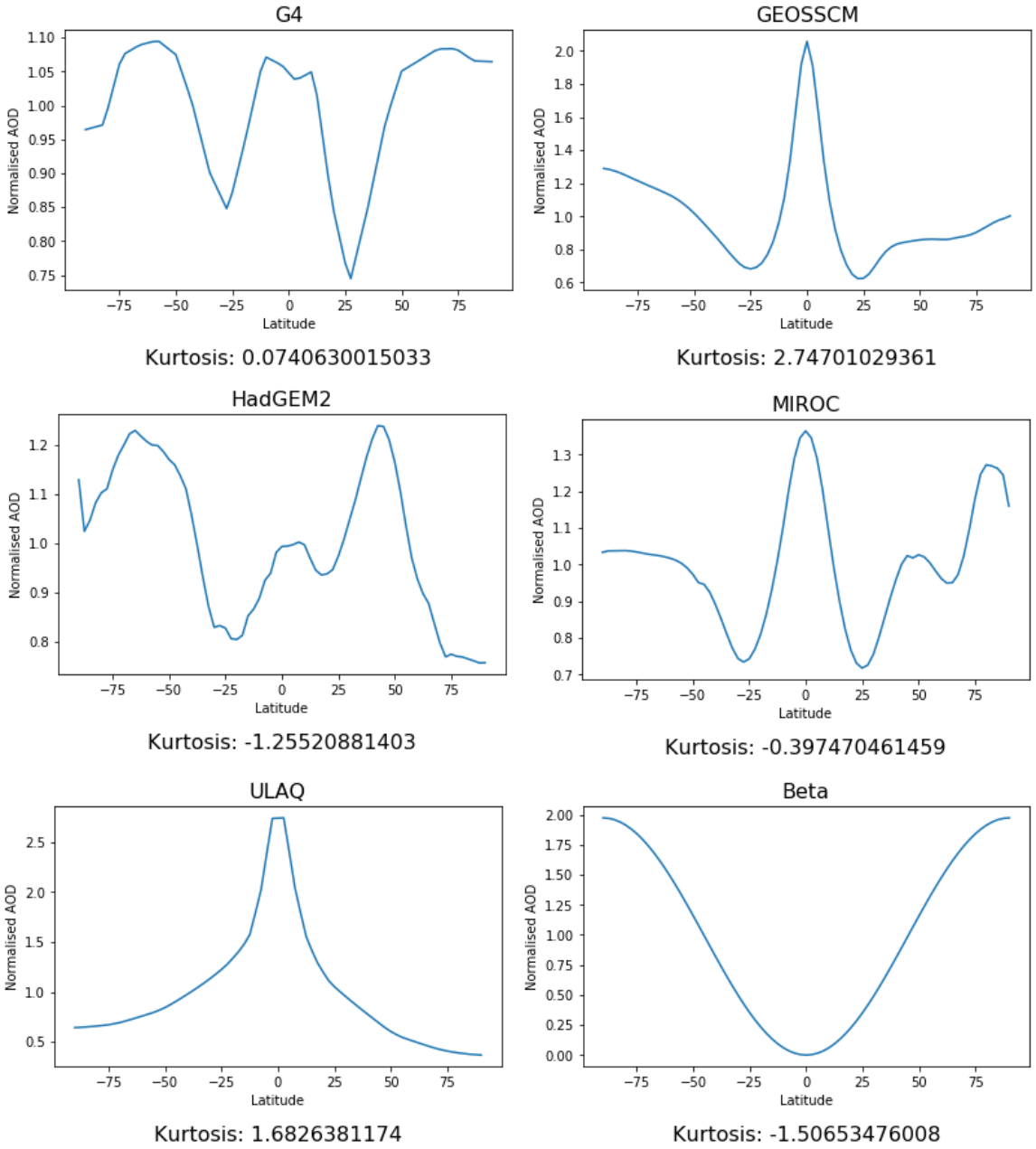

Figure 3.3: **Annual, zonal averages of AOD from the multimodel comparison of simulated aerosol distributions which result from the G4 equatorial injections.**

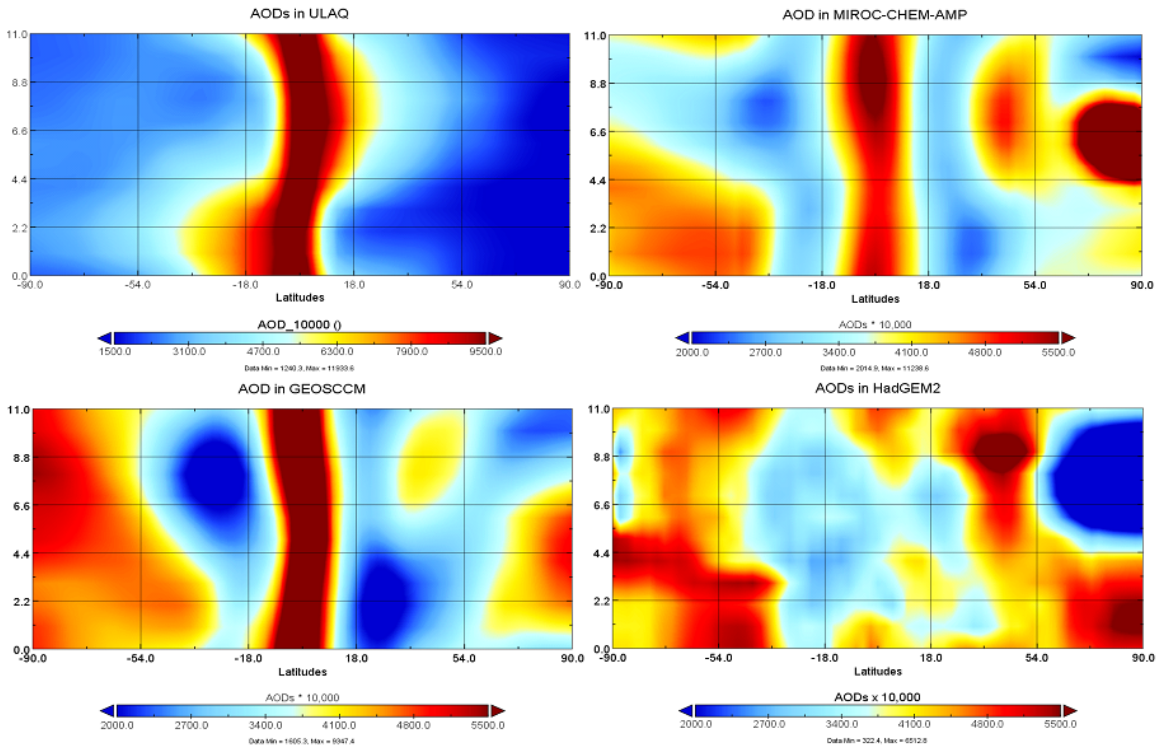

Figure 3.4: **AOD-time plots for the annual cycle from the multimodel comparison of G4 equatorial injections.** Clockwise from top left: ULAQ, MIROC, HadGEM2, GEOSCCM.

Simulation	**Ensemble size**
Preindustrial	27
4xCO_2	21
Uniform SRM	22
NH SRM	24
SH SRM	23
GEOSCCM	26
HadGEM	21
MIROC	23
ULAQ	14
Pinatubo (G4-SSA)	21
Beta	33

Table 3.1: Simulations and ensemble sizes considered in the first part of this chapter

3.3 Results

3.3.1 Characterising the response to uniform SRM and CO2 experiments

3.3.1.1 Temperature anomalies

Analysis of the abrupt4xCO_2 experiment in HadCM3 indicates that temperature increases approximately as the square root of time over the first decade, with a slower rate of warming subsequently. We observe this first decade of change in Figure 3.5. This corresponds to the two main timescales to the response - the fast one associated with the atmospheric response, which is effectively complete after a decade, and the slow one associated with the response of the oceans, which take centuries to fully equilibriate with abruptly imposed forcings. A spin-up period of a decade allows the atmospheric response to take place, while the ocean continues to warm slowly. [147] For SRM designed to cancel out global mean temperature the global mean temperature anomaly and net radiative forcing are both low by design, so there is not substantial net warming of the ocean later on, allowing us to use a shorter spin-up. (For example, GeoMIP used 10 years for G1).

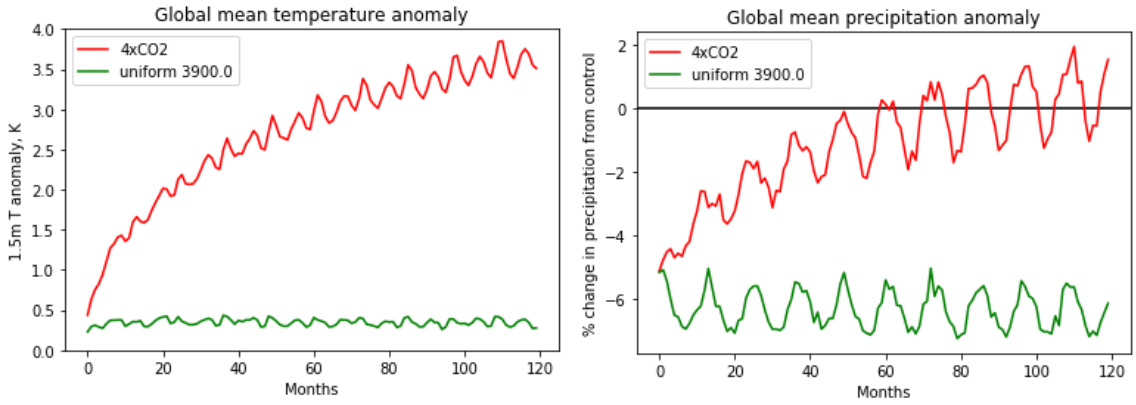

Figure 3.5: **Temperature and precipitation anomalies under SRM and abrupt 4xCO_2**

Temperature anomalies (Figure 3.6) persist where aerosol reflection is least effective: in the polar winters where there is insufficient incoming shortwave radiation for

aerosols to reflect to compensate for the CO_2 forcing. There are also notable positive anomalies over high Northern latitudes during their summer. These are areas prone to albedo feedbacks and Arctic amplification of temperature anomalies under greenhouse-gas induced warming. [148] For this reason, we would expect larger temperature anomalies in the Arctic even without the additional influence of the imperfect cancellation of radiative forcing by SRM.

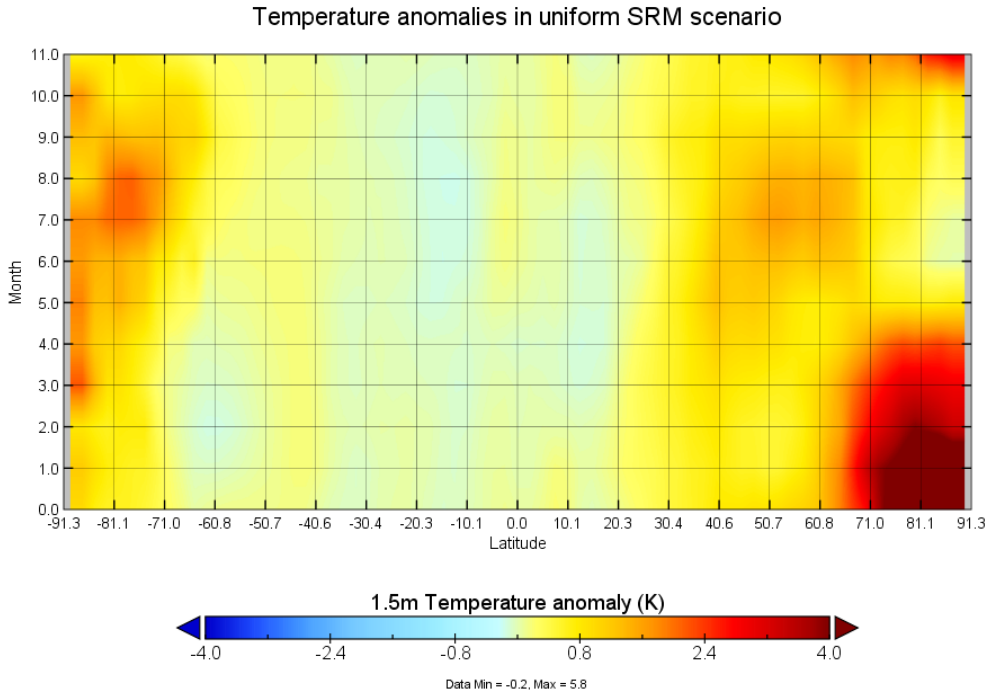

Figure 3.6: **Latitude-time plot of temperature anomalies and residual warming under uniform SRM.**

The large land fraction of these latitude bands likely also contributes to the larger temperature anomalies. The average residual warming over land was 0.5K, compared to 0.06K over sea: an enhanced land-sea contrast in residual warming, with a land-sea warming ratio of 8.33, compared to a land-sea warming ratio of 1.64 under the abrupt4xCO_2 scenario.

Meanwhile in the tropical summers where reflecting shortwave radiation is most effective, there is a slight overcooling. The effect of this uniform aerosol distribution is therefore to undercool the poles and overcool the tropics, in accordance with previous work. [54]

3.3.1.2 Precipitation anomalies

For precipitation, the global picture is similar to that noted in the G1 experiment. Abruptly changing the shortwave flux to the surface with SRM results in an initial reduction in evaporation and precipitation. (Figure 3.5) In 4xCO_2, the slow feedback from increasing temperature eventually drives global mean precipitation back above preindustrial levels, although there are regional differences, but in SRM, the temperature change and feedback are suppressed, and global mean precipitation decreases by 6.4%.

This effect was noted in the GeoMIP G1 experiment, where HadCM3 simulated a 4.6% decrease in precipitation (compared to the multimodel mean of 4.5%) in response to geoengineering by solar constant reduction. [57] As discussed, a more realistic implementation of geoengineering with stratospheric aerosols results in a greater precipitation reduction than solar constant reduction, in part because the shortwave flux to earth must be reduced by a greater amount to compensate for longwave heating from the aerosols.

We plot the structure of the decadal average changes to precipitiation, in percentage terms, in Figures 3.7 and 3.8 for the 4xCO_2 and SRM experiments respectively. Note that the 4xCO_2 rainfall is taken from years 10-20, so that temperature feedback has driven global mean precipitation back above control levels; SRM has a suppressed slow response, so global mean precipitation is reduced. The abrupt quadrupling of CO_2 simulation show wetting over much of the world, particularly over the oceans and in the tropics, and at high latitudes, but with drying over many land regions such as the Mediterranean, North Africa, Central America, and the Amazon. There is significant relative wetting over the Middle East although in absolute terms this is not as significant.

Many regions most affected by SRM are also sensitive to an abrupt quadrupling of CO_2 (as seen in Figures 3.7, 3.8 and 3.10). For example, both SRM and 4xCO_2 result in drying over the North of South America, Southern Europe and North Africa, while SRM experiences additional drying at high latitudes and over the Pacific compared to 4xCO_2.

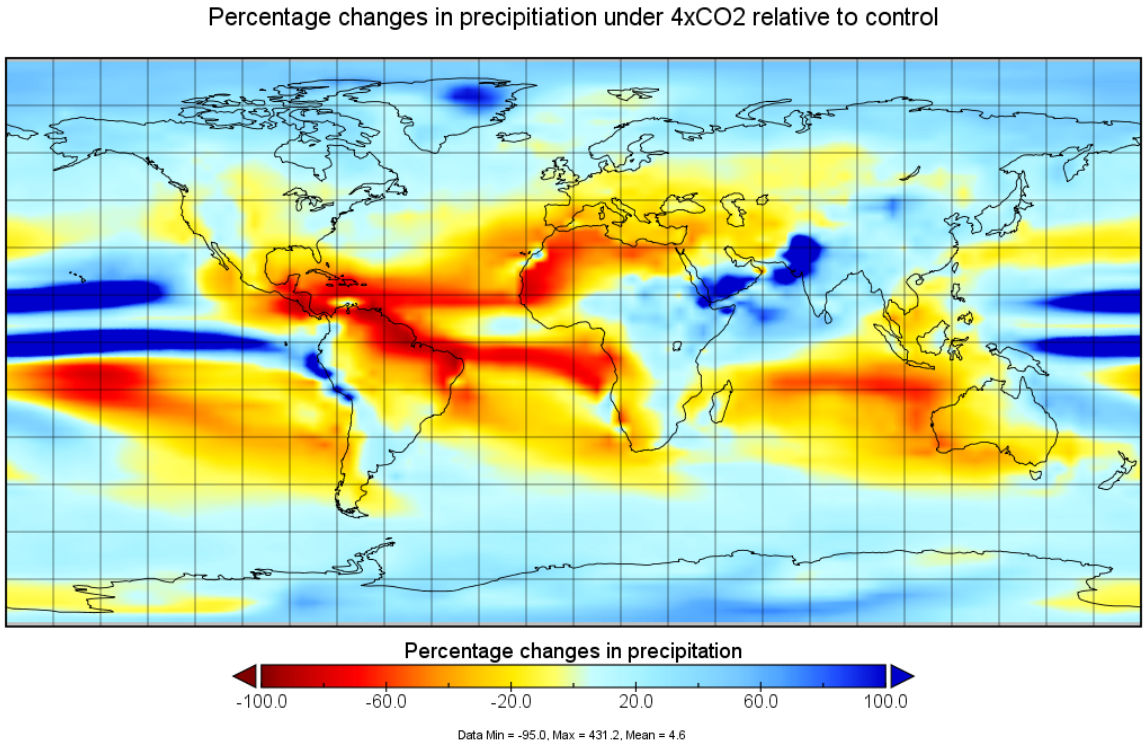

Figure 3.7: **Precipitation changes, in the decadal average, in percentages, under 4xCO2 relative to the control experiment.**

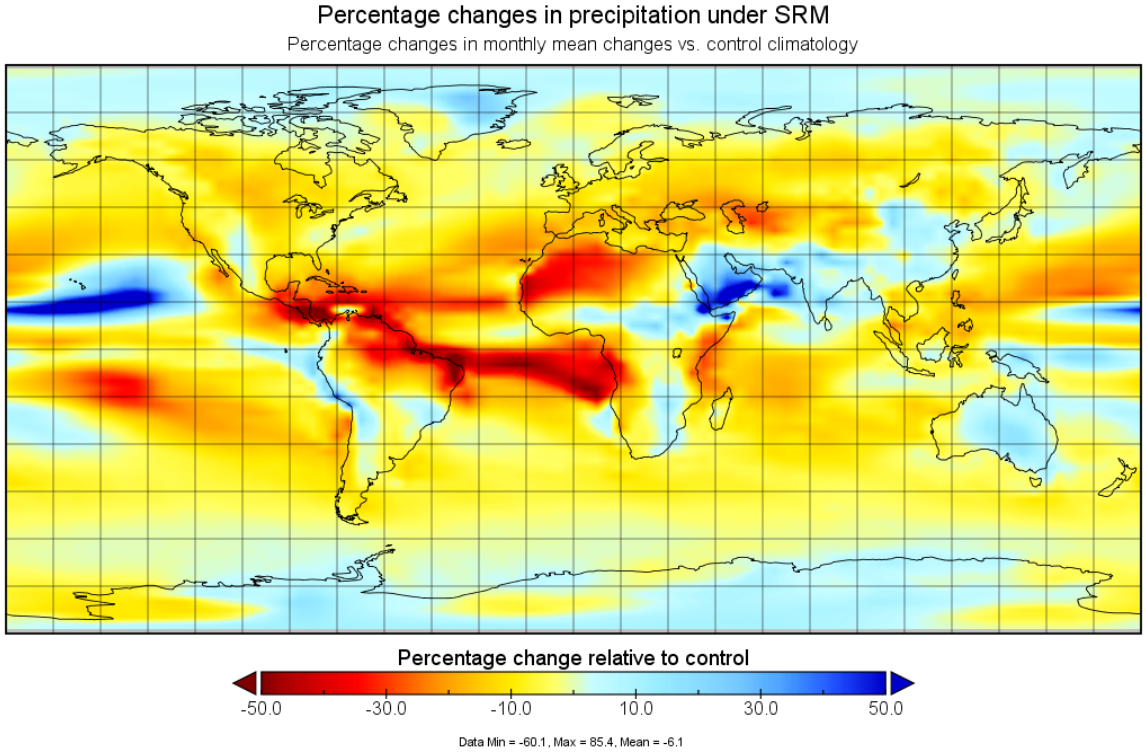

Figure 3.8: **Precipitation changes, in the decadal average, in percentages, under SRM relative to the control experiment.**

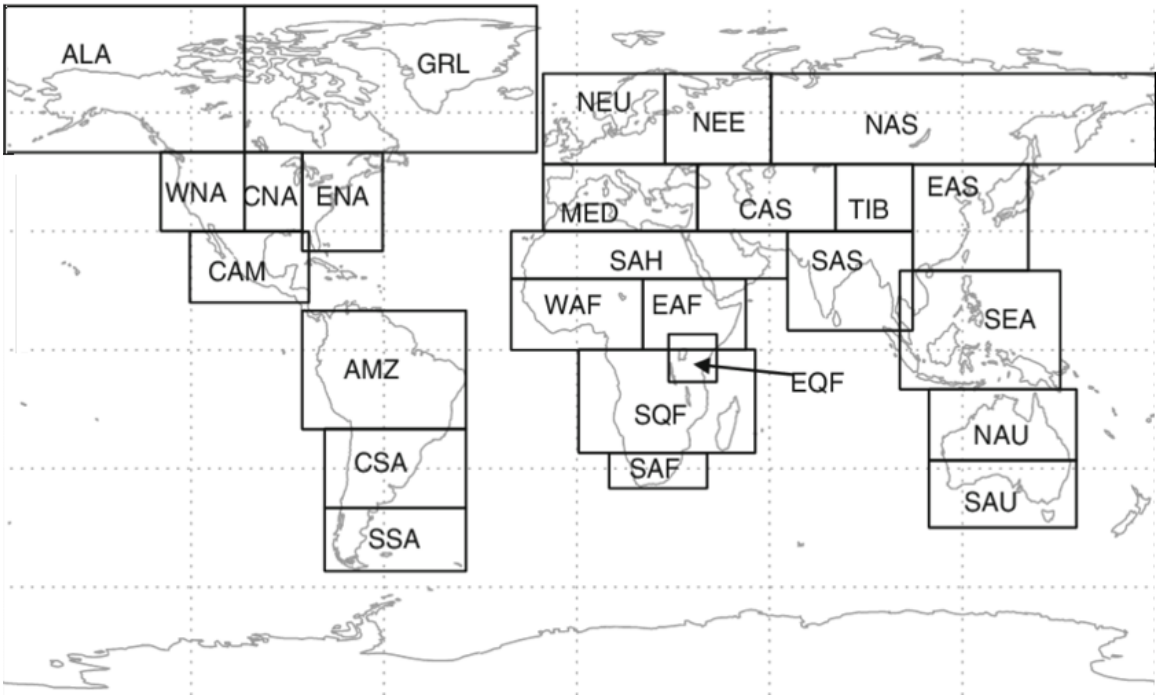

Figure 3.9: **Giorgi Regions** (after Giorgi and Bi 2005, common regional subdivisions for climate modelling.)

The precipitation shifts under SRM relative to the control are complex and highly variable but include a general drying, especially over Europe, North and Southern Africa, North America, and a drying over Brazil, along with a relative increase over the Middle East, parts of India and Australia, and a slight increase over China - alongside Northward shifts and general drying over the oceans. Note that more regions see drying under SRM, especially at higher latitudes, and the percentage change is illustrated on a half-scale compared to $4xCO_2$, showing that the changes are generally more moderate than in the global warming simulation.

To analyse changes to regional rainfall in more detail, we calculate the annual average precipitation across land points in each Giorgi region (Figure 3.9): percentage changes relative to the preindustrial are plotted below.

The regional picture over land is complex and varied (Figure 3.10). Australia, the Amazon, Central America, South and West Africa, the Sahel, and the Mediterranean (8/22 regions) experience drying trends in $4xCO_2$ which are ameliorated by SRM. Large increases in precipitation in snowy regions such as Tibet, Siberia, Alaska and Greenland, which have more moisture available under $4xCO_2$ as they heat up, are

reversed in SRM. North America, North Europe, the rest of South America, and South East Asia are drier than the control under SRM and wetter than the control under 4xCO_2. For India, China and East Africa, precipitation increases under both scenarios, but substantially less under SRM. While some regions may benefit from changes to precipitation, and not all regions are equally sensitive to changes, previous studies [80] define geoengineering's effectiveness as its ability to return regions to preindustrial conditions. By this rough metric, for annual averages, 17/22 Giorgi regions are closer to preindustrial conditions under SRM than 4xCO_2.

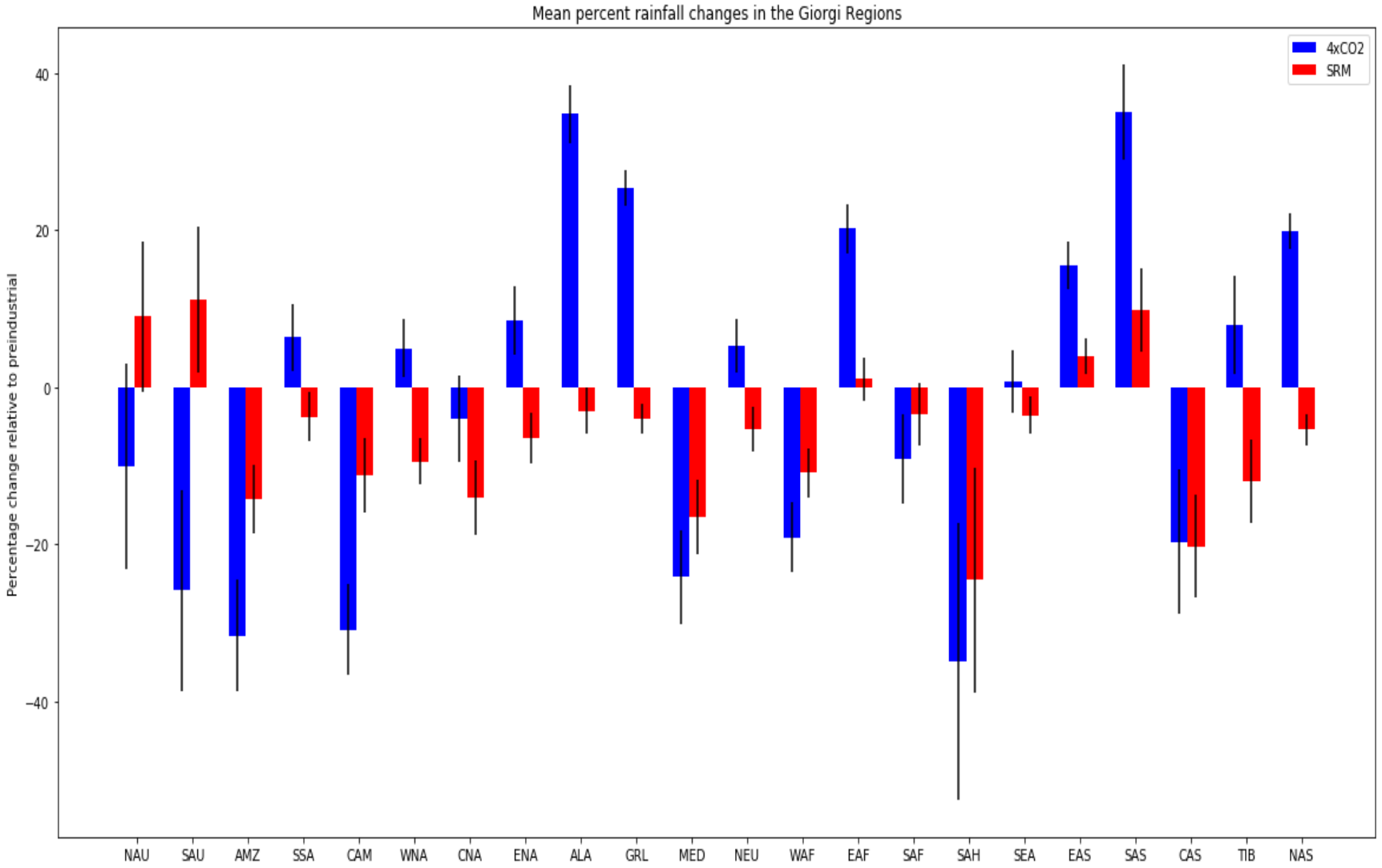

Figure 3.10: **Giorgi regional analysis of precipitation shifts under SRM**. Plotted are percentage changes to the preindustrial precipitation in the annual mean for 4xCO2 (blue) and SRM (red) cases for each of the Giorgi region. Error bars show a standard deviation in the annual precipitation under the 4xCO2/SRM scenario; the baseline for calculating percentages is the average across all preindustrial model years.

P - E and soil moisture show much smaller net changes under SRM than precipitation as reducing shortwave radiative flux to the surface reduces both evaporation and precipitation. P - E changes are moderated by SRM relative to $4xCO_2$ for all but 4 Giorgi regions. (Figure 3.11) If P - E and soil moisture are the most appropriate metrics for impact studies, and the most relevant for water availability, they may indicate less propensity for drought in a given region under SRM, and SRM may return them closer to preindustrial levels than unmitigated CO_2 warming - although regional inequalities persist. This is part of a general result in SRM modelling which suggests geoengineered climates may have less rainfall, but still be wetter and greener. [149]

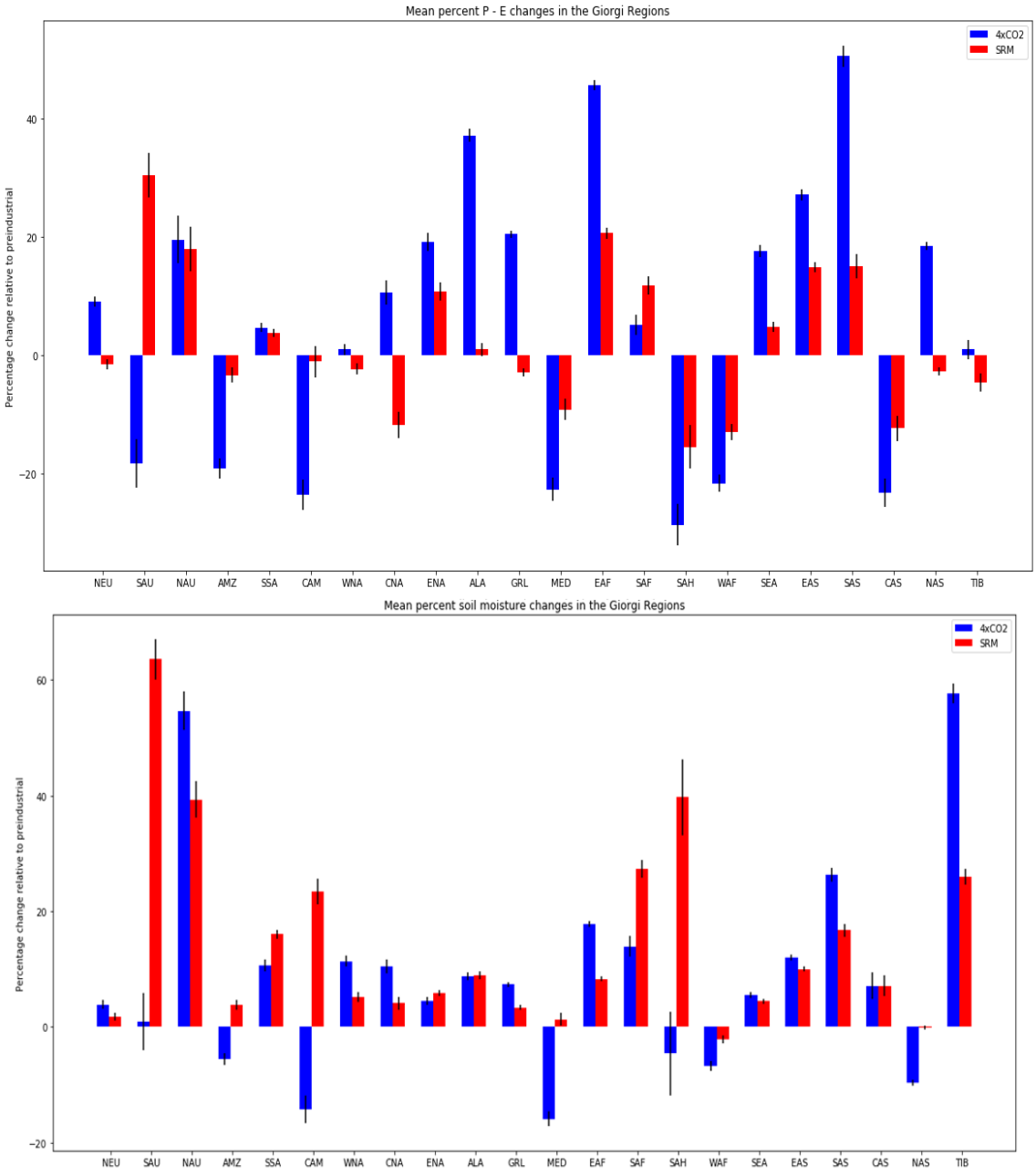

Figure 3.11: **Giorgi regional analysis of P - E (top) and soil moisture shifts under SRM.** As in the previous figure, but plotting P-E and soil moisture.

3.3.1.3 Longwave, shortwave, and net radiative forcing under uniform SRM

Examining the longwave, shortwave, and net radiative forcings in the uniform SRM case relative to the control run illustrates where exact cancellation cannot be obtained (Figure 3.12). The additional shortwave reflected radiation predominantly depends on the incoming solar radiation, while the longwave forcing predominantly depends on the surface temperature, as the quadrupled CO_2 is well-mixed. The radiative influence of the cloud banks in the ITCZ, shifting into the summer hemisphere, can be seen in both longwave and shortwave forcing.

In the polar night, no additional shortwave radiation is reflected by additional AOD, despite a positive net forcing due to CO_2. Consequently, for these regions and times, adjusting local AOD will not be sufficient to balance the radiative forcing; adjusting AOD elsewhere may help to reduce these radiative forcing anomalies, but risks overcompensating in other regions. (For example, one could overcool the tropics to attain preindustrial temperatures at the poles by increasing the AOD of a uniform layer of SRM.)

When uniform SRM is used to cancel the radiative forcing from CO_2, the resulting net radiative forcing (Figure 3.13) relative to the control is greater than zero during the NH summer and less than zero for the NH winter: in the annual mean, it is less than zero at high latitudes and less than zero in the tropics (Figure 3.14), contributing to the characteristic pattern of residual warming. Note the "overforcing" in the tropics where the shortwave reflected radiation exceeds the reduction in outgoing longwave radiation due to greenhouse gases, and the reverse in the midlatitudes.

The resulting net radiative forcing (Figure 3.13) is characterised by a slight negative radiative forcing in the tropics and a strong positive radiative forcing over Antarctica and the Southern Ocean in the SH winter, and a radiative forcing that is negative over Northern polar latitudes in the summer and positive in the winter. It seems likely that there are regions which no aerosol distribution can restore to net zero radiative forcing in each month, particularly as it is not possible to use aerosols to reflect radiation in the polar night, but that but that better cancellation could be

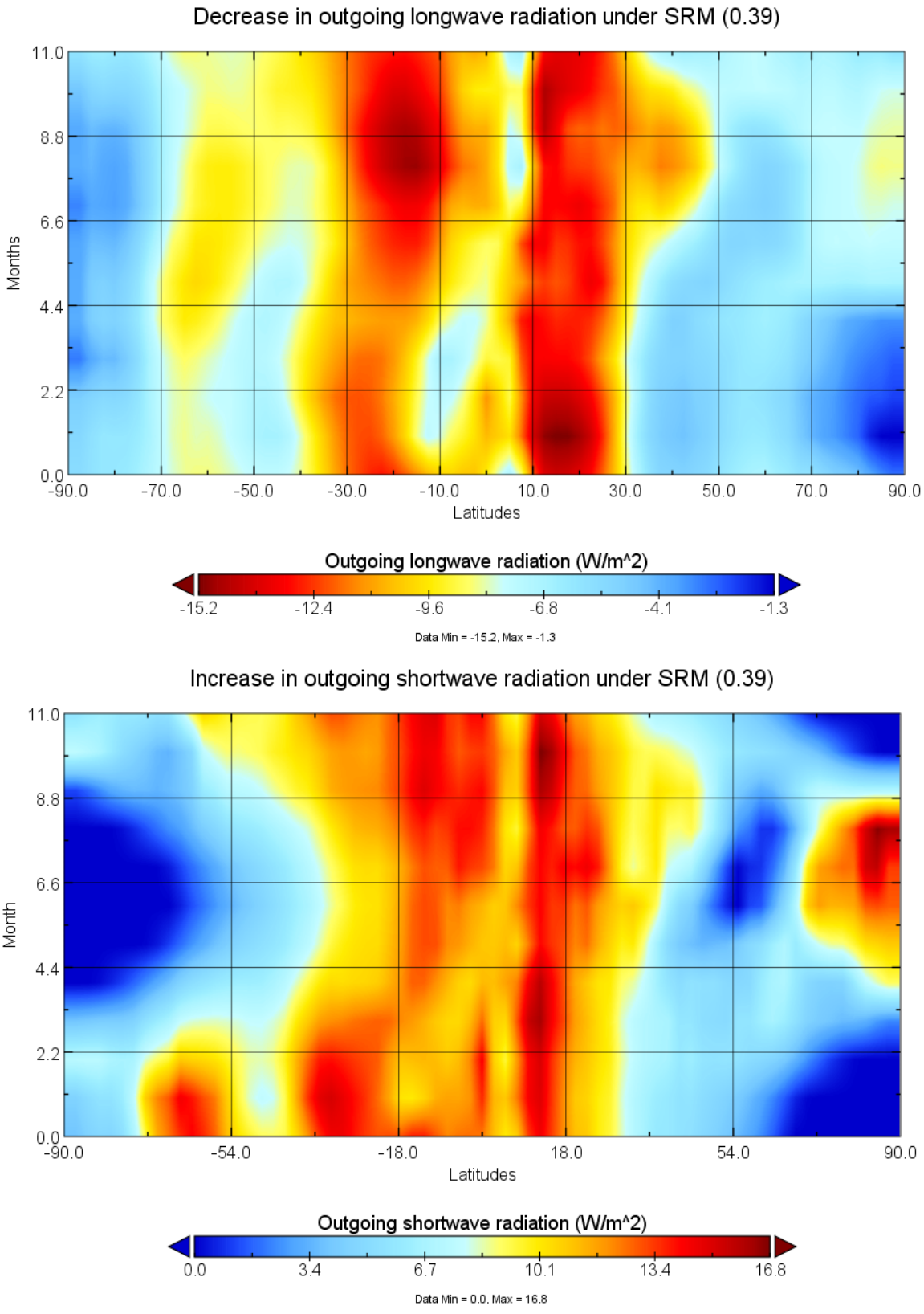

Figure 3.12: **Latitude-time plot for the annual cycle of longwave and shortwave forcings from uniform SRM+ 4xCO_2.**

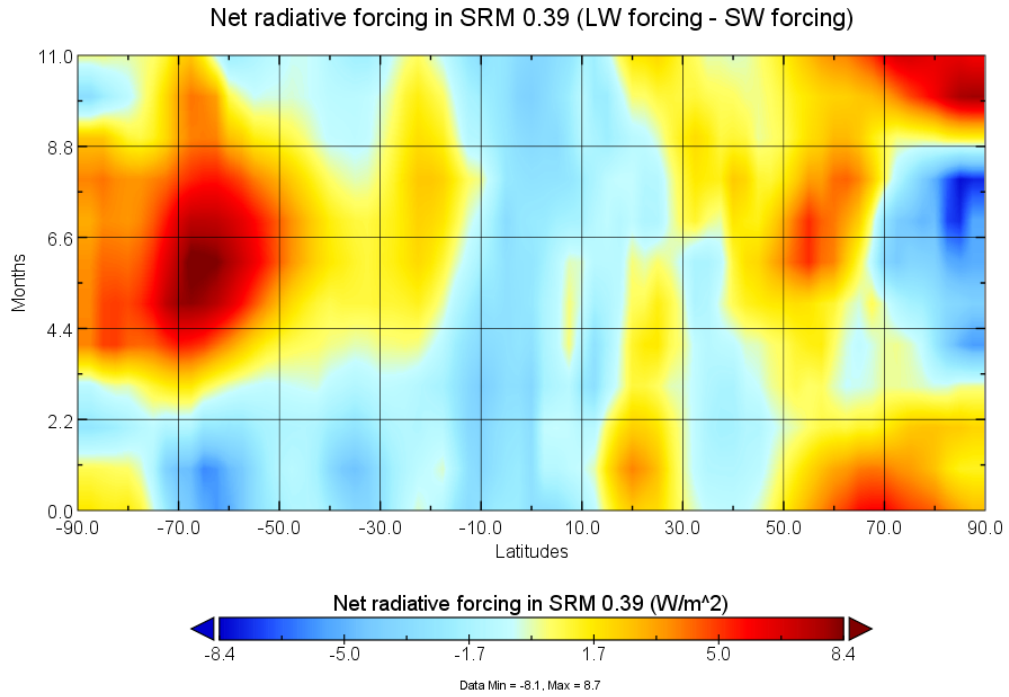

Figure 3.13: **Latitude-time plot for the annual cycle of net radiative forcing from uniform SRM used to cancel out 4xCO_2.**

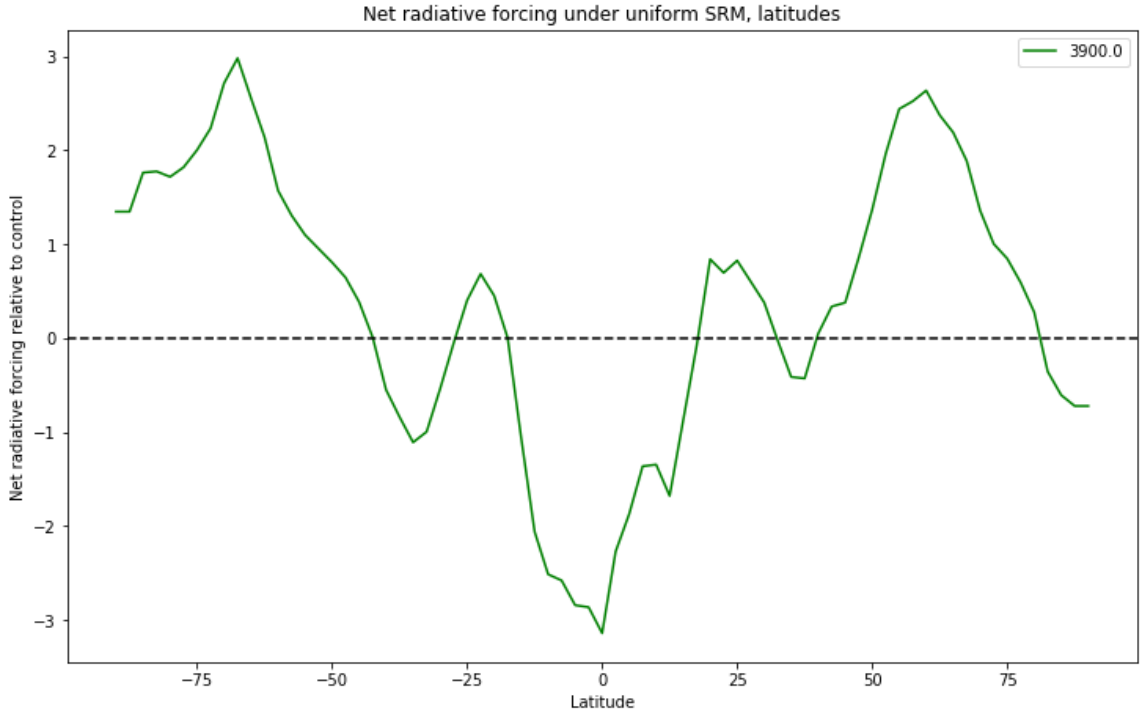

Figure 3.14: **Zonal annual mean net radiative forcing under uniform SRM cancelling out 4xCO_2, in green.**

achieved through some level of optimisation and iterating on earlier results from the same model, which we will explore in Chapter 4.

3.3.2 Investigating highly asymmetric radiative forcing - NH and SH SRM

We plot temperature anomalies under NH and SH SRM simulations, as described in the methods section.

Global mean temperature changes are predominantly influenced by the total AOD: under SRM in different hemispheres, with the same total AOD used, the average temperature difference is 0.3K (Figure 3.15). Geoengineering in the Southern hemisphere is less effective at controlling global mean temperature due to factors associated with Arctic amplification of global warming - snow and ice albedo effects, land fraction differences, circulation changes, and the lapse rate feedback [150]; consequently, the residual temperature is higher under the southern hemisphere geoengineered scenario.

By shifting the energy balance between the hemispheres, hemispheric geoengineering results in a large, robust shift to the intertropical convergence zone (ITCZ), where the Hadley cells of large-scale atmospheric circulation converge. SRM in the Northern hemisphere shifts the thermal equator, and hence the ITCZ, to the South, while SRM in the Southern Hemisphere shifts it to the North. (Figure 3.15). We can measure the ITCZ by the position of its centroid, which I define as the centre (average position) of the zonal, annual average precipitation between 30°N and 30°S. (Table 3.2) We can see from the data in the table that uniform SRM shifts the ITCZ centroid back towards its position in the preindustrial case, as compared to the shift around 1.1°to the North under 4xCO_2. It can also be seen that the latitudinal position of the ITCZ centroid under NH SRM and SH SRM average to equal the response of the ITCZ centroid to uniform geoengineering; however, the overall tropical precipitation response is less linear, as we will examine later on in this chapter. It is also the case the NH SRM shifts the ITCZ 4.74°to the South compared to the preindustrial case, while SH SRM shifts the ITCZ 3.58°to the North relative to the preindustrial case; this is due to the fact that the ITCZ centroid was already slightly in the Northern hemisphere in this model, and likely due to the fact that the Northern hemisphere

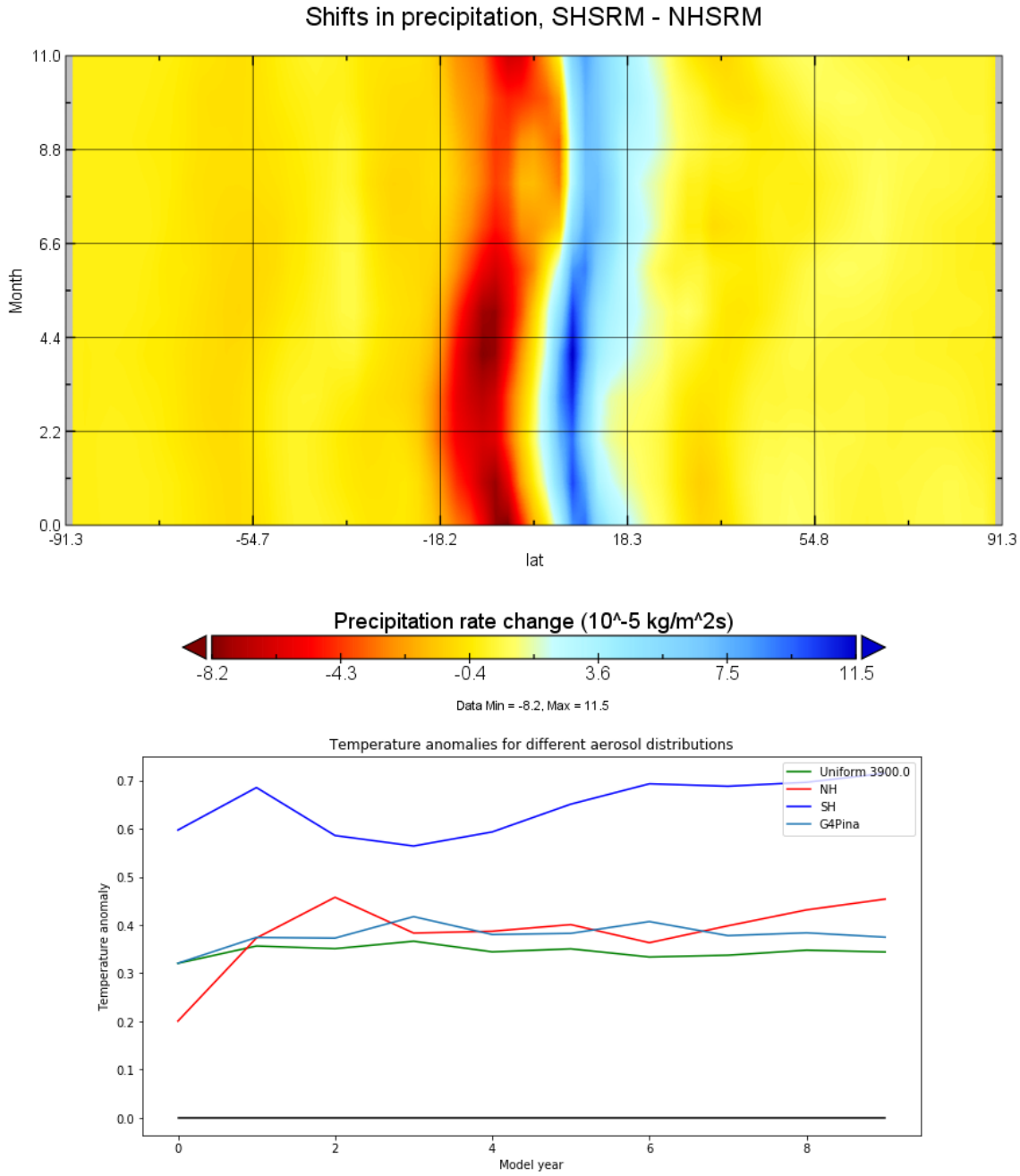

Figure 3.15: **Above: Latitude-time plot of precipitation for the annual cycle of SH SRM - NH SRM**, illustrating shifts in the Intertropical Convergence Zone (ITCZ) under hemispheric forcing; **Below: Annual global mean temperature anomalies compared across simulations**, including the uniform, NH, SH, and Pinatubo-like (G4) forcing patterns.

is cooled more than the Southern hemisphere by hemispheric SRM due to its greater land fraction and Arctic amplification of global warming, which means the resulting shift to the thermal equator is greater.

Simulation	**ITCZ centroid position**
Control	1.08 °N
4xCO_2	2.17 °N
Uniform SRM	1.00 °N
NH SRM	3.66 °S
SH SRM	4.66 °N

Table 3.2: ITCZ centroid positions under various control and geoengineering scenarios

Data from the hemispheric geoengineering scenarios provides the extreme case, but any distribution of SRM that changes the cross-equatorial energy flow will likely shift the ITCZ position. The hemispheric case also provides a first test of the linearity of the response of key climate variables with respect to the latitudinal distribution of AOD.

The plots (Figures 3.16, 3.17) show ensemble averages across the final year of the ten-year simulations. When comparing the effects of aerosols in the Northern and Southern hemispheres relative to the CO_2 baseline, the results in temperature and radiative forcing are quite close to linear. There are exceptions at the poles where Arctic amplification leads to nonlinearities in temperature: the seasonal effects of the polar night, especially on polar temperature, are partly masked by the annual average.

Temperature-driven shifts in tropical cloud take place under asymmetric SRM that are suppressed under uniform SRM: these shifts result in, for example, a reduced longwave cloud forcing and an increased shortwave cloud forcing North of the equator under SH SRM (and the opposite effect in NH SRM), which corresponds to more, lower clouds shifting to the North along with the ITCZ.

Due to the large-scale shifts to the ITCZ and the dynamical response, asymmetric SRM has much larger impacts on regional precipitation. (Figure 3.18) In particular,

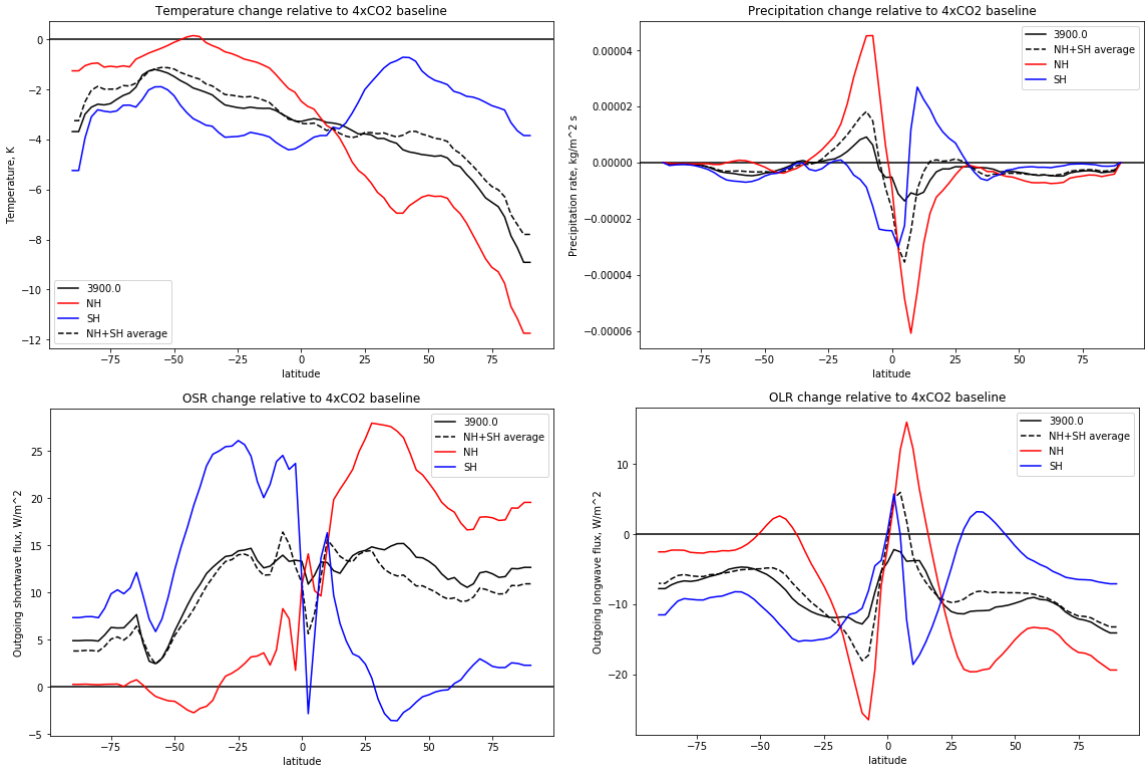

Figure 3.16: **Annual zonal mean temperature, precipitation, and outgoing shortwave and longwave flux changes under NH and SH SRM relative to $4xCO_2$ baseline.** The dotted line shows the average of the NHSRM and SHSRM simulations, and thus provides a test of the linearity of the response to inhomogenous forcings. The solid black line shows the actual result of uniform SRM.

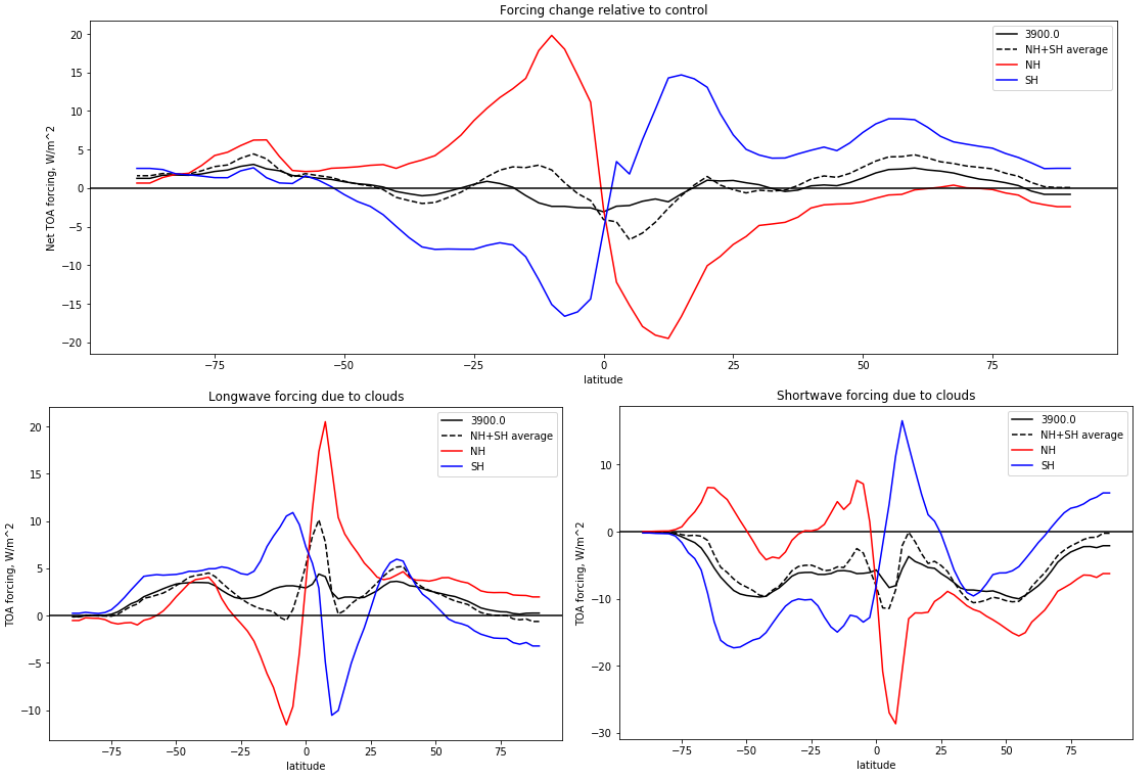

Figure 3.17: **Changes in annual zonal mean radiative forcing in NH and SH SRM with respect to the preindustrial simulation.** Top: net (longwave+shortwave radiative forcing); bottom left, longwave forcing due to clouds (clear-sky forcing minus net forcing); bottom right, shortwave forcing due to clouds. The dotted line shows the average of NH and SH SRM. The solid black line shows the actual result of uniform SRM.

regions such as Central America, Sahel and West Africa, China and India are affected by the shifting ITCZ with large changes in precipitation, with sign depending on the geoengineered hemisphere, as previous studies have shown for the Sahel. [30].

Some of these changes in precipitation can be attributed to the projection of changes onto modes of climate variability. Hemispheric geoengineering has a strong influence on the Indian Monsoon (rainfall in the Giorgi region with SAS as shorthand), at least in part via its influence on SSTs in the Pacific. Under NH SRM, the SSTs project strongly onto an El Nino pattern in the Pacific, which in turn suppresses the Indian Monsoon via the well-documented teleconnection. [151] In addition, NH SRM increases Eurasian snow cover relative to the control and relative to SH SRM, which acts to suppress monsoon precipitation. [152] Analysis was conducted which linked the hemispheric geoengineering simulations to ENSO-type SSTs in the NH case, and suppression of ENSO in the SH case, but it has been omitted from this thesis for space; however, we note that there has been work elsewhere which has linked volcanic eruptions which resulted in hemispherically asymmetric forcing to ENSO-like SSTs. [153]

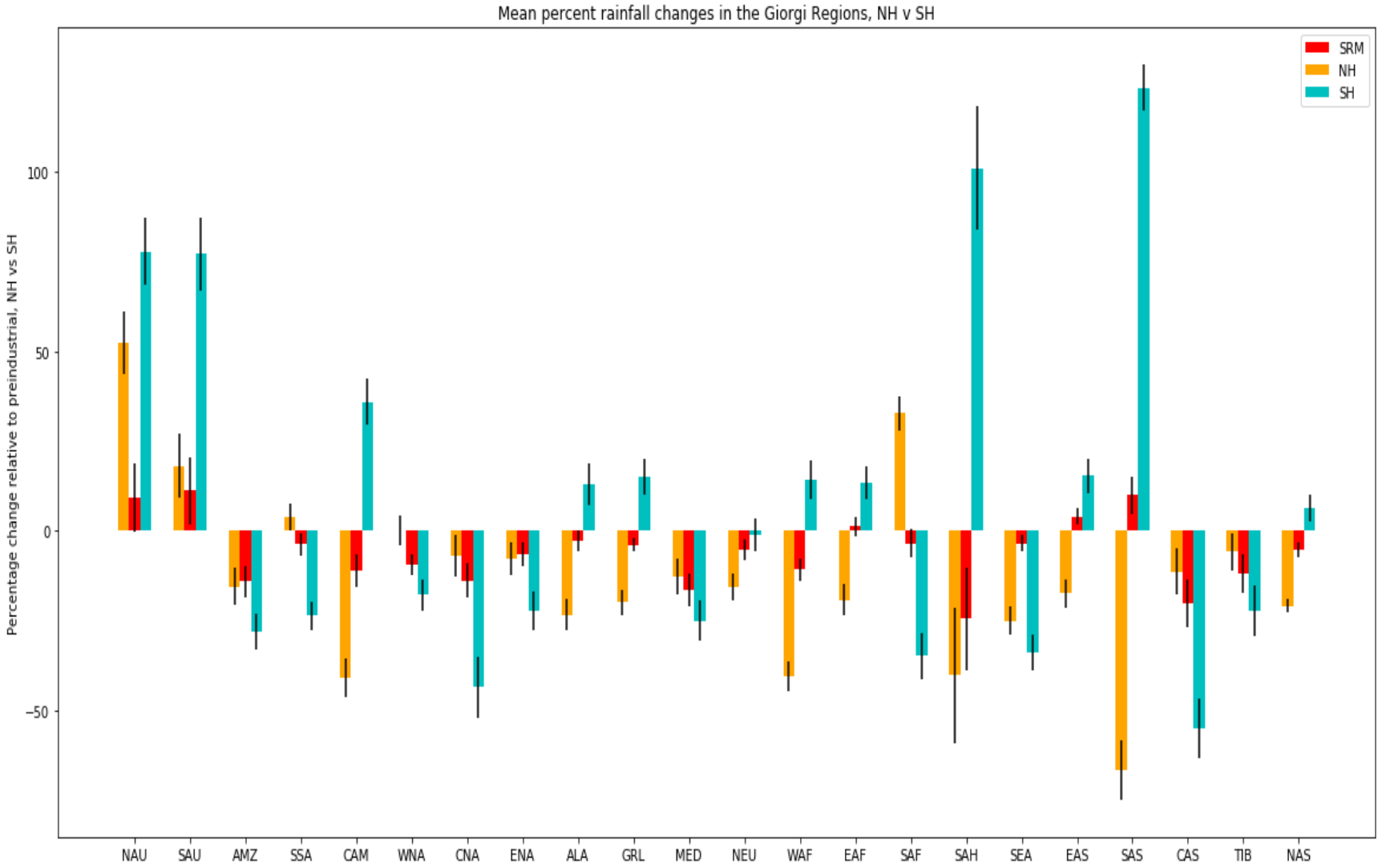

Figure 3.18: **Giorgi regional analysis of precipitation shifts under hemispherically asymmetric SRM**. Plotted are percentage changes to the preindustrial precipitation in the annual mean for each region. Error bars show a standard deviation in the annual precipitation under the SRM or CO2 scenario; the baseline for calculating percentages is the average across all preindustrial model years.

3.3.3 Characterising the response from the G4 multimodel distributions in HadCM3

Looking at a comparison of global mean temperature anomalies in the response to different distributions (not shown), we can see that for distributions where tropical confinement is more pronounced, global mean cooling is greater, as more insolation is reflected. However, regional temperature anomalies and disparities, and the undercooling of poles relative to tropics, are worse. The Giorgi regional analysis of temperature anomalies is plotted in Figure 3.19, which shows that SRM with the (in theory more realistic) G4 multimodel comparisons is often less effective at reducing regional temperature anomalies than uniform SRM. In particular, there are substantial differences in temperature between the different simulations for regions like the Amazon and Africa, although these are largely dominated by the outlier that is the beta scenario, which was intentionally designed to show the effects of large amounts of aerosol optical depth at higher latitudes, rather than the tropical confinement under G4 scenarios, and so is overall generally worse at cooling, especially in the tropics, where we see larger temperature anomalies in this scenario than any of the others.

Analysis of precipitiaton changes in the Giorgi regions reveals that, for most regions, equatorial injection are less effective than uniform SRM at restoring preindustrial precipitation. (Figure 3.20) Concentrating aerosol in the tropics results in decreased precipitation over the Amazon, North America, East Africa and South East Asia relative to other geoengineering scenarios. For other regions, such as Europe, Central America and high Northern latitudes, high-kurtosis distributions perform better, while they results in larger increases in rainfall over India and Australia. The Pinatubo and HadGEM cases resemble the uniform case more closely than distributions with strong confinement.

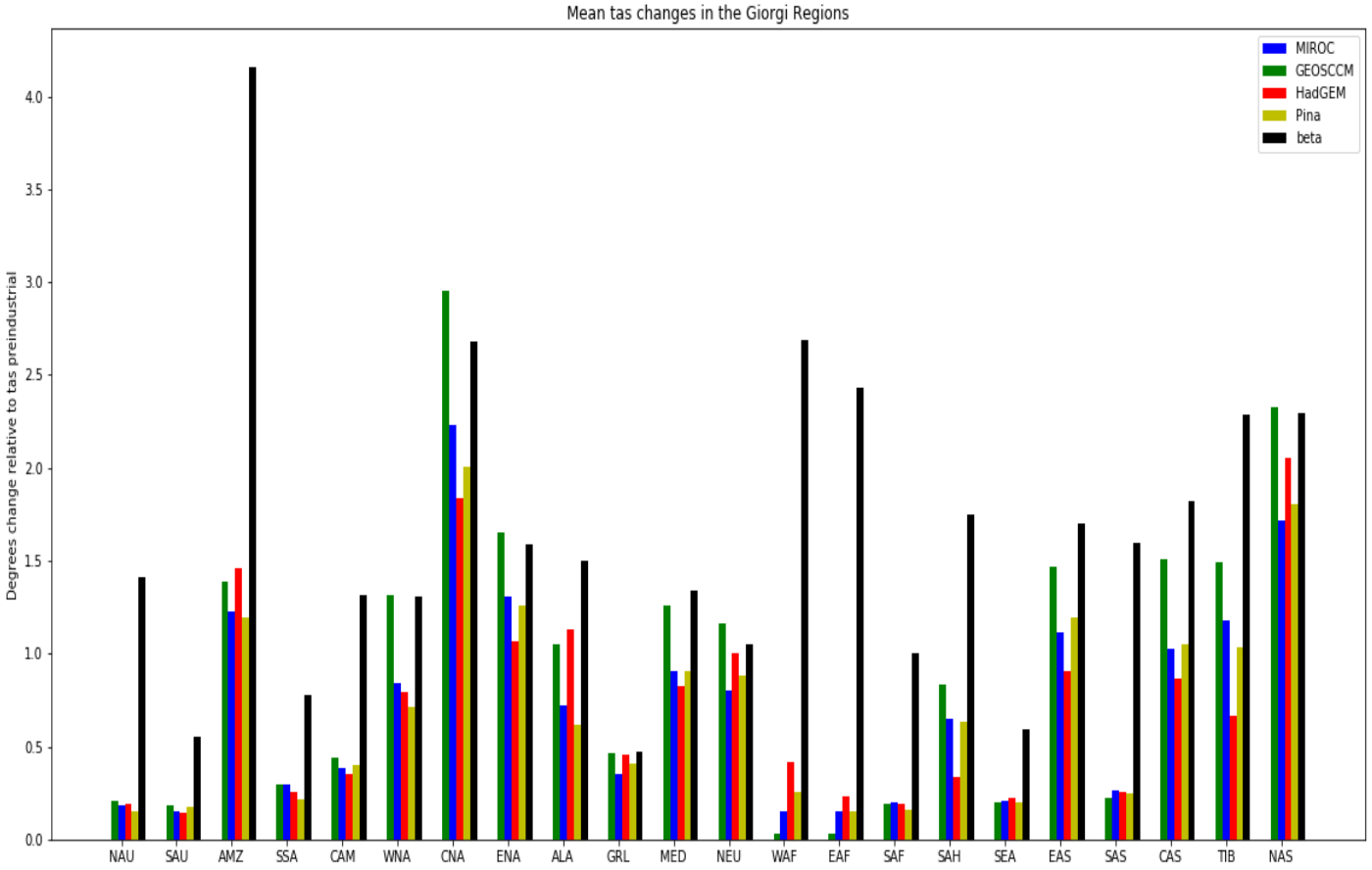

Figure 3.19: **Giorgi regional comparison of temperature anomalies shifts under different distributions from G4.** ULAQ simulations were not included in this plot due to data availability. Annual average 1.5m surface temperature anomalies over the grid cells in the Giorgi region is plotted, with colours indicating different distributions, as indicated in the key.

Differences in aerosol distribution alone - even arising from simulations of same injection strategy - have a large impact on changes to regional precipitation patterns. In some Giorgi regions (e.g. CAM, EAF, SAF, SAH, SEA, SAS and CAS) the differences between distributions are comparable to the difference between 4xCO_2 and uniform SRM. We can also see that for most regions, and sometimes for more than one pairwise comparison of simulations, the difference between the ensemble averages for a different distribution of aerosol exceeds the 67% confidence interval, suggesting that the different distributions modelled under the G4 scenario would drive statistically significant shifts to precipitation in these regions.

Trade-offs will therefore exist where regions disagree about their preferred distribution of aerosol, as well as the overall amount of geoengineering, as explored further in Chapter 4. Furthermore, the result implies that practical limitations in the latitudinal distribution of aerosol that can be obtained by equatorial injections may result in worse outcomes and greater regional inequalities than idealised, uniform SRM, especially if stratospheric models that simulate strong tropical confinement are accurate.

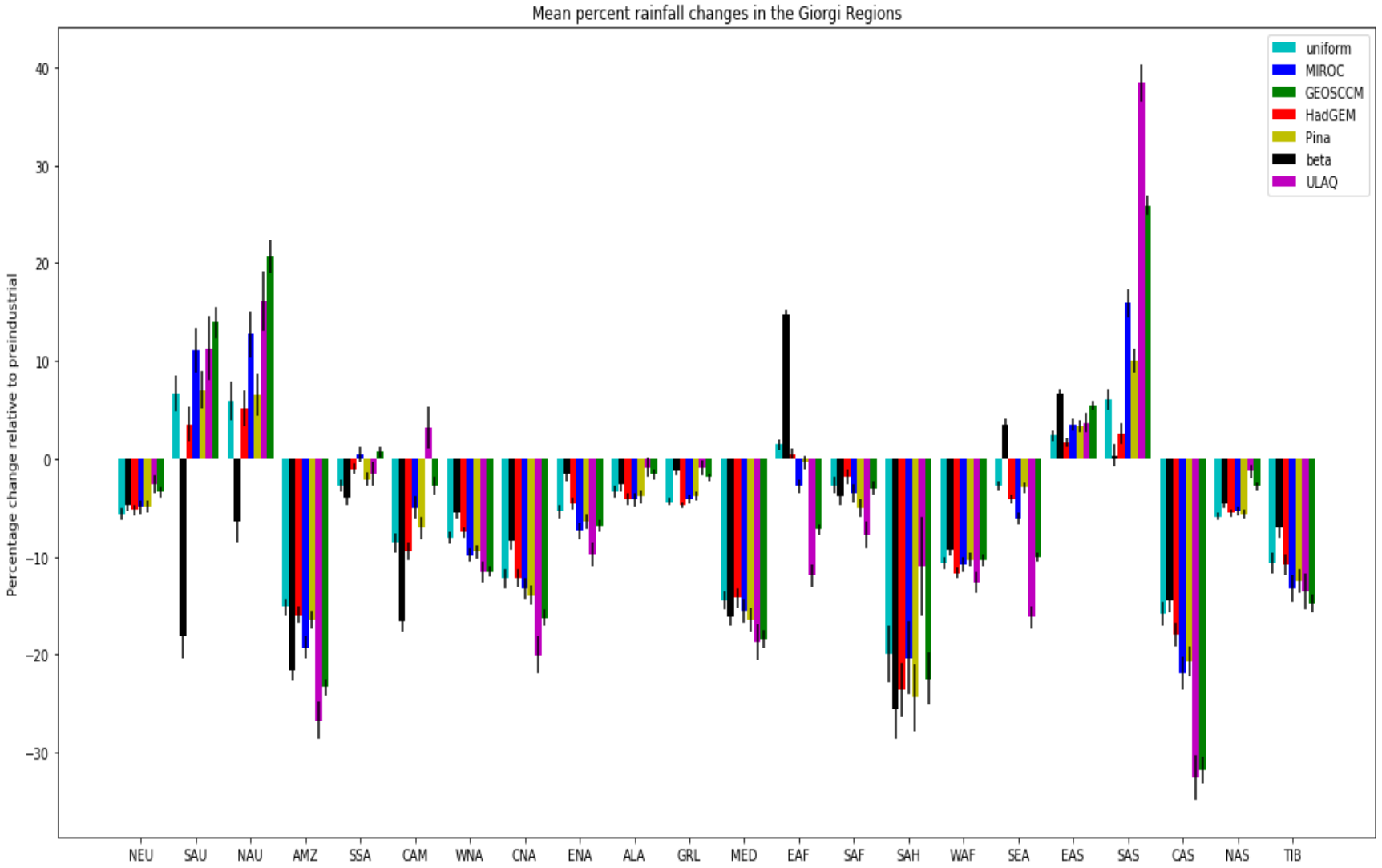

Figure 3.20: **Giorgi regional comparison of precipitation shifts under different distributions from G4.** Ordered from left to right by kurtosis of the distribution. Annual average precipitation over the grid cells in the Giorgi region is plotted, with colours indicating different distributions, as indicated in the key. Error bars indicate a 67% confidence interval across the standard deviation from the ensemble simulation.

3.3.4 Investigating stripe simulations

To investigate linearity a series of geoengineering simulations where the aerosol was confined to ten adjacent model latitude bands (22.5 degrees of latitude). In each stripe, the AOD was set to 0.39, and zero elsewhere, against a background of $4xCO_2$. The AOD was fixed at 0.39 for three latitude bands at the poles (two in the NH and one in the SH), so the data shown here excludes this region.

By calculating the effect that each aerosol injection has relative to $4xCO_2$, we can test for linearity and additivity by comparing to uniform SRM of the same optical depth. In Figures 3.21 and 3.22, "forcing" in a given simulation is defined as the additional outgoing radiation (OLR+OSR) in the SRM simulation relative to the $4xCO_2$ case.

Despite changes to the circulation and cloud feedbacks due to the imposed "stripes" of forcing, the changes in annual average temperature as a function of latitude are remarkably linear over the course of the first decade. This is true of the longwave and shortwave radiative forcing to a slightly lesser extent.

Notably, in the radiative forcing case, the approximate linearity arises due to cancellations between shifts in clouds, especially in the tropics - where both OSR and OLR are non-linear - but also in the mid-latitudes. Equatorial and tropical SRM tends to shift low clouds - which reflect shortwave radiation - away from the midlatitudes, while this is compensated by cloud shifts in simulations with higher-latitude SRM. Future ensembles, discussed in Chapter 4, will use sums of "basis vectors", to attempt to construct an AOD distribution where the radiative forcing cancels out that of the quadrupled CO_2 relative to the preindustrial control.

The response of precipitation is more complex and not linear: the sum of the effects of each stripe is greater than the uniform aerosol layer in the tropics, although the precipitation response is close to linear at higher latitudes due to cancellations. (Figure 3.23)

More in-depth characterisation and analysis of the precipitation changes under these striped simulations was conducted, but has been omitted here for space.

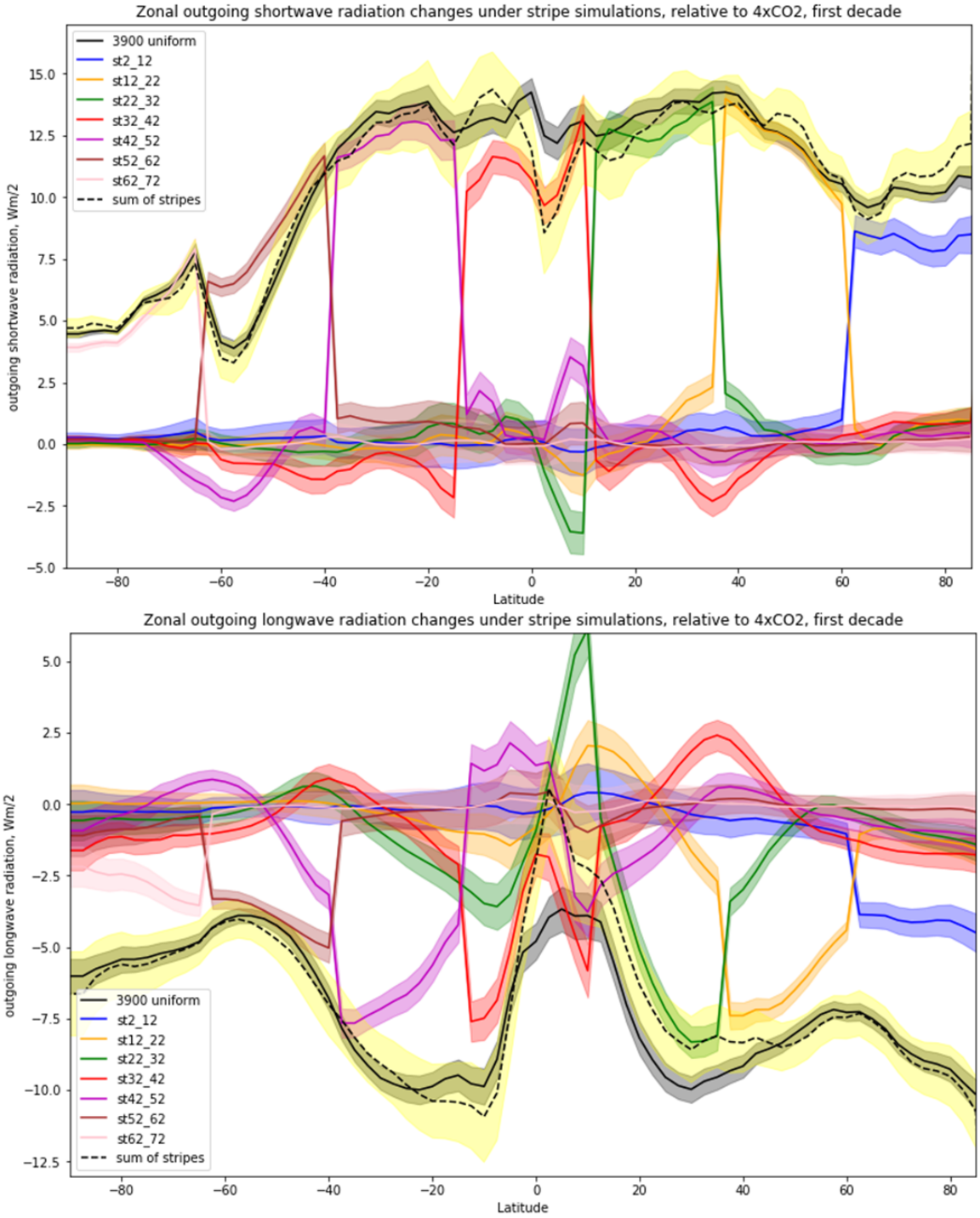

Figure 3.21: **Linearity testing in longwave and shortwave forcing.** Decadal, zonal averages of the variables are plotted. The black line indicates the uniform SRM; the dashed line is the sum of the response to each of the stripe simulations (due to the scaling, the sum of the stripe distributions is the uniform distribution.) In the linear case, the lines would map onto each other. Shading indicates a 67% confidence interval, based on the ensemble variability.

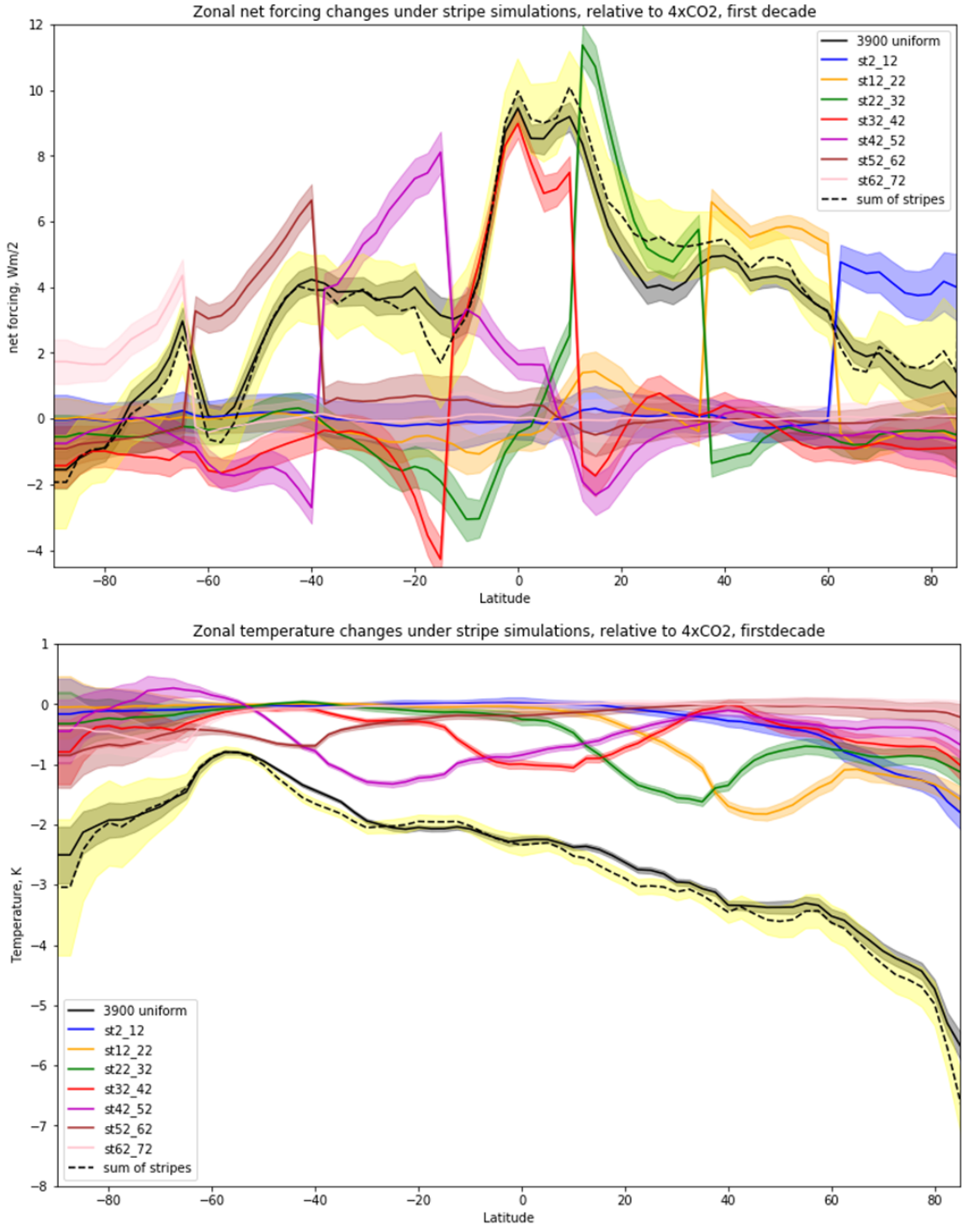

Figure 3.22: **Linearity testing in net radiative forcing and temperature** Decadal, zonal averages of the variables are plotted. The black line indicates the uniform SRM; the dashed line is the sum of the response to each of the stripe simulations (due to the scaling, the sum of the stripe distributions is the uniform distribution.) In the linear case, the lines would map onto each other. Shading indicates a 67% confidence interval, based on the ensemble variability.

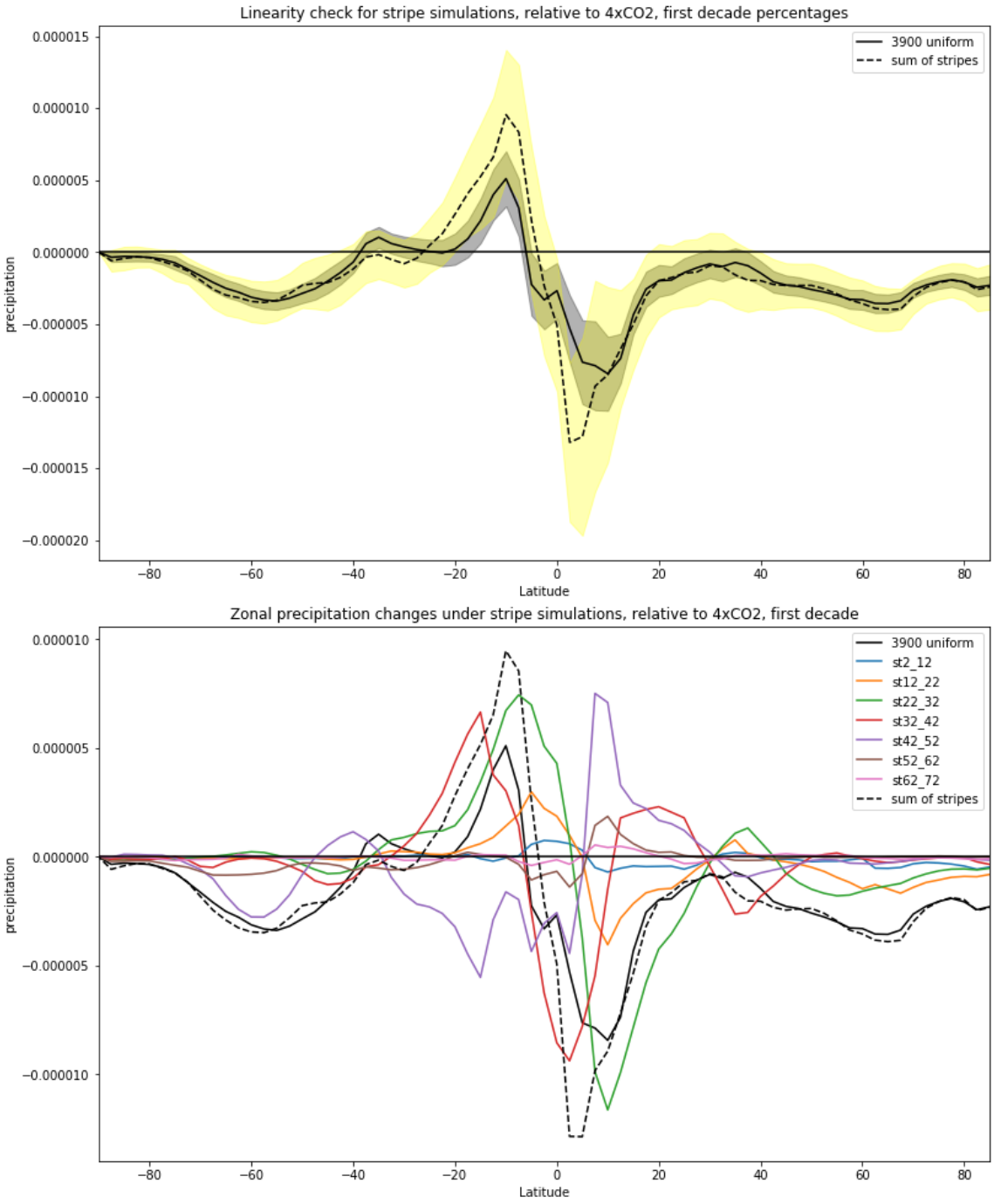

Figure 3.23: **Linearity testing in precipitation, absolute changes relative to $4xCO_2$.** Decadal, zonal averages of the variables are plotted. The black line indicates the uniform SRM; the dashed line is the sum of the response to each of the stripe simulations (due to the scaling, the sum of the stripe distributions is the uniform distribution.) Multicoloured lines indicate each of the latitude bands, in the legend from North to South. In the linear case, the lines would map onto each other. Shading indicates a 67% confidence interval, based on the ensemble variability.

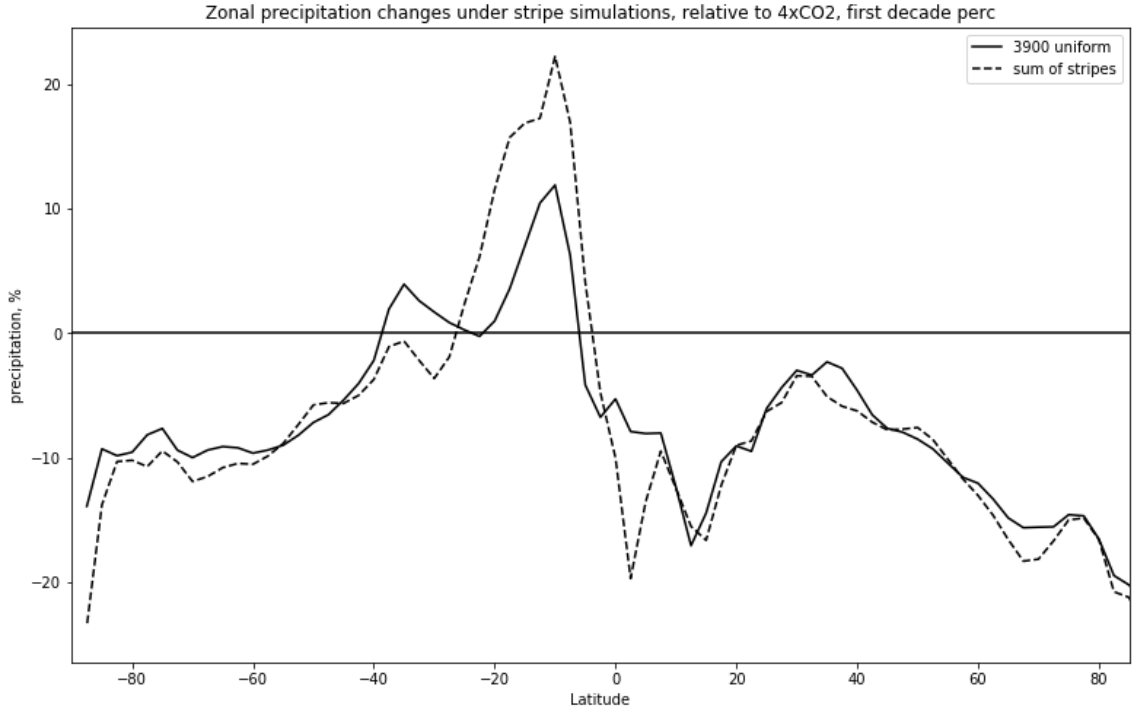

Figure 3.24: **Linearity testing in precipitation, percentage changes relative to $4xCO_2$**. Decadal, zonal averages of the variables are plotted. The black line indicates the uniform SRM; the dashed line is the sum of the response to each of the stripe simulations (due to the scaling, the sum of the stripe distributions is the uniform distribution.)

3.4 Discussion and conclusions

3.4.1 Usefulness of HadCM3 and CPDN

One objection that might be made to the results in this thesis is that HadCM3 is a relatively old and low-resolution climate model - first described in 1999, its 73 x 96 resolution atmosphere is much lower resolution than modern-day Earth System Models. Fewer processes in the climate are accurately modelled, and many are parameterised, while more and more climate-relevant processes and better representation of physical processes have been developed in the intervening years.

While this is undoubtedly true, there is reason to believe that HadCM3 is a relatively good climate model for these purposes even today. HadCM3 was first described in 1999 and used as one of the major models in the IPCC's Third Assessment Report (TAR), it proved an exceptionally good model for its generation - the first such model not to require flux adjustments. [138] For this reason, and its computational speed - up to 1,000x faster than CMIP6 models - it is still extensively used by modelling groups [154] and has been run as part of model comparisons alongside CMIP5 models, including in GeoMIP [60] [83] [22], as well as in standalone SRM studies. [155] We can see in Kravitz et al. (2013) [58] that HadCM3's response, both to abrupt 4xCO_2 and to SRM, is close to the ensemble mean for zonal surface temperature anomalies and P - E across most regions, albeit models differ substantially on how they represent tropical precipitaiton. Similarly, Tilmes et al. (2013) [57] contains a multimodel comparison of the response of HadCM3 to other models in the GeoMIP (CMIP5) ensemble, and we can see that HadCM3 is again not a significant outlier in terms of its simulation of precipitation changes in the annual, global average, or its simulation of the fast or slow response, and it performs reasonably well in the regional analysis. Indeed, this good performance is why HadCM3 continued to be used in multimodel comparison GeoMIP efforts alongside climate models from the next generation after it, such as HadGEM2-ES.

One of the known flaws of HadCM3 is a warm bias over the Amazon rainforest [156] while the drizzle problem, whereby the model undersimulates the number of days with no precipitation and oversimulates the number of days with a low level of

precipitation, is also known and discussed in Chapter 5. These model biases should be borne in mind when evaluating model results in this thesis, but it is also worth remembering that it is only the fact that HadCM3 is a relatively low-resolution climate model that allows it to be run on volunteer computers and to generate the large ensembles and investigate the large array of different radiative forcing patterns that we are able to here.

The use of CPDN also entails trade-offs. As simulations are run on volunteer computers, we cannot implement any monitoring to check if the simulation is running properly; we can only analyse output once it has been uploaded. Simulations can take some weeks or months to run, which inevitably meant over the course of the thesis that delays were incurred. Volunteer computers can only run models on spare computing power, so the size and resolution of the model we could use was limited, as were the output variables that we were able to obtain, which were restricted to monthly means for most variables except temperature and precipitation due to the bottleneck from users uploading data to the CPDN servers. The ensembles that were returned were not all the same size, because of the random element concerning which volunteers successfully completed and uploaded their runs. Because generating these ensembles relies on a unique, one-of-a-kind assembly of volunteer computers, there is a limit to how independently reproducible our results are without using the same network (or setting up a similar one, which would be costly and time-consuming). Finally, the version of the HadCM3 model that is used on CPDN volunteer computers, although widespread amongst the volunteers, is not open-source, with some of the modifications including the latitudinally-varying aerosol modification and triple call to the radiation code not present in the standard version, which again limits reproduciblity. The trade-off is that we can submit batches of hundreds to thousands of experiments on CPDN and generate larger ensembles and more model years with this low-resolution model than many experimental setups would be capable of.

3.4.2 Limitations in experimental set-up

In addition to the climate model we are using, with its own biases, there are certain unavoidable limitations in our experimental setup that should be acknowledged.

These are highly idealised simulations; they impose idealised radiative forcing onto an abrupt 4xCO_2 background, which is necessarily unrealistic. However, these idealised experiments are still widely studied and have a great deal to tell us about the expected climate response to forcings: running more realistic scenarios of climate change or SRM might result in a weaker signal to noise ratio, a greater role played by internal climate model variability, and consequently a lesser ability to distinguish the impact of different latitudinal distribution of aerosol. Conversely, if the effect size is still relatively small even in a relatively extreme scenario like an abrupt quadrupling of carbon dioxide and SRM to match, then we are more justified in concluding that it is likely to be small in more realistic scenarios with smaller perturbations to the climate system.

There are limitations in terms of our ability to represent stratospheric aerosol injection. HadCM3 does not have a stratosphere; it has no representation of aerosol microphysics or stratospheric dynamics; consequently, we cannot explicitly simulate the injection of sulphates into the stratosphere and the atmospheric physics and chemical processes that result in a distribution of stratospheric aerosol optical depth. We rely on prescribing the aerosol optical depths, which for most of our simulations are assumed to be the same in each model year. There is also a limited time resolution over which the distribution changes - we prescribe aerosol optical depths for each month of the year, whereas more advanced models would include an aerosol optical depth that changed each day or perhaps even more frequently. In reality, there would be feedbacks between the troposphere and the stratosphere, so we might expect the patterns of warming at the surface and in the troposphere to in turn alter the distribution of aerosols over time. HadCM3 does not include this coupling and there is no dynamic adjustment of the stratospheric aerosol optical depth on a year-to-year basis; it remains as prescribed at the beginning of the experimental run. However, it is not clear that these second-order effects would be particularly significant given that we accept our simulations are already highly idealised, and we are coupling distributions generated from one stratospheric model to a different model of the troposphere, or holding them constant to represent an idealised case, so it is unclear how helpful or significant it would be to attempt to simulate these explicitly. While it might be possible to exert some control over the distribution or maintain its broad features by

adjusting injection strategy, as discussed in Chapter 1, it is unlikely that we would be able to attain such a degree of control that the distribution was identical each year, nor would it be possible to obtain a perfectly uniform layer of aerosol optical depth.

Some of the inputs are highly idealised - the zonal "stripe" simulations are obviously not possible to achieve due to the Brewer-Dobson circulation which would move particles meridionally in the stratosphere, which is why we use them only as a thought experiment to test the limits of linearity and do not interrogate their impacts on surface climate significantly. Where possible, I have attempted to examine choices of aerosol optical depth patterns with some justification - including obvious idealised cases such as uniform aerosol optical depth. Hemispheric geoengineering may seem idealised but we have seen in Chapter 1 that volcanoes have had similar effects, when the injection of material is predominantly in one hemisphere and one branch of the Brewer-Dobson circulation. Finally, with the exception of the beta pattern, the other G4 multimodel patterns and their broad features, including meridional confinement in the tropics and asymmetry between the Northern and Southern hemispheres, as well as seasonality, were drawn from models that did have stratospheric dynamics and aerosol microphysics, and so might feasibly represent some of the features of an aerosol distribution resulting from an equatorial injection strategy. This is also true of the injection basis functions from the GLENS experiment used in Chapter 4.

Finally, there are also obvious limitations in the aerosol scheme itself - there is a relatively low resolution in the 71 latitude bands; there is no resolution in the zonal case, so we have to assume that the aerosol is well-mixed by zonal winds in the stratosphere, which may not be the case; there are bands at the poles that the radiation scheme does not interact with, which would not be the case in reality; and we discussed the limitation posed by having an odd number of bands for generating truly hemispheric geoengineering simulations earlier in the chapter.

3.4.3 Conclusions

In this chapter, we have demonstrated that HadCM3 can be used to simulate various scenarios for SRM, including uniform SRM that compensates for the temperature change under an abrupt quadrupling of CO_2, hemispheric geoengineering, and a range

of distributions taken from the multimodel G4 experiments which compared the effects of equatorial injections of aerosols. The model has been calibrated, and the appropriate globally uniform aerosol optical depth to compensate for 4xCO_2 determined. We have characterised large-scale features of the precipitation and temperature changes under SRM, showing the regional effects of uniform SRM, and reproducing the well-known result that SRM overcools the tropics and undercools the poles relative to the preindustrial case when used to compensate for global mean temperature increases.

We have characterised the regional effects of SRM on precipitation, showing that for most regions the magnitude of the average precipitation anomaly under SRM is smaller than under abrupt 4xCO_2, although the SRM simulation is generally drier (at least in terms of precipitation, not P - E) and for many regions the change is of a different sign (i.e. the region is wetter under 4xCO_2 and drier under SRM.) We have also shown that P - E and soil moisture show smaller net changes than precipitation, and in some cases regions that appear drier when considering only precipitation are actually shown to have more moisture availability under SRM than the preindustrial case when considering P - E and soil moisture. We show that these changes are moderated by SRM relative to 4xCO_2 for all but 4 Giorgi regions, although regional inequalities persist, in line with other work that suggests these metrics are less affected by SRM than precipitation alone.

We also investigate hemispherially asymmetric SRM, finding that there are significant shifts to tropical precipitation and the ITCZ in the opposite direction to the geoengineered hemisphere, in line with work elsewhere and the effects of volcanoes that predominantly resulted in stratospheric aerosols being injected into one hemisphere. We find that the zonal, decadal temperature changes are remarkably linear and additive, in that the average of the effects of geoengineering each hemisphere is approximately the effect of a uniform layer of stratospheric aerosol. We extend this analysis to highly inhomogenous, highly idealised "stripe" simulations and show that linearity is still a remarkably good approximation, even in this very extreme case of latitudinally confined geoengineering. This result will motivate much of the optimisation work that follows in Chapter 4, which defines "basis vectors" and exploits the assumption of linearity to construct "ideal" distributions to cancel out zonal, decadal

temperature anomalies, as well as using pattern-scaling to determine the results of optimising for different figures of merit.

We also briefly compare the simulated climate response to a range of different aerosol optical depth distributions drawn from the geoengineering literature, based on multimodel simulations of stratospheric aerosol injection under the G4 experiment. While these simulations are arguably under-analysed and it may have been fruitful to compare the differences in greater depth, we demonstrate that there are statistically significant differences in regional precipitation response between the simulations of different imposed patterns of aerosol optical depth. The G4 experiment compared different climate models, so when the climate response differed between models, it was not necessarily possible to distinguish whether this was due to differences in simulated distribution of aerosol optical depth that resulted from the same injection strategy, or other differences in the climate models, such as their tropospheric response. My work, which finds significant differences when the same climate model is driven by patterns taken from the G4 ensemble, suggests that the differences in the simulated patterns - and thus the differences in stratospheric dynamics and aerosol microphysics between models - are important, as the same injection strategy could lead to significantly different climate outcomes depending on the range of aerosol distributions that could result from that strategy across the range of models used.

The next chapter will build on the calibration, methods, simulations, and linearity findings of this chapter to investigate the potential for "optimising" SRM to achieve a range of different objectives.

Chapter 4

Exploring optimisation of zonal temperature and precipitation using SRM in HadCM3

4.1 Introduction

As discussed in earlier chapters, solar radiation management (SRM) has been explored in models [54] as a possible technique for reducing the negative impacts of climate change, most notably by the Geoengineering Model Intercomparison Project (GeoMIP). [55] The technique that is considered most feasible for large-scale solar radiation management purposes is stratospheric aerosol injection (SAI). Many modelling studies use a simplified implementation of SRM designed to represent the main features of SAI: for example, the G1sulfur and G1solar experiments in GeoMIP either reduce the solar constant or implement a uniform aerosol optical depth layer, calibrated to cancel out the warming effects of CO_2.

In reality, there are likely to be more degrees of freedom in a real-world implementation of SRM than a single uniform aerosol optical depth. [89] Changes to the timing and location of injections of aerosol precursor, typically SO_2 which nucleates into H_2SO_4 in the atmosphere, will change the resultant spatio-temporal distribution of aerosol, and hence the climatic effects that SRM will have. [157] Fast zonal winds

in the stratosphere will tend to result in a rapid zonal equilibriation, and thus an approximately zonally uniform distribution of aerosol over the timescale of days after an injection. However, the slower Brewer-Dobson meridional circulation - which has a characteristic timescale of months - gives rise to the possibility that some degree of control over the latitudinal and seasonal distribiution of aerosol could be achieved by injecting aerosol at different latitudes, altitudes, or times of year.

In this chapter, we seek to explore whether changing the zonal distribution of AOD can in principle result in significant advantages over the uniform implementation of SRM. We further aim to identify and explore any fundamental constraints on the climate system that we might expect would reduce the efficacy of this kind of latitudinal optimisation of SRM. To what extent can the climate system respond - for example, through changing patterns of circulation - to "smooth out" the climate response to imposed inhomogeneities in the aerosol distribution and its radiative forcing, and so reduce the efficacy of climate control even if we can choose our distribution? We aim to explore different methods and metrics for "optimising" the aerosol distribution under SRM to minimise a range of different anomalies under climate change. This has been achieved using modelling results from the climateprediction.net modelling project [136], which enables large ensembles of a robust GCM (HadCM3) [138] to be run in parallel: this, in turn, allows us to explore a wide range of different distributions and obtain estimates for the uncertainty, and the model's internal variability, across ensembles.

4.2 Methods

4.2.1 The implementation of, and response to, uniform SRM in HadCM3

To maximise the signal-to-noise ratio of our simulations, as well as for the sake of comparison to the widely conducted G1 simulation, we choose to superimpose aerosol optical depth distributions against an instantaneous quadrupling of preindustrial CO_2 ($4xCO_2$). Each ensemble member is generated from an initial 140-year spin-up "start file" of HadCM3 held at preindustrial CO_2. Initial conditions perturbations, ensuring

a separate dynamical evolution of the system, were applied to each ensemble member as described in Chapter 3.

In HadCM3, we simulate the effect of stratospheric aerosol injection by varying a uniform aerosol optical depth in each of the model's 73 latitude bands, which are equally spaced in latitude. This model has been modified from the standard Met Office version to allow the aerosol optical depth to be varied in each of the latitude bands, in an improvement on the standard volcanic forcing described in Tett (2001) [138], which only allowed for hemispheric variations. The aerosol optical depth is uniform across individual latitude bands, which corresponds to an assumption that aerosol is zonally well-mixed. As described in Chapter 3, the two bands covering the North and South poles are not used in the radiation scheme, so these values for AOD do not actually influence the simulation. The AOD distribution is allowed to vary in each month, so each annual distribution that we refer to is 12 x 73 (effectively 12 x 71) aerosol optical depth values for each latitude band and month.

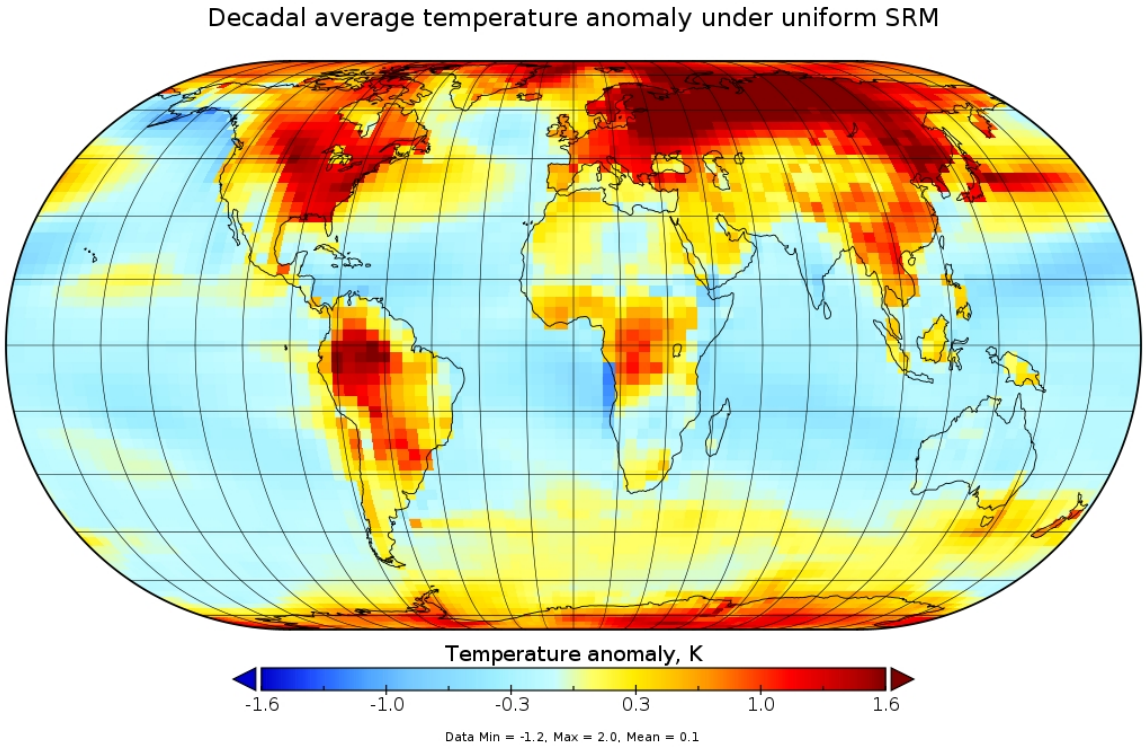

Figure 4.1: **The decadal average temperature anomaly, relative to preindustrial, in degrees Centigrade, under uniform SRM designed to cancel out global mean temperature increases.**

The implementation of the sulfate aerosol in HadCM3 both reflects incoming radiation in the shortwave and absorbs outgoing radiation in the longwave; consequently it reduces the outgoing longwave radiation as well as the incoming shortwave radiation. It was therefore necessary to run a series of simulations that allowed us to calibrate the appropriate AOD for the model to cancel out the temperature increase from $4xCO_2$ which would result in a stable global mean temperature equivalent to the preindustrial case, as described in Chapter 3.

This establishes a uniform aerosol optical depth sufficient to cancel the temperature increase from $4xCO_2$, which we use as a basis for comparison. Figure 4.1 illustrates the decadal average temperature anomaly under this uniform aerosol optical depth layer. Note that the average global mean temperature anomaly across this decade is kept to around 0.1K, but regional undercooling and overcooling well exceeds this. Note the general undercooling of the poles, slight overcooling of the tropics (especially over the ocean), and undercooling of the land relative to the ocean. It is often said that using a uniform layer of aerosol, or solar dimming, to compensate for CO_2-induced warming will overcool the tropics while undercooling the poles, owing to the differing effects of direct insolation and Arctic amplification on these regions, but it is also true that the land is undercooled relative to the ocean. This is due to a combination of the underlying AOD of the oceans relative to the land - which, in turn, determines how much additional reflectivity you can obtain in applying a layer of reflective aerosol above the underlying surface - and the different heat capacities of the land and ocean.

At least in HadCM3, there are notable exceptional regions in the tropics that are not overcooled relative to higher latitudes, including the Amazon and Congo basin regions. It is likely that in these regions, feedbacks including changes to evapotranspiration under the higher CO_2 concentration of the SRM scenario, which retains quadrupled preindustrial CO_2 concentrations, dominate surface temperature rather than the direct radiative balance, as found by previous studies. [62]

Many factors influence the structure of the temperature response to uniform SRM. The oceans are darker than the land, so the change to the stratosphere's albedo has a greater influence over the amount of radiation absorbed by the oceans relative to the land. The heat capacity and hence timescale for response to forcing is different over

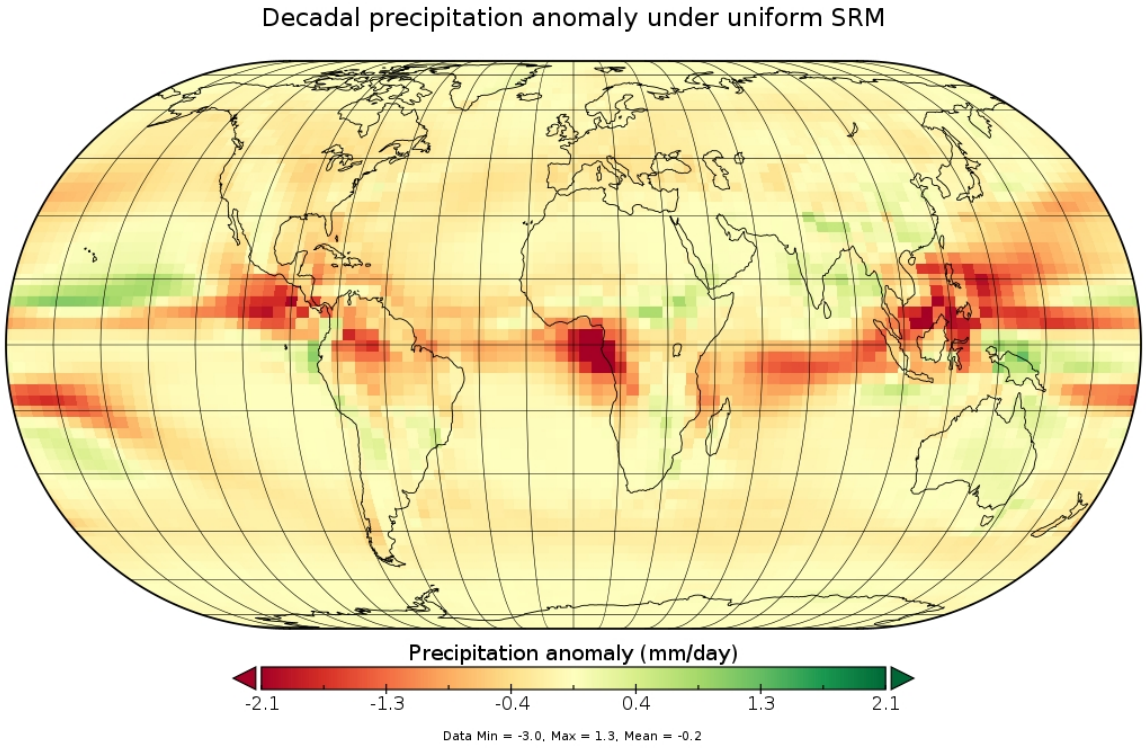

Figure 4.2: **The decadal average precipitation anomaly, in mm/day, under uniform SRM.**

the oceans and the land. At high latitudes, Arctic amplification of residual warming alongside the fact that the solar radiation management is ineffective during the polar night leads to residual warming of polar regions in their respective winters.

The precipitation response to uniform SRM is plotted in Figure 4.2. Note there is now substantial reduction of precipitation in the tropics and over the tropical oceans, but no substantial dipole shift as in the $4xCO_2$ case as discussed in Chapter 3. There is residual drying over Amazonia and the Congo basin, but smaller than that under $4xCO_2$, while the strong increases in precipitation over Central Africa, China, and India are all modulated. The increases in precipitation at high latitudes are overcompensated, resulting in drying. We can see that SRM designed to cancel global mean temperature anomalies still results in substantial regional precipitation anomalies

The question we seek to answer is the extent to which these residual temperature anomalies could be corrected by a latitudinally varying, but zonally uniform, application of SRM.

4.2.2 The use of basis functions from the Geoengineering Large Ensemble experiment

In order to implement latitudinally-varying SRM in a meaningful way, we adopt a semi-idealised approach. Given that HadCM3 does not have its own stratospheric model with stratospheric dynamics and chemistry, and therefore cannot determine the concentration of AOD in different locations from simulated injections of aerosol precursor, we instead prescribe aerosol optical depths to implement SRM.

We could have adopted the approach of ignoring any physical constraints on the aerosol optical depth distribution that can be obtained with a realistic pattern of injections or in a realistic stratosphere, and instead attempted to calculate a "perfect" AOD distribution by varying each of the 12 x 71 degrees of freedom available in the monthly zonal mean AODs that are prescribed. However, determining the model's response to each of these degrees of freedom would require an even larger amount of computational resources than the CPDN project can provide, and would also presume an unrealistic degree of control over the AOD distributions that can be obtained.

Instead, we opted to reduce the number of degrees of freedom in controlling the distribution, and to adopt more realistic aerosol optical depth patterns, derived from a stratospheric model with interactive chemistry and circulation, as "basis functions" for constructing partially idealised distributions of AOD. The (12 x 71) degrees of freedom are then reduced to a much more manageable number - and they now determine the coefficients for these basis functions, representing different magnitudes of aerosol precursor injected at each of the chosen injection latitudes. This gives rise to a relatively smooth set of aerosol distributions which are used as basis functions.

The basis functions for aerosol optical depth were taken from the Geoengineering Large Ensemble (GLENS) experiment [95], which simulated the effects of injecting SO2 daily at 15N, 15S, 30N, 30S, 50N, and 50S in the CESM-WACCM6 stratospheric model. [158] We determined an average aerosol optical depth pattern, or basis function, for injections at each of these injection sites. This involved taking the ensemble mean optical depth pattern (n=20 for the CESM-WACCM ensemble), and then averaging over a decade of the simulation to determine the average annual pattern of aerosol optical depth that results from these injections (once the distribution has

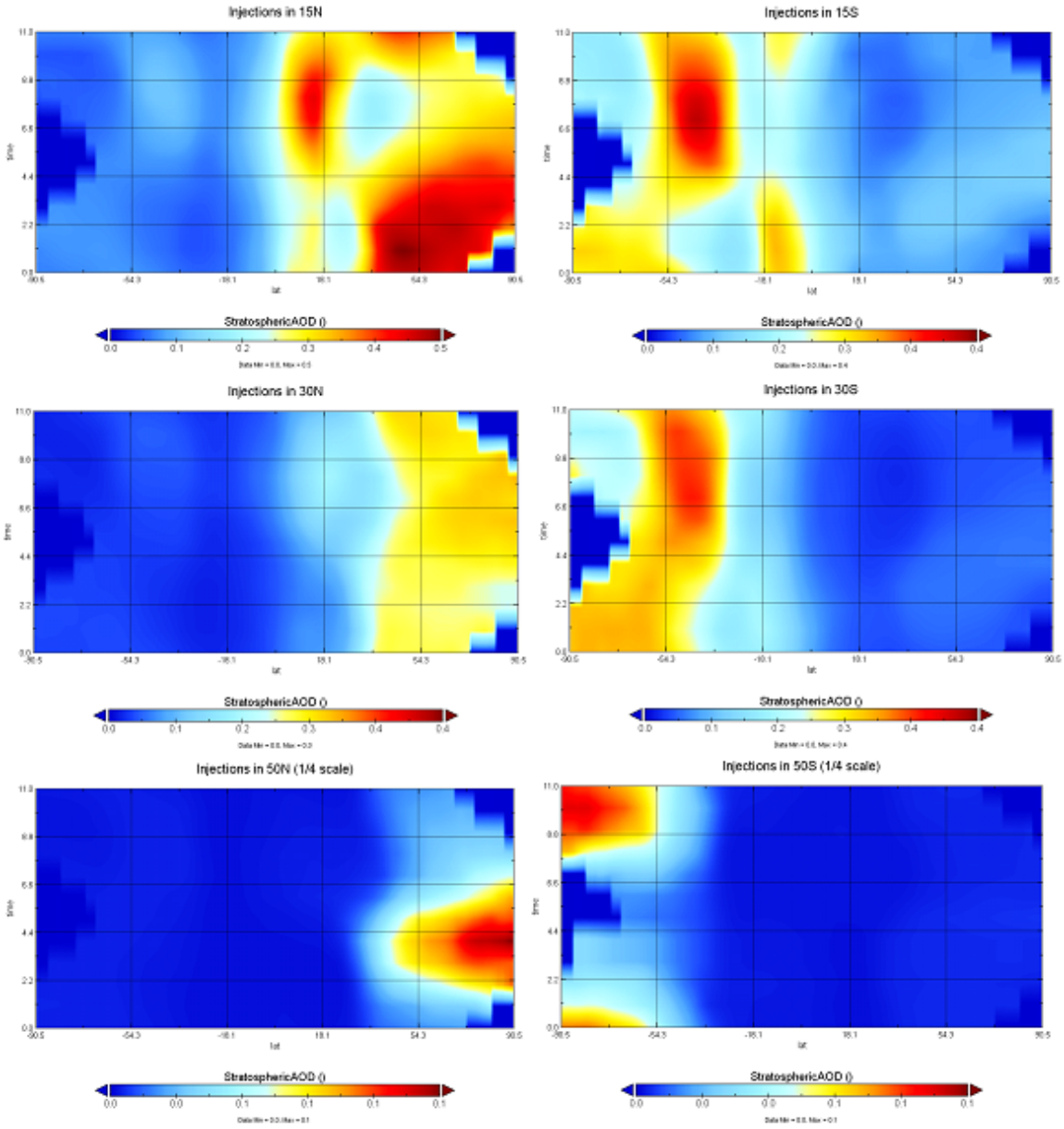

Figure 4.3: **Latitude-time plots for the injection basis functions of aerosol in GLENS.** Reading order: 15N, 15S, 30N, 30S, 50N, 50S. This illustrates the spatio-temporal distributions of aerosol that we impose as our "basis functions" for optimisation.

reached a relatively stable state). These annual patterns of aerosol optical depth are plotted in Figure 4.3.

These patterns were regridded into the 12 x 73 grid for HadCM3, and then scaled up so that they had the same global mean aerosol optical depth as was required to cancel out 4xCO_2 forcing when applied uniformly. We then simulated the effects of

each "basis function" against a background of quadrupled CO_2 in an ensemble of HadCM3 simulations to determine the basis function responses.

Taking the ensemble, decadal, zonal averages of surface temperature for these basis function responses gives us a set of patterns for zonal mean temperatures which characterise, to first order, the temperature response to each of the different distributions.

Finally, we assume that the responses are linear and additive. In other words, the response to a linear combination of the 15N and 15S aerosol optical depth distributions will be that same linear combination of the responses to the individual distributions. The GLENS experiment had its own linearity test, injecting both at 15N and 15S, and we also simulate the distribution that results from these combined injections as a further test of linearity. Results in Chapter 3 illustrated that, even for extremely inhomogenous "stripe" simulations, linearity holds surprisingly well. The validity of this linearity assumption is explored further in the Appendix, where we illustrate the percentage errors from assuming linearity. Figure 4.5 illustrates that the linear approximation holds well for decadal mean temperature.

With this assumption, we are now in a position to conduct a limited optimisation by finding the coefficients for our set of basis functions which is predicted to minimise a given figure of merit, assuming linearity. We can then test both the assumption of linearity and the efficacy of the optimization by simulating a set of distributions constructed in this way.

4.2.3 Constructing optimal distributions from basis vectors

The first optimisation that we consider attempts to minimize the predicted zonal, decadal temperature anomaly under SRM, by cancelling out as much of the $4xCO_2$ zonal, decadal temperature anomaly as possible. The linear optimisation method used was the L-BFGS-B algorithm, which allows for bounds to be specified on the parameters that are being minimised. [159] We selected bounds for the coefficients to be between 0 and 10, although in practice coefficients rarely exceed 1. Initial seed values for the coefficients were chosen as random floats between 0 and 1. To avoid any potential for the optimisation algorithm accidentally settling on local minima, and thus giving inconsistent results, we repeat the optimisation 100 times for

different randomised initial seeds, and then select the set of coefficients with the minimum predicted RMS difference between the targeted temperature distribution and the predicted temperature distribution. This "optimisation" procedure, with some small adaptations, was used throughout the rest of this chapter to construct optimal distributions for a variety of different figures of merit on the basis of optimising the coefficients and using 100 random initial conditions to avoid settling on local minima; hereafter we use optimisation as a shorthand for this procedure of determining coefficients for our basis vectors.

The results of this optimisation attempt for zonal temperature are illustrated in the distribution in Figure 4.4. The optimiser predominantly favours injections in 15S and 30N, with smaller contributions from the 50S, 15N, and 50N injections respectively. It favours greater aerosol optical depth at higher latitudes, counteracting the overcooling of the tropics and undercooling of the poles that is seen in uniform SRM. Within the available parameter space for optimisation, the distribution is relatively hemispherically symmetric, with perhaps a slightly higher peak of AOD in the Northern hemisphere.

The second subplot in Figure 4.4 illustrates how this optimised distribution varies across the seasons; note that as the AOD is not read in by the model in the top and bottom latitude bands, we can omit them from these graphs.

4.2.4 Simulations and ensemble sizes

The table below lists the ensemble sizes for completed simulations that are analysed in this chapter. Note that these ensemble sizes are larger than the ensemble sizes which were considered in Chapter 3, as more model runs were available when this analysis was conducted. The preindustrial, 4xCO_2, and uniform SRM simulations are as described earlier in this chapter, as well as Chapter 3. The standard length of runs, spin-ups used, years analysed, and process for analysis are as described in Chapter 3.

The new simulations considered in this chapter are the GLENS basis vectors, as described earlier in this methods section, and simulations that use these basis vectors, plus the optimisation process described above with an assumption of linearity, to optimise for zonal, decadal temperature (referred to as GLENS LAtiudinal Temperature

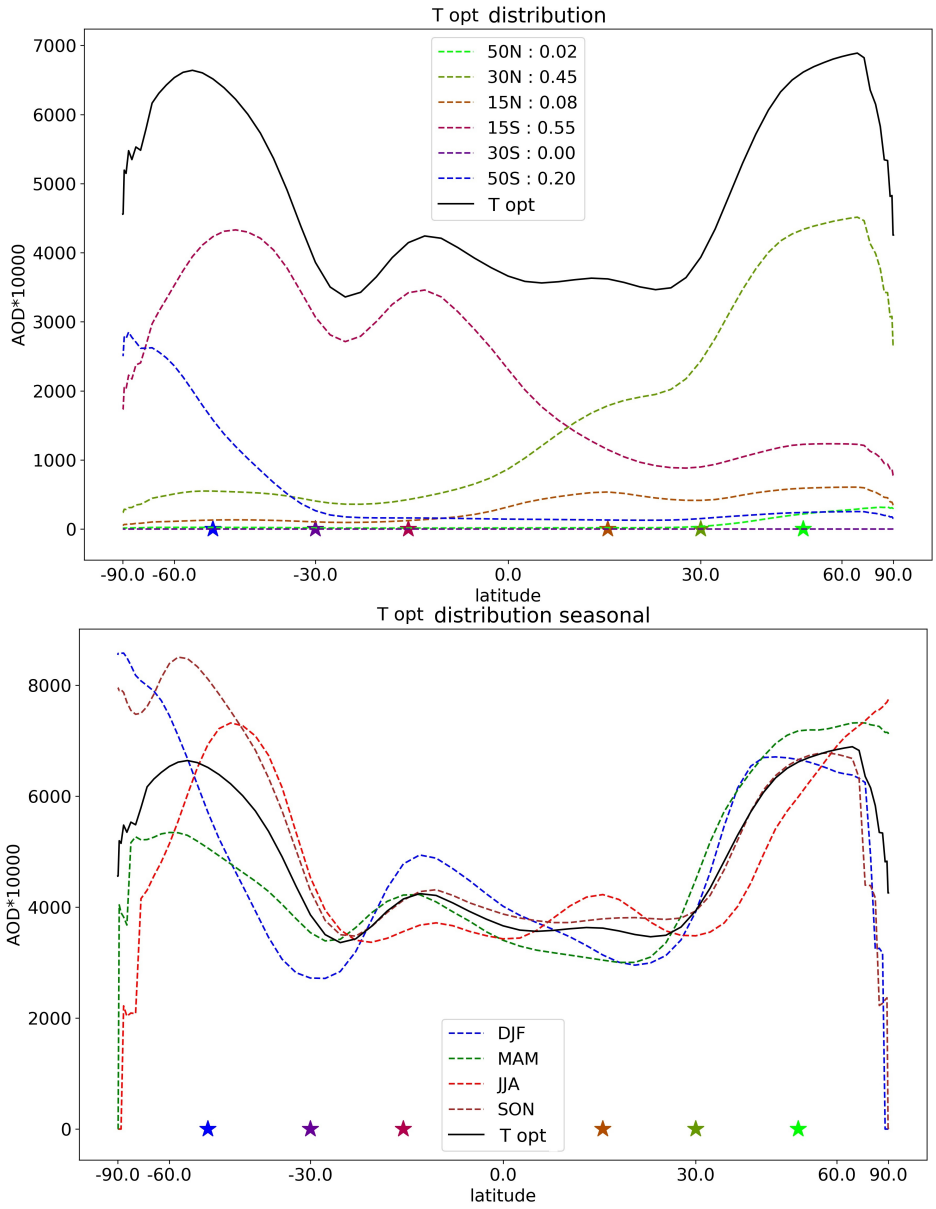

Figure 4.4: **Optimal AOD distribution for minimising zonal temperature anomalies (T Opt or GLAT distribution).** Top: Annual average; bottom: seasonal breakdown. The aerosol optical depth distribution that is predicted to minimise zonal, decadal temperature anomalies, assuming linearity and additivity in zonal temperature anomalies. The distribution is plotted against the sine of latitude, with the latitudes labelled. The relative scaling factors for each basis component in this distribution are displayed in the legend, with the contributions from each AOD pattern illustrated with dashed lines.

(GLAT) or T Opt) and precipitation (PROPT). These simulations were explicitly run on CPDN - other analysis for optimisation, such as optimising for temperatures on land, or optimising for individual Giorgi regions, was conducted on the basis of

assuming linearity and using pattern-scaling to calculate an estimate for the result, rather than explicitly simulating these runs. This allowed us to quickly calculate likely distributions and coefficients for each individual Giorgi region.

Simulation	**Ensemble size**
Preindustrial	56
4xCO_2	44
Uniform SRM	32
15N	26
15S	27
30N	21
30S	24
50N	30
50S	28
GLAT or T Opt	12
PROPT	14

Table 4.1: Simulations and ensemble sizes considered in Chapter 4

4.3 Results

4.3.1 Optimisation for zonal temperature

Figure 4.5 illustrates the effectiveness of the optimisation for zonal temperature, as well as the efficacy of the linear approximation and the contribution to the temperature anomalies from each of the underlying basis vectors. We can see that across most regions, zonal temperature is close to the prediction from linearity - linearity seems to slightly underesimate cooling across the Southern hemisphere, while it slightly overestimates cooling over the Northern hemisphere tropics. The average RMS difference between the predicted and actual zonal mean temperature in any given latitude band is approximately 0.1K, which is a similar magnitude to the average difference between the optimisation result and the target, suggesting that this approaches the limits of what this type of approximate optimisation can achieve.

Figures 4.32 and 4.33 in the appendix show both the absolute and percentage differences between the multimodel ensemble mean and the prediction assuming linearity in the linear optimisation for temperature.

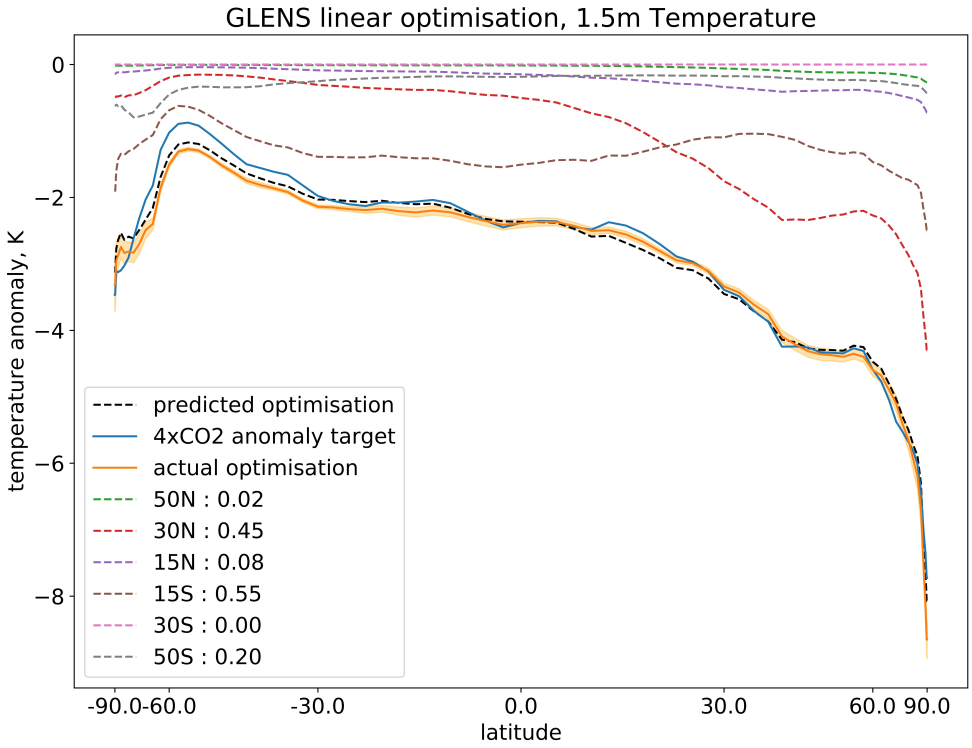

Figure 4.5: **Temperature response to optimal distribution and test of linearity.** Decadal, zonal, ensemble averages for 1.5m temperature anomalies, relative to the 4xCO_2 baseline. The blue line denotes the difference between the preindustrial and the 4xCO_2 case, which we target. The dashed black line is the predicted effect of the linear optimisation, assuming temperature anomalies are linear and additive; the individual anomalies are also plotted and the legend includes the relative components of each injection site. The orange line is the result of an ensemble average over simulations of the optimised distribution, while shading denotes a 67% confidence interval across the ensemble. Temperatures are plotted against the sine of latitude, with latitudes labelled.

The figures in 4.6 illustrate the seasonal effects of the linear, zonal temperature optimisation. This illustrates that the optimisation generally performs worse on a seasonal basis compared to in the annual average case due to compensating errors. The situation is worst at the poles, where the polar night - and hence the lack of ability to reflect sunlight and cool - is important. For example, for the South Pole in summer, and for the North Pole in autumn. We see small contributions to the overcooling of the Southern hemisphere in each of the seasons, while high Northern latitudes benefit from compensating errors of undercooling in the summer and overcooling in autumn.

Linearity holds reasonably well in each of the seasons considered separately, although it is not as accurate an approximation as it is in the annual average. The comparitive magnitudes of the anomalies can be seen in Figure 4.7 which plots the four seasonal anomalies on the same axes for the purposes of comparison.

In Figure 4.8, we compare the zonal temperature anomalies under the zonal temperature optimisation to the uniform case, and to the 4xCO_2 baseline. It is clear that the zonal optimisation reduces the residual temperature anomalies from the uniform case on average, especially in the Northern hemisphere, at the expense of undercooling in the Southern hemisphere. This in part owes to the fact that injection patterns which result in AOD predominantly in the Southern hemisphere still have a significant cooling impact on the Northern hemisphere, as illustrated in Figure 4.8's individual contributions from each of the basis patterns. There are also significant deviations at the poles where, owing to HadCM3's radiation scheme, it is more difficult to implement SRM to compensate for temperature anomalies. In general, we can see that the linear optimisation for zonal temperatures slightly improves the situation globally, and for most regions, but is imperfect. In both cases the residual temperature anomalies are far smaller than the unmitigated 4xCO_2 case.

Figure 4.8 illustrates the decadal average temperature anomaly under this first optimisation, from an n=12 ensemble of HadCM3 model runs. We see that the linearly-optimised zonal distribution slightly improves the decadal average temperatures, compensating for the undercooling at high latitudes, but both residual anomalies are small compared to the unmitigated 4xCO_2 case. The full map of these temperature anomalies is plotted in the upper panel of Figure 4.9. Here, we see the inherent deficiencies in the approach of optimising according to zonal mean temperatures, with only zonally uniform degrees of freedom in the applied aerosol optical depth. It is possible to achieve reduced zonal temperature anomalies, but at the expense of overcooling the oceans and undercooling the land. The temperature anomaly over high Northern latitudes, especially over the Arctic ocean, is substantially reduced compared to the uniform SRM case illustrated in Figure 4.1. However, the zonal optimisation does little to address the regions of land where warm anomalies exist, such as the Congo and Amazon basins, as well as the contrasting effects on temperature between India and China, and the Eastern versus Western United States regions. We see that for most

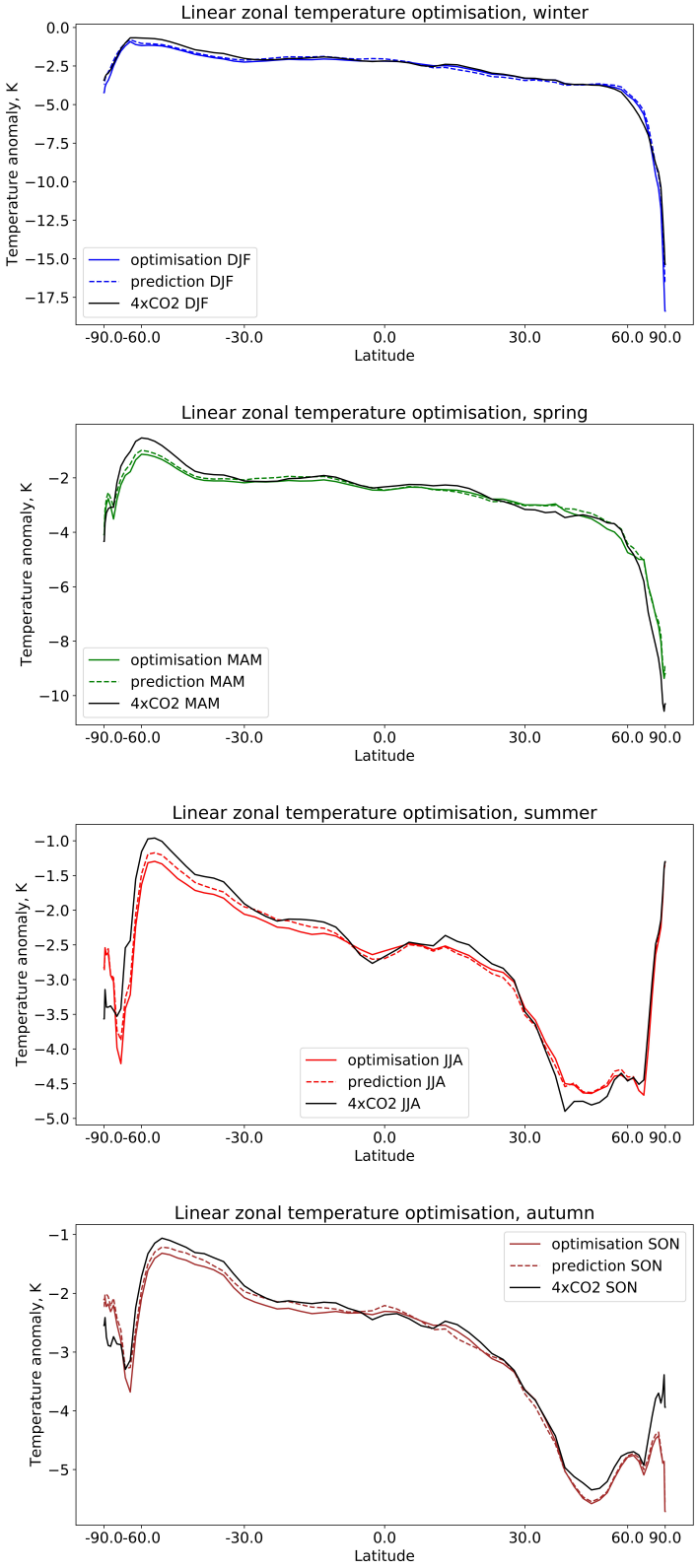

Figure 4.6: **Seasonal temperature responses to the zonal temperature optimisation distribution.** Dashed lines show the prediction from the linearity assumption; solid lines show the temperature response; black lines show the 'target' 4xCO_2 anomalies we aim to cancel. Top to bottom: DJF, MAM, JJA, SON.

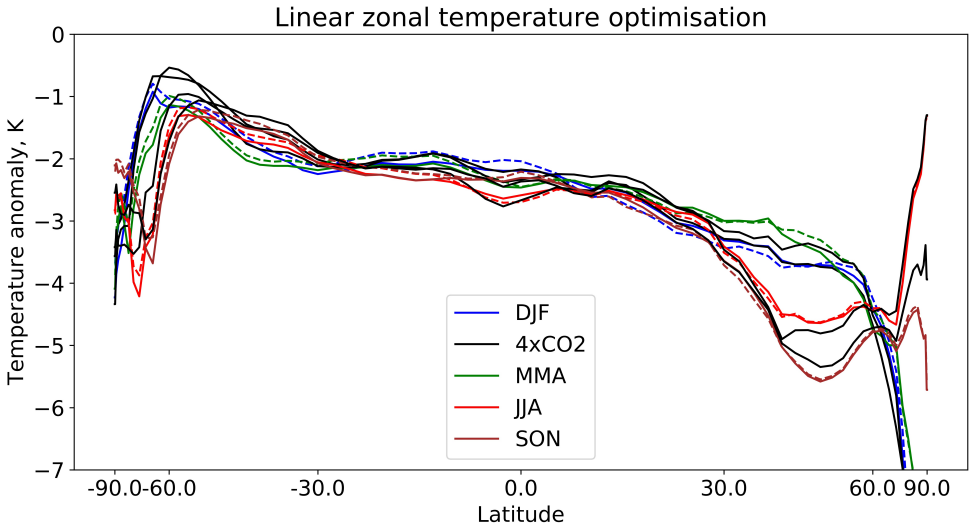

Figure 4.7: **Seasonal temperature responses to the zonal temperature optimisation distribution.** Dashed lines show the prediction from the linearity assumption; solid lines show the temperature response; black lines show the 'target' $4xCO_2$ anomalies we aim to cancel.

Giorgi regions over land, optimising for zonal temperatures does not substantially address these regional effects; significant regional temperature anomalies remain.

In order to determine the extent to which these anomalies can be reduced, we re-ran the optimisation with an alternative figure of merit that the algorithm must minimise: the root mean square (RMS) temperature anomaly over land grid cells in the decadal mean. This is a reasonable figure of merit on the assumption that geoengineers might be more concerned with minimising temperature anomalies over land and populated regions, as opposed to over the oceans, which would probably benefit from overcooling due to the significant heat they have absorbed already. Comparisons between the resulting coefficients selected, as well as the distributions used for each optimisation, can be seen in Figure 4.12 and Table 4.2. The distribution which minimises RMS land temperature in grid cells closely resembles the distribution which minimises zonal-mean temperature, but with an increased magnitude of AOD, particularly at high latitudes. Figure 4.9 illustrates the residual temperature pattern as predicted by the linear superposition. Attempting to minimise the land temperature anomalies results in overcooling the Earth as a whole - the global mean temperature anomaly is -0.3K compared to -0.1K in the zonal mean temperature anomaly

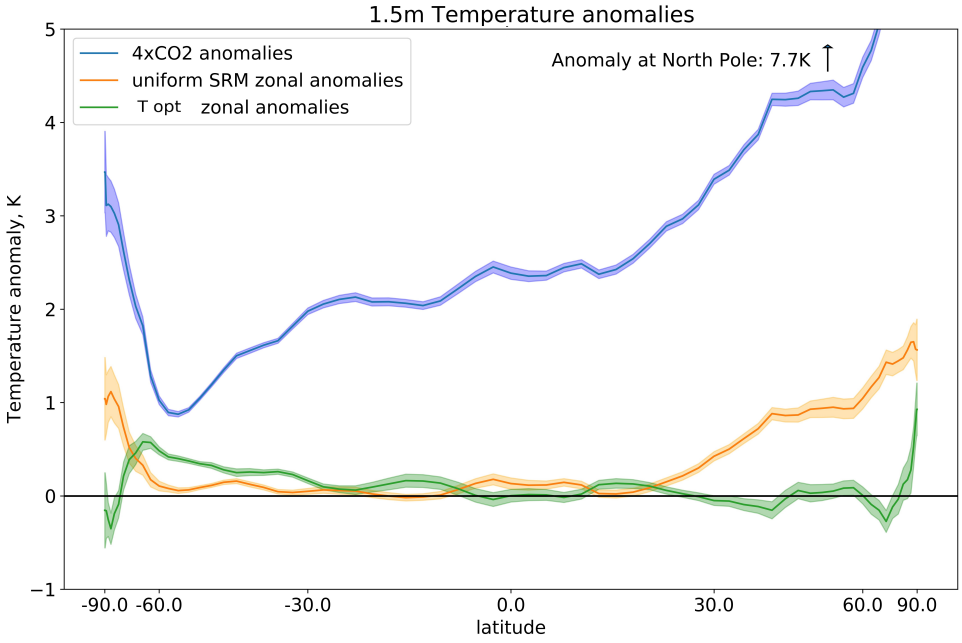

Figure 4.8: **Temperature response to uniform and optimised SRM and $4xCO_2$.** Decadal, zonal, ensemble averages for 1.5m temperature anomalies, relative to a preindustrial baseline. The blue line denotes the $4xCO_2$ case; the orange indicates the anomalies under the uniform SRM case, with the characteristic undercooling of the poles and slight overcooling in tropical regions. The green line plots the linearly-optimised distribution intended to minimise zonal average temperature anomalies. Shading denotes a 67% confidence interval across the ensemble. Temperatures are plotted against the sine of latitude, with latitudes labelled.

minimising case.

It's clear that the same pattern of temperature anomalies over land is still broadly present in this case, with temperature anomalies remaining high over rainforest regions, like the Amazon and Congo basins, as well as the Eastern United States, China, and Siberia. For each of these regions, there are other regions - the Western US, the Middle East and North Africa, India and Australasia - which appear to be more sensitive to cooling under SRM.

Densely forested regions have a substantial drying and heating feedback under $4xCO_2$ in HadCM3. This is in part directly due to the increase in CO_2 concentrations, which result in a modelled stomatal closing and therefore decreased transpiration from the foliage. As these high concentrations persist under the SRM scenario, it

may help explain why these regions experience higher temperatures than the rest of the tropics under SRM. This Amazon dieback feedback effect has been documented to be particularly and probably unrealistically severe in HadCM3 [160] which may cast some doubt on the magnitude of the simulated warming over this region in this model, but even in more recent Hadley Centre models and in scenarios with lower levels of anthropogenic radiative forcing from SRM and CO_2, we see persistent drying and warming in the Amazon which uniform SRM to limit global mean temperature rise struggles to ameliorate. [62]

As a result of the different levels of sensitivity to SRM in these different regions and the resulting trade-offs, the solution that minimises the RMS temperature undercools some regions and overcools others. It can reduce, but not eliminate, the warm and cool anomalies that persist over land in the zonally-optimised case. For reference, the average land-area weighted decadal anomaly in a grid cell for the 4xCO_2 case is 4.47K; in the zonally-optimised case, it is -0.3K; in the RMS optimised case, it is -0.07K: the residual error can be cut by two-thirds, but is small compared to the 4xCO2 case. The norm of the anomaly vector (i.e. the parameter that the optimiser seeks to minimise) reduces from 36.1 in zonally-optimised case to 31.5 in the RMS optimised case, showing that the ability of the optimisation process to reduce the average anomaly magnitude in a given grid cell.

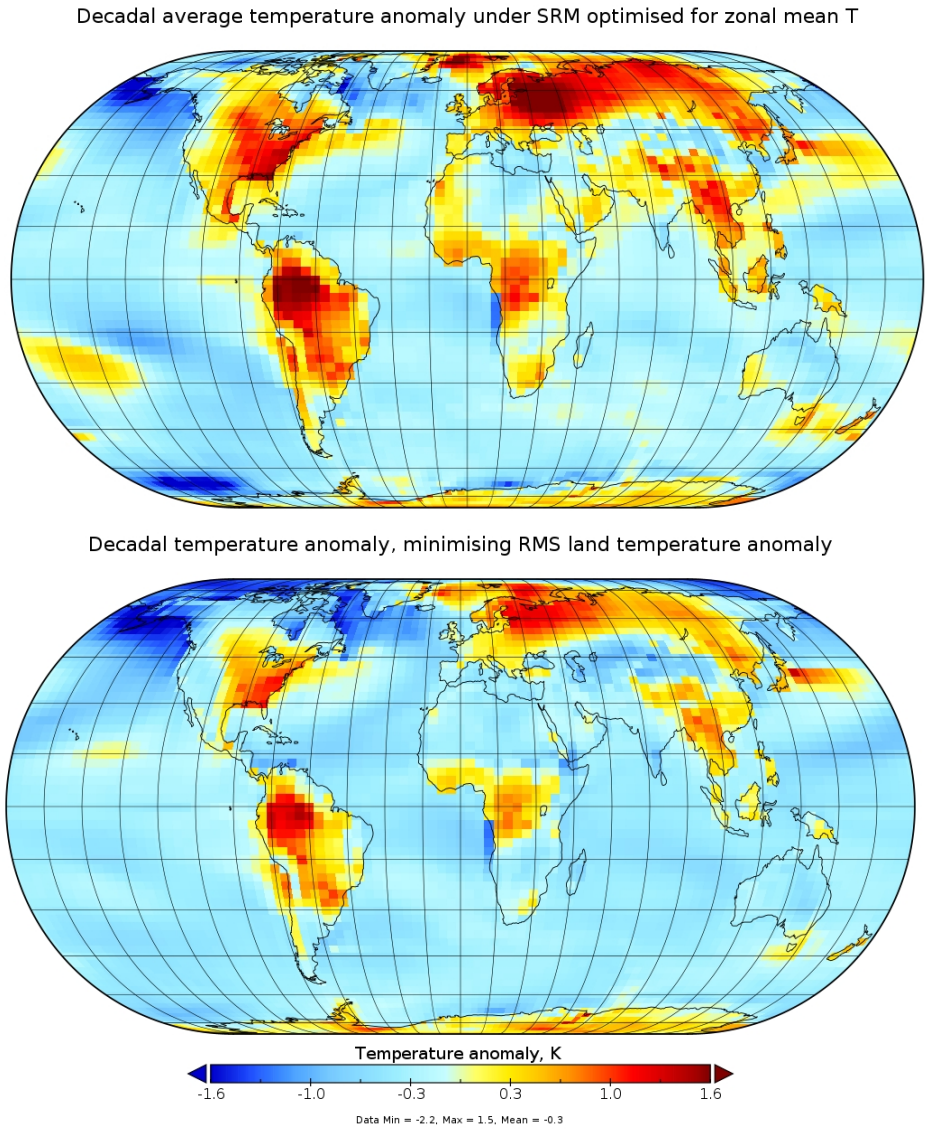

Figure 4.9: **Top: Residual temperature anomaly under the zonal temperature optimisation scenario. Bottom: Residual temperature anomalies under the distribution that minimises RMS temperature anomalies.** The maps illustrate decadal average temperature anomalies relative to the preindustrial.

4.3.2 Optimising for precipitation

4.3.2.1 Effects of optimising for temperature on precipitation

Temperature is not the only variable of interest in a geoengineered climate: precipitation is also crucial, and generally of much more concern under SRM scenarios. This section will investigate two questions: . Firstly, whether optimising for zonal temperature has a significant effect in reducing the precipitation anomalies that we see under uniform SRM - for example, by reducing anomalous meridional heat and moisture fluxes, which might result from residual zonal temperature anomalies under uniform SRM, and then go on to impact the distribution of precipitation compared to the preindustrial case. Secondly, whether it would be possible to "optimise for precipitation" in the same way as we had optimised for temperature.

In Figure 4.10, we plot at the zonal, decadal, ensemble mean precipitation anomalies, relative to the preindustrial simulations, comparing the 4xCO_2 case, the uniform SRM case, and the zonal temperature optimisation case. The uncertainties illustrated show the 67% confidence interval across the ensembles (n=44, n=32, and n=12 respectively).

We can see from this that the main effect of the 4xCO_2 scenario is to shift the Intertropical Convergence Zone (ITCZ) and the associated band of precipitation to the North, which leads to a tropical dipole in precipitation anomalies. Secondarily, by heating high latitudes, where evaporation and precipitation is often energy-limited rather than moisture-limited, the 4xCO_2 simulation increases precipitation towards the poles above the preindustrial case.

By comparison, both implementations of SRM reduce precipitation in every latitude band to some degree. Despite the additional aerosol optical depth at high latitudes in the zonal-optimisation temperature scenario, the net effect is quite similar, with a slight additional decrease in precipitation towards the poles in the zonal-optimised temperature case compared to the uniform.

Uniform SRM and the zonal simulation both shift the centroid of the ITCZ's precipitation back towards its preindustrial position compared to its Northward migration under 4xCO_2. The location of the ITCZ is determined by cross-hemispheric energy flux; and correlated with to the latitude of zero net meridional energy flux. [161]

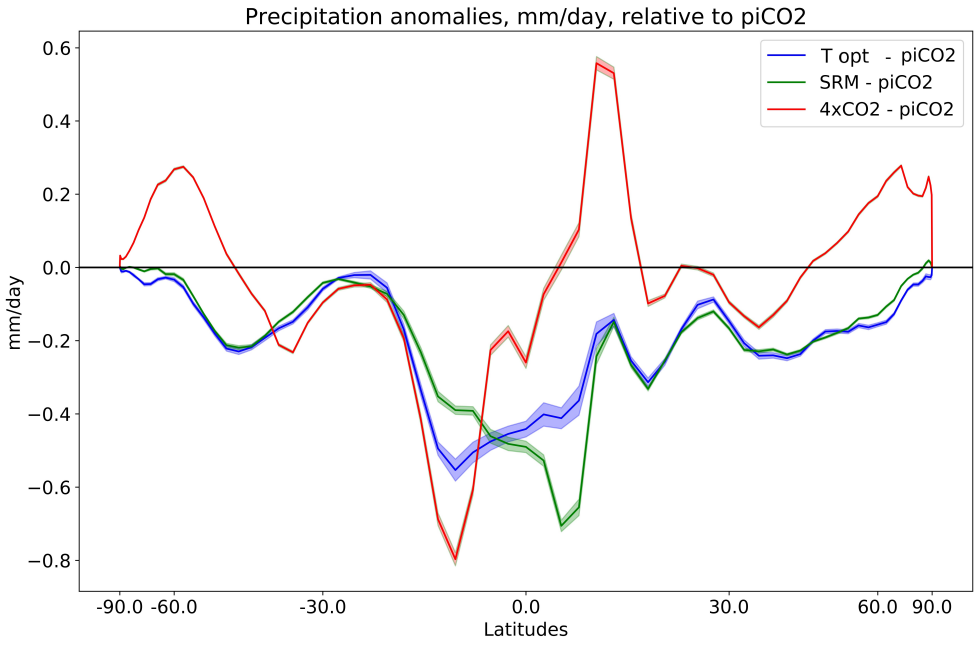

Figure 4.10: **Zonal, decadal precipitation anomalies relative to the preindustrial case in the 4xCO_2, zonal temperature optimisation, and uniform SRM cases are plotted.** Shading indicates a 67% confidence interval across the ensemble.

Consequently, 4xCO_2, which asymmetrically warms the Northern hemisphere, drags the ITCZ to the North: under either SRM case, there is substantially less hemispheric asymmetry in the residual warming, so the ITCZ is not substantially shifted from its predindustrial position; consequently there is no dipolar precipitation anomaly under the SRM simulations.

However, in both cases, there is a substantial decrease in rainfall over tropical latitudes, driven by reduced short-wave flux and evaporation at the surface, and increased atmospheric stability as the upper tropopshere warms due to the longwave absorption effect of CO_2 and stratospheric aerosol. This effect is more than enough to overcompensate for any increase in precipitation due to warming at high latitudes, resulting in a decrease over most regions. The biggest difference between the two SRM cases is in the tropics - the maximum drying occurs at around 12 °S in the zonal temperature optimisation case, compared to 6 °N in the uniform SRM case. This may be due to the slight asymmetry in tropical AOD under the zonal temperature

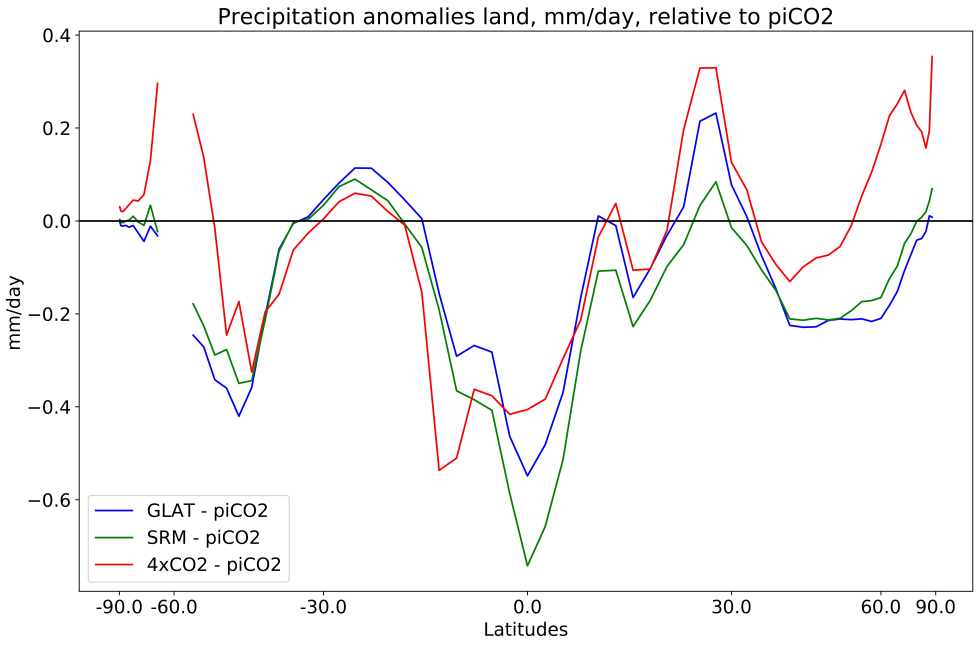

Figure 4.11: **Precipitation anomalies over land under zonal optimisation for temperature (GLAT), uniform SRM, and 4xCO_2.**

optimisation simulation (see Figure 4.12). It is clear from this that optimising for zonal mean temperature does little to optimise for precipitation, and that this is generally dominated by the overall tendency of SRM towards drying.

Figure 4.10 illustrates the zonal precipitation anomalies over land grid cells only; this removes the strong influence of the shifts to precipitation bands over the ocean. The pattern of anomalies is very similar over land across all three simulations. Both SRM simulations have drying at high latitudes, overcompensating for the increased precipitation in the 4xCO_2 simulation at these latitudes, and both modulate the increase in precipitation over 20-30 °N. It is arguable that the zonal-temperature optimised simulation slightly moderates the tropical precipitation anomalies of the uniform case. However, the difference between the SRM simulations is relatively small compared to the overall pattern of anomalies, and does little to modulate or erase the changes to zonal land precipitation in general under 4xCO_2: there are tradeoffs in each case. We can clearly see from this that optimising SRM for zonal temperature does not clearly also optimise for precipitation, and the lack of linearity in the response, as

well as the patterns of response in precipitation that are available to you (a general decrease), makes optimising for especially tropical precipitation difficult.

4.3.2.2 Attempting to construct a linear optimisation for precipitation

The assumption that temperature is linear and additive is based on the assumption that the predominant influence on surface temperature will be the direct radiation balance in a given region; consequently, cutting shortwave flux to the surface through applying AOD directly overhead is expected to shade regions directly below. Non-linearities in temperature may owe to second-order effects and feedbacks, such as the response of surface albedo due to snow or vegetation cover.

Precipitation is much more complex, and changes in precipitation owe to both dynamical and thermodynamical components. [162] For example, the overall amount of tropical precipitation, and the strength of associated overturning circulations, are likely to be reduced both by $4xCO_2$ and the application of SRM in any region which results in cooling. However, the position of the band of tropical precipitation (ITCZ) depends strongly on the global, rather than local, AOD distribution. A uniform distribution of AOD from 20 °N to the South Pole would shift the ITCZ to the North, while a uniform distribution from 20 °S to the North Pole would shift it to the South, resulting in very different distributions of precipitation in the region 20 °N to 20 °S, even though both distributions may have identical AOD locally.

For this reason, it is expected that the linear and additive approximation in precipitation response would be substantially less accurate, particularly for less smooth and symmetric distributions of AOD. However, to test the limits of our optimisation method and the linearity of the precipitation response, we constructed a zonal precipitation optimisation distribution using the same method described in the methods section, intended to minimise the anomaly relative to $4xCO_2$.

Figure 4.12 shows the comparitive ideal aerosol optical depth distributions for optimising with different figures of merit, assuming linearity. The coefficients of the respective GLENS aerosol optical depth basis vectors for the different distributions are listed in Table 4.2. Note that minimising the predicted precipitation anomaly involves substantially less aerosol overall, but the resulting distribution is also more asymmetric, with substantial additional AOD in the Northern hemisphere. This

strikes a balance between minimising the global drying effect, which is correlated with the overall magnitude of the aerosol optical depth for the SRM, while also shifting the ITCZ and its associated precipitation back towards the equator, counteracting the large Northward shift of precipitation expected under an unmitigated $4xCO_2$ scenario. We see that additional cooling over the tropics is required to minimise RMSE land temperatures, and the asymmetric distribution is chosen to minimise zonal precipitation anomalies by shifting the ITCZ back to the South, while reducing the overall optical depth to minimise the reduction in precipitation magnitude.

Optimisation	Zonal temperature	Land temperature RMSE	Zonal precipitation
15N	0.079	0.445	0.167
30N	0.454	0.051	0.144
50N	0.024	0.162	0.111
15S	0.549	0.441	0.069
30S	0.0	0.092	0.073
50S	0.200	0.167	0.080

Table 4.2: Coefficients of the GLENS basis vector distributions at different injection sites for the different optimisation figures of merit in the linear optimisation. These illustrate the relative contribution of each component to the distribution.

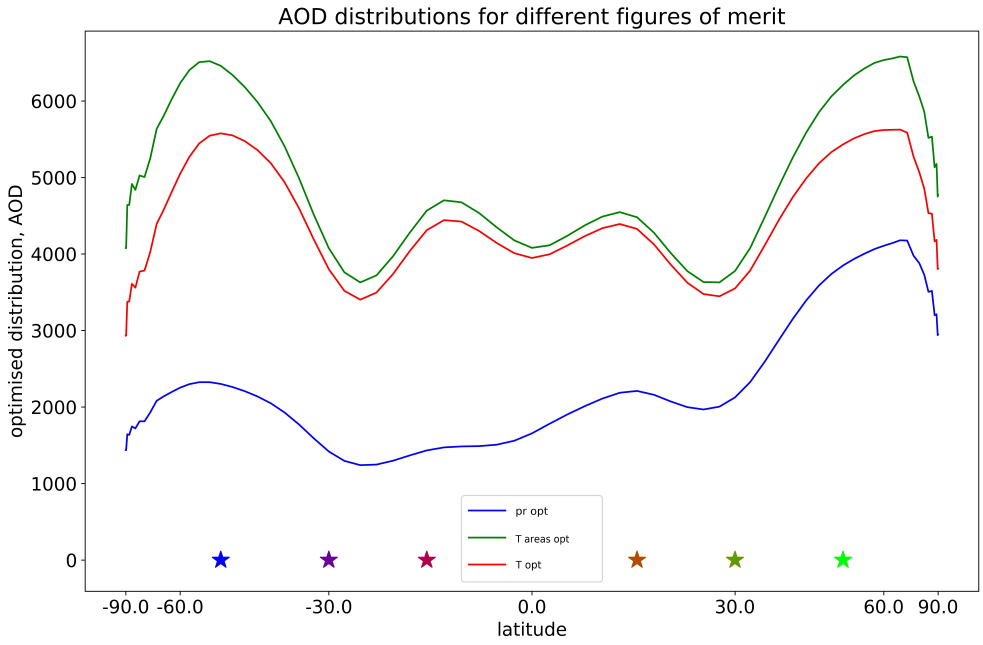

Figure 4.12: **Comparitive optimal AOD distributions for minimising precipitation, land-only RMS temperature, and zonal temperature anomalies.** In red, the zonal temperature optimisation distribution aims to minimise zonal and land RMSE temperature anomalies; in green, the area temperature optimisation distribution minimises RMSE temperature anomalies over land only, while the blue line shows the linear-optimiser's attempt to minimise the zonal average precipitation anomaly.

Figure 4.13 illustrates both the effectiveness of this precipitation optimisation and how well the linear and additive approximation holds. The precipitation anomalies under uniform SRM are reproduced for comparison. The uncertainty shown is the 67% confidence interval with one standard deviation across the ensemble (n = 10). The linear approximation performs somewhat well at predicting residual anomalies in precipitation, but much less well than in the temperature case. The linear approximation is outside the confidence interval for the actual simulation for around half of the latitude bands we consider.

As expected from the distribution, we can see that compared to the $4xCO_2$ case, the precipitation-optimised SRM reduces or overcompensate for the increase in precipitation at high latitudes, in the region where precipitation changes are dominated by thermodynamic effects. We can also see that the $4xCO_2$ dipole in precipitation is reduced. It is not clear that the pattern of tropical precipitation can be decom-

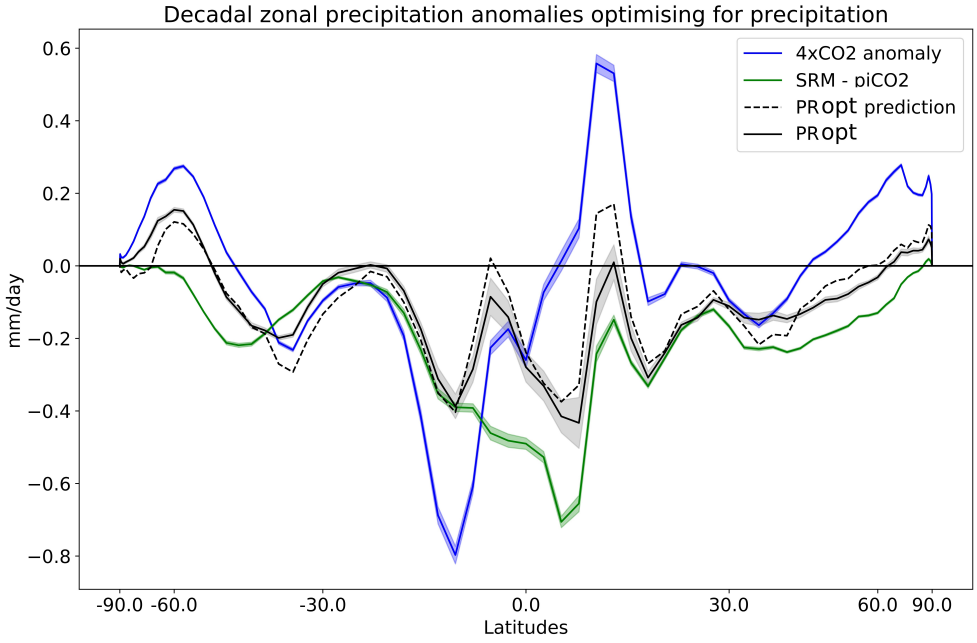

Figure 4.13: **Precipitation response and test of linearity in the zonal precipitation optimisation scenario (PR OPT).** The dashed line illustrates the predicted decadal, zonal precipitation anomalies relative to the preindustrial ensemble average under this optimisation; the solid black line shows the multi-run ensemble average simulation; the blue line shows the "target" anomaly - i.e. the $4xCO_2$ anomaly we are aiming to cancel out. The green line, for comparison, illustrates the precipitation anomalies under uniform SRM which cancels out global mean temperature.

posed into a reduced dipole shift alongside general drying, but there is drying to some degree across most of the tropics. It is clear that the "optimised for precipitation" distribution does perform better than the uniform AOD layer SRM which minimises the global mean temperature anomaly, across almost all latitude bands - generally because the overall reduction in precipitation is lower. This illustrates a trade-off between optimising for zonal mean temperature and zonal-mean precipitation; however, the effect is not that substantial. One metric for how much precipitation differs from preindustrial would be the length of the latitude-weighted anomaly vector, i.e. the weighted sum of squares of the deviation from the preindustrial case. This metric gives 1.66 for the $4xCO_2$ case, 1.77 for the uniform SRM case, and 1.17 for the precipitation optimised case - a reduction of around 30% in the RMS anomaly for the precipitation optimisation relative to $4xCO_2$.

Figures 4.14 and 4.15 illustrate the decadal precipitation anomalies for each of the $4xCO_2$, the uniform SRM, and the zonal precipitation optimisation cases.

In the $4xCO_2$ case (the upper plot in Figure 4.14), we see several features: drying over the Mediterranean, with increased precipitation at high latitudes, as well as over the Sahel, India, China, and particularly over the tropical Pacific. There are much more substantial decreases of precipitation over the Congo basin, the Amazonian region, and regions of the tropical Pacific to the West of the Congo and Mexico. The overall effect is a thermodynamically driven intensification of precipitation at high latitudes, and a shift to the North as well as an intensification of the precipitation associated with the ITCZ, with drying at midlatitudes.

Uniform SRM (the lower plot in Figure 4.14) substantially reduces the magnitudes of many of these precipitation anomalies - with reduced wetting across India, China, and the Sahel, and reduced drying over the Amazon and Congo basin. It also counteracts the high-latitude thermodynamically driven increase in precipitation in the Arctic, which is a decrease instead. In many regions, the pattern of anomalies under SRM is similar to that under $4xCO_2$, but reduced in magnitude. Under SRM, there is substantial reduction of precipitation in the tropics and over the tropical oceans, but no substantial dipole shift as in the $4xCO_2$ case. Residual drying over Amazonia and the Congo basin, strong increases in precipitation over Central Africa, China, and India are all modulated compared to 4xCO2. High-latitude precipitation increases are overcompensated, resulting in drying.

Comparing the SRM cases, we note that reducing zonal mean precipitation anomalies with the optimised scenario (Figure 4.15) - and the global mean precipitation anomaly - by no means reduces anomalies in general. In many latitude bands, especially over the oceans, this is achieved through compensating positive and negative anomalies in different regions along the same latitude band. At the poles and over the Southern Ocean, slight negative precipitation anomalies are replaced by slight positive precipitation anomalies - as we might expect given the reduced overall aerosol optical depth in the precipitation-optimised simulation, compared to the uniform case. Over land, it is difficult to discern any improvements in any regions, and there is exacerbated drying over the Amazon compared to the uniform case. As with the temperature optimisation, this illustrates that, even to the extent that zonal

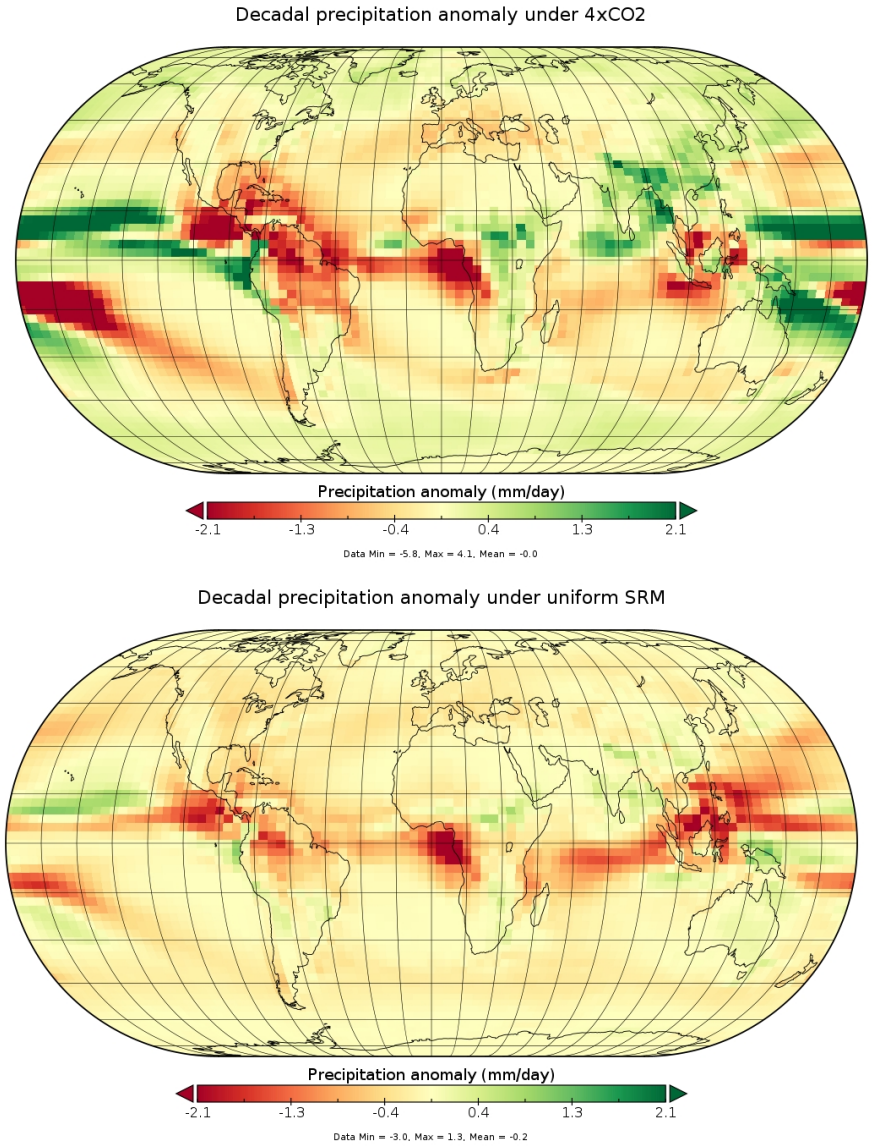

Figure 4.14: **Top: The decadal precipitation anomaly under 4xCO_2. Bottom: The decadal precipitation anomaly under uniform SRM.**

decadal precipitation can be finely-tuned and "optimised" with a tailored distribution of aerosol optical depth, it is unclear that this results in benefits within these latitude bands, and over key regions.

In this section, we have examined the potential of latitudinally optimising SRM to achieve a better cancellation of temperature and precipitation anomalies under 4xCO_2 with large ensembles of HadCM3 simulations. In general, we find that it is possible to achieve significantly better cancellation of zonal, decadal anomalies, particularly in

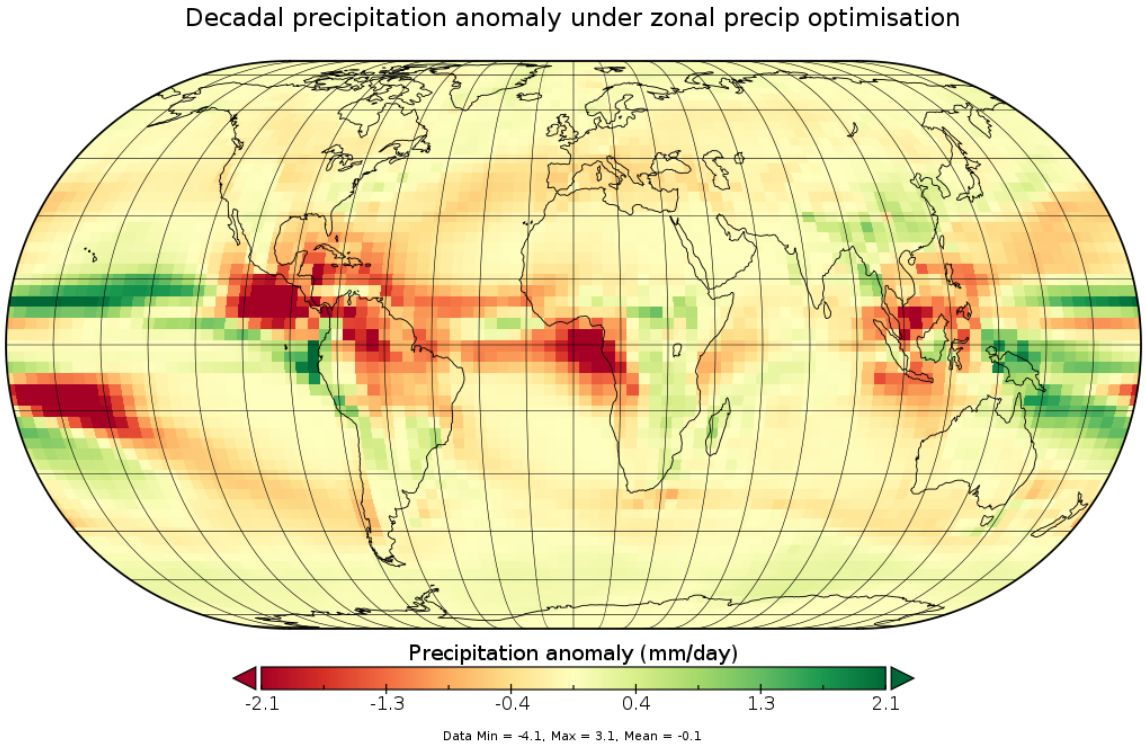

Figure 4.15: **The decadal precipitation anomaly under SRM linearly optimised to minimise the zonal precipitation anomaly.**

temperature, and to a lesser extent in precipitation, by imposing an AOD distribution that is a linear superposition over distributions obtained from injections at different locations in the stratosphere. Particularly in the case of temperature, the zonal temperature response can be well approximated by assuming the response is linear and additive; this approximation holds less well for precipitation owing to dynamical effects.

Despite this, it is unclear that zonally optimising SRM to reduce zonal anomalies in temperature and precipitation in this way will necessarily result in outcomes that society will find preferable. This more effective zonal optimisation is often achieved through compensatory errors within the latitude band. For temperature, this can arise due to contrasting effects of the aerosol optical depth at cooling over the oceans relative to the land. In the case of precipitation, much of the apparent improvement in precipitation appears to arise due to changes in the ITCZ over the Pacific Ocean.

4.3.3 Empirical orthogonal function analysis: Temperature

A further aim of investigating the impacts of a range of different distributions is to try to determine the extent to which different temperature responses can be excited by those different distributions. In other words, how many degrees of freedom for the response for temperature exist in the climate model? Can we obtain substantial latitudinal tailoring of the temperature by applying inhomogenous patterns of forcing, or is it instead the case that the atmospheric circulation responds well to inhomogenous forcings and projects the response onto several primary modes of variation? To explore these questions, we undertook an empirical orthogonal function (EOF) analysis for the temperature repsonse across a range of different distributions.

To explore this question, we choose a wide range of different decadal temperature response patterns to find the empirical orthogonal functions that span their space. This range of eighteen imposed distributions consists of:

- The seven highly inhomogenous "stripe" functions from earlier tests, i.e. seven thick bands of AOD across ten of HadCM3's adjacent latitude bands. (These were illustrated in Chapter 3.)
- Each of the six GLENS basis functions - year-round injections in 15N, 15S, 30N, 30S, 50N and 50S. (These were illustrated earlier in this Chapter.)
- The idealised alpha and beta functions - distributions of AOD proportional to $sin^2\varphi$ and $1 + sin^2\varphi$.
- The AOD distribution that results from equatorial injections in the MIROC-CHEM-AMP model.
- The AOD distribution that results from equatorial injections in the ULAQ model.
- The idealised AOD distribution that is prescribed to represent the Pinatubo eruption (G4 AOD distribution).

Each of these aerosol optical depth patterns produces a decadal temperature anomaly which we add to the collection for analysis. This then provides the basis set of functions which we look at to determine the EOFs that span the space.

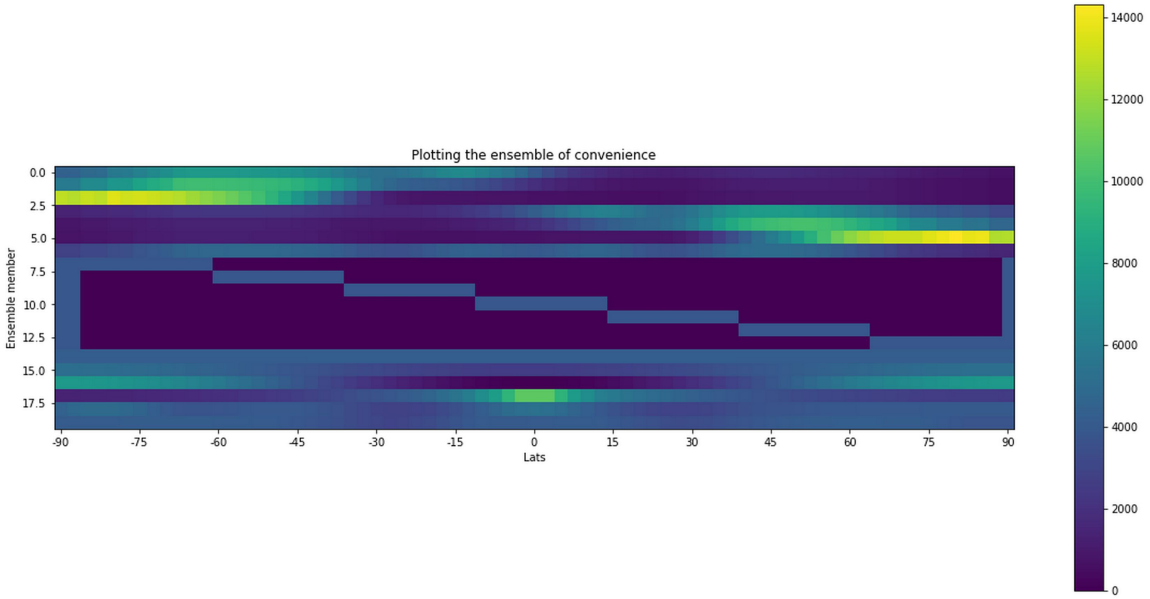

Figure 4.16: **Latitude plots for the spectrum of AOD distributions in the ensemble tested in HadCM3.** Ordered from top to bottom: the GLENS basis functions, uniform, the thick stripes, the alpha and beta distributions, the ULAQ, MIROC and G4 distributions.

The EOFs are illustrated in Figures 4.17 and 4.18. Across the first four EOFs, 98.7% of the total variance in temperature is explained. In the temperature analysis, the four principal EOFs have physically intuitive interpretations: the first (75%) represents the overall global warming function; the second (19%) represents a broadly hemispherically asymmetric warming pattern, with colder patches over Alaska, India, South Asia, and the Sahara desert. The third EOF, explaining 3.3% of the variance, is broadly a pole-to-equator temperature gradient warming pattern, with cooling across the continents and warming at high latitudes. The fourth EOF, explaining 1.7% of the variance, resembles the second (sinusoidal) harmonic in latitude.

This indicates that the vast majority of the variation in responses to a range of different forcings can be explained with a small number of empirical orthogonal functions with an clear physical interpretation. This suggests that, at least in the decadal temperature response, regardless of the complexity or nature of our distribution, we are fundamentally exciting a limited number of degrees of freedom in the climate model's response that project onto relatively simple patterns: global warming, hemispherical asymmetry, pole-to-equator temperature gradient and so on. These modes

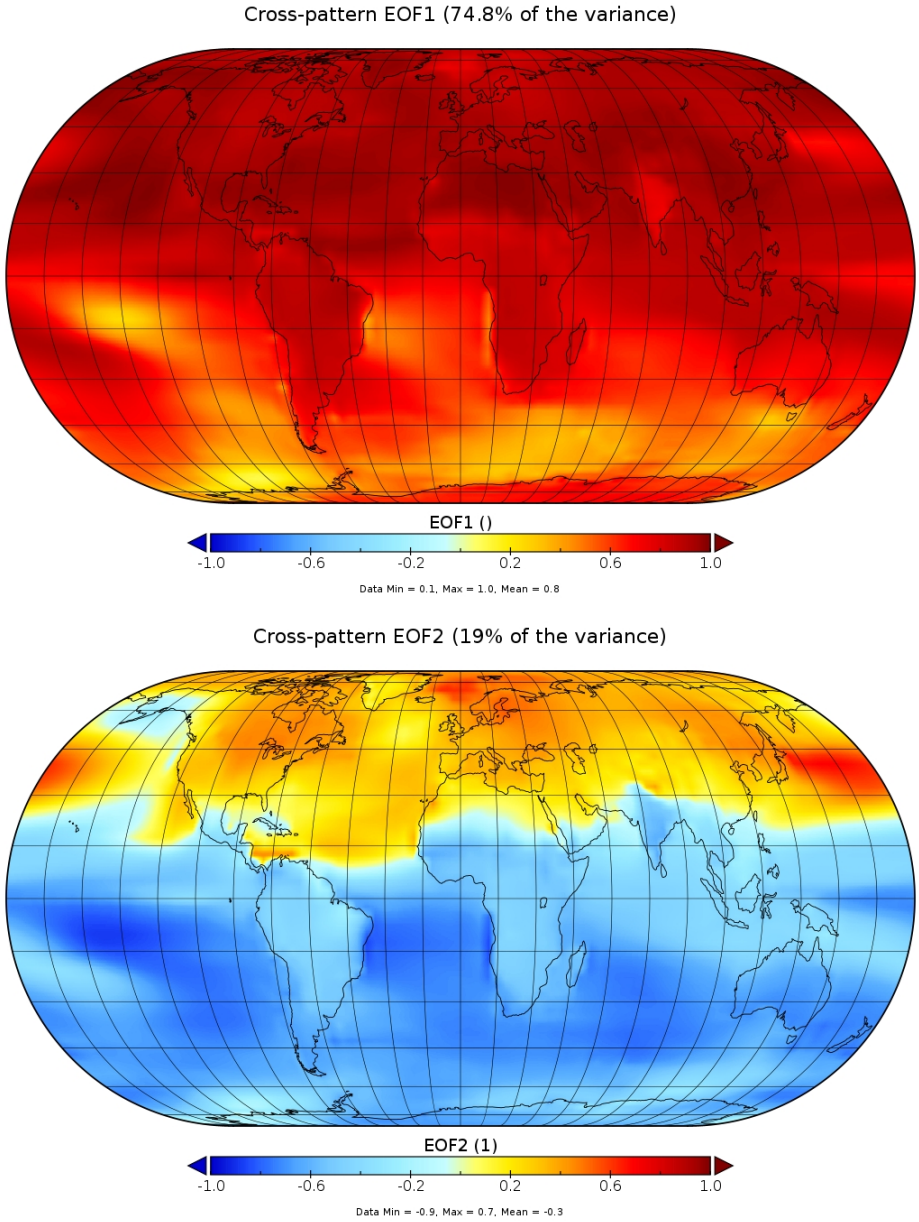

Figure 4.17: **The first and second EOFs for decadal temperature.**

correspond to modes of circulation in the atmosphere which adjust to redistribute heat around the world - the Hadley Cells transporting heat poleward and within a hemisphere, and the Arctic-amplified response we see to global warming as the fundamental pattern which explains most of the variation between the responses to different distributions.

We next want to determine to what extent we can predict the temperature response from the imposed AOD pattern through these projections on the EOFs. This

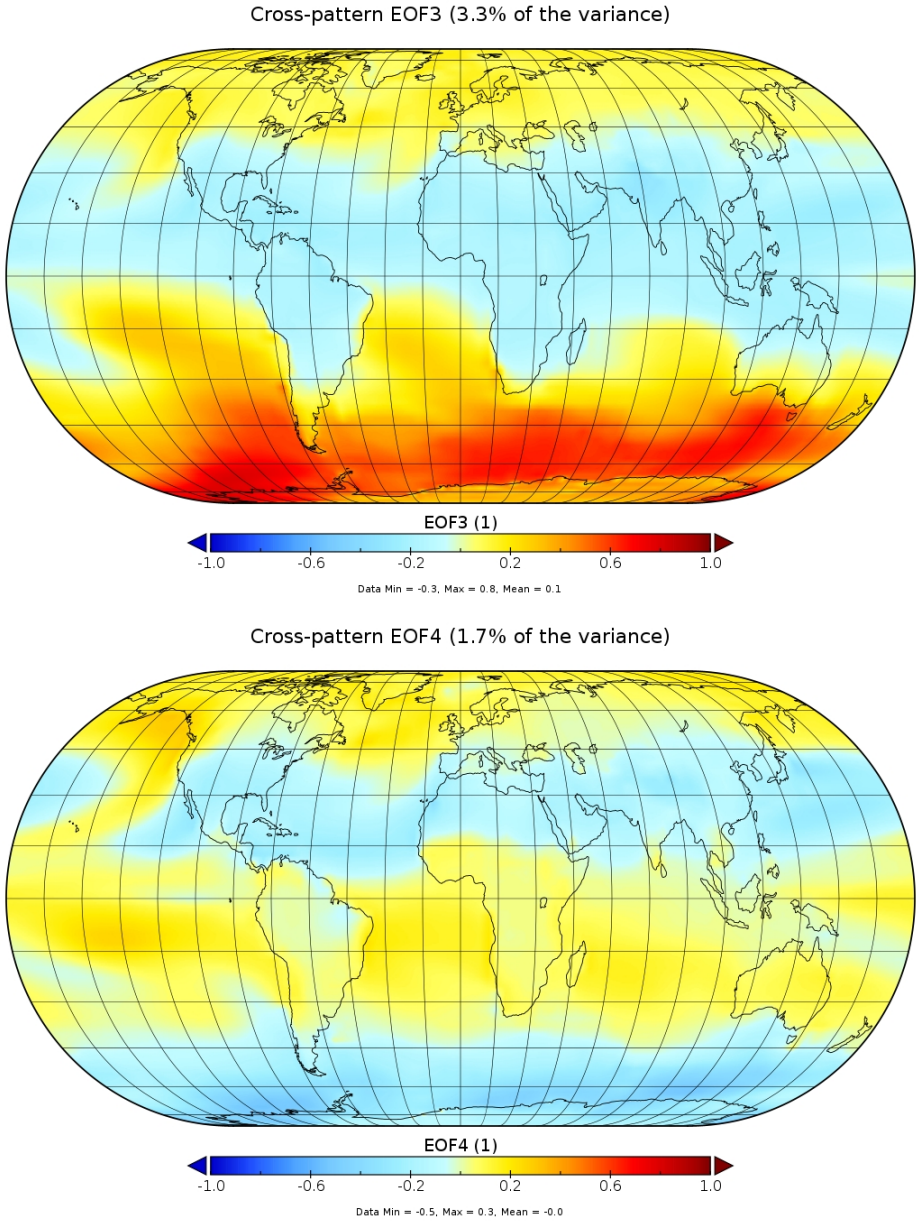

Figure 4.18: **The third and fourth EOFs for decadal temperature.**

is explored in Figure 4.19. In these figures, we determine the correlation between the principal component associated with the EOF and properties of the AOD distribution. Figure 4.19 shows a strong correlation between the net amount of latitude-weighted aerosol optical depth injected - i.e. roughly the net solar radiation reflected - and the principal component associated with the overall global warming pattern in the top panel. In other words, to first order, the total cooling effect scales as the total weighted amount of aerosol optical depth applied. Similar correlations are seen

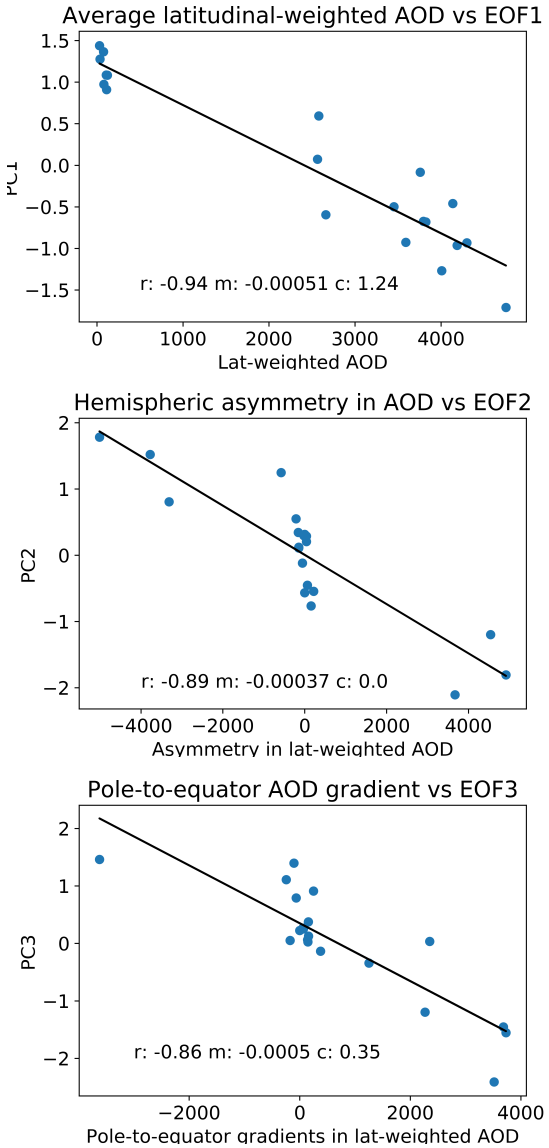

Figure 4.19: **Identifying EOFs of decadal temperature and their principal components with properties of the AOD distributions.**

in the subfigures where the hemispheric asymmetry of the aerosol distribution, and the pole-to-equator gradient of aerosol optical depth, are correlated with the same features in the response.

4.3.4 Precipitation: EOFs analysis

The analysis in the previous section, across the 18 distributions, was repeated for precipitation - but interpreting the EOFs in precipitation is more difficult. The first four EOFs again explain a high fraction of the variance - 95% - but this is slightly more spread out amongst the EOFs, with lower-order modes accounting for a higher fraction of the variance, as we can see in Figure 4.20. We can see that the precipitation pattern eigenvalues generally (up to the 3rd mode) decline more steeply than the eigenvalues that are associated with the AOD patterns; this suggests that a larger number of AOD patterns are being "compressed" by the response of the climate system to a much smaller number of precipitation pattern responses.

The EOFs are plotted in Figures 4.21 and 4.22. The first EOF arguably resembles a hemispheric asymmetry in precipitation - with, generally, increases in the Northern hemisphere and at high latitudes (with exceptions over midlatitude land) and decreases in precipitation in the Southern hemisphere, particularly just below the equator over the Pacific Ocean. It is possible that this mode of variation corresponds to the change in precipitation associated with global warming (generally increased precipitation at high latitudes and a Northward shift of the ITCZ, but a tendency for dry regions to get drier as we see in the Mediterranean) - the "wet gets wetter, dry gets drier" response. The first EOF, which appears to be predominantly a hemispherically asymmetric precipitation response, explains 61% of the variance. The second EOF, which predominantly seems to represent precipitation responses of a different sign in the tropics and at high latitudes - based on the overall reduction to the shortwave flux to the surface, and the moisture-limited/energy-limited regimes for precipitation in the high latitudes/tropics respectively, explains 26% of the variance.

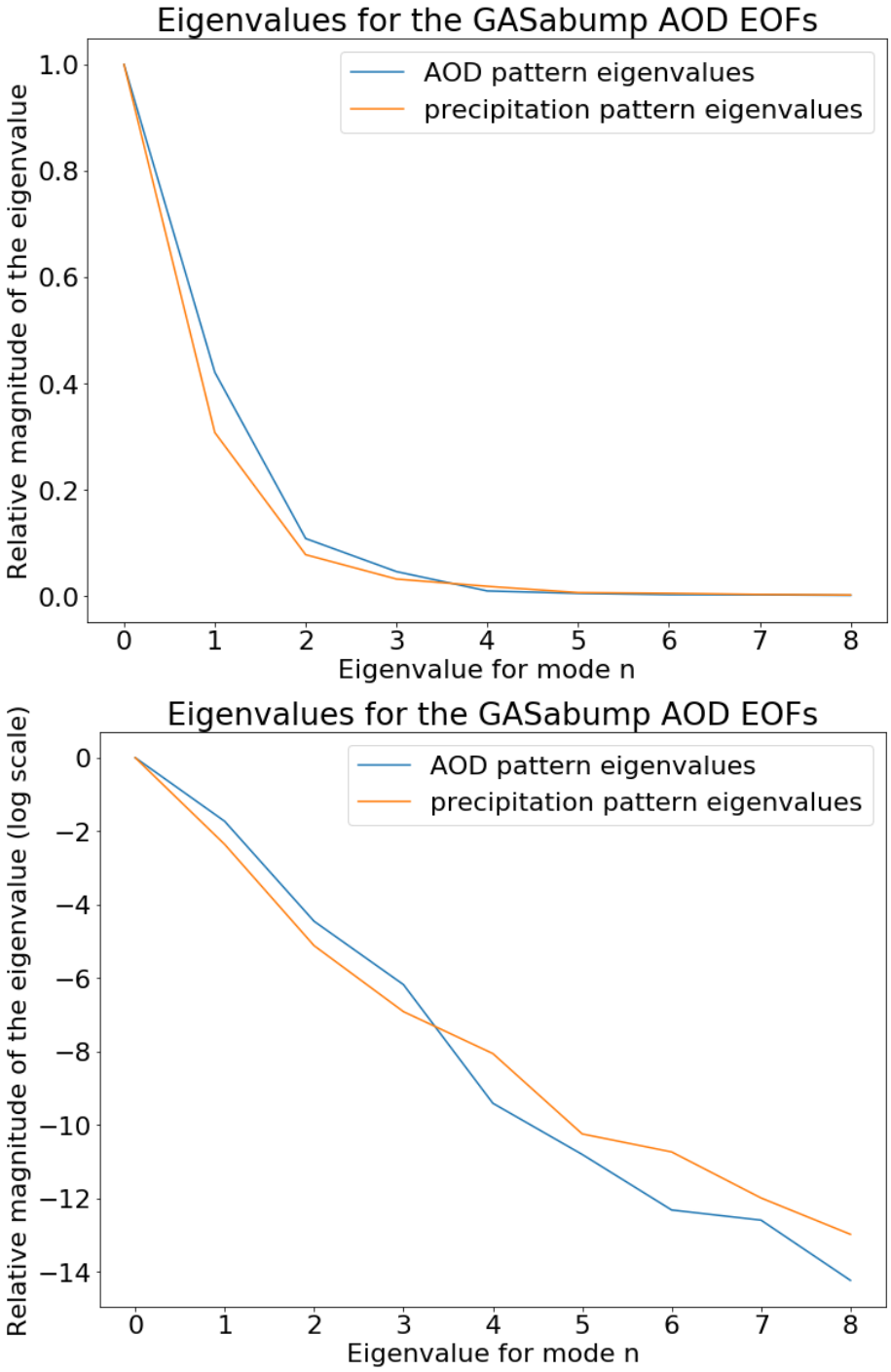

Figure 4.20: **Illustrations of the eigenvalue spectrum.** First (in blue) across the range of EOFs for aerosol optical depths, and then (orange) across the range of EOFs for the precipitation patterns that are responsive to those AOD EOFs. GASabump is shorthand for the ensemble of convenience described above.

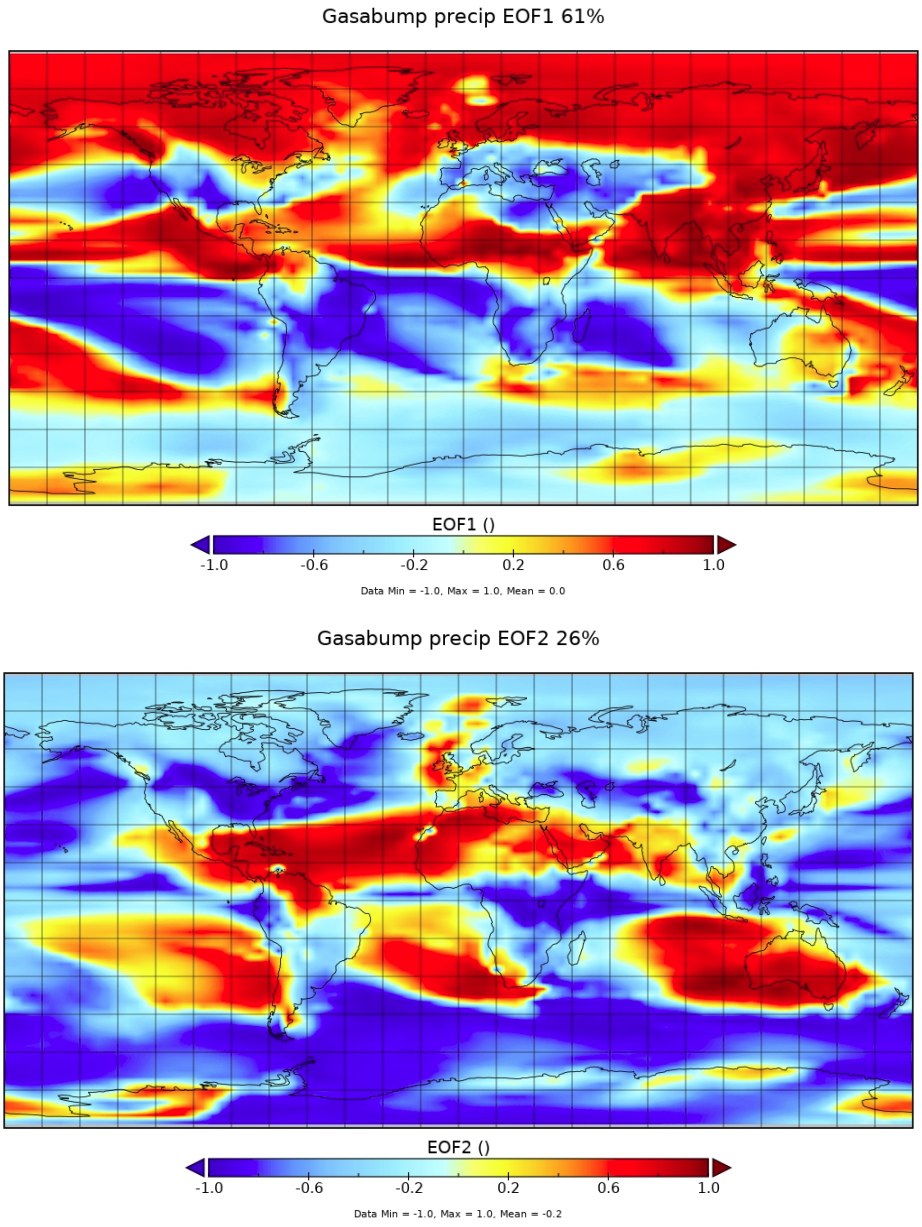

Figure 4.21: **The first and second EOFs for decadal precipitation.** Across a range of different injection patterns, we can calculate the empirical orthogonal functions (EOFs) for the decadal precipitation anomaly.

Interpreting the remaining EOFs and their physical significance is more difficult. EOF2 and EOF3 are close to being inverses of each other. Note that the fourth EOF seems to correspond to the strong correlation between precipitation anomalies in the Amazonian and Congo Basin regions, associated with the rainforests in these regions. Under increased CO_2 concentrations, the effect of decreased stomatal opening reduces

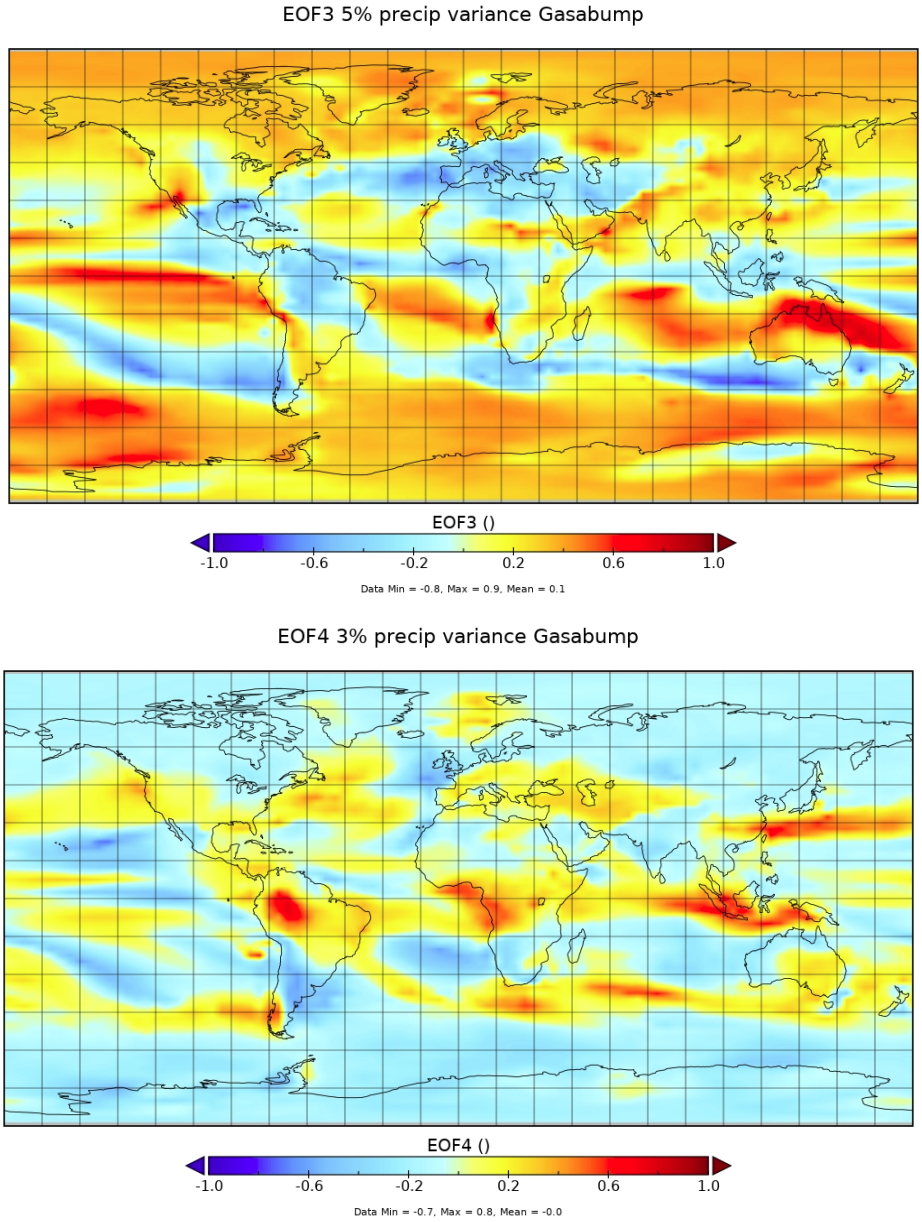

Figure 4.22: **The third and fourth EOFs for decadal precipitation.** Across a range of different injection patterns, we can calculate the empirical orthogonal functions (EOFs) for the decadal precipitation anomaly. Across the first 4 EOFs, 95% of the variance is explained.

evapotranspiration from the rainforested regions, and this results in drying over these regions independent of the applied aerosol optical depth. The third and fourth EOFs look like almost exact inverse patterns from each other, and between them they explain, explains 8% of the variance (5% for EOF3 and 3% for EOF4). However, in the absence of a clear physical interpretation of the EOFs for precipitation, it is

important not to overinterpret their physical significance, as they may be statistical artefacts that arise from the orthogonal decomposition.

It is unsurprising that the response of precipitation to inhomogenous forcings is more complicated in structure than the temperature response. The precipitation response is a second-order response to the imposed aerosol optical depth. There are direct thermodynamic elements to the response which depend on the temperature change in a given region, and this is particularly dominant at high latitudes; but, for the bulk of precipitation in the tropics and midlatitudes, the dynamical effects - due to shifting storm tracks and ITCZ and changing patterns of circulation in the atmosphere - are more important, and it is these dynamical effects and shifts to circulation that are more likely to be related to the precise spatiotemporal distribution of inhomogenous forcings than the smaller, direct, thermodynamic effects.

4.3.5 Precipitation: fast vs slow response

Another lens through which to look at precipitation is to analyse the patterns of precipitation that result from simpler scenarios - global warming and a globally applied uniform aerosol optical depth.

Studies of the standard abrupt $4xCO_2$ experiment in climate models illustrate that it can be decomposed into a fast and slow response. [163]

The fast response occurs on a timescale of weeks to months - before the global mean surface temperature has time to respond - and is generally characterised by drying, especially over land and in the tropics. [164] This occurs as a response to the initial heating of the upper troposphere in the presence of CO_2 and the reduced ability for longwave radiation to escape from cloudtops; this decreases the temperature gradient between the surface and the troposphere, resulting in suppressed convection and a reduction in precipitation.

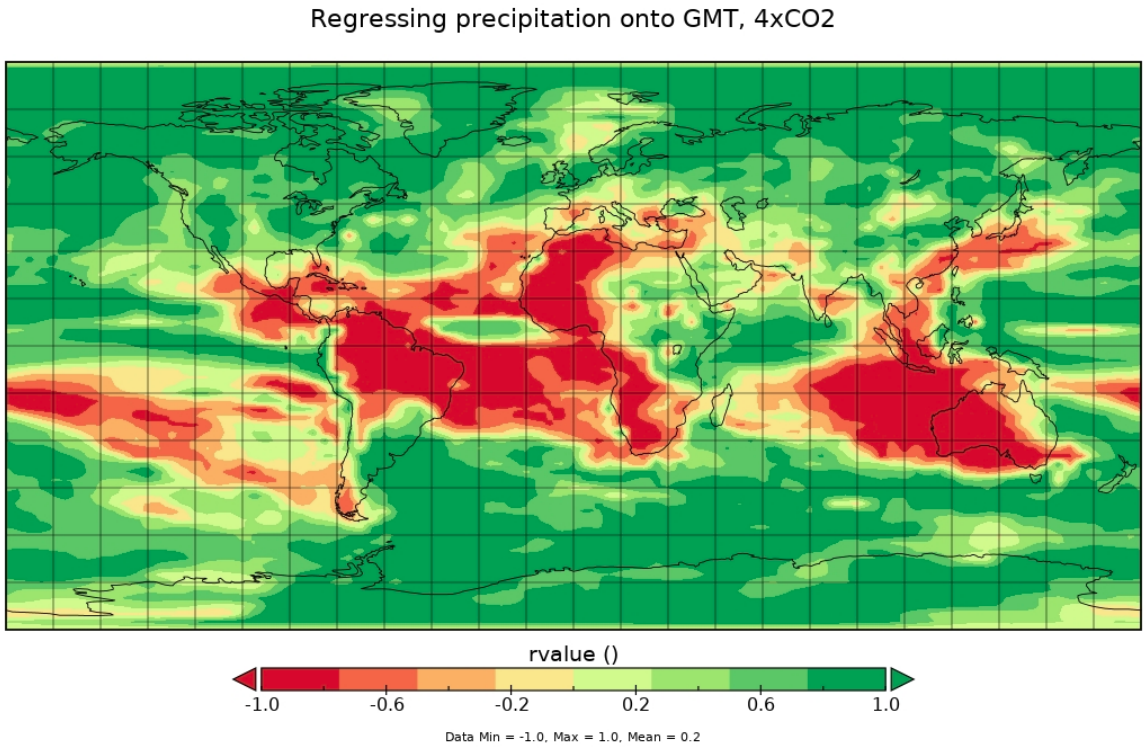

Figure 4.23: **Correlation coefficient for precipitation on global mean temperature under $4xCO_2$.**

Then, as surface temperatures begin to respond to the radiative imbalance and

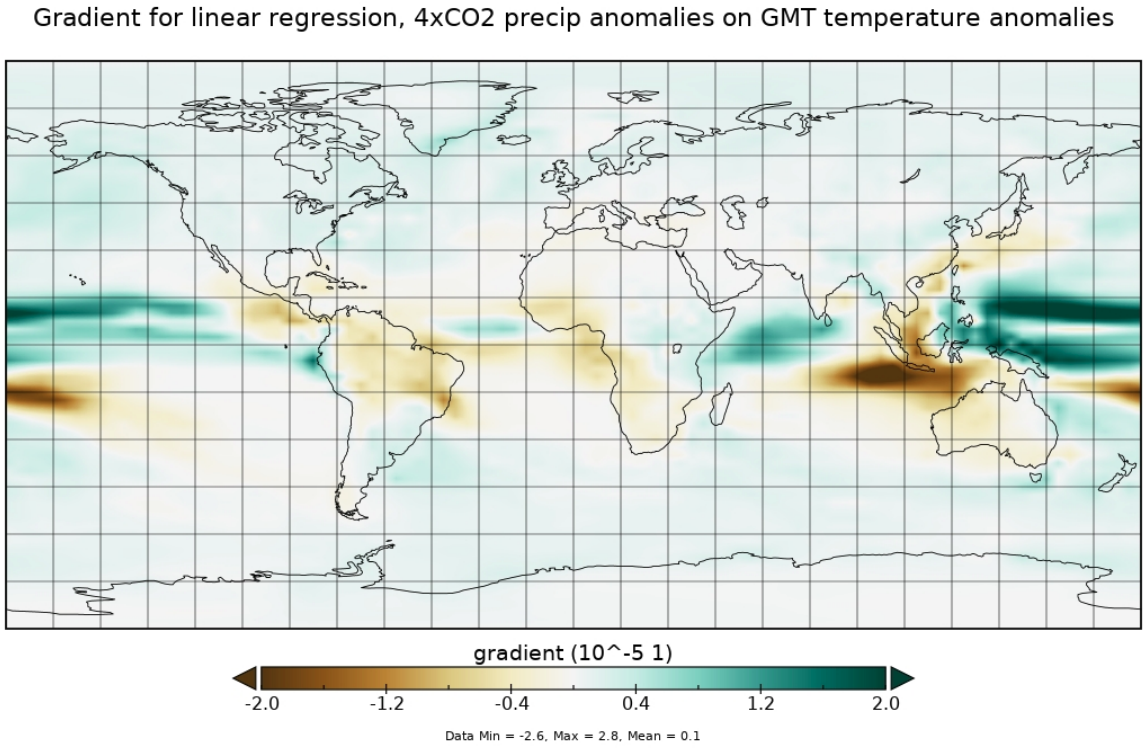

Figure 4.24: **The gradient of precipitation changes when precipitation is regressed on global mean T.** This aims to illustrate the "slow response" of precipitation to temperature as the 4xCO_2 simulation warms.

the world warms, the slow response is a general increase in precipitation as higher temperatures lead to more evaporation at the surface. In other words, the direct effect of CO_2 on atmospheric stability and the ability of cloud-tops to radiate away latent heat, and the slower warming of the land surface combine to generally reduce precipitation in the short term as the fast response, while the response of the general circulation to rising surface temperatures and global warming under quadrupled CO_2 characterises the slow response. The fast response for precipitation therefore follows changes in atmospheric absorption while the slow response follows changes in surface temperature. We can see that the slow response involves a shift north of the ITCZ, especially over the oceans, but also some additional drying over the Amazon region, parts of China, Australia, and Africa - i.e. tropical rainforest regions additionally dry as temperature increases, while precipitation generally increases everywhere else.

For a simple metric like global mean precipitation (GMP), this is characterised by an initial global decrease in GMP, followed by a gradual increase of GMP in line with

global mean temperature before eventually reaching an equilbrium with higher GMP. In the quadruple CO_2 case, this pattern corresponds to the zonal wet-gets-wetter dry-gets-drier general pattern that is expected to manifest under global warming. The main contributions to this global change are over the oceans.

Studies of the G1 geoengineering experiment, where a uniform layer of aerosol optical depth is used to cancel out the warming due to $4xCO_2$, note that there is a general reduction in precipitation below the preindustrial levels. This is in part due to the effect of the aerosol layer in reducing the shortwave radiation that reaches the surface, reducing evaporation and therefore precipitation in turn - for this reason, several have suggested that P - E is a more important metric for determining moisture availability and the effect of these changes on civilisation.

One way of conceptualising precipitation changes under the G1 scenario is that it represents the fast response to CO_2, alongside a fast response due to the aerosol forcing and reduction in shortwave flux to the surface, without a compensatory slow response due to increasing temperature. To the extent that this is the case, we may be able to infer the response to different levels of aerosol optical depth, and also determine whether it will be possible to find an optimal temperature or level of uniform aerosol forcing to minimise temperature anomalies. Characterising these features of the temperature response can allow us to distinguish aspects of the response to increased CO_2 and the direct effects of the AOD layer, as well as the differences in precipitation that owe to temperature being kept constant and not allowed to increase as in the global warming scenario.

Figure 4.23 and 4.24 analyse the $4xCO_2$ scenario, performing a linear regression on precipitation onto the increasing annual average global mean temperature in each grid cell of the model. Figure 4.23 illustrates the R-coefficient for the regression, giving a sense of both the strength and direction of the correlation between global mean temperature and precipitation in the given grid cells. Plotted is the Pearson's R coefficient for regressing the annual average precipitation in each grid cell on the global mean temperature for that year in the $4xCO_2$ scenario. In other words, positive coefficients indicate where warming under $4xCO_2$ increases precipitation, while negative coefficients indicate where this warming decreases precipitation. We see that the correlation is strong across most regions, between 0.8 and 1 in magnitude for most grid

cells examined. Broadly speaking, we see precipitation increasing at high latitudes, but decreasing over the South Atlantic and areas of the Pacific Ocean just below the equator (due to a Northward shift of the ITCZ under warming overall), alongside decreases over the Mediterranean, the Amazon and Central America, Australia, regions of south east Asia including China, and much of Africa (particularly the West of the continent.) In general, then, we see drying across much of tropical and midlatitude land and increased precipitation at high latitudes as the world warms.

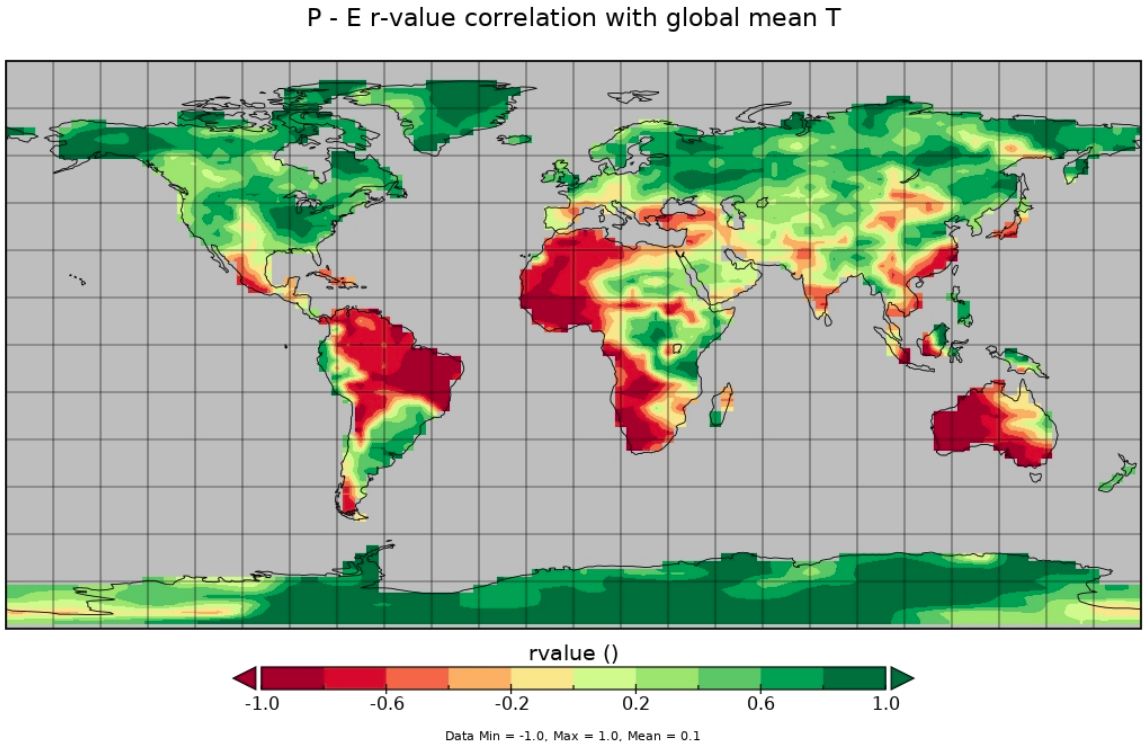

Figure 4.25: **Correlation coefficient for P - E over land on global mean temperature under 4xCO_2.**

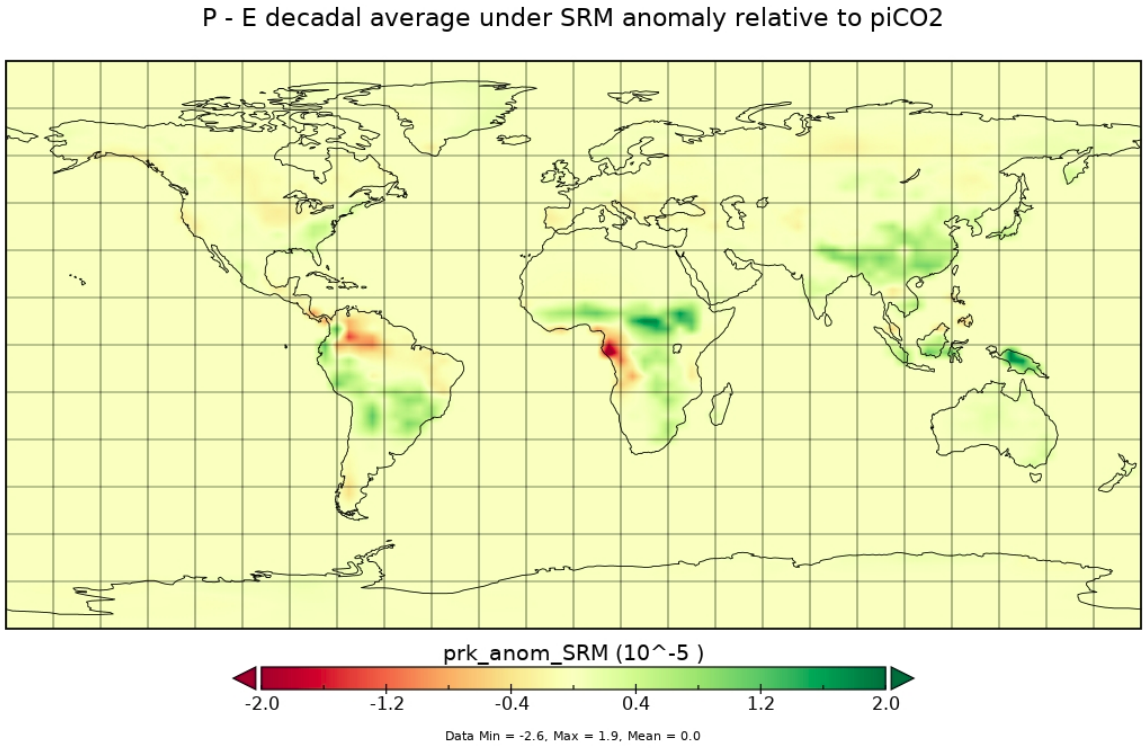

Figure 4.26: **The gradient of P - E changes when precipitation is regressed on global mean T.**

Figure 4.25 and 4.26 do the same for P - E, as a metric for moisture availability, in order to compare this to the raw precipitation response. Notably, over land, we can see that the R-value pattern is very similar in P - E as it is for precipitation alone. In other words, for these regions - where evaporation is often moisture-limited - it is not simply a case that the hydrological cycle's intensity is decreased, with a decrease in precipitation compensated by a similar decrease in evaporation. Instead, moisture availability decreases in these regions as well under global warming, although to a lesser extent than precipiation. We see when looking at P - E that in many regions, the apparent drying under SRM is more than compensated for in P - E owing to a reduction to evaporation from the reduced shortwave flux to the surface; the only regions where this is not true are regions with dense rainforest, which dry possibly owing in part to the CO_2 effect. This is in accordance with the model finding that under SRM, we would see "less rain, but a wetter and greener" climate. [149]

A direct comparison between the gradient in the linear regression of precipitation on global mean temperature and the precipitation anomalies under uniform SRM

illustrates some of the areas where additional warming might be hoped to help compensate for the sharp drying under SRM - i.e. where some amount of the "slow response" may be able to compensate the "fast response". We can certainly see that this is the case at high latitudes - particularly over the high-latitude continents in the Northern hemisphere; here, additional warming can compensate for the drying under SRM. However, there are regions of West Africa and Amazonia where both the initial application of SRM and continued warming under 4xCO_2 result in additional drying; for such regions we can anticipate that any strategy that involves "allowing" some warming so that the precipitation increases can compensate for decreases will be unsuccessful, as the signs of the fast and slow response are the same. Other regions - like China and areas of India - appear to see increases in precipitation both under uniform SRM and with increased temperature, although this effect is not uniform across China. It is again notable that the vast majority of changes to precipitation which can be observed are over the tropical Pacific ocean - with large increases in precipitation as the world warms and sharp decreases under uniform SRM.

This is in accordance with results from the Precipitation Drive and Model Response Project (PDRMIP) [59] which found, across a range of different forcers (including CO_2, sulfate aerosol, and reduced insolation) that there were certain land regions across the multi-model ensemble where the fast response consistently dominated over the slow, temperature-driven response.

4.3.6 Giorgi regional optimisation for temperature using pattern-scaling

One of the concerns raised surrounding the prospect of latitudinally tailoring geoengineering efforts is the potential for international conflicts to arise over the optimal distribution. If powerful nations benefit substantially from "tailoring" the distribution in their favour, there is a heightened risk of unequal outcomes and further governance issues in an area already fraught with them.

One way we can attempt to shed some light on this problem through climate modelling is by trying to determine the advantages that individual regions may be

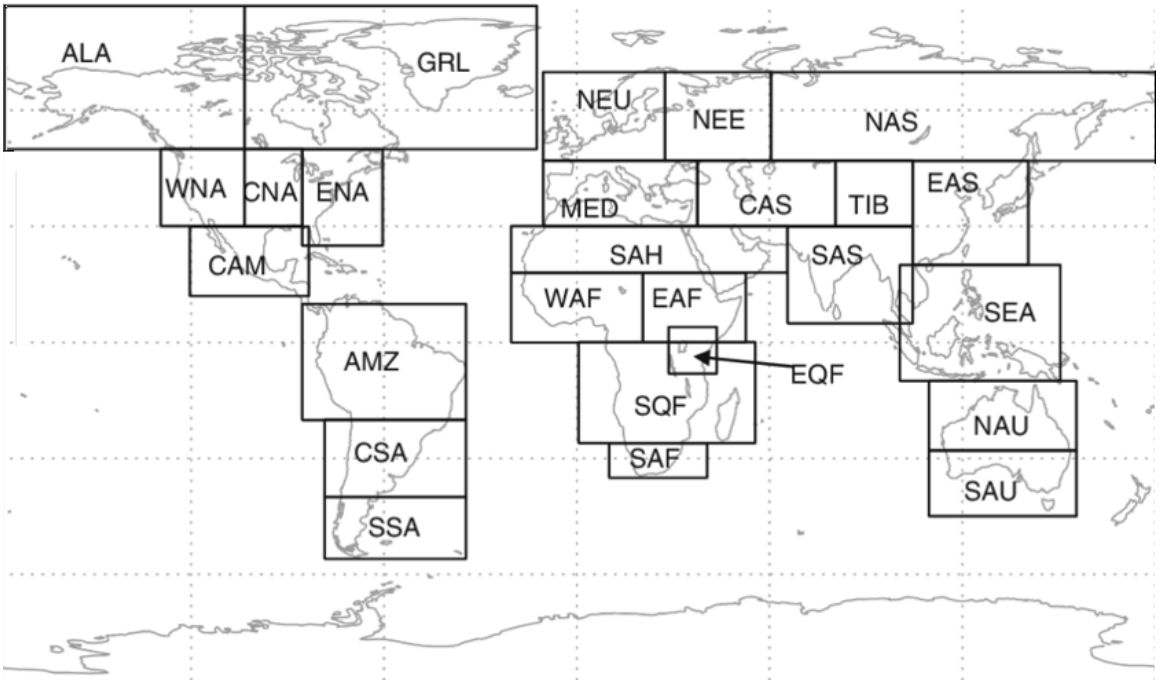

Figure 4.27: **The Giorgi regions.** (Our regions merge SAF, SQF and EQF into SAF, NEU and NEE into NEU, and CSA + SSA into SSA.)

able to obtain through selfish "steering" of the global aerosol optical depth distribution. For this analysis, we used the linear, additive approximation to determine the decadal average temperatures in each of the given Giorgi regions (adapted from those illustrated in Figure 4.27), depending on the coefficients of the basis injection functions that are used to determine the overall distribution. We then seek the optimal coefficients to minimise fitness functions based on the RMS temperature difference between the geoengineered scenario and the preindustrial, but with different weightings for each Giorgi region. We can then determine which coefficients each Giorgi region may favour with different levels of "selfishness", i.e. weightings for the particular region.

In Figure 4.28, we plot the fractional residual RMS temperature anomaly in the given Giorgi regions against the fractional residual RMS temperature anomaly in the rest of the world under two different scenarios: the distribution that minimises global RMS temperature, with only a slight weighting towards the given Giorgi region, and the distribution which weights this region as much as the rest of the world combined. We plot the percentage residual result for RMSE temperature on the land points in the given Giorgi region against the percentage residual result in the rest of the world.

The black data point denotes the residual percentage anomaly under the optimisation to minimise the global RMSE temperature with no weighting for any region. The residual is normalised to the anomaly with no SRM, i.e. under 4xCO_2, and the line y=x is plotted for convenience, as it denotes when the temperature anomaly in the region is the same as that in the rest of the world. Points above the line therefore have a greater fractional temperature anomaly than the rest of the world, while points below the line have a lower regional fractional temperature anomaly than the rest of the world.

The blue dataset shows the effect of optimisations that weight the given Giorgi region twice as much as other regions, or 'moderate selfishness', the red dataset shows the more selfish case where the Giorgi region is weighted as much as the rest of the world in the optimisation. The lines connect the same Giorgi regions before and after the heavier weighting - so a line with a shallow gradient represents a region which benefits marginally from selfishness, in exchange for a substantial increase in temperature anomalies for the rest of the world, while regions with steeper line gradients can benefit more at less expense to others. We can see that most Giorgi regions can see reductions in their temperature anomaly through this optimisation procedure - the red points are lower on the y-axis - but generally at a significant expense to temperature anomalies across the rest of the world - the red points are significantly further along on the x-axis. Higher selfishness can reduce the fractional residual anomaly in some regions by 3-5%, but often at a cost of increasing the anomaly for the rest of the world by 10-20%. It seems unlikely, given the constrained optimisation that we consider in this scenario, that individual Giorgi regions would be so sensitive to a few tenths of a degree in their average temperature so as to wish to "steer" the geoengineering substantially away from a Pareto-optimal or global RMSE minimising solution for cooling.

We note, however, that while this is likely true for temperature - where linearity is expected to hold - the situation is much more sensitive for precipitation. Concerns surrounding hemispherically asymmetric applications of SRM which could, for example, substantially disrupt monsoons in agriculturally sensitive areas, as discussed in Chapter 1 and modelled in Chapter 3, remain a much more concerning threat for disputes over the optimal distribution of SRM.

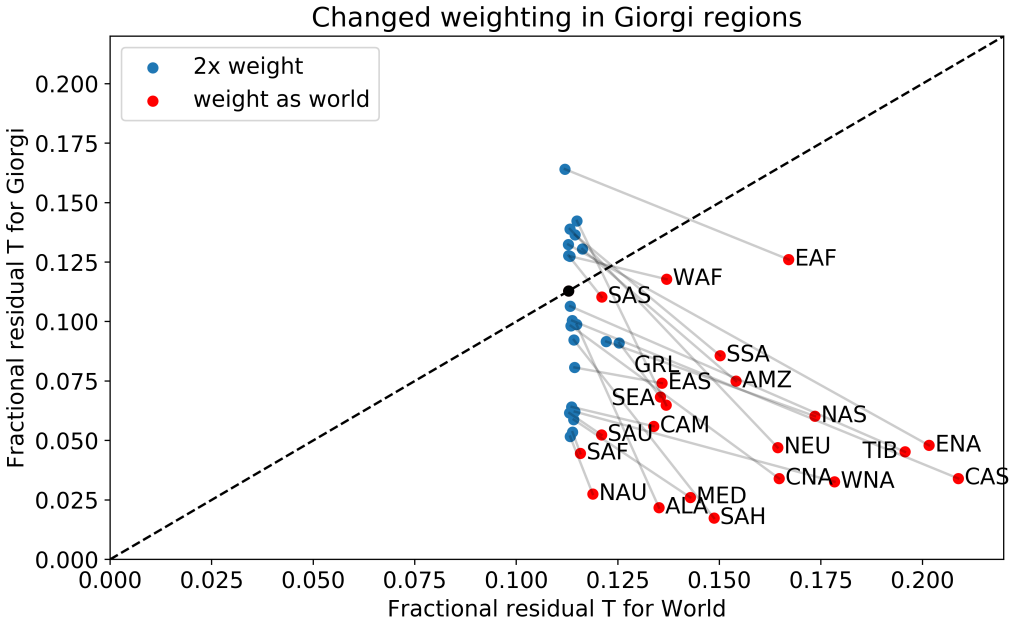

Figure 4.28: **Scatter plot of residual temperature anomalies in the Giorgi region optimisation scenarios.** Each point indicates a different optimisation for a different Giorgi region.

4.3.7 Giorgi regional optimisation for precipitation using uniform layers of AOD

Figure 4.29 illustrates the effect of changing the uniform AOD parameter on the RMS precipitation anomaly in given Giorgi regions. Given the strength of the linear regression for precipitation on aerosol optical depth in most regions shown by the high correlation coefficients for the linear fit of precipitation on AOD, we can ask what the optimal aerosol optical depth for each Giorgi region to return its RMSE precipitation as close to the preindustrial as possible. In all cases, we plot the decadal RMS precipitation anomaly over land for the given Giorgi region against the uniform aerosol optical depth applied.

When plotted against the level of the uniform aerosol optical depth, the RMS precipitation for each Giorgi region gives a general parabolic function for these precipitation anomalies. In other words, there is an AOD that minimises the precipitation anomaly; increasing or decreasing the AOD from this level increases the precipitation anomaly in that region.

The box-whisker diagram for each region is plotted with its location at the minimum of each curve in Figure 4.30, illustrating the extent to which the precipitation anomaly in this region can be reduced from its error at zero AOD, by changing the uniform aerosol optical depth; meanwhile, the colour of the box denotes the magnitude of the RMS error in precipitation under 4xCO_2 relative to the preindustrial precipitation (i.e. up to 50% increases or reductions of the net RMS preindustrial precipitation.) The whiskers indicate the relative width of these parabolae, and therefore how sensitive the regional precipitation in that region is to changing aerosol optical depth. The height of the box-whiskers shows the minimum precipitation anomaly which can occur across the range of aerosol optical depth, relative to the precipitation anomaly at zero AOD. (Therefore all regions would converge on the X at 1.0 for 0 AOD.) The x-position of the box-whisker indicates the AOD that needs to be imposed (assuming linearity) to minimise the RMS precipitation anomaly in the given Giorgi region. The shade of the box illustrates the fractional precipitation anomaly in that given Giorgi region for the 4xCO_2 case (0 AOD) relative to the preindustrial, i.e. giving a measure of how relatively bad the error was under unmitigated 4xCO_2 in that Giorgi region before applying any AOD.

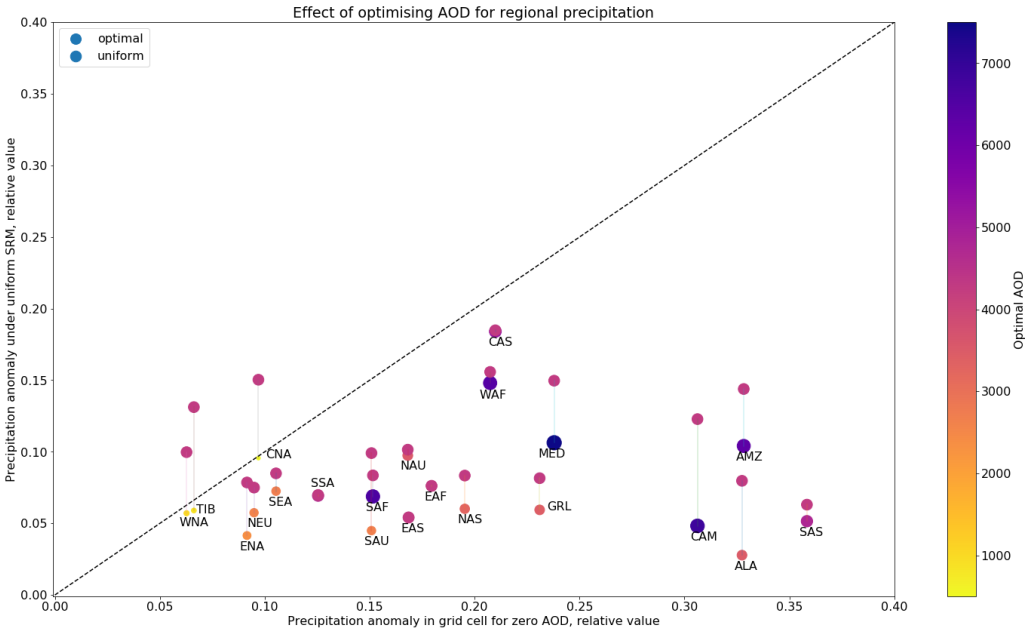

Figure 4.29: **Optimising relative precipitation anomalies for Giorgi regions by changing uniform AOD.** As the previous figure, but with all precipitation RMSE anomalies measured as a percentage of the preindustrial precipitation in the given Giorgi region (indicating how relatively important the effects of tailoring the uniform AOD of geoengineering and 4xCO_2 are to minimising the precipitation anomaly.)

The y-axis (fractional precipitation anomaly) is normalised to the case with zero AOD (4xCO_2). The purpose of this is illustrating which Giorgi regions would prefer greater or reduced aerosol optical depth beneath the two cases we have examined in most detail - the uniform AOD (G1) case and the 4xCO_2 (zero SRM) case, as well as giving a sense for the size of the fractional improvement in the anomaly that can be obtained by this intervention. We note, for example, that no regions can be reduced to precisely zero anomaly by any uniform AOD.

All regions benefit, with reduced precipitation anomaly, from a nonzero application of AOD relative to the unmitigated 4xCO_2 case. Some, rather than no, aerosol optical depth results in reduced precipitation anomalies across most regions. However, regions disagree about the ideal depth of aerosol that would be needed to minimise their precipitation anomaly, and for some regions - such as the Sahel and Tibet - increasing AOD substantially above their low, optimal level rapidly results in worsening outcomes and larger precipitation anomalies than under 4xCO_2. We see from the

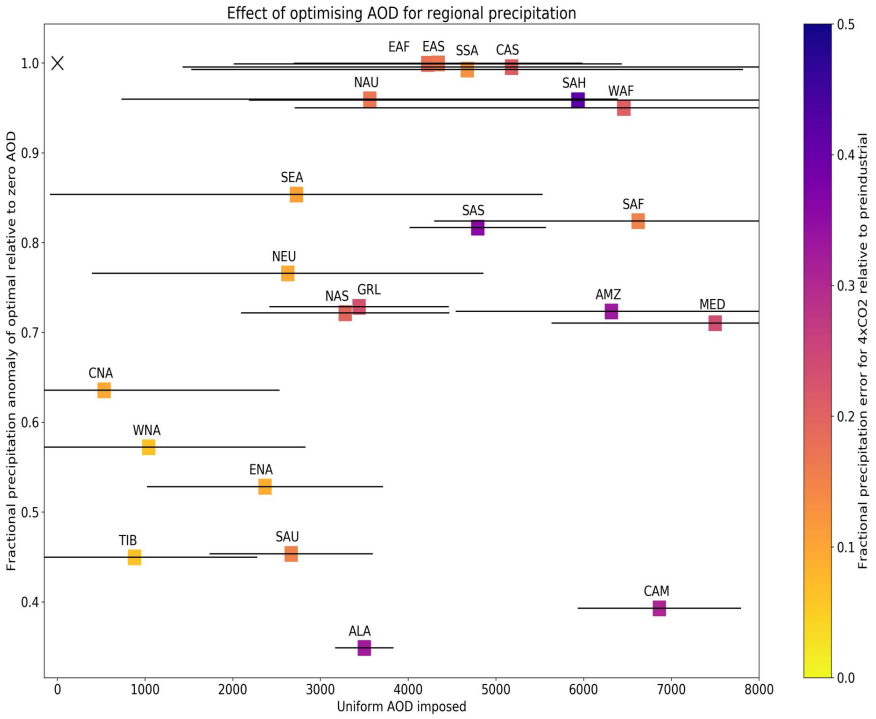

Figure 4.30: **The effect of changing the uniform aerosol optical depth parameter on the RMS precipitation error in given Giorgi regions.**

different levels of curvature that greater fractional reductions in precipitation can be achieved by fine-tuning the global AOD for some regions, while others have persistent anomalies that are relatively insensitive to this.

Compared to the uniform AOD that cancels global mean temperature case, most regions benefit from a reduced aerosol optical depth; there is clearly a conflict for individual regions between the uniform intervention that would "optimise" for their temperature and which would "optimise" for the precipitation within that region. However, for at least three regions, including North Asia (NAS), Central America (CAM), and the Amazon (AMZ), we see that the precipitation anomalies could be reduced further by overcooling the globe. We will discuss this further in the conclusions section.

Appendix 2 explores the different distributions that individual Giorgi regions may prefer by calculating the optimal coefficients for each of the "basis vectors", which

would minimise land temperature anomalies in a given Giorgi region, assuming linearity. This illustrates that there are substantial differences in the preferred distribution of aerosol optical depth between Giorgi regions.

4.4 Discussion and conclusions

4.4.1 Limitations

There are a number of limitations to this work which we should address. Firstly, the limitations that apply to HadCM3 and CPDN itself, which are discussed at length in Chapter 3, all apply to the modelling results in this chapter as well. In particular, concerns around the fidelity and accuracy of the precipitation response in the Amazon are important to the caveats I have expressed around optimising for precipitation. Again, the trade-off is that we are able to quite easily generate large ensembles for each of the basis vectors considered. I will not repeat the discussion from Chapter 3 in full here, but it should be borne in mind.

Another major limitation is in the nature of the optimisation. We only considered six basis functions and superpositions of these six patterns. Additional work (omitted here for space) showed that, when we considered the same optimisation process, but using the full suite of 18 simulations described in this chapter and Chapter 3, there were slight differences in the distribution, suggesting that a greater number of degrees of freedom might result in different "optimal" aerosol optical depth distributions. However, the differences between the zonal temperature optimisation attainable with this distribution and with the one used in this chapter were small, and because the distribution made use of some of the "stripe" simulations, it was less smooth and arguably less physically attainable. In reality, there would inevitably be limits in how much control one could attain over the aerosol distribution. In addition, the work presented here on EOFs and other work such as Kravitz et al. (2016) [94] suggest that there may be a limited number of degrees of freedom in the climate response in any case.

An arguably bigger limitation is that we were not able to deploy the feedback and control methods outlined in Kravitz et al. (2016) [94] and also used by the

GLENS experiment [95] whereby the coefficients of injection are adjusted in each year according to the changes in temperature that are observed in the model. This was not possible for us, as it would have required a way to modify the version of HadCM3 that volunteers were experimenting with to read in temperature output, perform the calculations, and then modify their own ancillary files in the middle of a model run, which goes beyond the modifications to HadCM3 that were possible on volunteer computers. It might have been possible to do this calculation "off-line" by setting up simulations to run for a single year, retrieving the data, and then calculating the coefficients ourselves, but this would have taken far too long to generate a decadal number of model years given the time we had available to conduct the project and the time taken to submit new batches to volunteers, which was outside of this experimenter's control. Instead, I had to specify an aerosol optical depth distribution at the beginning and then proceed with that distribution throughout the length of our model runs, without dynamic adjustment. It seems likely that, if I had been able to implement the closed-loop feedback, optimisation for the goals we had considered would have been more effective. I attempted various other ways of optimising (for example, attempting distributions that would cancel the predicted radiative forcing in the zonal mean under $4xCO_2$ for each latitude band) during the course of the DPhil, which were unsuccessful and have been omitted for space.

Other limitations arise in the Giorgi regional analysis, and in the optimisation for alternative figures of merit. Here, we relied on assumptions of linearity - first, that the temperature response in individual grid cells is linear with respect to the patterns imposed, allowing us to predict this response by pattern-scaling, and secondly, that the precipitation response is linear in aerosol optical depth. We justify this to some extent in the chapter - the linearity of temperature is explored in the Appendix, while the linearity of precipitation in AOD is justified by the finding that the R-value for precipitation regressed on the uniform aerosol optical depth applied in a range of scenarios with different AOD is close to 1 or -1 in most regions, suggesting linearity holds reasonably well across most grid cells, as illustrated in Figure 4.31. However, linearity is not a perfect approximation, which means that our optimisations in the individual Giorgi regions are not exact results.

It could further be argued, as we discuss below in the conclusions section, that

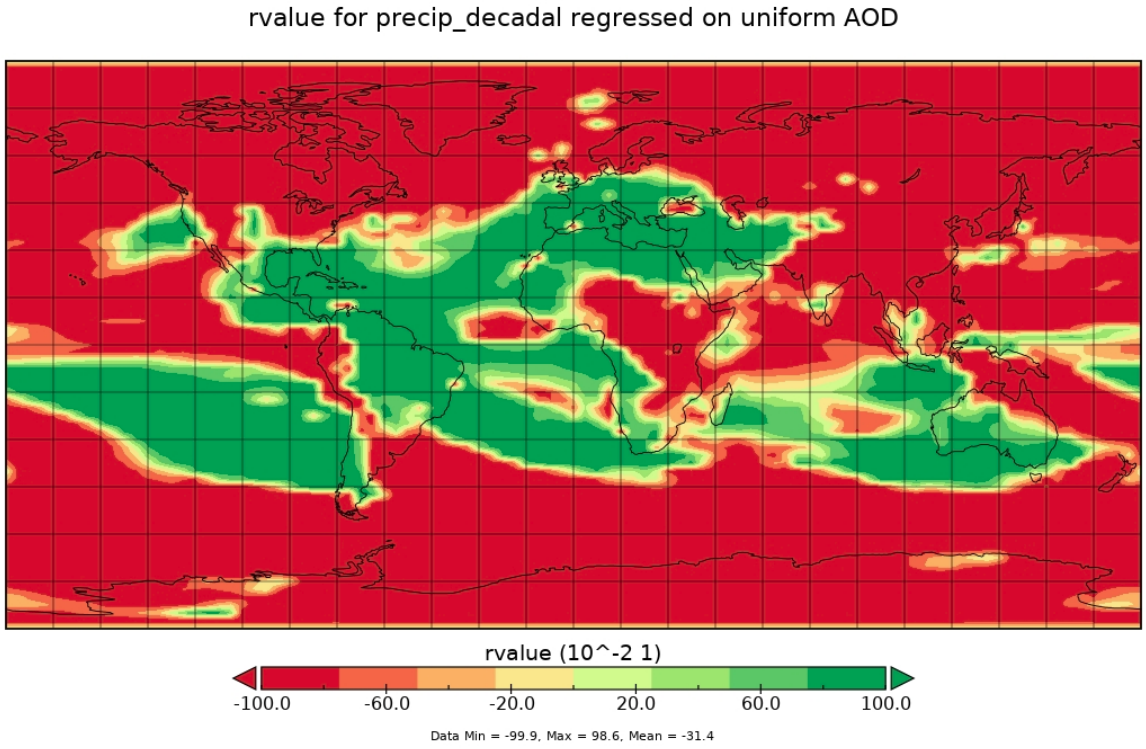

Figure 4.31: **Correlation coefficient for precipitation regressed on global mean aerosol optical depth, for a range of simulations with different uniform aerosol optical depths between 0 and 0.41 applied.**

the Giorgi regional optimisation analysis is not that instructive concerning potential political conflicts over geoengineering, as it assumes that every region wants to return to its preindustrial temperature and precipitation. Without studying in detail why particular regions might prefer particular climates, it is difficult to conduct an analysis that is consistent and meaningful without making this assumption. Our results here should be taken to outline that some regionally "selfish" optimisation is possible, and to give an indication as to how successful it might be, without being considered exhaustive as to how would-be geoengineers might seek to alter climate. It is also worth noting that, especially for temperature, returning the global mean temperature to preindustrial levels brings most regions much closer to preindustrial temperatures than $4xCO_2$, as set out in Chapter 3, but if individual regions preferred being (say) 2K warmer, this would probably entail a substantial departure for the global mean temperature anomaly. In other words, preindustrial conditions, or at the very least targets set by global mean temperature anomaly, have the advantage of being a well-

defined baseline which might see widespread agreement.

4.4.2 Conclusions

Our results in this chapter illustrate the significant variety of ways in which would-be geoengineers could attempt to optimise for different figures of merit, whether decadal, zonal average quantities or regional temperature and precipitation. In general, we show that some improvement is possible over uniform SRM for quantities such as zonal decadal temperature means, and arguably precipitation means as well, although the benefits of optimisation are less clear for precipitation, and the effect is generally quite small. We show that the temperature-optimised geoengineering simulations do not significantly improve decadal zonal precipitation anomalies. We also show that, due to the structure of the "fast" and "slow" response to precipitation, there are some regions where allowing the world to warm slightly to compensate for the fast drying effect of CO_2 and SRM is unlikely to compensate for persistent regional drying across scenarios, especially in regions like the Congo and Amazon basin. Finally, we explore whether Giorgi regions might prefer different distributions (for temperature) or different aerosol optical depths (for precipitation). It is the case that regions disagree, sometimes significantly, on their preferred distribution or thickness of uniform AOD layer, so a "selfish" region that targets preindustrial conditions might wish to have a significantly different distribution than the globally optimal ones. However, we also show that the relative benefits of selfishness are often quite small and come at a significant cost to the rest of the world, so it is unclear that pursuing such a lopsided policy would really be worthwhile, at least if returning a region to preindustrial conditions is the goal - which can also be disputed. We will discuss our results further in following subsections.

4.4.2.1 Optimising for temperature

We outlined the results of attempts to use the GLENS basis vectors to construct distributions that would minimise temperature anomalies. We compared optimisations for two figures of merit, one explicitly simulated, and the other constructed using pattern-scaling assuming linearity. These were distributions that aimed to minimise zonal decadal temperature anomalies, and also RMSE temperature anomalies over land specifically. These are compared in Figure 4.9. In summary, while zonal mean temperatures are minimised - with smaller temperature anomalies over the poles -

optimising for zonal temperature anomalies still results in considerable undercooling of the land, particularly over Russia and in regions like Amazonia where transpiration changes are important, and overcooling of the oceans. Minimising RMSE temperature over land still results in regions that SRM is more and less effective at cooling - but we also see that - by overcooling the globe by about 0.3K below the preindustrial - it is possible to further reduce average land temperature anomalies.

4.4.2.2 Optimising for precipitation

In Figure 4.10, we compared the precipitation under temperature-optimised SRM, to the precipitation under uniform SRM and under 4xCO_2, attempting to determine whether optimising for temperature also optimises for precipitation. Comparing the SRM precipitation anomalies to the 4xCO_2 anomaly, we see that the ITCZ is shifted to the North by the polar warming in the 4xCO_2 scenario, resulting in a precipitation dipole. This is partially compensated for by any SRM, but SRM also reduces tropical precipitation overall. The increase in extratropical, high-latitude precipitation under 4xCO_2 is more than cancelled out by SRM, resulting in a decrease. In both SRM distributions, zonal precipitation decreases in the tropics, likely due to a weakening in tropical circulation around the ITCZ. The extratropical response is similar in both cases, with a slightly greater reduction in precipitation at high latitudes under zonal temperature optimisation likely due to the additional AOD at these latitudes compared to the uniform case, which cuts shortwave flux to the surface and hence evaporation in regions where precipitation is energy-limited rather than moisture-limited. From this, we can clearly see that optimising SRM to cancel out zonal temperature increases does not also optimise to reduce the precipitation anomalies under uniform SRM.

Looking only at precipitation over land, we see that the zonal anomalies take a similar pattern under each scenario, with the main differences in the magnitude of the anomaly. Extratropical increases in precipitation are reduced and often overcompensated by SRM. Zonal temperature optimisation slightly ameliorates the reduction in precipitation over the equatorial region compared to SRM. We can see from this, however, that precipitation anomalies over land are not particularly ameliorated by either distribution of SRM.

When we consider the optimisation attempt for zonal precipitation, we note that substantial non-linearities, especially over the tropics, mean that our attempt to construct a distribution to minimise zonal precipitation anomalies is much less effective than that for temperature. Even where zonal precipitation anomalies are indeed reduced, this often occurs with compensating errors within a given latitude band, and the precipitation anomalies in many regions over land are not improved by this optimisation.

4.4.2.3 Empirical orthogonal functions, fast vs slow precipitation response

Across a range of different injection patterns and resulting temperatures, one can calculate the empirical orthogonal functions (EOFs) for the decadal temperature anomaly to determine whether there are a number of fundamental "modes" of response for the climate system regardless of the imposed aerosol optical depth distribution. Across the first 4 EOFs, 98.7% of the variance is explained. The first EOF, explaining 75% of the variance, corresponds to the (Arctic amplified) global warming pattern when global mean surface temperature increases. The second EOF, explaining 19% of the variance, corresponds to a hemispheric asymmetry in temperatures. The third EOF, explaining 3.3% of the variance, corresponds to a pole-to-equator temperature gradient. The fourth EOF, explaining 1.7% of the variance, resembles a second harmonic of temperature as a function of latitude, and may be an artefact of orthogonality. We also find that the components of variation explained by these EOFs for a given distribution reassuringly correspond to properties of the imposed distribution. For example, the EOF that corresponds to the global warming pattern is more prominent in distributions with a greater overall cooling effect, while the hemispherically asymmetric EOF is more prominent in asymmetric AOD distributions. Repeating this analysis for precipitation, we find a more varied picture - the EOFs are more difficult to physically interpret, although the dominant one appears again to correspond to the pattern of precipitation changes under global warming.

I also conducted some analysis of the "fast" vs "slow" precipitation responses under SRM and 4xCO_2 in order to determine whether allowing some warming, and thus some slow precipitation response, could compensate for the initial decrease in precipitation that arises due to a reduction in shortwave radiation reaching the surface,

and consequently evaporation, as well as the effects of CO_2 on transpiration. I find that this may be effective at high latitudes, but is unlikely to be effective over regions such as the Amazon or the Congo which show persistent drying in HadCM3.

4.4.2.4 Giorgi regional optimisation for temperature and precipitation

One question with potential geopolitical relevance is the degree to which individual regions might benefit from skewing the aerosol optical depth distribution in their favour. The advantages they could gain by doing this, in terms of regional temperature, are illustrated in Figure 4.28. We can see here that most regions can reduce their own fractional anomaly to be less than the world's on average (below the red point). The high-selfishness case seems to show that the trade-off is relatively small, however, with modest gains in reducing local temperature anomalies in the optimal case at the expense of significant losses elsewhere.

In Figure 4.30 and Figure 4.29, we considered optimising for precipitation in different Giorgi regions by changing the amount of uniform aerosol optical depth applied. Overall, it is not clear if there will be a Pareto-optimal solution for minimising the precipitation anomaly across the range of regions. What's more, it's also unclear if this would be the desirable solution to aim for, with a number of different concievable weightings for anomalies (population exposed, agricultural importance, etc.). Some regions are relatively insensitive to changing the total amount of SRM, and therefore it would arguably be better to slightly increase their anomaly, deviating further from their ideal uniform AOD, in favour of gains elsewhere. Finally, these relative anomalies obscure the differing magnitude of the different anomalies in various regions. Figure 4.30 illustrates some of these discrepancies between different regions.

The results presented in this chapter clearly do not illustrate the full range of possible optimisation scenarios or patterns that could potentially be constructed. In particular, we assume for convenience that individual regions would "steer" the climate back to their preindustrial conditions, but this may not be the case - desertified regions may want more precipitation to enable agriculture, and some regions may benefit from overcooling or from allowing some warming (as has been argued for Siberia, for example.) They do not fully characterise the tradeoffs that could arise between regions, but they are sufficient to illustrate that it is plausible there would

be conflicts over the amount of uniform SRM, or the pattern of latitudinally-tailored SRM, that different Giorgi regions would prefer for their precipitation.

4.5 Appendix

4.5.1 Maps of non-linearities in temperature anomalies

Figures 4.32 and 4.33 show the error in assuming that the temperature response to individual injection patterns is linear and additive. The mean across all grid cells is less than 0.1K in the decadal average when compensating for the temperature anomaly from 4xCO2, but in individual grid cells it can be as high as 1.5K (with notable anomalies around Greenland and in the Middle East and Sahel regions where significant temperature anomalies arise due to second-order effects such as ice melting and rainfall pattern shifts which are not linear.) For most land grid cells, the percentage error in the prediction from linearity is less than around 10% in magnitude.

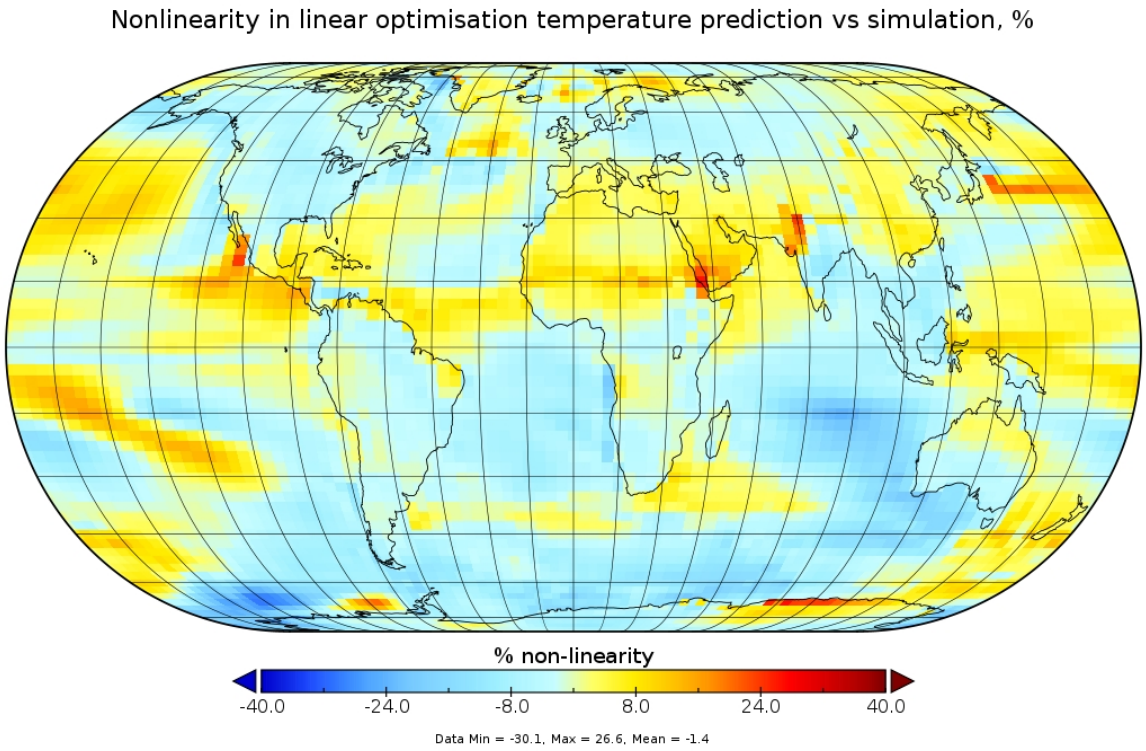

Figure 4.32: **Non-linearity in trying to use the linear, additive model to predict the temperature anomalies in the linear-optimisation for temperature (GLAT) simulation.** The non-linearity is listed as a percentage of the residual anomaly in the GLAT simulation. For most regions, the error is below 10% of the residual anomaly.

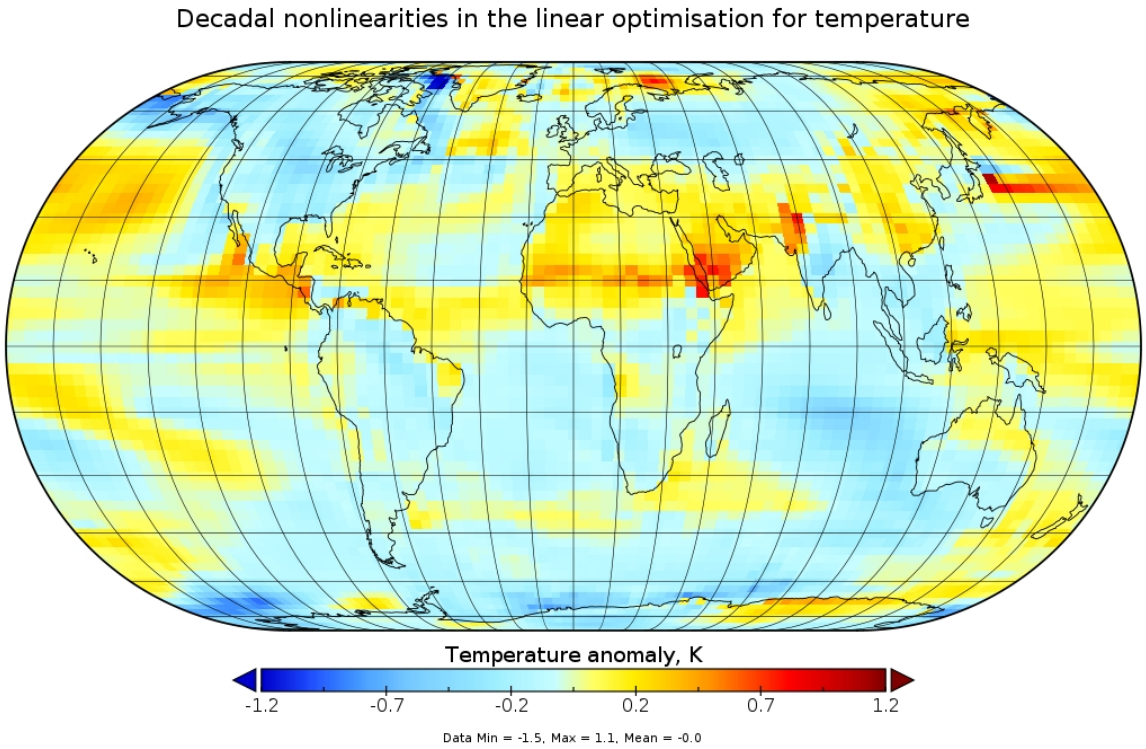

Figure 4.33: **The absolute error in using the linear approximation to predict the results of the linear-optimisation for temperature (GLAT) simulation.**

4.5.2 Preferred coefficients for Giorgi regional optimisation

In this analysis, we calculate the optimal AOD distribution, assuming linearity, with two weighting functions of predicted temperature anomalies to minimise. In the "low selfishness" case, the temperature anomaly in the Giorgi region is given twice as much weighting in the optimisation function as other grid cells. In the "high selfishness" case, the temperature anomaly is weighted as much as the rest of the world's grid cells combined. We then plot the coefficients for the basis functions (15N, 15S, 30N, 30S, 50N, 50S) in the low and high selfishness scenarios, to determine how much the optimal distribution varies.

We can see that for a relatively low selfishness (2x weighting of a given Giorgi region relative to other land grid points) there is a relatively high degree of consensus among regions about the preferred coefficients for the distribution - which are all, in turn, close to preferred coefficients for the distrbution that minimises global mean land temperature anomaly indicated by a solid horizontal black line. As selfishness increases, there is a greater divergence - in particular between regions in the global South, like Africa, Central America, and India and, for example, Central and Eastern North America, which have opposing incentives for the coefficients in the Southern hemisphere. For example, we can see in Figure 4.35 that while the ideal coefficient for the 15S pattern for the global mean is 0.4, it is more than double this at high-selfishness for the Amazon region, and drops to zero for the high-selfishness cases for Western, Central North America and North Europe, while remaining close to 0.4 for the Indian subcontinent.

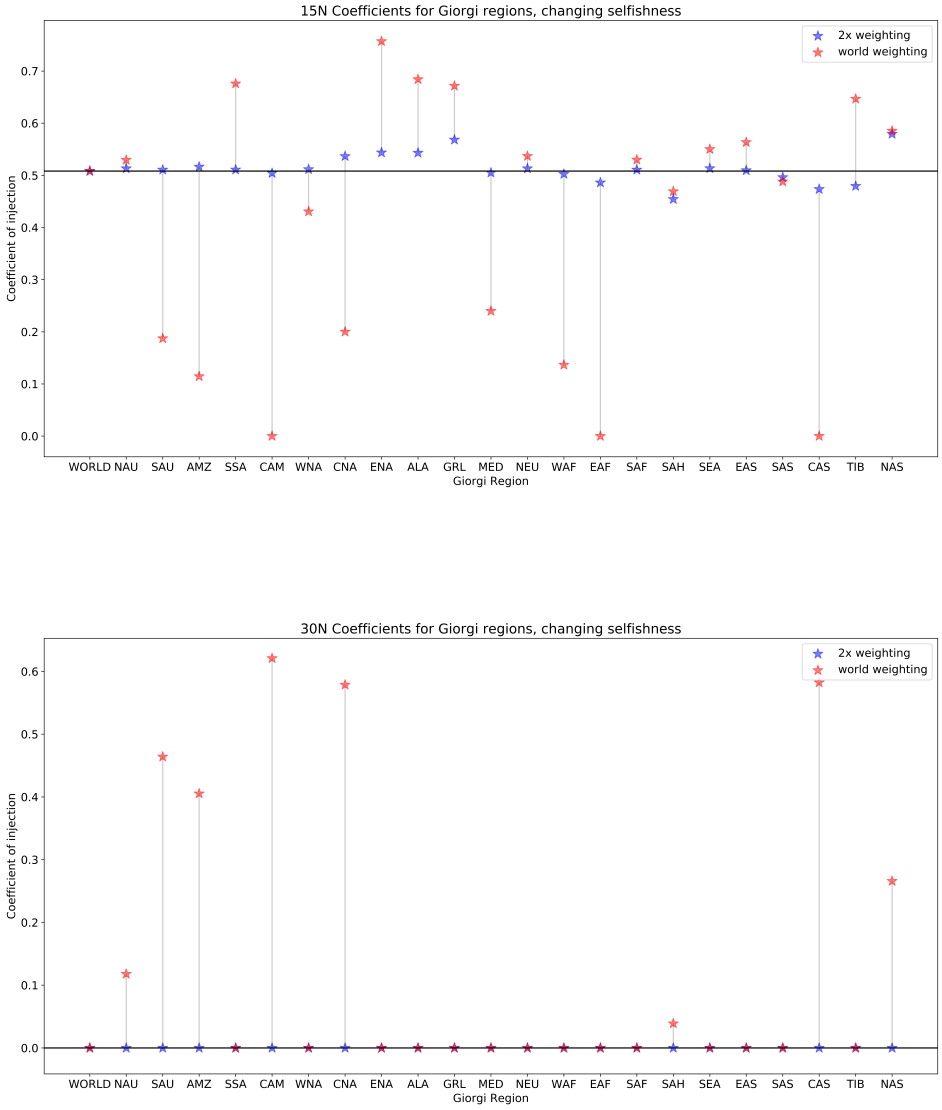

Figure 4.34: **Preferred coefficients for each of the Giorgi regions.** Assuming linearity, we attempt to use the GLENS basis linear optimiser to minimise the RMS anomaly for temperature on land in different Giorgi regions. Top (15N), bottom (30N). The different stars indicate (blue) weighting each land point in the preferred Giorgi region twice as much as those outside and (red) weighting the Giorgi region as much as the rest of the world. The height represents the relative magnitudes of the coefficients for each coordinate of injection; and each column represents prioritising a different Giorgi region. Horizontal lines denote the coefficient that minimise the global RMS temperature anomaly over land.

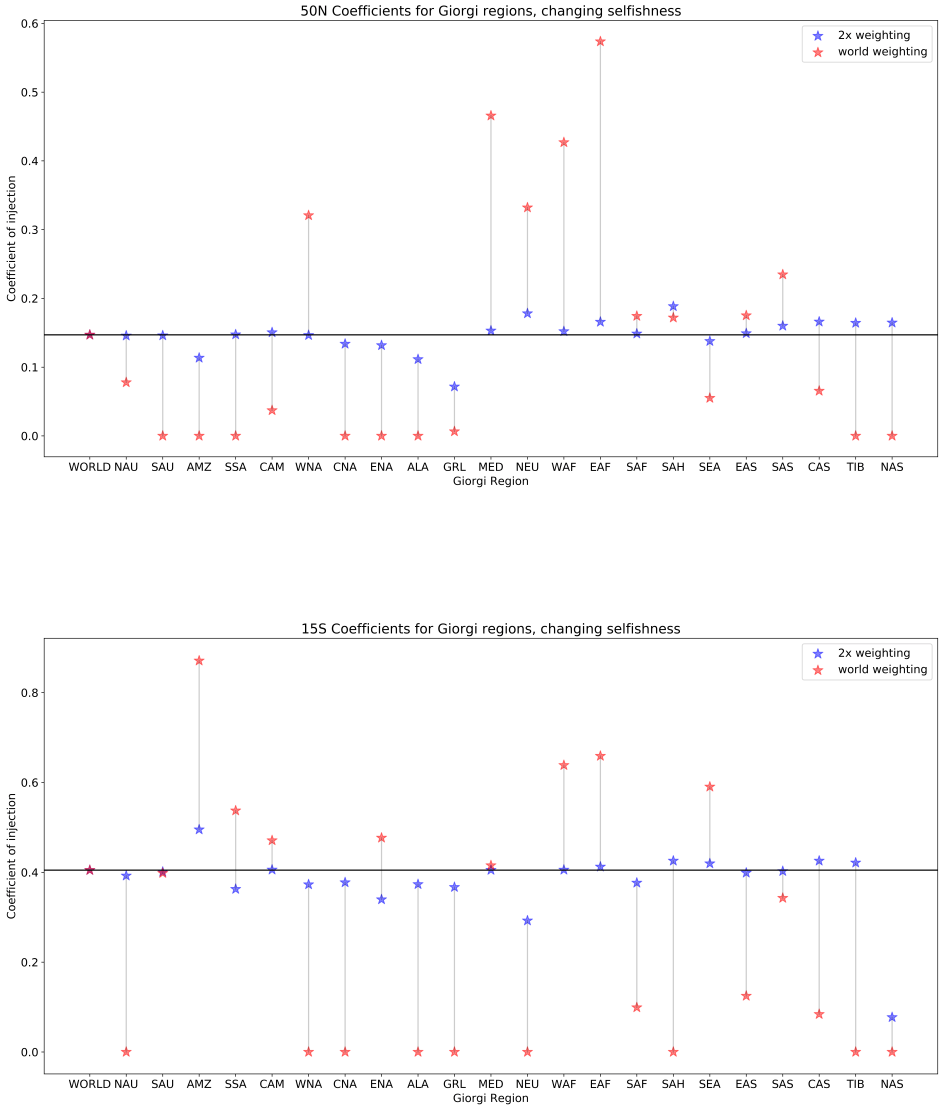

Figure 4.35: **Preferred coefficients for each of the Giorgi regions.** Assuming linearity, we attempt to use the GLENS basis linear optimiser to minimise the RMS anomaly for temperature on land in different Giorgi regions. Top (50N), bottom (15S). The different stars indicate (blue) weighting each land point in the preferred Giorgi region twice as much as those outside and (red) weighting the Giorgi region as much as the rest of the world. The height represents the relative magnitudes of the coefficients for each coordinate of injection; and each column represents prioritising a different Giorgi region. Horizontal lines denote the coefficient that minimise the global RMS temperature anomaly over land.

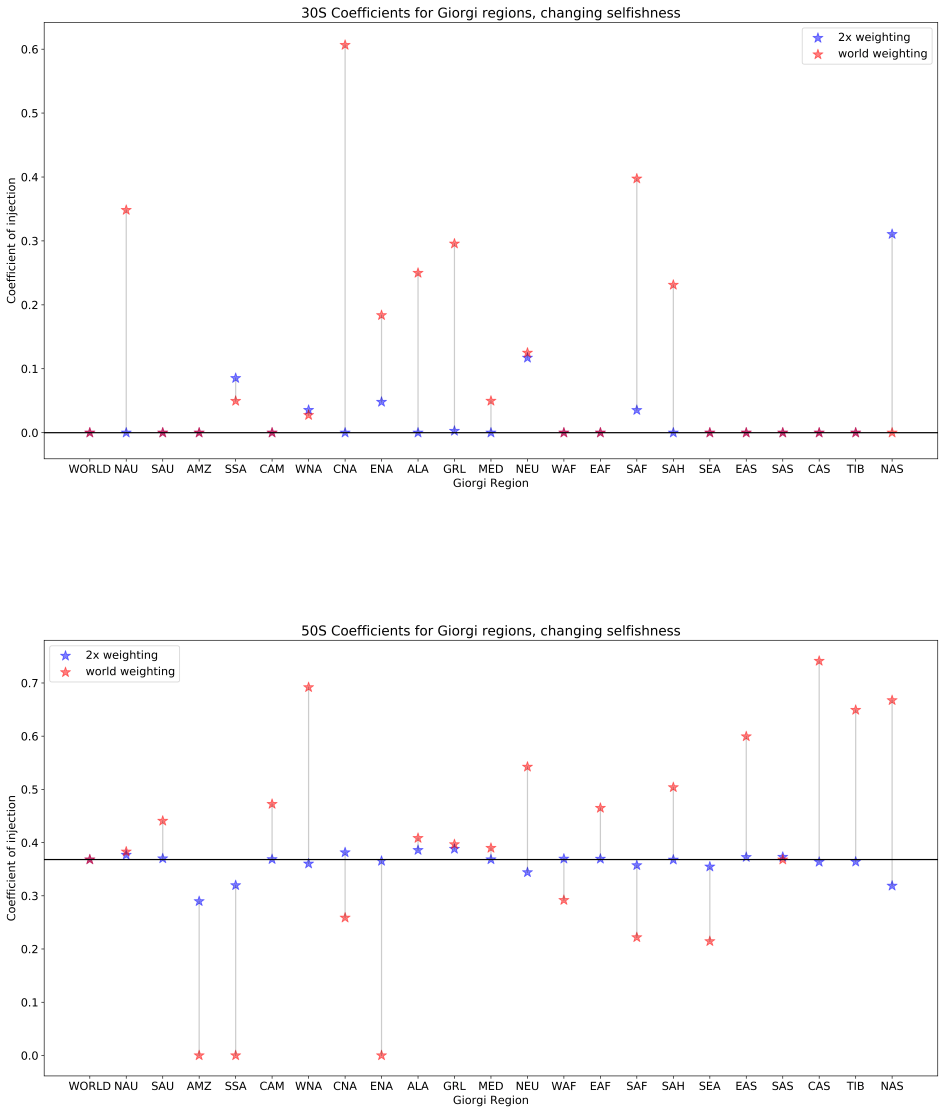

Figure 4.36: **Preferred coefficients for each of the Giorgi regions.** Assuming linearity, we attempt to use the GLENS basis linear optimiser to minimise the RMS anomaly for temperature on land in different Giorgi regions. Top (30S), bottom (50S). The different stars indicate (blue) weighting each land point in the preferred Giorgi region twice as much as those outside and (red) weighting the Giorgi region as much as the rest of the world. The height represents the relative magnitudes of the coefficients for each coordinate of injection; and each column represents prioritising a different Giorgi region. Horizontal lines denote the coefficient that minimise the global RMS temperature anomaly over land.

Chapter 5

The impact of geoengineering on climate and weather extremes

5.1 Introduction

Examining changes to the mean climatology under SRM gives an incomplete picture of the potential societal and environmental impacts of geoengineering. Many of the most costly and damaging impacts expected to arise from anthropogenic climate change owe to the change in spatio-temporal distribution, severity, and likelihood, of the extreme weather events that result from a changing climate. The large ensembles of climate models provided by CPDN allow us to investigate the ways in which SRM changes the likelihood and probability distribution for extreme weather events.

In this way, we can look beyond global mean changes and ask whether SRM designed to return global mean temperature to its preindustrial level also restores the preindustrial distribution of extreme weather events, and whether the variability of weather and climatic extremes is affected by SRM. For example, are dry spells or cold spells more likely in a climate stabilised by SRM than in the preindustrial climate that is targeted by an intervention?

One question to explore is to what extent changes to temperature and precipitation extremes can be attributed to the residual changes to the mean climate under SRM we explored in Chapter 4. For example, a null hypothesis to test for SRM's effect on

extremes might be that the distribution of temperature extremes in a grid cell under SRM is simply a translation of the preindustrial distribution, depending on the mean climatic changes in that grid cell. Alternatively, SRM may also affect the variability of the extreme distribution, with implications for the frequency of extreme events, or the relative width of the extreme event probability distribution.

5.2 Methods

5.2.1 Metrics for extreme events

In order to investigate this, we select a number of extreme temperature and precipitation metrics available from HadCM3 model output to stand in as proxies for extreme weather events of interest. Specifically, we choose three metrics for initial analysis:

- Three-day average daily temperature anomaly (analysing the annual summer (JJA) maxima, as a proxy for heat-waves).
- Daily precipitation (analysing annual maxima as a proxy for intense rainfall events and flooding).
- Monthly precipitation (analysing annual minima as a proxy for drought and dry spells).

Data for the uniform SRM, 4xCO_2, and piCO2 cases are generated in the same way as described in Chapters 3 and Chapter 4 for the HadCM3 models under CPDN. By the time this analysis was conducted, a greater number of model years were available, which allowed us to analyse 1079 model years for the preindustrial simulation and 614 model years for the SRM simulation, although there are different numbers of extreme events available because some quantities are daily, some are 3-day averages over JJA, and some are monthly mean outputs.

In each case, before the analysis begins, we detrended the data by subtracting the appropriate smoothed annual or seasonal cycle from the average over preindustrial simulations in order to obtain detrended, deseasonalised anomalies for the given metrics. The SRM data was also detrended by subtracting the average preindustrial

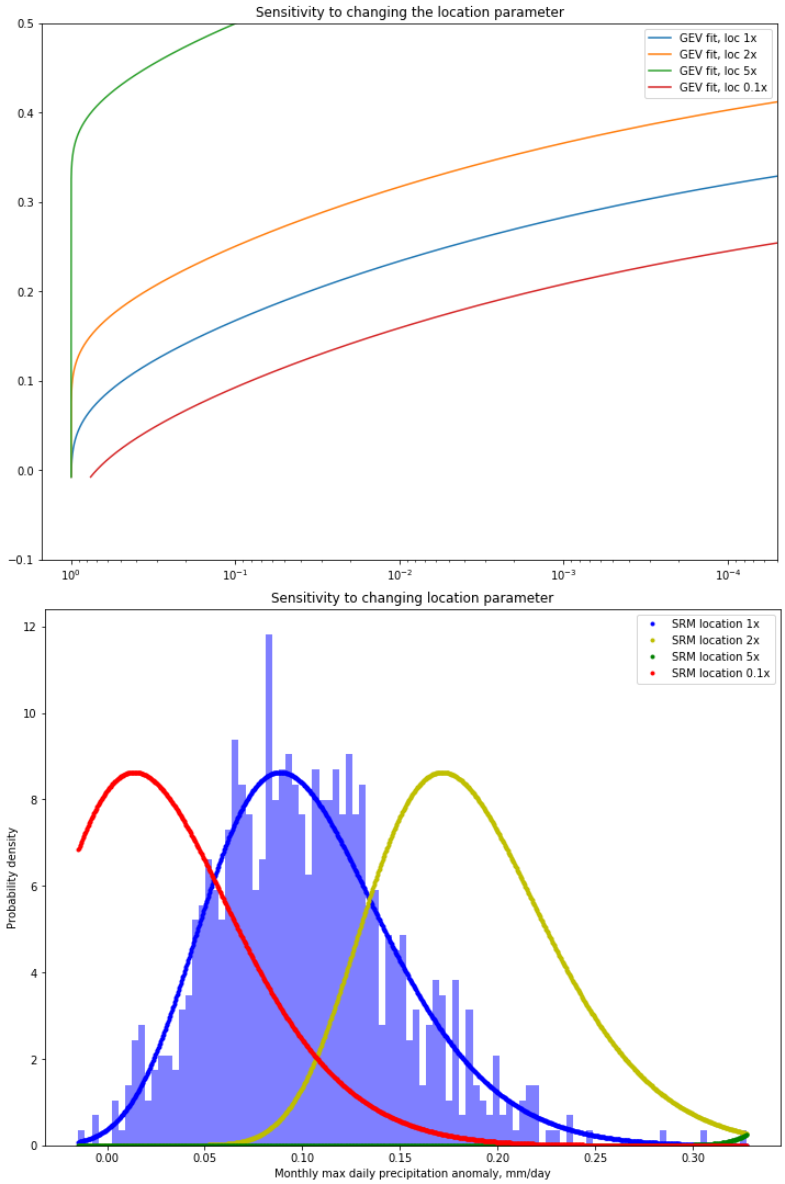

Figure 5.1: **Illustrative extreme values analysis and return times plots, illustrating the effects of varying the location parameter.** Bottom: Extreme values analysis histogram plot, based on annual maximum daily precipitation under SRM, showing the full probability distribution function for the GEV distribution. Top: Quasi-return time plot, based on annual maximum daily precipitation under SRM, showing the cumulative distribution survival function for the GEV fit. Both plots illustrate the effects of varying the generalised extreme value distribution's location parameter - translating the distribution along the axis of the extreme variable.

annual cycle for each of the given variables. We detrend using the preindustrial seasonal cycle for the SRM case in order to capture any unusual shifts to the seasonal

cycle. This deseasonalisation is necessary in order to make sure, for example, that when we study extremes in daily precipitation, we are not just analysing the monsoon season in monsoon grid cells where most of the rain falls within a few days of the year, and that we are actually determining the most significant extreme anomalous events for a given grid cell relative to the oridinary annual cycle. For example, we might not analyse the driest month in each year, but we will analyse the month with the biggest dry anomaly in each year.

In selecting these metrics, we have to take into consideration the strengths and weaknesses of our model and the diagnostics it can return.

For example, consecutive dry days is typically used as a proxy for drought [165], but HadCM3 - along with many other models of its generation - suffers from the "drizzle problem", where days with low levels of precipitation are over-represented in the distribution of simulated results. [166]

Consequently, we determined that the consecutive dry days metric is inappropriate for this model as it will underestimate the prevelance of drought and precipitation lows. This guided our selection of monthly low precipitation as a proxy for drought and extreme negative precipitation anomalies.

For each extreme metric, we analyse the data across available simulation results in several ways. First, we plot histograms and quasi-return time plots for the metrics in a sample grid cell which we know the HadCM3 model simulates well. For the purpose of this exercise, we chose the (52.5N, 0E) grid cell over the United Kingdom. Then, we fit these distributions to a generalised extreme value (GEV) distribution (see section below); analysis of the fit parameters allows us to determine how and whether the extreme distributions have changed under the influence of SRM.

The quasi-return time plots plot exceedance thresholds for the extreme metric against the probability of that threshold being exceeded within a given time period. For example, in the case of three-day temperature, the x-value (probability) can be interpreted as the probability of observing a three-day temperature exceeding the y-value threshold temperature in any given three-day interval.

This is similar to a return time plot. The difference is that return time plots account for autocorrelation in the data - either because they are based on an ensemble

of simulations of the same time period, rather than a generic year, or by removing autocorrelations - for example, by requiring individual classifications of events to be non-adjacent. For example, if a given heatwave is characterised by six days of excessively high temperature anomalies - and a similar event does not occur for another thirty years - then the appropriate return time for such an extreme event is thirty years. However, calculating the annual maximum of a three-day average temperature metric, as we do, might fail to take into account that two of the threshold exceedance events are adjacent to each other, and are part of the same event. Therefore our plots are more properly expressed in terms of probability that neglects any autocorrelation, rather than return time.

This analysis, including the GEV fits, can then be repeated for the values of the extreme metrics in each HadCM3 model grid cell. We can then plot maps of the three parameters for the GEV fit - location, shape, and scale - to illustrate how the distribution of extreme events has changed in different locations. This allows us to investigate whether the spatial pattern of changes to extremes significantly differs from the changes to mean quantities investigated in Chapter 4, and also whether the distribution of extreme events is significantly different in a world where global temperatures are stabilised by SRM but CO_2 concentrations remain high, as was found by Baker et al. (2018) [32]. Potential physical interpretations for changes to the GEV parameters are explored in the next section.

5.2.2 Generalised extreme value distributions

Extreme value distributions are the limiting distribution for the extreme values (minima or maxima, taken over a given period) of a sufficiently large ensemble of values from an underlying initial distribution, which can be arbitrary. This will be the limiting distribution for the extreme values sampled from any underlying distribution which is continuous and invertible. [167]

The generalised extreme value distribution formula allows for the fitting of extreme value tails to standardised distributions which are characterised by three parameters - location, shape, and scale. First, we define a standard variable:

$$s = \frac{(x - \mu)}{\sigma} \tag{5.1}$$

where μ is the location parameter and σ is the scale parameter.

Then, we can express the GEV cumulative distribution function as follows. For a general shape parameter ξ and $\xi s > -1$:

$$F(s, \xi) = exp(-(1 + \xi s)^{-\frac{1}{\xi}}) \tag{5.2}$$

Where $\xi s < -1$, for positive ξ the cumulative distribution function is 0 and for negative ξ it is 1. In other words excessively negative s-values too far from the location parameter are not sampled, and by the time excessively positive s values are reached for a negative ξ we are also beyond the distribution.

The probability density function can be obtained by differentiating this function with respect to x. Finally, in the special case that $\xi = 0$ (Gumbel or extreme value Type I distribution), the cumulative distribution function is:

$$F(s) = exp(exp(-s)) \tag{5.3}$$

For the sake of illustrating how these different parameters affect the GEV distribution's functional form, we have produced a number of plots where the location (Figure 5.1), scale (Figure 5.2), and shape (Figure 5.3) parameters are varied individually, using as a basis the actual quasi-return time and histogram plots for annual

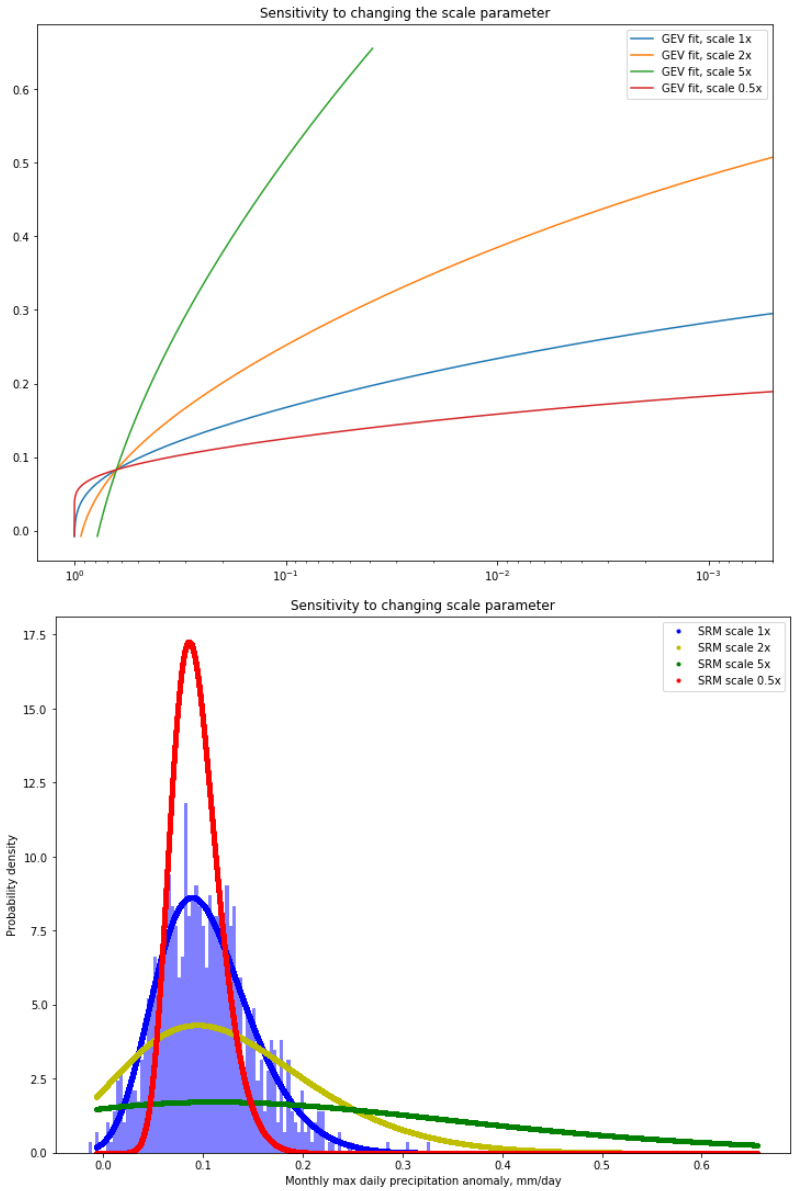

Figure 5.2: **Illustrative extreme values analysis and return times plots, illustrating the effects of varying the scale parameter.** Top: Quasi-return time plot, based on annual maximum daily precipitation under SRM, showing the cumulative distribution survival function for the GEV fit and the effects of varying the scale parameter. Bottom: Extreme values analysis histogram plot, based on annual maximum daily precipitation under SRM, showing the full probability distribution function for the GEV distribution and the effects of varying the scale parameter. The scale parameter represents the typical width of the extreme value distribution - smaller scale parameters give a probability distribution more tightly concentrated around the median.

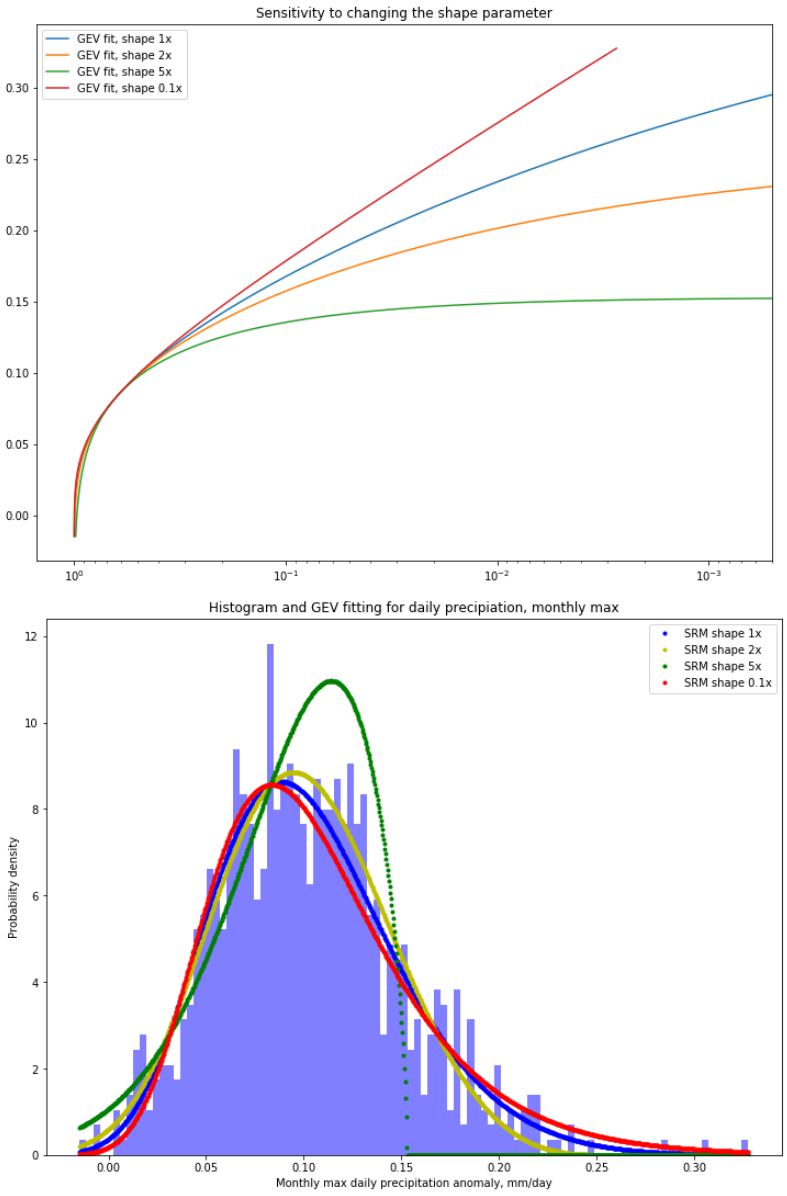

Figure 5.3: **Illustrative extreme values analysis and return times plots, illustrating the effects of varying the shape parameter.** Top: Quasi-return time plot, based on the curve for annual maximum daily precipitation under SRM, showing the cumulative distribution survival function for the GEV fit and the effects of varying the shape parameter. Bottom: Extreme values analysis histogram plot, based on annual maximum daily precipitation under SRM, showing the full probability distribution function for the GEV distribution and the effects of varying the shape parameter. The shape parameter determines the shape of the extreme distribution's tail; smaller shape parameters mean a fatter-tailed probability distribution, while larger shape parameters indicate a more truncated tail - in the limit of a shape parameter much larger than 1, the tail of the distribution shifts to the other side.

maximum precipiation (the 1x location parameter plots in Figure 5.1) show these plots.)

5.2.3 Varying parameters

With these basis plots in mind, we can illustrate for the sake of clarity the effect of varying each of the parameters in the GEV distribution, using as an example function the SRM extreme event distribution for annual maximum daily precipitation.

As can be observed from the functional form of the distribution, changing the location parameter translates the curve along the x-axis. Shifts to the location parameter indicate that the average of the maxima have shifted - this represents "climatological" changes to the extremes; all else being equal, the relative distribution is unchanged.

The scale parameter determines the width of the distribution, giving the typical x-axis scale of the probability density function. Smaller scale parameters mean less variability in the extreme distribution, with distributions of extreme values more tightly concentrated around their mean, while larger scale parameters indicate the distribution has a much larger variability.

Increases to the shape parameter corresponds to a sharper cut-off in the extremes distribution. The shape parameter broadly determines the shape of the tail – a lower shape parameter means fatter tail for the probability distribution, with rare events more likely, while shifting the in the opposite direction; a higher shape parameter gives a more truncated tail and, in extremis, shifts the tail to the other side.

It is worth noting that while the scale and location parameters share the units of the underlying variable in question, the shape parameter, governing the tail behaviour, is dimensionless. The generalised extreme value distribution unites the Gumbel, Frechet, and Weibull extreme value distributions, which correspond to the case where the shape parameter is 0, positive, and negative respectively (sometimes also referred to as Type I, II, and III extreme value distributions).

In the following sections, we conduct GEV fitting for the three extrema proxies - annual three day temperature maxima, annual maximum daily precipitation, and annual minimum monthly precipitation - in the UK grid cell, and then in each grid cell,

plotting maps of the parameters. The applicability of GEV fitting and the importance of any discernible shifts to the GEV fit parameters, and hence the distribution of extreme events, are discussed.

5.3 Results

5.3.1 Three-day temperature maxima

5.3.1.1 Single grid-cell generalised extreme value fitting

In this section, we examine the distribution of deseasonalised summer maxima in three-day temperature anomalies (as a proxy for heatwaves) under piCO2 and uniform SRM. For the illustrative plots in Figures 5.4 and 5.5, the 3-day temperature maxima are taken from a sample grid cell - the (0.0, 52.5) HadCM3 grid cell over the United Kingdom.

The data used are the distribution of summer (JJA) maxima for the three-day temperature distribution, after the data has been de-seasonalised by subtracting the average seasonal cycle from HadCM3's preindustrial annual runs. The first two plots show the extreme value distribution for this quantity; for the maps, we fit GEV distribution parameters in each HadCM3 grid cell. In total, we used 1079 model years for the preindustrial simulation and 614 model years for the SRM simulation. The parameters required to fit the 3-day maximum temperature parameter to a GEV distribution in the preindustrial and SRM cases are tabulated in Table 5.1.

Predominantly, the difference in the three-day temperature distribution in this grid cell is a shift in the location parameter by approximately 1.6K - this follows the "climatological" change to the extremes, suggesting the residual warming under the SRM scenario is the most significant factor in changing the distribution of extremes. The residual warming in summer in this grid cell under SRM was 1.29K for the same ensemble - so, although the extreme (maximum) distribution shift slightly exceeds the mean climatological shift, they are on a similar scale. However, we must note that the extreme values are not a perfect fit to the GEV distribution - for example, although most points are within the 5 - 95% confidence interval for the GEV parameter fits, the events around 1 in 50 years are slightly underestimated by the GEV fit. We also note that the shape parameter increases by around 0.05 and becomes positive in the SRM case compared to the piCO2 case, illustrating a more truncated tail and higher curvature in the SRM case - suggesting that the most extreme temperature events

in this grid cell in summer are suppressed by the uniform SRM, even as the average 3-day temperature maximum of a summer is 1.6K warmer.

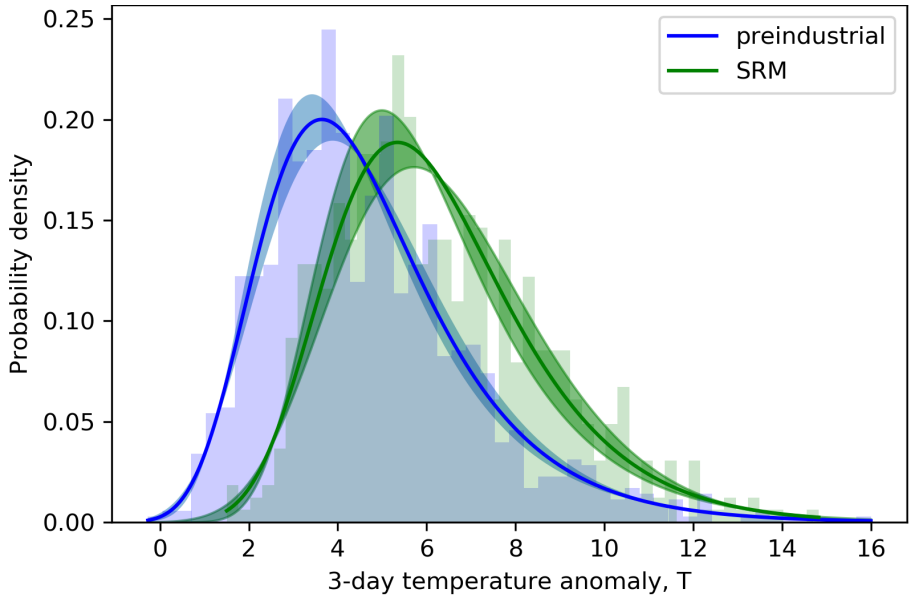

Figure 5.4: **Extreme values analysis histogram plot for 3-day summer temperature maxima in the UK.**. The data shown is for the 52.5N, 0E grid cell, in the piCO2 and SRM cases. The GEV fit is illustrated with solid lines, with the 95% fit parameter confidence interval shown with shading.

Parameters	piCO2	SRM
Location	3.64	5.24
Shape	-0.003	0.050
Scale	1.839	1.952

Table 5.1: **Extreme value distribution parameters.** Changes to the fitted generalised extreme value distribution parameters between piCO2 and SRM for the annual maximum summer 3-day temperature anomalies in the United Kingdom grid cell.

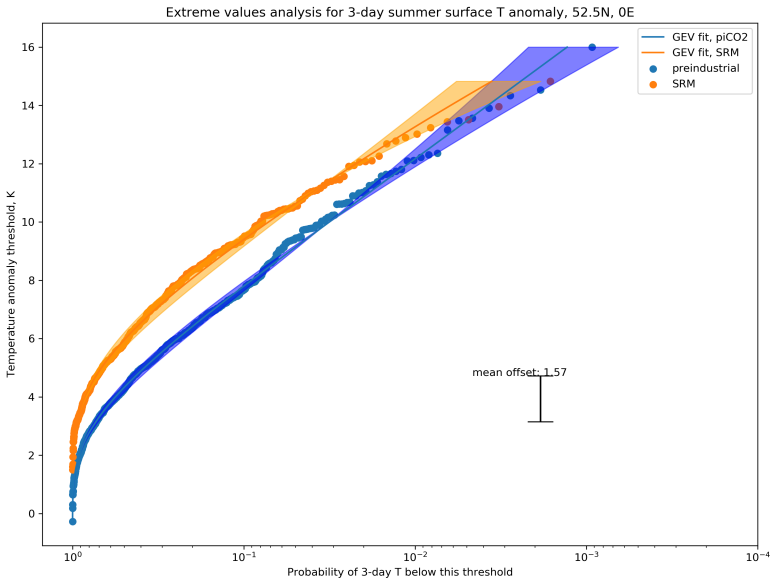

Figure 5.5: **Quasi-return time plot showing the survival function for annual three-day summer temperature maxima**. The data shown is the three-day summer temperature maxima for the 52.5N, 0E grid cell, in the piCO2 and SRM cases. The GEV fit is illustrated with solid lines, with the 95% fit parameter confidence interval shown with shading. The mean offset between the average extreme value under piCO2 and SRM is plotted on the figure.

5.3.1.2 Global maps of GEV parameters

In this subsection and the corresponding subsections, we construct global maps of the parameters for the GEV fits to our extreme event proxies. The GEV fitting procedure is applied to data from each grid cell. Since this produces 6,912 GEV fits, it is necessary to create filters to determine the applicability of GEV fitting in each grid cell.

This was done using a "residuals test". For each grid cell, we calculated the fit residuals for events that were rarer than one in ten years, normalised by the average extreme value in the grid cell. This allows us to determine whether GEV fitting is applicable for the tail end of rare events in this distribution. When the average of these normalised fit residuals exceeded 0.1, GEV fitting is not considered to be applicable. This is the equivalent of saying that the average rare event differs from the GEV fit prediction by more than 10% of the average annual extremum for that

grid cell, so GEV fitting is a poor predictor for these rare events. Grid cells which fail the residual test are denoted by grey cells in the parameter maps.

As before, location parameter for 3-day T summer extremes follows the regional residual mean summer temperature changes under SRM, analysed extensively in Chapter 4, quite well - in Figure 5.6, we see substantial warming across the Arctic landmasses and Northern hemisphere landmasses, with overcooling of the tropical oceans especially. Similar contrasts between China and India temperature extreme responses - cooling in India, warming in China - and increases in temperature extremes in tropical rainforests in the Amazon and the Congo, are also seen in the extremes shift to location parameters as in the mean temperature response.

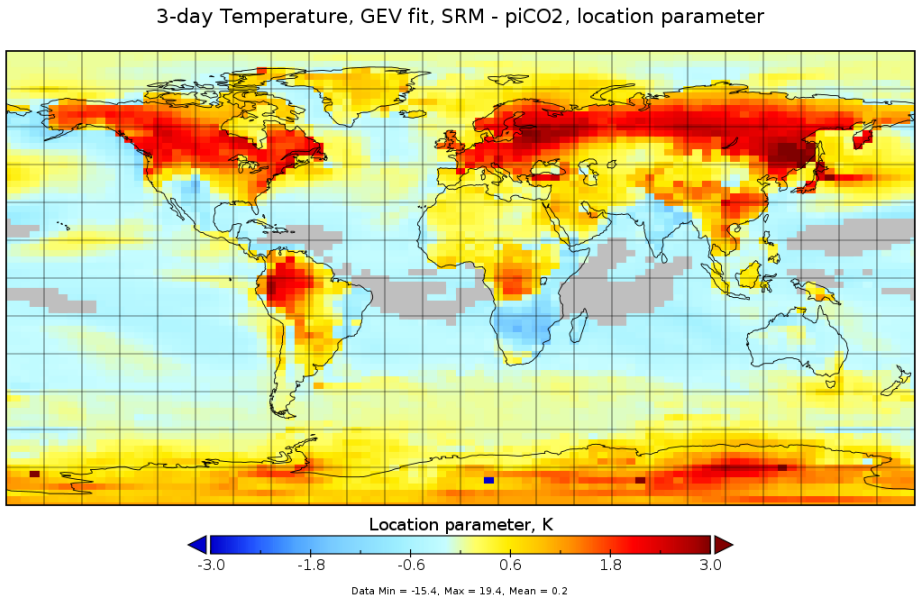

Figure 5.6: **Map of location parameters for GEV fits to 3-day summer temperature anomaly annual maxima, fitted in each grid cell.** The map illustrates the difference between the SRM and piCO2 simulations. Grid cells marked in grey are "missing data" where there was a poor fit to the generalised extreme value distribution, defined by the fit residuals exceeding a certain threshold.

As described, increases to the shape parameter means more abruptly truncated tails, while decreases mean fatter tails relative to the rest of the distribution. When analysing the variation in the shape parameter in Figure 5.7, we can see a much less coherent set of changes here, but in general we see increases in the shape parameter under SRM - particularly in regions that saw high shifts to the location parameter,

such as the Amazon, the Congo, China, and Central Europe. This suggests that even as these regions warm, and the average summer maximum temperature increases, tail events that reach even hotter temperatures are suppressed relative to the preindustrial case, following from the narrower distribution and smaller variability in extremes under SRM relative to piCO2.

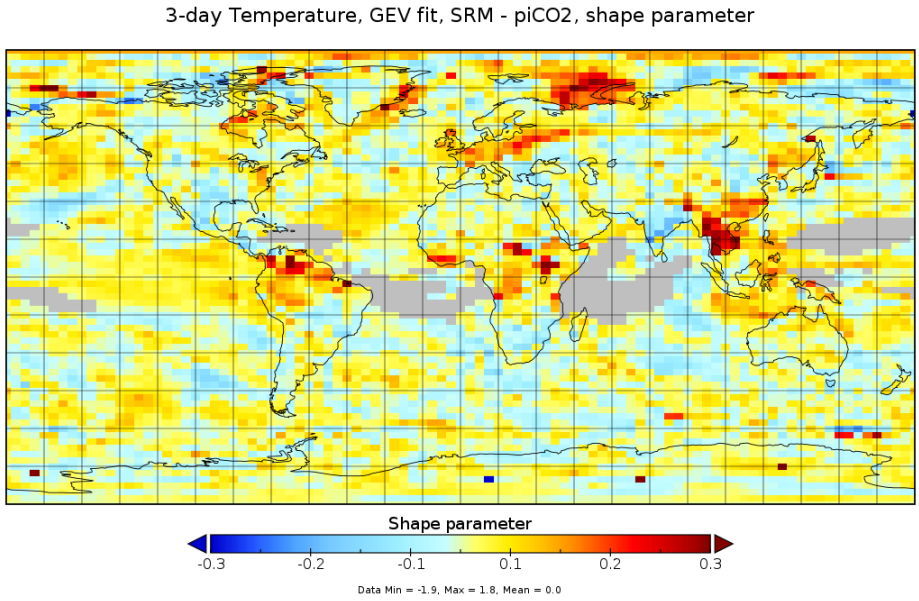

Figure 5.7: **Map of shape parameters for GEV fits to 3-day summer temperature anomaly annual maxima, fitted in each grid cell.** The map illustrates the difference between the SRM and piCO2 simulations. Grid cells marked in grey are "missing data" where there was a poor fit to the generalised extreme value distribution, defined by the fit residuals exceeding a certain threshold.

In Figure 5.8 we can see a general decrease of the scale parameter across most of the globe, particularly over land (although to a lesser degree at high latitudes) this suggests SRM generally suppresses extreme temperature variability while shifting the mean (with narrower temperature anomaly distributions concentrated around a higher mean owing to the shift in the location parameter). A notable exception is the Indian subcontinent, where we see suppression of temperature extremes in numerous ways: the average temperature extreme is cooler, the tails are slightly truncated, and the overall scale of the extreme distribution is reduced with large negative anomalies to the scale parameter under the SRM case compared to the preindustrial. Summer temperatures in this region are therefore both significantly colder and less variable

overall. We can also see that there is a limited correlation between the variations in the location and the scale parameters for temperature; in other words, the effect of SRM on extremes is not simply to increase the magnitude and variability of temperature extremes, as we would expect unmitigated global warming to do - even within the regions, like the Arctic, which see residual warming and shifts to the location parameter for the temperature extreme distributions.

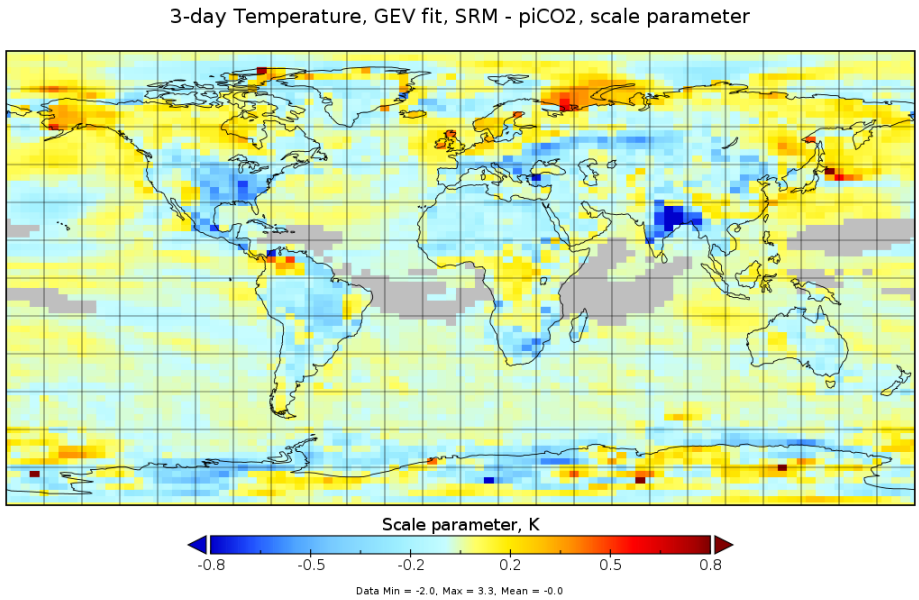

Figure 5.8: **Map of scale parameters for GEV fits to 3-day summer temperature anomaly annual maxima, fitted in each grid cell.** The map illustrates the difference between the SRM and piCO2 simulations. Grid cells marked in grey are "missing data" where there was a poor fit to the generalised extreme value distribution, defined by the fit residuals exceeding a certain threshold.

5.3.2 Daily precipitation maxima

5.3.2.1 Single grid-cell generalised extreme value fitting

The cumulative distribution function and probability distribution function histogram comparison for annual daily precipitation maxima were plotted in Figure 5.1, the demonstration figure which illustrates the impact of varying GEV fit parameters, and are reproduced separately in Figures 5.9 and 5.10. The annual daily precipitation maxima are taken from a sample grid cell - the (0.0, 52.5) HadCM3 grid cell over the United Kingdom. In total, 1080 model years were used from the preindustrial simulations, and 840 model years from the uniform SRM simulations. For this particular grid cell, the location parameter is shifted down by approximately 1mm/day under SRM - the wettest annual days are drier on average under SRM - but the shape and scale parameters only show very slight decreases, with only a slight reduction to precipitation variability in the SRM case relative to the preindustrial distribution of maxima in this grid cell.

Parameters	piCO2	SRM
Location	17.804	16.784
Shape	0.022	0.020
Scale	3.793	3.590

Table 5.2: **Fitted generalised extreme value distribution parameters for daily precipitation maxima in the UK grid cell.**

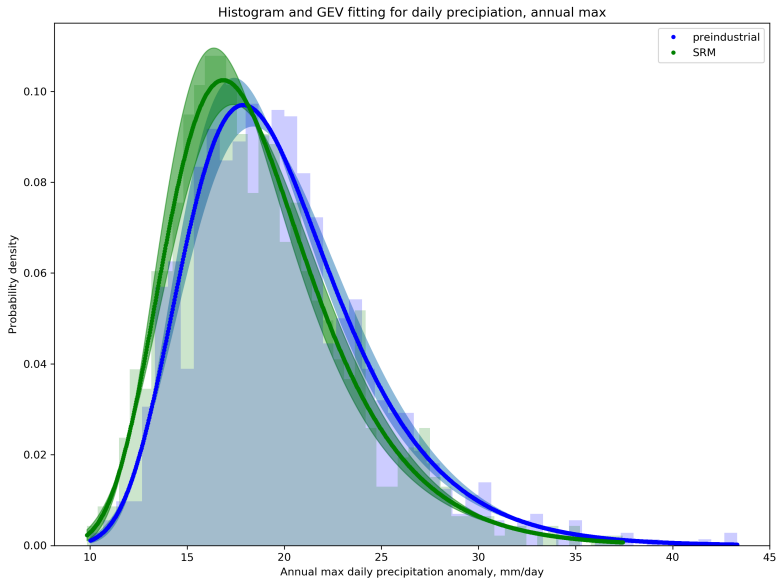

Figure 5.9: **Extreme values analysis plot for annual daily precipitation maxima in the United Kingdom grid cell.** The figure shows the full probability distribution function for the GEV distribution as well as the different distributions under preindustrial and SRM scenarios. Shading illustrates the 95% confidence interval for the GEV fit parameters.

5.3.2.2 Global maps of GEV parameters

Examining changes to the location parameter (Figure 5.11), we can see that this broadly follows the mean drying pattern for precipitation under uniform SRM – we see reductions in the location parameter, especially in the overcooled and overdried tropics, and to a lesser extent at high latitudes. We also see the contrast in precipitation response between India and China, where India sees a general increase in precipitation and China a general decrease under uniform SRM.

The scale parameter changes (Figure 5.12) for annual maximum daily precipitation under SRM vs piCO2 correspond closely to the precipitation location parameter changes. This suggests that, in general, the effect on daily maximum precipitation is for its average to shift to a greater value as its variability increases. Across many of the regions that are dried by uniform SRM, therefore, we see reductions to the location and scale parameter, suggesting that the average annual precipitation maximum is reduced and the variability of maxima suppressed by SRM in these regions,

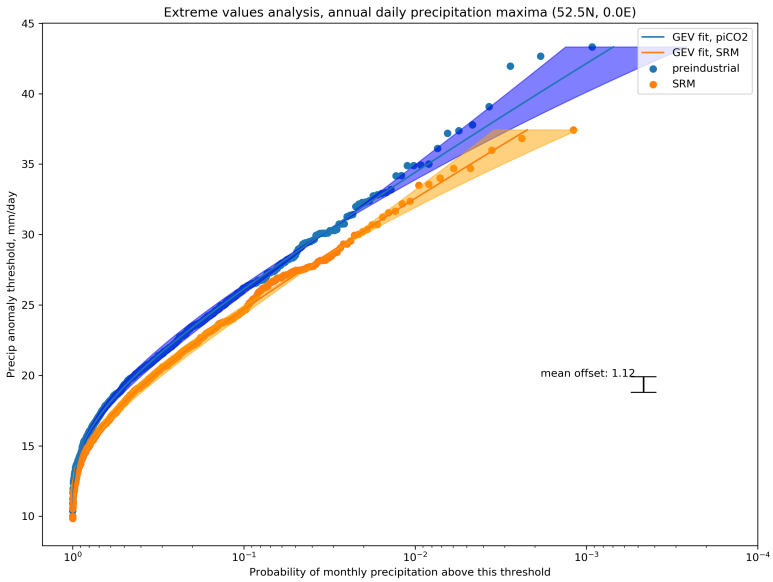

Figure 5.10: **Quasi-return time plot for annual daily precipitation maxima in the United Kingdom grid cell.** The figure shows the data and GEV survival function for the GEV distributions under preindustrial and SRM scenarios. Shading illustrates the 95% confidence interval for the GEV fit parameters.

in accordance with an overall suppression of the intensity of the hydrological cycle in these regions. However the scale parameter changes are not particularly large for most of the grid cells - while the location parameter shifts are significant for almost every grid cell.

Changes to the shape parameter (Figure 5.13) are harder to interpret and arguably less spatially coherent; nor do they correlate particularly consistently with changes to the location and scale parameters. For example, the shape parameter noticably increases, suggesting more abruptly truncated tails, for regions like Australia which get wetter overall on average. The Sahel and Northern Amazon both see fatter tails with a reduced shape parameter, suggesting an increased probability for extreme wet precipitation events within the shifted precipitation distributions, but the Sahel gets wetter while the Amazon is dried out under uniform SRM.

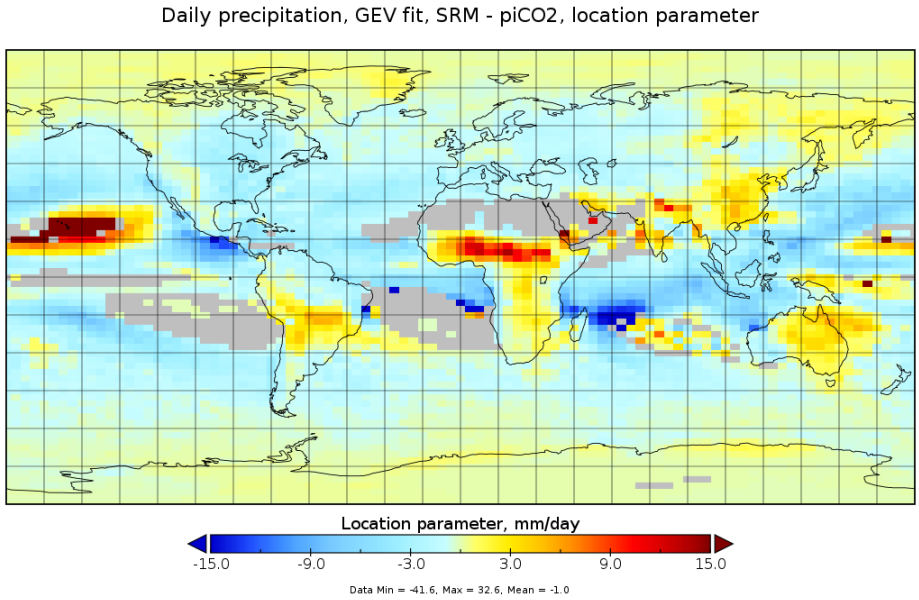

Figure 5.11: **Map of location parameters for GEV fits to annual daily precipitation maxima, fitted in each grid cell.** The map illustrates the difference between the SRM and piCO2 simulations. Grid cells marked in grey are "missing data" where there was a poor fit to the generalised extreme value distribution, defined by the fit residuals exceeding a certain threshold.

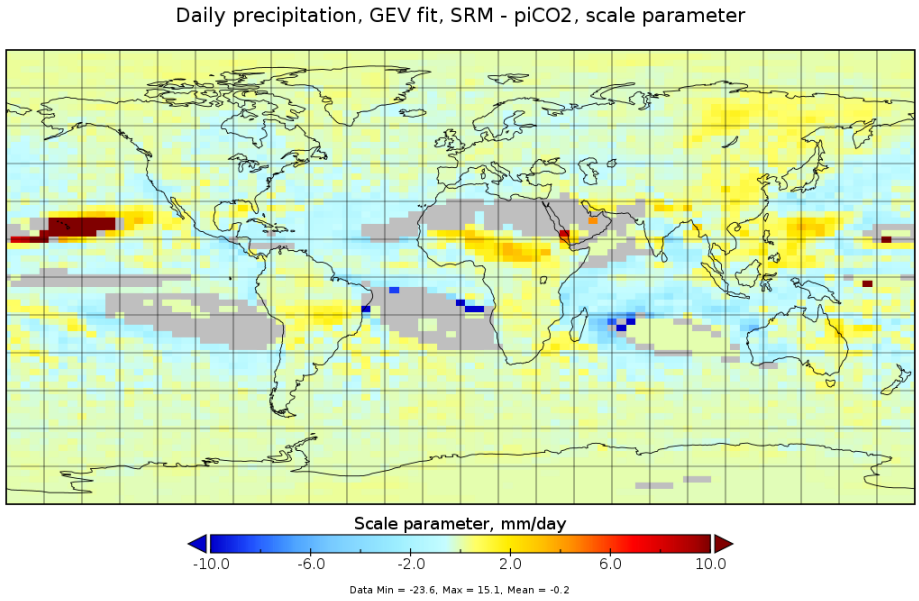

Figure 5.12: **Map of scale parameters for GEV fits to annual daily precipitation maxima, fitted in each grid cell.** The map illustrates the difference between the SRM and piCO2 simulations. Grid cells marked in grey are "missing data" where there was a poor fit to the generalised extreme value distribution, defined by the fit residuals exceeding a certain threshold.

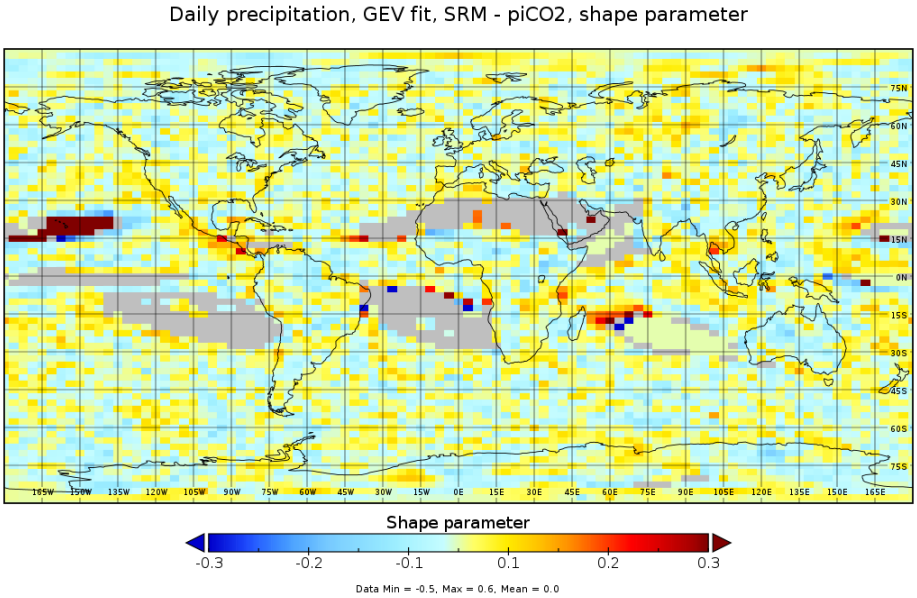

Figure 5.13: **Map of shape parameters for GEV fits to annual daily precipitation maxima, fitted in each grid cell.** The map illustrates the difference between the SRM and piCO2 simulations. Grid cells marked in grey are "missing data" where there was a poor fit to the generalised extreme value distribution, defined by the fit residuals exceeding a certain threshold.

5.3.3 Monthly precipitation minima

5.3.3.1 Single grid-cell generalised extreme value fitting

We can use the same generalised extreme value formula to fit distributions for minima simply by replacing x with -x and subtracting cumulative probability functions from 1. Again, we fit monthly precipitation annual minima from a sample grid cell - the (0.0, 52.5) HadCM3 grid cell over the United Kingdom. In total, 2075 model years are analysed for the preindustrial simulations, and 854 model years are analysed for the SRM simulations. We see that over this grid cell, the main impact of SRM on the extremes is a reduction in the location parameter, and hence the average driest month of a year gets drier - as would be expected with the general drying over this region induced by uniform SRM. There is also a slight increase to the shape parameter and a negligble decrease to the scale parameter (Table 5.3). The histogram and fit (Figure 5.14) and quasi-return time plot (Figure 5.15) are below.

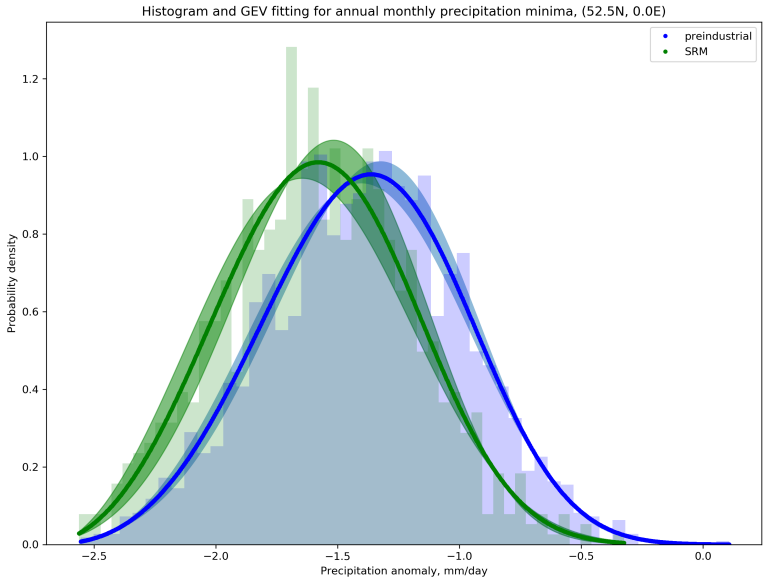

Figure 5.14: **Extreme values plot for annual monthly precipitation minima.** The plot illustrates precipitation minima in the piCO2 and SRM cases, in the United Kingdom grid cell, with the GEV fit shown. Shading illustrates the 95% confidence interval for the GEV fit parameters.

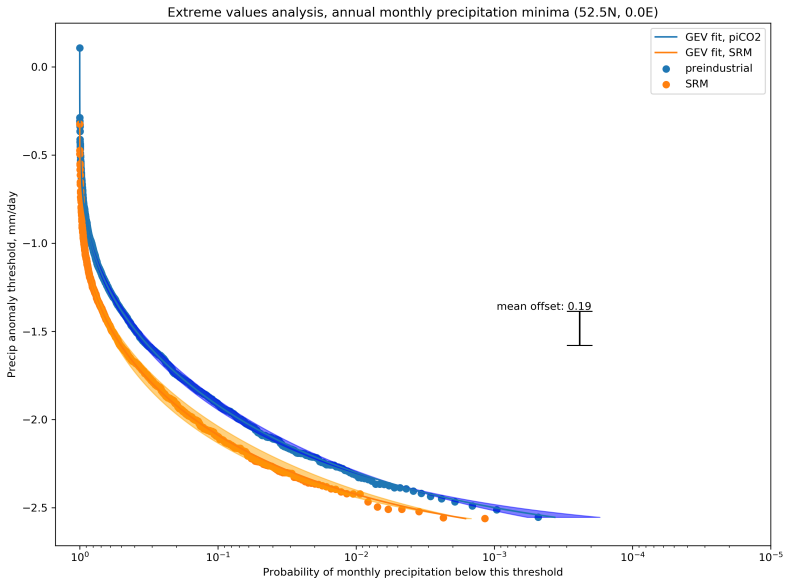

Figure 5.15: **Quasi-return time plot for annual monthly precipitation minima.** The plot shows the survival function for the GEV fit of annual monthly precipitation minima in the piCO2 and SRM cases, in the United Kingdom grid cell. The GEV fit is shown with solid lines. Shading illustrates the 95% confidence interval for the GEV fit parameters.

Parameters	piCO2	SRM
Location	-1.243	-1.446
Shape	0.270	0.301
Scale	0.401	0.393

Table 5.3: **Generalised extreme value distribution parameters for minimum monthly precipitation.**

5.3.3.2 Global maps of GEV parameters

For the dry month minima, residuals for the annual driest monthly precipitation extreme events are fitted in the same way as described in subsection 5.3.2 - using residuals for events rarer than 1 in 10 years, and normalising by the average extreme value in this grid cell. Those cells with residuals that exceeded 0.1 are greyed-out in

the parameter maps. No additional filters were applied.

We can see from the map of location parameter changes for the dry month minima (Figure 5.16) that they broadly follow the mean drying pattern under SRM described in Chapter 4 – the location parameter is shifted downwards, especially in overcooled and dried tropics and at high latitudes owing to the thermodynamic component of the precipitation response, where precipitation is predominantly energy-limited and not moisture-limited and therefore sensitive to cooling. This results in marked decreases in precipitation at high latitudes where this energetic availability is the limiting factor.

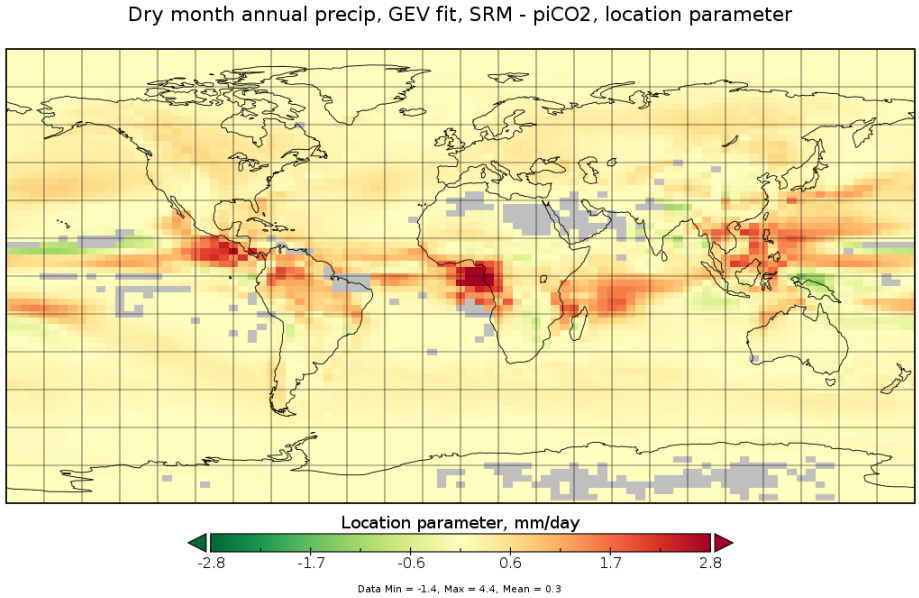

Figure 5.16: **Map of location parameters for GEV fits to annual monthly precipitation minima, fitted in each grid cell.** The map illustrates the difference between the SRM and piCO2 simulations. Grid cells marked in grey are "missing data" where there was a poor fit to the generalised extreme value distribution, defined by the fit residuals exceeding a certain threshold. The location parameter scale is inverted as we fit to -x, so red regions indicated drying under SRM

Examining the scale parameters in Figure 5.17, we can see that there are generally less coherent changes and smaller in magnitude changes to the precipitation distrbution's scale parameter. However, it arguably generally decreases, especially across the tropics – suggesting that there is less extreme variability in places where the shift to mean precipitation is predominantly drying.

This is also observed in the contrast between China and India as described for the location parameter. China's precipitation extremes get more variable on a drier

mean, while India's are narrower but focused around a wetter mean. For the same regions in the Congo basin where we see substantial drying, we also see a widening of the extreme distribution. In general, then, changes to the scale parameter seem to be anticorrelated with changes to the location parameter.

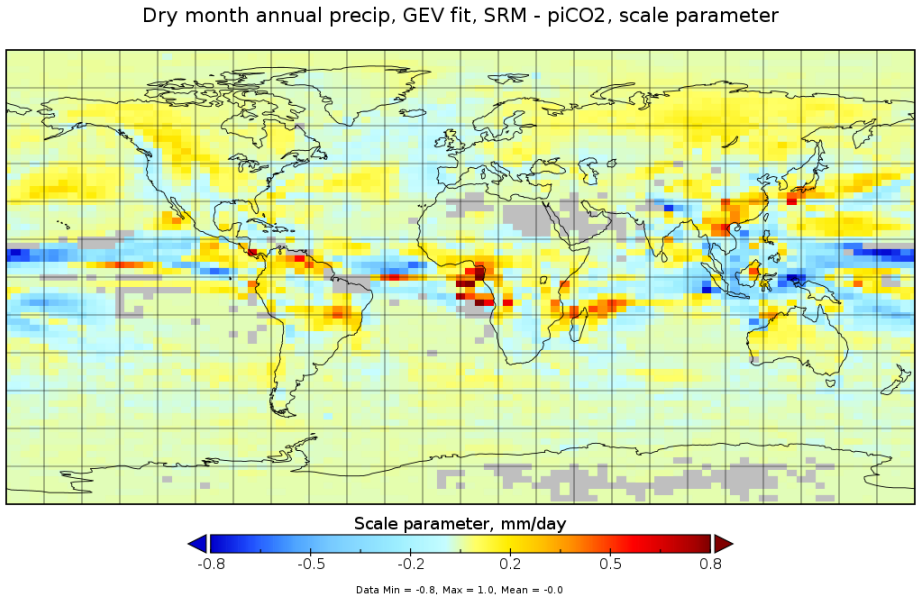

Figure 5.17: **Map of scale parameters for GEV fits to annual monthly precipitation minima, fitted in each grid cell.** The map illustrates the difference between the SRM and piCO2 simulations. Grid cells marked in grey are "missing data" where there was a poor fit to the generalised extreme value distribution, defined by the fit residuals exceeding a certain threshold.

Examining changes to the shape parameter for the GEV distribution for annual monthly minimum precipitation (Figure 5.18), we can see that again there are less coherent changes which are generally smaller in magnitude, and are not clearly related to the mean change to the precipitation distribution. However, there are generally decreases in shape parameter across most land regions and over the tropical Mid-Atlantic. This suggests fatter tails for the annual minimum in monthly precipitation under SRM compared to piCO2 - so, with shifts to the mean and variability of the distribution taken into account, tail (extreme dry month) events within these shifted distributions will be comparitively more likely to occur.

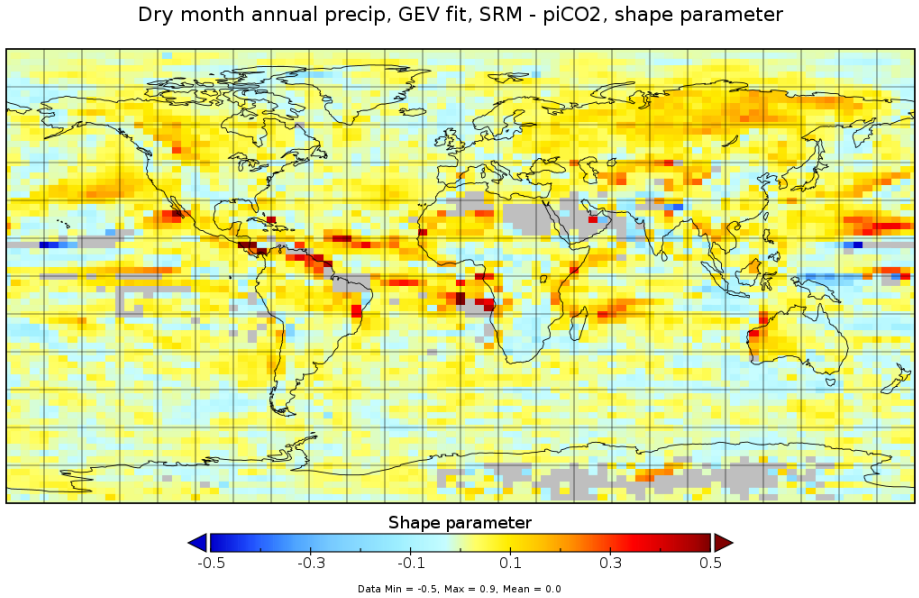

Figure 5.18: **Map of shape parameters for GEV fits to annual monthly precipitation minima, fitted in each grid cell.** The map illustrates the difference between the SRM and piCO2 simulations. Grid cells marked in grey are "missing data" where there was a poor fit to the generalised extreme value distribution, defined by the fit residuals exceeding a certain threshold.

5.4 Discussions and conclusions

5.4.1 Limitations

There are a number of limitations to the approach we have taken here. The limitations of HadCM3 as a model, in particular for extreme events as characterised by the drizzle problem, have been discussed in this Chapter and also in Chapter 3 and Chapter 4, so I do not propose to reiterate those arguments here.

There are some further limitations of our analysis in this chapter. It would have been preferable to conduct some statistical significance testing on the GEV parameter fits in order to determine whether the differences in each of the parameters between the preindustrial and SRM cases were statistically significant, particularly for the shape and scale parameters where the changes were often small and hard to interpret. This was attempted, and seemed to suggest that the changes were generally statistically significant for the location parameter, but not for the scale and shape parameters, for a significant proportion of grid cells in each of the three cases. However,

due to inconsistencies in the statistical significance tests, they have been omitted here. To indicate the nature of these significance differences, Figure 5.19 shows histograms produced by fitting the location and scale parameters for the 3-day T anomaly in a particular grid cell. There is significant overlap between the histograms for the scale parameter, indicating that these differences may not be statistically significant, while the location parameter clearly undergoes a very significant shift under SRM compared to the preindustrial case, which suggests that the changes to the location parameter are much more likely to be statistically significant than those for the scale parameter. To confirm that this holds true for most grid cells, however, would require additional significance testing. Of course, it would also be better to analyse even more data for each of the variables we considered, in order to have some more data for 1-in-100 year events and similarly uncommon but still politically relevant extremes in different grid cells.

Another limitation is in the metrics that we chose to analyse. It would be preferable if HadCM3 had a greater range of model output so that we could choose more recognisable extreme metrics - in particular, the driest month is not a particularly satisfactory metric for whether or not a drought is occurring. In this we were limited by the model output which we had the largest range of data for, which was daily temperature and precipitation outputs. It is also regrettable that analysis focused on the 3-day temperature maxima for June, July, and August. This might be said to accurately capture summer heatwaves in the Northern hemisphere, but of course it does not in the Southern hemisphere, where it instead analyses unusally warm periods in winter. A more comprehensive analysis would focus on seasonal extremes in summer and winter, and also examine extreme cold. Finally, maximum daily precipitation is not a perfect proxy for flooding, particularly in light of the effects of SRM on the hydrological cycle - it would have been preferable to use P - E, but again, we were unable to obtain daily evaporation data from HadCM3.

An additional concern arises from the means we used to exclude the grid cells where the GEV fit was not particularly effective. This was done in a relatively common-sense way by examining the residuals in each fit and their relative values, but this author was not able to find a formal goodness-of-fit metric which could be applied to the extreme value distributions that would have allowed us to exclude the

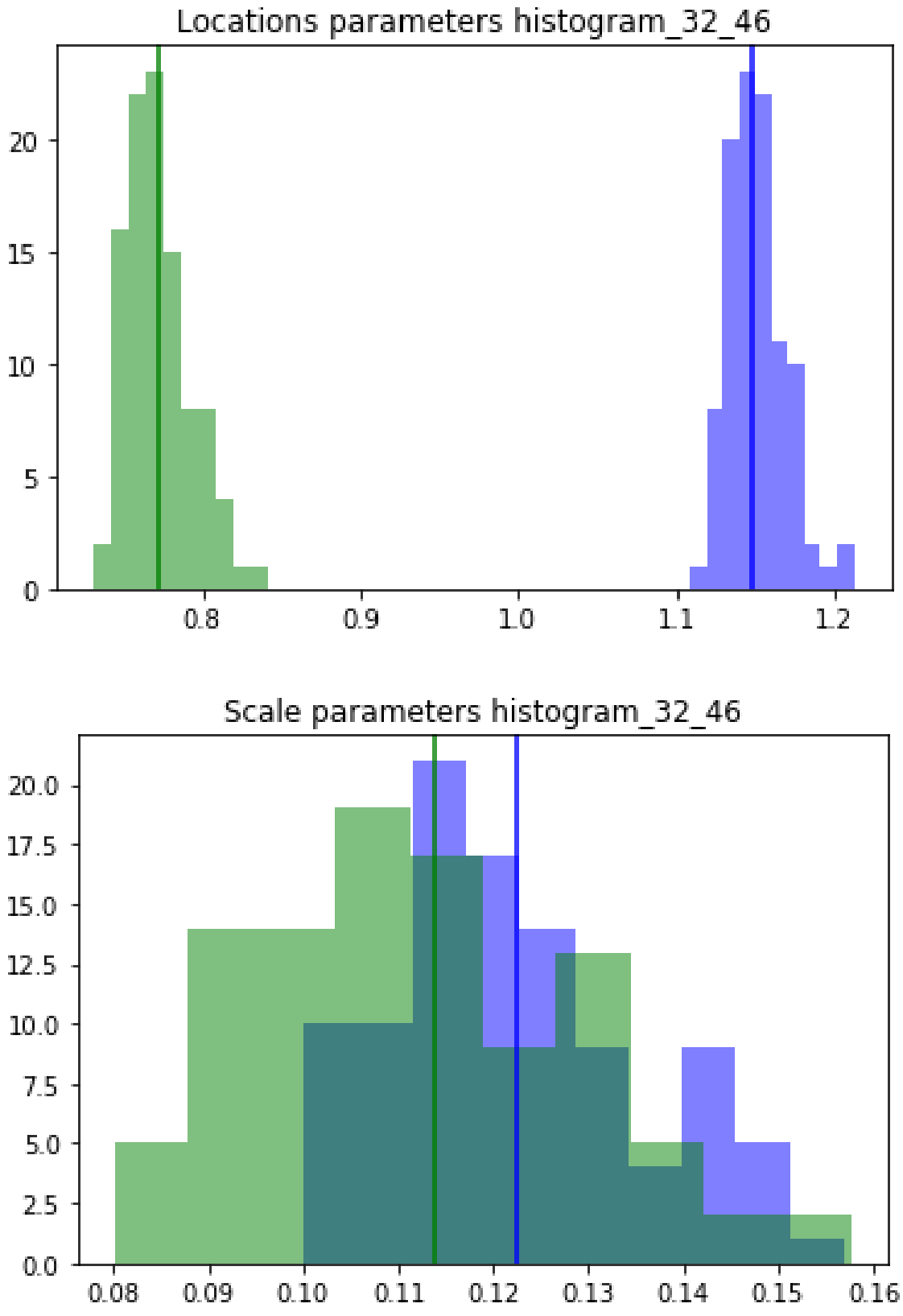

Figure 5.19: **Histograms of location (top) and scale (bottom) parameters for a single grid cell, as fitted for individual model runs. The green histogram shows the value of the parameters under SRM, while the blue shows the parameters under the preindustrial simulation.**

grid cells where no GEV fit was a particularly good distribution for the metrics that we examined.

One point that is not made in the analysis above, but we should make here, is that the analysis chose to focus on the differences between SRM and the preindustrial case. In this, we were trying to determine whether these extreme metrics were any worse under geoengineering that returned global mean temperatures to the preindustrial control, and whether geoengineering that superficially controlled global mean temperatures might still result in a greater frequency of extreme weather events (or, alternatively, if it might reduce them). However, it is worth noting that, compared to the 4xCO_2 that SRM aims to compensate for, SRM generally significantly reduced the incidence of temperature extremes, even in regions that saw more extreme temperature events under SRM than the preindustrial case. Figure 5.20 as an example shows you the very significant differences between the 4xCO_2 and SRM cases for extreme temperatures in one grid cell, indicating the expected outcome that SRM can also substantially reduce the incidence of extreme events relative to the unmitigated case when it involves substantial reductions in decadal average temperature anomalies. Based on our analysis of the distinction between SRM and preindustrial extreme event distributions, we might expect the location parameter to shift in accordance with the climatological shifts to temperature and precipitation under 4xCO_2 relative to the preindustrial case. However, it would have been instructive to analyse whether the extreme events in 4xCO_2 had fatter tails or a different skew to the distribution, i.e. significant changes to the scale or shape parameter, which were then compensated for by SRM.

A final concern with the analysis presented here is whether we were overfitting to choose the extreme value distribution with three parameters, rather than just two. Given the lack of easy interpretation for the shape and scale parameters, it may be the case that there is some degeneracy in these parameters which fitting a two-parameter extreme value distribution (e.g., with the shape parameter set to zero) would have avoided. When deciding on which distribution to use, I considered that the shape and scale parameters represented different changes to the extreme distribution, as set out in the methods section. A difference to the scale parameter would indicate fatter (or thinner) tails, i.e. that the extreme events were becoming more likely to be even more

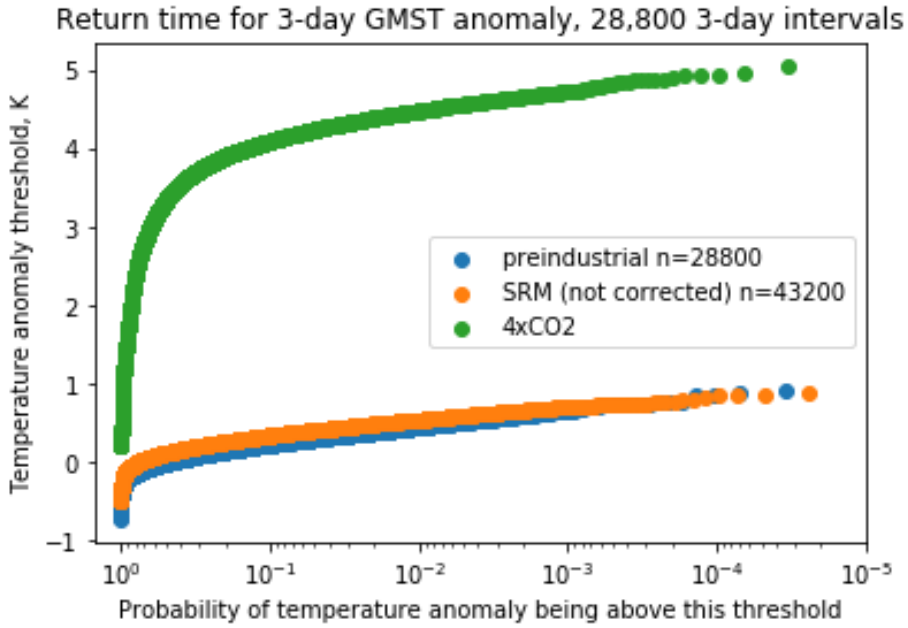

Figure 5.20: **Quasi return time plots for extreme temperatures in the (0.0, 52.5) UK grid cell under the 4xCO_2, SRM, and preindustrial cases.**

extreme relative to the average extreme event, or else clustering around the average extreme event. A difference to the shape parameter, on the other hand, would alter the skewness of the distribution, and illustrate whether the extreme events were becoming more "one-sided" or not - i.e., if there was evidence for a very long tail of extreme events in one direction, or a truncation of certain extreme events, which would merit further investigation. These both seemed like plausible changes to the extreme event distribution when the analysis was first conducted, and so I attempted to test for either explanation at the same time. Again, if time permitted statistical significance testing, future work could include testing whether a two-parameter extreme value fit was statistically significant across more grid cells than the three-parameter test used here, and therefore which one was more appropriate to use for these extreme values.

5.4.2 Conclusions

In this chapter, we have undertaken extreme values analysis of three different proxies for extreme weather events - three-day temperature maxima, as a proxy for heatwaves, daily precipitation maxima, as a proxy for flooding, and monthly precipitation min-

ima, as a proxy for droughts. The process of fitting extreme value distributions to these parameters worked reasonably well across most grid cells, although some were excluded due to poor fits, and in most of them, shifts to the extreme value distribution were seen. The most obvious shifts were to the location parameter, which reflected the pattern of residual warming and drying under SRM that imperfectly compensates for the preindustrial - for example, regions with residual warming saw a corresponding increase to the location paramete for 3-day temperature, with substantial increases to this parameter over land at high latitudes, and in the Amazon and Congo regions, while there were decreases in India and over tropical regions of the ocean, just as was seen for the mean temperature changes under SRM in HadCM3. Similarly, regions that were dried under SRM saw a significant decrease to the location parameter for minimum precipitation. However, as with changes to the mean quantities, for most grid cells and most considered parameters, these changes were substantially smaller than those in unmitigated 4xCO_2.

It was more difficult to interpret changes to the shape and scale parameters, which would have allowed us to interpret whether the skewness or tails of the distribution, relative to its shifted mean, were significantly altered by SRM. These changes formed less spatially coherent and easy to interpret patterns compared to those of the location parameter, which clearly reflected robust shifts to the mean quantities. However, there were some results we could interpret - for example, looking at maximum precipitation, the scale parameter tended to decrease, especially in drying regions, indicating reduced variability in precipitation extremes in regions that were dried overall. For temperature, the scale parameter generally decreases, indicating slightly reduced variability in the extreme temperatures, except in some regions like India where it is even more substantially reduced where these regions also cooled. In this case, the shape parameter often varied inversely, which could be seen to suggest that the most extreme events were more common even as the overall variability decreased, or it could potentially be a product of overfitting, as discussed in the discussion section. The regionally coherent nuances in these changes are more apparent for the 3-day temperature and daily precipitation variables; here, we had more data, so it may be that the extreme-value fitting approach worked more effectively for these variables than for the monthly minima.

A null hypothesis to test in considering how SRM affects extremes in temperature and precipitation would be to suggest that they can be characterised entirely by shifts to the location parameter which represent the residual changes in the mean quantities of temperature and precipitation, leaving the variability or skew of the distribution unchanged. This would still be perceived as extreme weather, in the same way that under unmitigated climate change, 1 in 50 year events (for example, in daily maximum temperature) can become 1 in 10 year events, but it would imply variability in the distribution of extremes that remained similar to preindustrial conditions. Looking at individual grid cells, it was clear that there were highly statistically significant shifts in the location parameters, while the data for the shape and scale parameters often overlapped, suggesting the possibility that the changes might be statistically insignificant in some areas. However, due to the limitations in statistical testing discussed above, we cannot categorically accept or reject the null hypothesis that shifts to the location parameter which follow residual temperature and precipitation changes are a good model for how extremes will shift under SRM.

Overall, the data analysis suggests that while SRM can reduce some of the most extreme temperature and precipitation events compared to unmitigated climate change, it does not eliminate significant shifts in the extremes which reflect the residual changes to the mean quantities.

Chapter 6

Conclusions

The preceding work has analysed solar geoengineering. It has used a number of idealised models, ranging in complexity from simple emissions-based impulse response models to large ensembles of the modified HadCM3 General Circulation Model, to investigate the effects of varying different parameters and scenarios for geoengineering. These have included the timing, duration, and emissions pathways (Chapter 2), plausible latitudinal distributions of aerosol optical depth (Chapters 3 and 4), the magnitude of the temperature increase that is cancelled out with geoengineering (Chapter 4), and the effect of geoengineering on climate and weather extremes with use of a large number of model years from the HadCM3 ensembles (Chapter 5.)

These are far from exhaustive; additional work on the timing and potential hysteresis of late-in-the-day application of SRM (to cool down a world that has already warmed) and on additional distributions (including further comparitive work between the G4 multimodel ensemble) was conducted, but cannot be reproduced here owing to space and time constraints on the behalf of the author. There is a vast potential parameter space which could be explored with geoengineering; if this is to be seriously considered, this would form an entire multidisciplinary field of study far beyond the scope of any single project. However, it is crucial at all times not to put too much faith in any model, let alone these idealised and simple models, in particular with few natural analogues for what would be an unprecedented (intentional) intervention in the climate system to give us confidence in the robustness of our results. John von Neumann reputedly said that "with four parameters, I can fit an elephant, and with

five, I can make him wiggle his trunk." And there are many variables and responses which even state-of-the-art climate models cannot predict and often disagree on owing to our continued uncertainties about the climate system and the inexactness of our ability to model it. In summarising the work, then, I will look to reiterate what was done and draw broad conclusions, but these caveats should be borne in mind.

Chapter 2 used the FAIR impulse-response simple climate model to investigate scenarios for geoengineering for "peak-shaving"; a potential temporary deployment of SRM to hold temperatures below a certain threshold, allowing time for decarbonisation and negative emissions to reduce the climate overshoot. Simple, approximate analytical formulae for approximate emissions profiles were developed to illustrate the factors which influence the duration SRM must be deployed for in different peak-shaving scenarios. It is clear that shaving a significant peak (of 0.5K) will require a commitment to geoengineering on the scale of around a century (70 years under highly optimistic assumptions), as well as the deployment of large-scale carbon removal (on the order of 10s of gigatonnes of CO_2 annually) which is only likely to be possible with direct air capture. These broad conclusions are confirmed, and the range of possible peak-shaving scenarios explored, through modelling various scenarios in fair: variants of the SSP5-34-OS scenario and scenarios from the IIASA database which involve a temperature peak and decline with negative emissions. The trade-off between the years of commitment to SRM and the peak in temperature avoided illustrates that, SRM that shaves a significant temperature peak will need to be deployed on centennial timescales.

Chapter 3 predominantly describes the methods and calibration of the HadCM3 model and the climateprediction.net computing project for the remaining chapters of the thesis. It sets out the process by which we determined the uniform aerosol optical depth that would cancel the global mean temperature increase from 4xCO2, which was then used as a scaling factor for other patterns of aerosol optical depth in order to maximise signal-to-noise ratio; we also re-derive the ECS for the system and the length of time simulations must be spun up for temperature drift to cease in the SRM scenarios.

The chapter then explored the use of the modified HadCM3 volcanic forcing implementation which allows for the aerosol optical depth to vary in each of the latitude

bands, to characterise the model's response to latitudinally-varying forcing. Highly asymmetric radiative forcing, and the subsequent shifts to the ITCZ, were demonstrated in the model. We illustrated significant, substantial differences in the regional climate response to the G4 scenarios - different latitudinal distributions of aerosol optical depth which were derived from equatorial injections of sulfate aerosol in different stratospheric models. This illustrates that the uncertainties as to how equatorial injections of aerosol would be distributed across models would result in substantially different regional climate responses. Finally, Chapter 3 shows results for zonally-averaged temperature and precipitation across both highly inhomogenous "stripe" aerosol distributions and more realistic profiles based on injections at different latitudes from the GLENS ensemble in CESM. In both cases, we show that the zonal temperature response in HadCM3 averaged over a decade is reasonably well modelled as linear - in other words, the sum of responses to imposed patterns of aerosol optical depths is approximately the response to a superposition of those patterns with the same linear coefficients, in the range investigated. This result allows for some of the analysis in Chapters 4 and 5, as well as the derivation of "optimal" aerosol optical depth patterns for reducing decadal zonal mean temperature anomalies. However, we find that the precipitation response is not linear, due to significant shifts to the ITCZ based on hemispheric forcing.

Chapter 4 contains the bulk of the analysis of the SRM runs in HadCM3. We find that linearity is a substantially better approximation for smooth-edged aerosol distributions, such as the GLENS "basis vectors" that are simulated, compared to the highly inhomogenous "stripe" simulations. We analyse simulations of the aerosol optical depth distribution calculated to minimise zonal temperature anomalies. While this provides a more exact cancellation of the zonal annual temperature anomaly, we find that it does so by overcooling the oceans and undercooling the land, and that regional temperature anomalies persist, particularly over the Amazon, Congo, and European Russia. These regional disparities in temperature cannot be cancelled out through purely zonal AOD optimisation. Zonally optimising for temperature is not the same as zonally optimising for precipitation, and we show that precipitation anomalies over land are not substantially reduced by either uniform or temperature-optimised SRM. Assuming linearity and attempting to optimise for precipitation also

shows that, while zonal anomalies can be reduced compared to uniform SRM and 4xCO2, they persist. Overall, we find that while there is significant scope to reduce zonal temperature anomalies and limited scope to reduce zonal precipitation anomalies through constructing these "optimal distributions" in the model, this may not lead to outcomes that society finds preferable as it often arises due to compensatory errors (undercooling and overcooling) within latitude bands.

One of the major concerns raised around geoengineering is that different nations or regions of the world may prefer different amounts or latitudinal distributions of aerosol in order to more closely restore their preindustrial climates, which could in turn lead to conflicts over how best to geoengineer between major powers. We investigated this in Chapter 4 in a number of ways. Assuming linearity, we can construct different profiles of aerosol which optimise temperature in each individual Giorgi region. We find that, for all Giorgi regions, it is possible to reduce the preindustrial temperature anomaly by some degree with an optimised distribution, but always at the expense of higher temperature anomalies for the rest of the world. We note that the preferred distributions between Giorgi regions to achieve these outcomes substantially differs between Giorgi regions. Although predicting how precipitation will shift in response to inhomogenous geoengineering is not simple, in the case of uniform aerosol optical depth, you can approximate the precipitation response by superimposing the "fast response" to 4xCO2+SRM with the "slow response" pattern as global mean temperature is allowed to increase. This analysis shows that there are regions of West Africa and the Amazon which are dried both by 4xCO2 and SRM, so any strategy which hopes to compensate for the precipitation reductions under SRM by allowing some warming to take place will not succeed. We also demonstrate that the Giorgi regions disagree about the preferred level of uniform SRM which would most closely return their precipitation to the preindustrial level.

Given that we had simulated a wide range of different latitudinal distributions in the model, we were also in a position to analyse the extent to which it was possible to "tailor" the response in individual latitude bands with inhomogenous forcing, or whether the climate response was ultimately determined by several primary modes of variation. We conducted an empirical orthogonal function (EOF) analysis on the climate responses to 18 different distributions of aerosol. 98.7% of the temperature

variance was explained by the first four EOFs. These EOFs generally had clear physical interpretations, with the first three corresponding to the arctic-amplified global warming pattern, a hemispheric asymmetry in warming, and a pole-to-equator temperature gradient. This suggests that the vast majority of decadal mean temperature changes, regardless of the imposed forcing, would consist of some sum of these patterns, and consequently that any tailoring that doesn't depend on a superposition of these patterns is likely to be ineffective. Similar analysis was conducted for precipitation changes; although the EOFs were more difficult to interpret physically in this case, 95% of the variation was explained by the first four, which again suggests that the ability to tailor local precipitation changes with latitudinally inhomgenous forcing is probably limited.

Chapter 5 examined the effect that uniform SRM to cancel the global mean temperature increase would have on climate and weather extremes. This was done through choosing three proxy metrics for extreme weather events of interest - three-day average daily temperature anomaly, as a proxy for heatwaves; daily precipitation, as a proxy for intense rainfall events and flooding; and monthly precipitation minima as a proxy for drought and dry spells. Generalised extreme value distributions were then used to characterise the distribution of extreme events by three parameters, allowing for comparison between the extreme event distribution under preindustrial and SRM conditions. In general, the dominant changes to extreme values appeared to be due to shifts to the location parameter in each case. In other words, there was evidence that average extreme events in precipitation and temperature followed the global patterns for average temperature and precipitation changes characterised in Chapter 3 for SRM in the model; extreme temperature events were warmer in the locations that had residual shifts to mean temperature, extreme dry events were drier in regions that were dried on average, and so on, although it is worth noting that these extremes would be ameliorated by SRM compared to unmitigated climate change in the regions where shifts to mean quantities were also ameliorated. Due to limitations in the analysis, discussed in the chapter, we were unable to conclusively determine whether there were other changes to the variability of extreme weather events - for example, fatter-tailed distributions.

There are substantial areas of potentially interesting work that we were unable to follow up on owing to time constraints. In particular, I would highlight some preliminary work that was done on simulations where SRM is imposed on an already-warmed world. Many simulations, such as the GeoMIP simulations, examine scenarios where SRM is imposed alongside increases in CO_2, but if SRM were imposed tomorrow, it would be on a world that had already experienced significant climate change. We saw evidence in preliminary simulations for "hysteresis" effects, where it took a long time for residual warming to be reduced to preindustrial levels, owing to the heat capacity of the oceans. The oceans would be warmer relative to the land under this scenario than under a preindustrial scenario, and this would likely have significant and underexplored implications for weather and climate, including monsoons and other rainfall patterns which are influenced by sea-surface temperatures and land-sea temperature contrasts. It is also the case that warmer oceans would drive thermosteric sea level rise and ice-sheet melt, and the ability of SRM applied late-in-the-day to reverse or avert "tipping points" in the climate system, themselves often not that well understood, deserves much more investigation. The same concern applies to scenarios - increasingly common - where large-scale carbon dioxide removal results in a substantial period of global mean cooling after a peak in temperatures sometime this century. Much more needs to be done to understand the impacts of late-in-the-day SRM or carbon dioxide removal on climate. Dynamical changes to the climate system as patterns change, particularly in dramatic cases such as the "termination shock", could also have been examined in more depth using a significant number of model years, to better understand the potential to trigger extreme events that might occur due to a termination shock or rapid imposition of forcing: this will be extremely politically relevant if geoengineering were ever to be deployed, as there may well need to be robust science around the attribution of any unusual weather or climate events to geoengineering in order to make any global framework politically viable.

In addition, most geoengineering simulations have used sulphate aerosols, which absorb longwave radiation as well as reflecting shortwave radiation, resulting in a heating of the stratosphere. Our model did not have a significantly resolved stratosphere and it did not simulate the effects of stratospheric dynamics and aerosol microphysics on the highly idealised scenarios that we simulated, so for us, this only manifested

as an extension of the calibration process described in Chapter 2. However, using alternative aerosols that do not absorb longwave radiation would result in a different radiative forcing profile and may allow for less aerosol optical depth to achieve the same cooling, and this could also be further investigated. Finally, a weakness of climateprediction.net is that while it can run a substantial number of models in parallel, running models for a long time (in excess of 40 years or so) is difficult with this system. Investigating the long-term consequences of any imposed patterns of SRM was therefore outside the scope of what we could accomplish, but will also be of interest.

It is clear that this has been a highly idealised investigation of an extremely complex set of questions, attempting to use the imperfect tools at our disposal to come to some broad conclusions about what the implications of geoengineering might be. The potential parameter space for interventions is huge, but so are our uncertainties about the climate consequences of intervention. In this context, humility is extremely important: there is still so much that we do not know and cannot say with confidence about the possible impacts of geoengineering. For many, the idea is, and remains, dangerous and unthinkable.

Yet we have seen dangerous and unthinkable ideas migrate from the fringes into the mainstream of climate politics regularly over the last few decades. At one point, it was considered that talking about adaptation to climate change was a moral hazard, because to imply that we could adapt to climate change might reduce the political urgency of mitigation. Now, many countries have national adaptation strategies, having recognised that some climate change is inevitable - indeed, much has already happened, and we will be living with the consequences for many decades. The 2C warming "target" of the Paris Agreement, discussed in Chapter 1, was once considered a 2C "upper limit" for dangerous warming which must be avoided at all costs. It would now be considered a huge success if we were able to limit warming to this level. Carbon capture and storage, and negative emissions technologies more generally, have moved from fringe suggestions to increasingly central to any scenarios that are compliant with the Paris Agreement. Many of these scenarios, which require massive near-terms emissions cuts and substantial commitment to negative emissions in order to remain Paris-compliant, evidence the gap between scenarios produced by economic

Integrated Assessment Models and what is likely to occur in reality. [129] It is not so far-fetched to think that, if reducing emissions and scaling up negative emissions continues to prove as politically and economically difficult as it has done over the last few decades, or if some sectors prove especially difficult to decarbonise under current economic and political paradigms, that solar radiation management will also move from the fringe and towards the mainstream. Indeed, nearly a decade on from the Paris Agreement, with global mitigation efforts increasingly derailed by geopolitical instability, global economic crises, and other catastrophic risks such as pandemics, there are signs that this has happened even over the course of writing this thesis.

The research in these chapters has illustrated and touched upon some of the political problems that would arise from attempted deployment, including the requirement for a centennial commitment to a programme of action with substantial negative emissions, and the potential that different regions of the world would disagree over how SRM should be deployed. One of the disadvantages of a future where geoengineering is needed is that it would present the world with a difficult global coordination problem in a geopolitical scenario where we have already failed at the difficult global coordination problem of mitigating and adapting to climate change: a gloomy prospect. There is near-universal agreement that the best solution to climate change is rapid mitigation and a reduction of emissions to net zero as soon as possible, a view also held by this author. Alternative paths are likely to be costlier and riskier.

In this context, it is often tempting for climate scientists to dismiss research into geoengineering as, in itself, dangerous and misguided. This is a complex moral issue that this author has struggled with throughout researching this topic. But ultimately, a studied, idealistic ignorance of the possibility of SRM on behalf of the mainstream scientific community leaves a gap that is more likely to be filled by panglossian charlatans than scientifically informed debate. As I write this, global greenhouse gas emissions have not yet even begun to decline. Decision-makers of the future will want the best evidence that we can provide.

Glossary

Aerosol microphysics involves the study of the physical properties of aerosols, such as size, composition, and interactions, which influence their behavior and effects on climate and air quality.

Afforestation involves planting trees on land that has not been forested for a long time, increasing forest cover to absorb atmospheric CO2 and combat global warming.

Albedo is a measure of the reflectivity of a surface, with higher values indicating greater reflection of incoming solar radiation, significantly affecting Earth's energy balance.

Atlantic Meridional Overturning Circulation (AMOC) is a system of ocean currents that transports warm water from the tropics northward into the North Atlantic, significantly influencing the climate of the surrounding continents.

Anthropogenic climate change refers to global climate change primarily caused by human activities, notably the emission of greenhouse gases from burning fossil fuels, deforestation, and industrial processes.

AOD (Aerosol Optical Depth) measures the degree to which aerosols prevent the transmission of sunlight through the atmosphere and is used to understand more about air quality and climate changes.

Arctic amplification describes the phenomenon where the Arctic warms faster than the global average, due to various factors including the loss of sea ice and subsequent increase in heat absorption by darker ocean surfaces, as well as the fact that colder regions require less radiative forcing to warm by a degree due to the non-linear nature of the radiative dependence on temperature.

Bio-energy with Carbon Capture and Storage (BECCS) refers to a climate change mitigation strategy that combines biomass energy production with carbon capture and storage to remove carbon dioxide from the atmosphere.

Brewer-Dobson circulation describes the large-scale movement of air in the stratosphere, characterized by upward motion of air in the tropics, poleward and downward motion in the higher latitudes, crucial for the distribution of ozone.

Climate change mitigation encompasses actions and strategies to reduce or prevent the emission of greenhouse gases and to enhance sinks that absorb these gases, thereby reducing the extent of global warming.

CMIP stands for Coupled Model Intercomparison Project, a standard experimental framework for studying the output of coupled atmosphere-ocean general circulation models (AOGCMs) to understand past, present, and future climate changes.

The **Cryosphere** encompasses all the areas on the Earth's surface where water is in solid form, including ice sheets, icebergs, glaciers, sea ice, snow cover, and frozen ground.

Detrended or deseasonalised data have had influences of trends and seasonal variations removed to analyze underlying patterns and anomalies in more detail.

Dipole refers to a pair of related anomalies with opposite characteristics, such as temperature or pressure differences, occurring over different geographical areas.

Eigenvalues are coefficients in a characteristic equation that represent the magnitude of vectors in linear transformations, used in climate science to analyze variance in data sets - for example, to express the relative size of the different EOFs in a pattern of temperature or precipitation.

ENSO (El Nino-Southern Oscillation) is a periodic variation in winds and sea surface temperatures over the tropical eastern Pacific Ocean, affecting climate and weather patterns globally. The positive phase of this oscillation, often called El Nino, is associated with a tongue of warm sea-surface temperatures in the Pacific and globally warmer temperatures, while the negative phase is called La Nina and has the opposite effect.

Empirical Orthogonal Functions (EOFs) are statistical tools used in climate science to analyze spatial patterns of variability in large datasets, like temperature or precipitation fields. They determine orthogonal (i.e. independent degrees of freedom) patterns across the variability of different fields, such that the fields can be constructed by linear combinations of these orthogonal basis functions. **Orthogonal basis functions** are mathematical functions used in various computations, characterized by their mutual orthogonality within a specified domain, important in simplifying complex systems into principal components.

Enhanced weathering is a geoengineering approach that involves spreading finely crushed minerals, for example olivine, which react with carbon dioxide, on the land surface to accelerate the natural process of carbon dioxide removal from the atmosphere through weathering.

Equilibrium Climate Sensitivity (ECS) is a quantity which is the change in the global mean surface temperature resulting from a doubling of the atmospheric carbon dioxide concentration, measuring the long-term response once the climate reaches equilibrium.

Generalised extreme value (GEV) distribution is used to model the largest or smallest value among a large set of independent, identically distributed random values representing extremes such as maximum one-day rainfalls.

Geoengineering refers to intentional, large-scale manipulation of Earth's climate, generally aimed at mitigating the adverse effects of climate change. Common strategies include solar radiation management, which reflects sunlight to cool the Earth, and greenhouse gas removal (GGR) or carbon dioxide removal (CDR), which would remove greenhouse gases from the atmosphere.

Hadley cell is a large-scale atmospheric circulation feature in which warm air rises near the equator, flows poleward at high altitudes, cools and sinks at subtropical latitudes, then returns equatorward near the surface, forming large "cells" of airflow in the atmosphere.

Hydrological cycle describes the continuous movement of water on, above, and below the surface of the Earth, involving processes such as evaporation, condensation, precipitation, and runoff.

Hysteresis in climate science refers to the delayed response or non-linear reactions of climate systems to environmental changes, potentially leading to irreversible effects.

The **Intertropical Convergence Zone (ITCZ)** is a belt of low pressure wrapping around the Earth near the equator where the trade winds of the Northern and Southern Hemispheres come together, driving major weather patterns and tropical rain belts.

IPCC stands for the Intergovernmental Panel on Climate Change, an international body responsible for assessing the science related to climate change and providing a scientific basis for governments at all levels to develop climate-related policies.

Kurtosis is a statistical measure that describes the shape of a probability distribution, particularly how sharp the peak is, indicating the presence of outliers in the data.

Marine cloud brightening is a geoengineering technique that aims to increase the albedo (reflectivity) of clouds by spraying fine sea water droplets into the atmosphere to enhance cloud brightness and potentially mitigate global warming.

Null hypothesis in statistical testing represents a statement or hypothesis that there is no significant effect or association between two measured phenomena, serving as the default assumption until evidence indicates otherwise.

The **Quasi-biennial oscillation** or QBO is a regular variation of the winds that occurs in the equatorial stratosphere, characterized by alternating easterlies and westerlies typically every 28 months.

Radiative forcing is a measure of the influence a factor has in altering the balance of incoming and outgoing energy in the Earth's atmosphere and is an important mechanism in climate dynamics. It is expressed in watts per square metre (W/m^2). Positive radiative forcing leads to warming, while negative forcing results in cooling. Anthropogenic radiative forcing over the Industrial Era was estimated at 2.72 watts per square metre in 2019 by the IPCC's Sixth Assessment Report.

RCP scenarios (Representative Concentration Pathways) are four greenhouse gas concentration trajectories adopted by the IPCC for its climate modeling and research, representing different climate futures based on varying levels of greenhouse

gas emissions. RCP8.5 is the highest emissions scenario and has sometimes been used as a "baseline" representing a scenario where no climate action was taken.

Realised Warming Fraction (RWF) quantifies the proportion of committed warming that has already been realized at a given time, considering the delay between emission of greenhouse gases and the resultant climate response. It can be defined as the ratio of the TCR to the ECS (see elsewhere in this glossary.)

The **Stratosphere** is the second layer of Earth's atmosphere, just above the troposphere and below the mesosphere, known for its ozone layer which absorbs and scatters incoming solar radiation. It extends up to about 45-50km above the Earth's surface.

Stratospheric dynamics refer to the study of movements and processes within the stratosphere, focusing on phenomena like temperature distributions, wind patterns, and interactions with the troposphere.

SRM refers to solar radiation management, which encompasses a range of techniques to reflect solar radiation back into space in order to cool the climate. **SAI** is short for stratospheric aerosol injection, the most studied of these techniques, which involves putting reflective aerosols - typically sulphate aerosols - into the stratosphere to reflect sunlight.

Teleconnection in climate science refers to climate anomalies being related to each other at large distances (typically thousands of kilometers) due to atmospheric wave patterns.

Transient Climate Response (TCR) is the increase in the global mean surface temperature, at the time when carbon dioxide concentrations have doubled, under a scenario of gradually increasing emissions.

Transient Climate Response to Cumulative Emissions (TCRE) quantifies the global average surface temperature increase resulting from a cumulative specified amount of carbon dioxide emissions, typically a thousand gigatonnes of carbon dioxide, measuring the responsiveness of the climate system to ongoing carbon emissions.

The **Troposphere** is the lowest layer of Earth's atmosphere, where nearly all weather conditions take place. It begins at the Earth's surface, but the height of the

troposphere varies. It is 18-20 km high at the equator, 9 km at 50°N and 50°S, and just under 6 km high at the poles.

List of Figures

Bibliography

[1] Hoesung Lee, Katherine Calvin, Dipak Dasgupta, Gerhard Krinner, Aditi Mukherji, Peter Thorne, Christopher Trisos, Jose Romero, Paulina Aldunce, Ko Barret, et al. Ipcc, 2023: Climate change 2023: Synthesis report, summary for policymakers. contribution of working groups i, ii and iii to the sixth assessment report of the intergovernmental panel on climate change [core writing team, h. lee and j. romero (eds.)]. ipcc, geneva, switzerland. 2023.

[2] Paris Agreement. Paris agreement. In *report of the conference of the parties to the United Nations framework convention on climate change (21st session, 2015: Paris). Retrived December*, volume 4, page 2. HeinOnline, 2015.

[3] United Nations Environment Programme. Emissions gap report 2023: Broken record. 2023.

[4] Hannah Ritchie, Pablo Rosado, and Max Roser. Co and greenhouse gas emissions. *Our World in Data*, 2023. https://ourworldindata.org/co2-and-greenhouse-gas-emissions.

[5] Paul J. Crutzen. Albedo enhancement by stratospheric sulfur injections: A contribution to resolve a policy dilemma? *Climatic Change*, 77(3):211, Jul 2006.

[6] Douglas G MacMartin, Katharine L Ricke, and David W Keith. Solar geoengineering as part of an overall strategy for meeting the 1.5 c paris target. *Phil. Trans. R. Soc. A*, 376(2119):20160454, 2018.

[7] J. G. Shepherd, K. Caldeira, P. Cox, J. Haigh, D. Keith, B. E. Launder, G. Mace, G. MacKerron, J. Pyle, S. Raynor, C. Redgwell, and A. Watson.

Geoengineering the Climate:: Science, governance and uncertainty. Royal Society policy document. The Royal Society Publishing, United Kingdom, 9 2009.

[8] David G Victor, M Granger Morgan, Jay Apt, John Steinbruner, and Katharine Ricke. The geoengineering option: a last resort against global warming? *Foreign Aff.*, 88:64, 2009.

[9] Aseem Mahajan, Dustin Tingley, and Gernot Wagner. Fast, cheap, and imperfect? us public opinion about solar geoengineering. *Environmental Politics*, 28(3):523–543, 2019.

[10] Alan Robock, Allison Marquardt, Ben Kravitz, and Georgiy Stenchikov. The practicality of geoengineering. *Geophysical Research Letters*, 2009.

[11] Timothy M Lenton, Johan Rockstrom, Owen Gaffney, Stefan Rahmstorf, Katherine Richardson, Will Steffen, and Hans Joachim Schellnhuber. Climate tipping points—too risky to bet against. *Nature*, 575(7784):592–595, 2019.

[12] Mark G Lawrence and Paul J Crutzen. Was breaking the taboo on research on climate engineering via albedo modification a moral hazard, or a moral imperative? *Earth's Future*, 5(2):136–143, 2017.

[13] Michael S Diamond. Detection of large-scale cloud microphysical changes within a major shipping corridor after implementation of the international maritime organization 2020 fuel sulfur regulations. *Atmospheric Chemistry and Physics*, 23(14):8259–8269, 2023.

[14] Mikhail Ivanovich Budyko. *Climatic changes*. Amer Geophysical Union, 1977.

[15] Ben Kravitz, Alan Robock, Olivier Boucher, Hauke Schmidt, Karl E. Taylor, Georgiy Stenchikov, and Michael Schulz. The geoengineering model intercomparison project (geomip). *Atmospheric Science Letters*, 12(2):162–167, 2011.

[16] Alan Robock. Volcanic eruptions and climate. *Reviews of Geophysics*, 38(2):191–219, 2000.

[17] P. Minnis, E. F. Harrison, L. L. Stowe, G. G. Gibson, F. M. Denn, D. R. Doelling, and W. L. Smith. Radiative climate forcing by the mount pinatubo eruption. *Science*, 259(5100):1411–1415, 1993.

[18] Georgiy L Stenchikov, Ingo Kirchner, Alan Robock, Hans-F Graf, Juan Carlos Antuña, RG Grainger, Alyn Lambert, and Larry Thomason. Radiative forcing from the 1991 mount pinatubo volcanic eruption. *Journal of Geophysical Research: Atmospheres*, 103(D12):13837–13857, 1998.

[19] Gregg JS Bluth, Scott D Doiron, Charles C Schnetzler, Arlin J Krueger, and Louis S Walter. Global tracking of the so2 clouds from the june, 1991 mount pinatubo eruptions. *Geophysical Research Letters*, 19(2):151–154, 1992.

[20] Katharine Ricke, M Morgan, and Myles R. Allen. Regional climate response to solar-radiation management. *Nature Geoscience*, 3, 07 2010.

[21] Andy Parker and Peter J. Irvine. The risk of termination shock from solar geoengineering. *Earth's Future*, 6(3):456–467, 2018.

[22] Andy Jones, Jim M. Haywood, Kari Alterskjaer, Olivier Boucher, Jason N. S. Cole, Charles L. Curry, Peter J. Irvine, Duoying Ji, Ben Kravitz, Jon Egill Kristjansson, John C. Moore, Ulrike Niemeier, Alan Robock, Hauke Schmidt, Balwinder Singh, Simone Tilmes, Shingo Watanabe, and Jin-Ho Yoon. The impact of abrupt suspension of solar radiation management (termination effect) in experiment g2 of the geoengineering model intercomparison project (geomip). *Journal of Geophysical Research: Atmospheres*, 118(17):9743–9752, 2013.

[23] Kevin E Trenberth and Aiguo Dai. Effects of mount pinatubo volcanic eruption on the hydrological cycle as an analog of geoengineering. *Geophysical Research Letters*, 34(15), 2007.

[24] Myles R. Allen and William J. Ingram. Constraints on future changes in climate and the hydrologic cycle. *Nature*, 419:224 EP –, Sep 2002.

[25] Matthew Toohey, Kirstin Kruger, Hauke Schmidt, Claudia Timmreck, Michael Sigl, Markus Stoffel, and Rob Wilson. Disproportionately strong climate forcing

from extratropical explosive volcanic eruptions. *Nature Geoscience*, 12(2):100–107, 2019.

[26] M. Sato, J. E. Hansen, M. P. McCormick, and J. B. Pollack. Stratospheric aerosol optical depths, 1850-1990. *J. Geophys. Res.*, 98:22987–22994, 1993.

[27] A. Laakso, H. Korhonen, S. Romakkaniemi, and H. Kokkola. Radiative and climate effects of stratospheric sulfur geoengineering using seasonally varying injection areas. *Atmospheric Chemistry and Physics*, 17(11):6957–6974, 2017.

[28] A. C. Jones, J. M. Haywood, and A. Jones. Climatic impacts of stratospheric geoengineering with sulfate, black carbon and titania injection. *Atmospheric Chemistry and Physics*, 16(5):2843–2862, 2016.

[29] L. S. Jackson, J. A. Crook, A. Jarvis, D. Leedal, A. Ridgwell, N. Vaughan, and P. M. Forster. Assessing the controllability of arctic sea ice extent by sulfate aerosol geoengineering. *Geophysical Research Letters*, 42(4):1223–1231, 2015.

[30] Jim M. Haywood, Andy Jones, Nicolas Bellouin, and David Stephenson. Asymmetric forcing from stratospheric aerosols impacts sahelian rainfall. *Nature Climate Change*, pages 660 EP –, Mar 2013.

[31] Aditya Nalam, Govindasamy Bala, and Angshuman Modak. Effects of arctic geoengineering on precipitation in the tropical monsoon regions. *Climate Dynamics*, 50(9):3375–3395, May 2018.

[32] Hugh Baker, Richard J. Millar, David Karoly, Urs Beyerle, Benoit Guillod, Dann Mitchell, Hideo Shiogama, Sarah Sparrow, Tim Woollings, and Myles R. Allen. Higher co2 concentrations increase extreme event risk in a 1.5 k world. *Nature Climate Change*, 8, 06 2018.

[33] Cao and Caldeira. Atmospheric carbon dioxide removal: long-term consequences and commitment. *Environmental Research Letters*, 5(2):024011, 2010.

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

[35] O Boucher, PR Halloran, EJ Burke, M Doutriaux-Boucher, CD Jones, J Lowe, MA Ringer, E Robertson, and P Wu. Reversibility in an earth system model in response to co2 concentration changes. *Environmental Research Letters*, 7(2):024013, 2012.

[36] Chris D Jones, Philippe Ciais, Steven J Davis, Pierre Friedlingstein, Thomas Gasser, Glen P Peters, Joeri Rogelj, Detlef P van Vuuren, Josep G Canadell, A Cowie, et al. Simulating the earth system response to negative emissions. *Environmental Research Letters*, 11(9):095012, 2016.

[37] Andrew H MacDougall. Reversing climate warming by artificial atmospheric carbon-dioxide removal: Can a holocene-like climate be restored? *Geophysical Research Letters*, 40(20):5480–5485, 2013.

[38] Peter Irvine, Kerry Emanuel, Jie He, Larry W. Horowitz, Gabriel Vecchi, and David Keith. Halving warming with idealized solar geoengineering moderates key climate hazards. *Nature Climate Change*, 9(4):295–299, 2019.

[39] John T. Fasullo, Simone Tilmes, Jadwiga H. Richter, Ben Kravitz, Douglas G. MacMartin, Michael J. Mills, and Isla R. Simpson. Persistent polar ocean warming in a strategically geoengineered climate. *Nature Geoscience*, 11(12):910–914, 2018.

[40] Ying Sun and YiHui Ding. Responses of south and east asian summer monsoons to different land-sea temperature increases under a warming scenario. *Chinese Science Bulletin*, 56(25):2718, Aug 2011.

[41] Hiroki Kashimura, Manabu Abe, Shingo Watanabe, Takashi Sekiya, Duoying Ji, John C Moore, Jason NS Cole, and Ben Kravitz. Shortwave radiative forcing, rapid adjustment, and feedback to the surface by sulfate geoengineering: analysis of the geoengineering model intercomparison project g4 scenario. *Atmospheric Chemistry and Physics*, 17(5):3339–3356, 2017.

[42] Giovanni Pitari, Valentina Aquila, Ben Kravitz, Alan Robock, Shingo Watanabe, Irene Cionni, Natalia De Luca, Glauco Di Genova, Eva Mancini, and Simone Tilmes. Stratospheric ozone response to sulfate geoengineering: Results

from the geoengineering model intercomparison project (geomip). *Journal of Geophysical Research: Atmospheres*, 119(5):2629–2653, 2014.

[43] H. Richter, Jadwiga, Simone Tilmes, Anne Glanville, Ben Kravitz, Douglas G. MacMartin, Michael J. Mills, Isla R. Simpson, Francis Vitt, Joseph J. Tribbia, and Jean-Francois Lamarque. Stratospheric response in the first geoengineering simulation meeting multiple surface climate objectives. *Journal of Geophysical Research: Atmospheres*, 123(11):5762–5782, 2018.

[44] Jadwiga H. Richter, Simone Tilmes, Michael J. Mills, Joseph J. Tribbia, Ben Kravitz, Douglas G. MacMartin, Francis Vitt, and Jean-Francois Lamarque. Stratospheric dynamical response and ozone feedbacks in the presence of so2 injections. *Journal of Geophysical Research: Atmospheres*, 122(23):12,557–12,573, 2017.

[45] U. Niemeier and H. Schmidt. Changing transport processes in the stratosphere by radiative heating of sulfate aerosols. *Atmospheric Chemistry and Physics*, 17(24):14871–14886, 2017.

[46] Simone Tilmes, Jadwiga H. Richter, Michael J. Mills, Ben Kravitz, Douglas G. MacMartin, Rolando R. Garcia, Douglas E. Kinnison, Jean-Francois Lamarque, Joseph Tribbia, and Francis Vitt. Effects of different stratospheric so2 injection altitudes on stratospheric chemistry and dynamics. *Journal of Geophysical Research: Atmospheres*, 123(9):4654–4673, 2017.

[47] Angus J Ferraro, Andrew J Charlton-Perez, and Eleanor J Highwood. Stratospheric dynamics and midlatitude jets under geoengineering with space mirrors and sulfate and titania aerosols. *Journal of Geophysical Research: Atmospheres*, 120(2):414–429, 2015.

[48] A. J. Ferraro, E. J. Highwood, and A. J. Charlton-Perez. Stratospheric heating by potential geoengineering aerosols. *Geophysical Research Letters*, 38(24), 2011.

[49] U. Niemeier and C. Timmreck. What is the limit of climate engineering by stratospheric injection of so2. *Atmospheric Chemistry and Physics*, 15(16):9129–9141, 2015.

[50] C. Kleinschmitt, O. Boucher, and U. Platt. Sensitivity of the radiative forcing by stratospheric sulfur geoengineering to the amount and strategy of the so2 injection studied with the lmdz-s3a model. *Atmospheric Chemistry and Physics*, 18(4):2769–2786, 2018.

[51] Kravitz Ben, Robock Alan, Forster Piers M., Haywood James M., Lawrence Mark G., and Schmidt Hauke. An overview of the geoengineering model intercomparison project (geomip). *Journal of Geophysical Research: Atmospheres*, 118(23):13,103–13,107, 2013.

[52] H. Kashimura, M. Abe, S. Watanabe, T. Sekiya, D. Ji, J. C. Moore, J. N. S. Cole, and B. Kravitz. Shortwave radiative forcing, rapid adjustment, and feedback to the surface by sulfate geoengineering: analysis of the geoengineering model intercomparison project g4 scenario. *Atmospheric Chemistry and Physics*, 17(5):3339–3356, 2017.

[53] E. M. Bednarz, A. H. Butler, D. Visioni, Y. Zhang, B. Kravitz, and D. G. MacMartin. Injection strategy – a driver of atmospheric circulation and ozone response to stratospheric aerosol geoengineering. *EGUsphere*, 2023:1–32, 2023.

[54] Bala Govindasamy and Ken Caldeira. Geoengineering earth's radiation balance to mitigate co2-induced climate change. *Geophysical Research Letters*, 27(14):2141–2144, 2000.

[55] B. Kravitz, A. Robock, S. Tilmes, O. Boucher, J. M. English, P. J. Irvine, A. Jones, M. G. Lawrence, M. MacCracken, H. Muri, J. C. Moore, U. Niemeier, S. J. Phipps, J. Sillmann, T. Storelvmo, H. Wang, and S. Watanabe. The geoengineering model intercomparison project phase 6 (geomip6): simulation design and preliminary results. *Geoscientific Model Development*, 8(10):3379–3392, 2015.

[56] R. D. Russotto and T. P. Ackerman. Changes in clouds and thermodynamics under solar geoengineering and implications for required solar reduction. *Atmospheric Chemistry and Physics*, 18(16):11905–11925, 2018.

[57] Simone Tilmes, John Fasullo, Jean-Francois Lamarque, Daniel R Marsh, Michael Mills, Kari Alterskjaer, Helene Muri, Jon E Kristjansson, Olivier Boucher, Michael Schulz, et al. The hydrological impact of geoengineering in the geoengineering model intercomparison project (geomip). *Journal of Geophysical Research: Atmospheres*, 118(19), 2013.

[58] Kravitz Ben, Caldeira Ken, Boucher Olivier, Robock Alan, Rasch Philip J., Alterskjær Kari, Karam Diana Bou, Cole Jason N. S., Curry Charles L., Haywood James M., Irvine Peter J., Ji Duoying, Jones Andy, Kristjánsson Jón Egill, Lunt Daniel J., Moore John C., Niemeier Ulrike, Schmidt Hauke, Schulz Michael, Singh Balwinder, Tilmes Simone, Watanabe Shingo, Yang Shuting, and Yoon Jin-Ho. Climate model response from the geoengineering model intercomparison project (geomip). *Journal of Geophysical Research: Atmospheres*, 118(15):8320–8332, 2013.

[59] B. H. Samset, G. Myhre, P. M. Forster, Ø. Hodnebrog, T. Andrews, G. Faluvegi, D. Fläschner, M. Kasoar, V. Kharin, A. Kirkevåg, J.-F. Lamarque, D. Olivié, T. Richardson, D. Shindell, K. P. Shine, T. Takemura, and A. Voulgarakis. Fast and slow precipitation responses to individual climate forcers: A pdrmip multimodel study. *Geophysical Research Letters*, 43(6):2782–2791, 2016.

[60] Kravitz Ben, Rasch Philip J., Forster Piers M., Andrews Timothy, Cole Jason N. S., Irvine Peter J., Ji Duoying, Kristjansson Jon Egill, Moore John C., Muri Helene, Niemeier Ulrike, Robock Alan, Singh Balwinder, Tilmes Simone, Watanabe Shingo, and Yoon Jin-Ho. An energetic perspective on hydrological cycle changes in the geoengineering model intercomparison project. *Journal of Geophysical Research: Atmospheres*, 118(23):13,087–13,102, 2013.

[61] Niemeier U., Schmidt H., Alterskjær K., and Kristjánsson J. E. Solar irradiance reduction via climate engineering: Impact of different techniques on the energy balance and the hydrological cycle. *Journal of Geophysical Research: Atmospheres*, 118(21):11,905–11,917, 2013.

[62] Anthony C. Jones, Matthew K. Hawcroft, James M. Haywood, Andy Jones, Xiaoran Guo, and John C. Moore. Regional climate impacts of stabilizing

global warming at 1.5 k using solar geoengineering. *Earth's Future*, 6(2):230–251, 2018.

[63] Peter J Irvine, Daniel J Lunt, Emma J Stone, and Andy Ridgwell. The fate of the greenland ice sheet in a geoengineered, high co 2 world. *Environmental Research Letters*, 4(4):045109, 2009.

[64] John C Moore, Svetlana Jevrejeva, and Aslak Grinsted. Efficacy of geoengineering to limit 21st century sea-level rise. *Proceedings of the National Academy of Sciences*, 107(36):15699–15703, 2010.

[65] P. J. Irvine, D. W. Keith, and J. Moore. Brief communication: Understanding solar geoengineering's potential to limit sea level rise requires attention from cryosphere experts. *The Cryosphere*, 12(7):2501–2513, 2018.

[66] K. E. McCusker, D. S. Battisti, and C. M. Bitz. Inability of stratospheric sulfate aerosol injections to preserve the west antarctic ice sheet. *Geophysical Research Letters*, 42(12):4989–4997, 2015.

[67] Patrick J Applegate and Klaus Keller. How effective is albedo modification (solar radiation management geoengineering) in preventing sea-level rise from the greenland ice sheet? *Environmental Research Letters*, 10(8):084018, 2015.

[68] Charles L. Curry, Jana Sillmann, David Bronaugh, Kari Alterskjaer, Jason N. S. Cole, Duoying Ji, Ben Kravitz, Jón Egill Kristjánsson, John C. Moore, Helene Muri, Ulrike Niemeier, Alan Robock, Simone Tilmes, and Shuting Yang. A multimodel examination of climate extremes in an idealized geoengineering experiment. *Journal of Geophysical Research: Atmospheres*, 119(7):3900–3923, 2013.

[69] D. Ji, S. Fang, C. L. Curry, H. Kashimura, S. Watanabe, J. N. S. Cole, A. Lenton, H. Muri, B. Kravitz, and J. C. Moore. Extreme temperature and precipitation response to solar dimming and stratospheric aerosol geoengineering. *Atmospheric Chemistry and Physics*, 18(14):10133–10156, 2018.

[70] Katherine Dagon and Daniel P. Schrag. Regional climate variability under model simulations of solar geoengineering. *Journal of Geophysical Research: Atmospheres*, 122(22):12,106–12,121, 2017.

[71] L. Wei, D. Ji, C. Miao, H. Muri, and J. C. Moore. Global streamflow and flood response to stratospheric aerosol geoengineering. *Atmospheric Chemistry and Physics*, 18(21):16033–16050, 2018.

[72] S. Tilmes, D. G. MacMartin, J. T. M. Lenaerts, L. van Kampenhout, L. Muntjewerf, L. Xia, C. S. Harrison, K. M. Krumhardt, M. J. Mills, B. Kravitz, and A. Robock. Reaching 1.5 and 2.0c global surface temperature targets using stratospheric aerosol geoengineering. *Earth System Dynamics*, 11(3):579–601, 2020.

[73] Daniele Visioni, Alan Robock, Alistair Duffey, and Ilaria Quaglia. Process-level experiments and policy-relevant scenarios in future geomip iterations. *Bulletin of the American Meteorological Society*, 104(2):E501–E503, 2023.

[74] DG MacMartin, D Visioni, B Kravitz, JH Richter, T Felgenhauer, WR Lee, DR Morrow, EA Parson, and M Sugiyama. Scenarios for modeling solar radiation modification. *Proceedings of the National Academy of Sciences*, 119(33):e2202230119, 2022.

[75] Q. Wang, J. C. Moore, and D. Ji. A statistical examination of the effects of stratospheric sulfate geoengineering on tropical storm genesis. *Atmospheric Chemistry and Physics*, 18(13):9173–9188, 2018.

[76] Anthony C. Jones, James M. Haywood, Nick Dunstone, Kerry Emanuel, Matthew K. Hawcroft, Kevin I. Hodges, and Andy Jones. Impacts of hemispheric solar geoengineering on tropical cyclone frequency. *Nature Communications*, 8(1):1382, 2017.

[77] M. R. Tye, K. Dagon, M. J. Molina, J. H. Richter, D. Visioni, B. Kravitz, and S. Tilmes. Indices of extremes: geographic patterns of change in extremes and associated vegetation impacts under climate intervention. *Earth System Dynamics*, 13(3):1233–1257, 2022.

[78] Jingrong Wang, Zhihua Zhang, M. James C. Crabbe, and Lipon Chandra Das. Responses of extreme climates in south asia under a g6sulfur scenario of climate engineering. *Atmosphere*, 14(10), 2023.

[79] Adechina Eric Alamou, Ezechiel Obada, Eliezer Iboukoun Biao, Esdras Babadjid Josue Zandagba, Casimir Y Da-Allada, Frederic K Bonou, Ezinvi Baloitcha, Simone Tilmes, and Peter J Irvine. Impact of stratospheric aerosol geoengineering on meteorological droughts in west africa. *Atmosphere*, 13(2):234, 2022.

[80] Ben Kravitz, Douglas G MacMartin, Alan Robock, Philip J Rasch, Katharine L Ricke, Jason N S Cole, Charles L Curry, Peter J Irvine, Duoying Ji, David W Keith, Jón Egill Kristjánsson, John C Moore, Helene Muri, Balwinder Singh, Simone Tilmes, Shingo Watanabe, Shuting Yang, and Jin-Ho Yoon. A multimodel assessment of regional climate disparities caused by solar geoengineering. *Environmental Research Letters*, 9(7):074013, 2014.

[81] Juan B. Moreno-Cruz, Katharine L. Ricke, and David W. Keith. A simple model to account for regional inequalities in the effectiveness of solar radiation management. *Climatic Change*, 110(3):649–668, Feb 2012.

[82] Douglas Macmartin, David Keith, Ben Kravitz, and Ken Caldeira. Managing trade-offs in geoengineering through optimal choice of non-uniform radiative forcing. *Nature Climate Change*, 3:365–368, 04 2013.

[83] Xiaoyong Yu, John C. Moore, Xuefeng Cui, Annette Rinke, Duoying Ji, Ben Kravitz, and Jin-Ho Yoon. Impacts, effectiveness and regional inequalities of the geomip g1 to g4 solar radiation management scenarios. *Global and Planetary Change*, 129:10 – 22, 2015.

[84] JA Crook, LS Jackson, SM Osprey, and PM Forster. A comparison of temperature and precipitation responses to different earth radiation management geoengineering schemes. *Journal of Geophysical Research: Atmospheres*, 120(18):9352–9373, 2015.

[85] Camilla W Stjern, Helene Muri, Lars Ahlm, Olivier Boucher, Jason NS Cole, Duoying Ji, Andy Jones, Jim Haywood, Ben Kravitz, Andrew Lenton, et al.

Response to marine cloud brightening in a multi-model ensemble. *Atmospheric Chemistry and Physics*, 18(2):621–634, 2018.

[86] Douglas G MacMartin, Ben Kravitz, and Paul B Goddard. Transboundary effects from idealized regional geoengineering. *Environmental Research Communications*, 5(9):091004, sep 2023.

[87] Tilmes Simone, Richter Jadwiga H., Mills Michael J., Kravitz Ben, MacMartin Douglas G., Vitt Francis, Tribbia Joseph J., and Lamarque Jean-Francois. Sensitivity of aerosol distribution and climate response to stratospheric so2 injection locations. *Journal of Geophysical Research: Atmospheres*, 122(23):12,591–12,615, 2017.

[88] Walker R Lee, Daniele Visioni, Ewa M Bednarz, Douglas G MacMartin, Ben Kravitz, and Simone Tilmes. Quantifying the efficiency of stratospheric aerosol geoengineering at different altitudes. *Geophysical Research Letters*, 50(14):e2023GL104417, 2023.

[89] MacMartin Douglas G., Kravitz Ben, Tilmes Simone, Richter Jadwiga H., Mills Michael J., Lamarque Jean-Francois, Tribbia Joseph J., and Vitt Francis. The climate response to stratospheric aerosol geoengineering can be tailored using multiple injection locations. *Journal of Geophysical Research: Atmospheres*, 122(23):12,574–12,590, 2017.

[90] Daniele Visioni, Alan Robock, Jim Haywood, Matthew Henry, and Alice Wells. A new era for the geoengineering model intercomparison project (geomip). *Bulletin of the American Meteorological Society*, 2023.

[91] George A Ban-Weiss and Ken Caldeira. Geoengineering as an optimization problem. *Environmental Research Letters*, 5(3):034009, 2010.

[92] Douglas G MacMartin, David W Keith, Ben Kravitz, and Ken Caldeira. Management of trade-offs in geoengineering through optimal choice of non-uniform radiative forcing. *Nature Climate Change*, 3(4):365, 2013.

[93] B Kravitz, DG MacMartin, H Wang, and PJ Rasch. Geoengineering as a design problem. *Earth System Dynamics Discussions*, 6(2), 2015.

[94] Ben Kravitz, Douglas G MacMartin, Hailong Wang, and Philip J Rasch. Geoengineering as a design problem. *Earth System Dynamics*, 7(2):469–497, 2016.

[95] Simone Tilmes, Jadwiga H. Richter, Ben Kravitz, Douglas G. MacMartin, Michael J. Mills, Isla R. Simpson, Anne S. Glanville, John T. Fasullo, Adam S. Phillips, Jean-Francois Lamarque, Joseph Tribbia, Jim Edwards, Sheri Mickelson, and Siddhartha Ghosh. Cesm1(waccm) stratospheric aerosol geoengineering large ensemble project. *Bulletin of the American Meteorological Society*, 99(11):2361–2371, 2018.

[96] Thomas Hornigold. Tradeoffs inherent in solar geoengineering peak-shaving strategies. *arXiv preprint arXiv:2108.00096*, 2021.

[97] Marlos Goes, Nancy Tuana, and Klaus Keller. The economics (or lack thereof) of aerosol geoengineering. *Climatic change*, 109(3-4):719–744, 2011.

[98] David W Keith and Peter J Irvine. Solar geoengineering could substantially reduce climate risks—a research hypothesis for the next decade. *Earth's Future*, 4(11):549–559, 2016.

[99] Juan B Moreno-Cruz, Katharine L Ricke, and David W Keith. A simple model to account for regional inequalities in the effectiveness of solar radiation management. *Climatic change*, 110(3-4):649–668, 2012.

[100] Phillip Williamson and Carol Turley. Ocean acidification in a geoengineering context. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1974):4317–4342, 2012.

[101] John T Fasullo, Simone Tilmes, Jadwiga H Richter, Ben Kravitz, Douglas G MacMartin, Michael J Mills, and Isla R Simpson. Persistent polar ocean warming in a strategically geoengineered climate. *Nature Geoscience*, 11(12):910, 2018.

[102] Andy Jones, Jim M Haywood, Kari Alterskjaer, Olivier Boucher, Jason NS Cole, Charles L Curry, Peter J Irvine, Duoying Ji, Ben Kravitz, Jon Egill

Kristjansson, et al. The impact of abrupt suspension of solar radiation management (termination effect) in experiment g2 of the geoengineering model intercomparison project (geomip). *Journal of Geophysical Research: Atmospheres*, 118(17):9743–9752, 2013.

[103] Ulrike Niemeier and Claudia Timmreck. What is the limit of climate engineering by stratospheric injection of so 2? *Atmospheric Chemistry and Physics*, 15(16):9129–9141, 2015.

[104] Christoph Kleinschmitt, Olivier Boucher, and Ulrich Platt. Sensitivity of the radiative forcing by stratospheric sulfur geoengineering to the amount and strategy of the so 2 injection studied with the lmdz-s3a model. *Atmospheric Chemistry and Physics*, 18(4):2769–2786, 2018.

[105] Jane CS Long and John G Shepherd. The strategic value of geoengineering research. *Global environmental change*, 1:757–770, 2014.

[106] Helen de Coninck, Aromar Revi, Mustafa Babiker, Paolo Bertoldi, Marcos Buckeridge, Anton Cartwright, Wenjie Dong, James Ford, Sabine Fuss, J-C Hourcade, et al. Strengthening and implementing the global response. *IPCC*, 2018.

[107] Peter Irvine, Kerry Emanuel, Jie He, Larry W Horowitz, Gabriel Vecchi, and David Keith. Halving warming with idealized solar geoengineering moderates key climate hazards. *Nature Climate Change*, 9(4):295, 2019.

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

[109] Victor Brovkin, Vladimir Petoukhov, Martin Claussen, Eva Bauer, David Archer, and Carlo Jaeger. Geoengineering climate by stratospheric sulfur injections: Earth system vulnerability to technological failure. *Climatic Change*, 92(3):243–259, Feb 2009.

[110] Steven J Smith and Philip J Rasch. The long-term policy context for solar radiation management. *Climatic Change*, 121(3):487–497, 2013.

[111] Arnulf Grubler, Charlie Wilson, Nuno Bento, Benigna Boza-Kiss, Volker Krey, David L McCollum, Narasimha D Rao, Keywan Riahi, Joeri Rogelj, Simon De Stercke, et al. A low energy demand scenario for meeting the 1.5 c target and sustainable development goals without negative emission technologies. *Nature Energy*, 3(6):515, 2018.

[112] Johan Rockstrom, Owen Gaffney, Joeri Rogelj, Malte Meinshausen, Nebojsa Nakicenovic, and Hans Joachim Schellnhuber. A roadmap for rapid decarbonization. *Science*, 355(6331):1269–1271, 2017.

[113] Frederike Neuber and Konrad Ott. The buying time argument within the solar radiation management discourse. *Applied Sciences*, 10(13), 2020.

[114] EASAC. Negative emissions technologies: what role in meeting paris agreement targets? *EASAC policy report 35*, 2018.

[115] Philipp Gunther and Felix Ekardt. Human rights and large-scale carbon dioxide removal: Potential limits to beccs and daccs deployment. *Land*, 11(12):2153, 2022.

[116] S. V. Hanssen, V. Daioglou, Z. J. N. Steinmann, J. C. Doelman, D. P. Van Vuuren, and M. A. J. Huijbregts. The climate change mitigation potential of bioenergy with carbon capture and storage. *Nature Climate Change*, 10(11):1023–1029, Nov 2020.

[117] MATHILDE Fajardy, ALEXANDRE Koberle, NIALL MacDowell, and ANDREA Fantuzzi. Beccs deployment: a reality check. *Grantham Institute Briefing Paper*, (28), 2019.

[118] S Tilmes, BM Sanderson, and BC O'Neill. Climate impacts of geoengineering in a delayed mitigation scenario. *Geophysical Research Letters*, 43(15):8222–8229, 2016.

[119] M. Meinshausen, Z. Nicholls, J. Lewis, M. J. Gidden, E. Vogel, M. Freund, U. Beyerle, C. Gessner, A. Nauels, N. Bauer, J. G. Canadell, J. S. Daniel, A. John, P. Krummel, G. Luderer, N. Meinshausen, S. A. Montzka, P. Rayner, S. Reimann, S. J. Smith, M. van den Berg, G. J. M. Velders, M. Vollmer, and H. J. Wang. The ssp greenhouse gas concentrations and their extensions to 2500. *Geoscientific Model Development Discussions*, 2019:1–77, 2019.

[120] Keywan Riahi, Detlef P. van Vuuren, Elmar Kriegler, Jae Edmonds, Brian C. O'Neill, Shinichiro Fujimori, Nico Bauer, Katherine Calvin, Rob Dellink, Oliver Fricko, Wolfgang Lutz, Alexander Popp, Jesus Crespo Cuaresma, Samir KC, Marian Leimbach, Leiwen Jiang, Tom Kram, Shilpa Rao, Johannes Emmerling, Kristie Ebi, Tomoko Hasegawa, Petr Havlik, Florian HumpenÃ¶der, Lara Aleluia Da Silva, Steve Smith, Elke Stehfest, Valentina Bosetti, Jiyong Eom, David Gernaat, Toshihiko Masui, Joeri Rogelj, Jessica Strefler, Laurent Drouet, Volker Krey, Gunnar Luderer, Mathijs Harmsen, Kiyoshi Takahashi, Lavinia Baumstark, Jonathan C. Doelman, Mikiko Kainuma, Zbigniew Klimont, Giacomo Marangoni, Hermann Lotze-Campen, Michael Obersteiner, Andrzej Tabeau, and Massimo Tavoni. The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, 42:153–168, jan 2017.

[121] J. T. Bacmeister, C. Hannay, B. Medeiros, A. Gettelman, R. Neale, H. B. Fredriksen, W. H. Lipscomb, I. Simpson, D. A. Bailey, M. Holland, K. Lindsay, and B. Otto-Bliesner. Co2 increase experiments using the cesm: Relationship to climate sensitivity and comparison of cesm1 to cesm2. *Journal of Advances in Modeling Earth Systems*, 12(11):e2020MS002120, 2020. e2020MS002120 10.1029/2020MS002120.

[122] James S Risbey, Michael R Grose, Didier P Monselesan, Terence J O'Kane, and Stephan Lewandowsky. Transient response of the global mean warming rate and its spatial variation. *Weather and climate extremes*, 18:55–64, 2017.

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

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

[125] Thomas F Stocker, Dahe Qin, Gian-Kasper Plattner, Melinda Tignor, Simon K Allen, Judith Boschung, Alexander Nauels, Yu Xia, Vincent Bex, Pauline M Midgley, et al. Climate change 2013: The physical science basis, 2013.

[126] Daniel Huppmann, Elmar Kriegler, Volker Krey, Keywan Riahi, Joeri Rogelj, Katherine Calvin, Florian Humpenoeder, Alexander Popp, Steven K. Rose, John Weyant, Nico Bauer, Christoph Bertram, Valentina Bosetti, Jonathan Doelman, Laurent Drouet, Johannes Emmerling, Stefan Frank, Shinichiro Fujimori, David Gernaat, Arnulf Grubler, Celine Guivarch, Martin Haigh, Christian Holz, Gokul Iyer, Etsushi Kato, Kimon Keramidas, Alban Kitous, Florian Leblanc, Jing-Yu Liu, Konstantin Löffler, Gunnar Luderer, Adriana Marcucci, David McCollum, Silvana Mima, Ronald D. Sands, Fuminori Sano, Jessica Strefler, Junichi Tsutsui, Detlef Van Vuuren, Zoi Vrontisi, Marshall Wise, and Runsen Zhang. IAMC 1.5°C Scenario Explorer and Data hosted by IIASA, August 2019.

[127] Giulia Realmonte, Laurent Drouet, Ajay Gambhir, James Glynn, Adam Hawkes, Alexandre C Koberle, and Massimo Tavoni. An inter-model assessment of the role of direct air capture in deep mitigation pathways. *Nature communications*, 10(1):3277, 2019.

[128] Alistair A. Sellar, Colin G. Jones, Jane Mulcahy, Yongming Tang, Andrew Yool, Andy Wiltshire, Fiona M. O'Connor, Marc Stringer, Richard Hill, Julien Palmieri, Stephanie Woodward, Lee de Mora, Till Kuhlbrodt, Steve Rumbold, Douglas I. Kelley, Rich Ellis, Colin E. Johnson, Jeremy Walton, Nathan Luke Abraham, Martin B. Andrews, Timothy Andrews, Alex T. Archibald, Ségolène Berthou, Eleanor Burke, Ed Blockley, Ken Carslaw, Mohit Dalvi, John Edwards, Gerd A. Folberth, Nicola Gedney, Paul T. Griffiths, Anna B. Harper, Maggie A. Hendry, Alan J. Hewitt, Ben Johnson, Andy Jones, Chris D. Jones, James Keeble, Spencer Liddicoat, Olaf Morgenstern, Robert J. Parker,

Valeriu Predoi, Eddy Robertson, Antony Siahaan, Robin S. Smith, Ranjini Swaminathan, Matthew T. Woodhouse, Guang Zeng, and Mohamed Zerroukat. Ukesm1: Description and evaluation of the uk earth system model. *Journal of Advances in Modeling Earth Systems*, 0(ja), 2019.

[129] Kevin Anderson and Glen Peters. The trouble with negative emissions. *Science*, 354(6309):182–183, 2016.

[130] Kirsten Zickfeld, Andrew H MacDougall, and H Damon Matthews. On the proportionality between global temperature change and cumulative co2 emissions during periods of net negative co2 emissions. *Environmental Research Letters*, 11(5):055006, 2016.

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

[132] Katarzyna B Tokarska, Kirsten Zickfeld, and Joeri Rogelj. Path independence of carbon budgets when meeting a stringent global mean temperature target after an overshoot. *Earth's Future*, 7(12):1283–1295, 2019.

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

[134] Patrik L Pfister and Thomas F Stocker. The realized warming fraction: a multi-model sensitivity study. *Environmental Research Letters*, 13(12):124024, 2018.

[135] Myles Allen. Do-it-yourself climate prediction. *Nature*, 401:642 EP –, Oct 1999.

[136] D.J Frame, T Aina, C.M Christensen, N.E Faull, S.H.E Knight, C Piani, S.M Rosier, K Yamazaki, Y Yamazaki, and M.R Allen. The climate¡i¿prediction¡/i¿.net bbc climate change experiment: design of the coupled

model ensemble. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1890):855–870, 2009.

[137] N. Massey, R. Jones, F. E. L. Otto, T. Aina, S. Wilson, J. M. Murphy, D. Hassell, Y. H. Yamazaki, and M. R. Allen. weather@home—development and validation of a very large ensemble modelling system for probabilistic event attribution. *Quarterly Journal of the Royal Meteorological Society*, 141(690):1528–1545, 2015.

[138] M Collins, S Tett, and C Cooper. The internal climate variability of hadcm3, a version of the hadley centre coupled model without flux adjustments. *Climate Dynamics*, 17:61–81, 01 2001.

[139] D. A. Stainforth, T. Aina, C. Christensen, M. Collins, N. Faull, D. J. Frame, J. A. Kettleborough, S. Knight, A. Martin, J. M. Murphy, C. Piani, D. Sexton, L. A. Smith, R. A. Spicer, A. J. Thorpe, and M. R. Allen. Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, 433:403 EP –, Jan 2005.

[140] Peter A. Stott, Nikolaos Christidis, Friederike E. L. Otto, Ying Sun, Jean-Paul Vanderlinden, Geert Jan van Oldenborgh, Robert Vautard, Hans von Storch, Peter Walton, Pascal Yiou, and Francis W. Zwiers. Attribution of extreme weather and climate-related events. *Wiley Interdisciplinary Reviews: Climate Change*, 7(1):23–41, 2016.

[141] Matthew Collins, Ben BB Booth, B Bhaskaran, Glen R Harris, James M Murphy, David MH Sexton, and Mark J Webb. Climate model errors, feedbacks and forcings: a comparison of perturbed physics and multi-model ensembles. *Climate Dynamics*, 36(9-10):1737–1766, 2011.

[142] Benjamin M Sanderson, C Piani, WJ Ingram, DA Stone, and MR Allen. Towards constraining climate sensitivity by linear analysis of feedback patterns in thousands of perturbed-physics gcm simulations. *Climate Dynamics*, 30(2-3):175–190, 2008.

[143] DJ Frame, T Aina, CM Christensen, NE Faull, SHE Knight, C Piani, SM Rosier, K Yamazaki, Y Yamazaki, and MR Allen. The climate prediction. net bbc climate change experiment: Design of the coupled model ensemble. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1890):855–870, 2008.

[144] Gunnar Myhre, Eleanor J. Highwood, Keith P. Shine, and Frode Stordal. New estimates of radiative forcing due to well mixed greenhouse gases. *Geophysical Research Letters*, 25(14):2715–2718, 1998.

[145] James Hansen, MKI Sato, R Ruedy, L Nazarenko, A Lacis, GA Schmidt, G Russell, I Aleinov, M Bauer, S Bauer, et al. Efficacy of climate forcings. *Journal of geophysical research: atmospheres*, 110(D18), 2005.

[146] J. M. Gregory, W. J. Ingram, M. A. Palmer, G. S. Jones, P. A. Stott, R. B. Thorpe, J. A. Lowe, T. C. Johns, and K. D. Williams. A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, 31(3), 2004.

[147] Timothy Andrews, Jonathan M. Gregory, and Mark J. Webb. The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *Journal of Climate*, 28:1630–1648, 02 2015.

[148] Mark C Serreze and Roger G Barry. Processes and impacts of arctic amplification: A research synthesis. *Global and planetary change*, 77(1-2):85–96, 2011.

[149] October 9. Less rain but still wetter and greener?, Oct 2018.

[150] Judah Cohen, James A. Screen, Jason C. Furtado, Mathew Barlow, David Whittleston, Dim Coumou, Jennifer Francis, Klaus Dethloff, Dara Entekhabi, James Overland, and Justin Jones. Recent arctic amplification and extreme mid-latitude weather. *Nature Geoscience*, 7:627 EP –, Aug 2014. Review Article.

[151] D. A. Mooley and B. Parthasarathy. Indian summer monsoon and el nino. *pure and applied geophysics*, 121(2):339–352, Apr 1983.

[152] RH Kripalani, JH Oh, Ashwini Kulkarni, SS Sabade, and HS Chaudhari. South asian summer monsoon precipitation variability: coupled climate model simulations and projections under ipcc ar4. *Theoretical and Applied Climatology*, 90(3-4):133–159, 2007.

[153] Shayne McGregor, Myriam Khodri, Nicola Maher, Masamichi Ohba, Francesco S. R. Pausata, and Samantha Stevenson. *The Effect of Strong Volcanic Eruptions on ENSO*, chapter 12, pages 267–287. American Geophysical Union (AGU), 2020.

[154] P. J. Valdes, E. Armstrong, M. P. S. Badger, C. D. Bradshaw, F. Bragg, M. Crucifix, T. Davies-Barnard, J. J. Day, A. Farnsworth, C. Gordon, P. O. Hopcroft, A. T. Kennedy, N. S. Lord, D. J. Lunt, A. Marzocchi, L. M. Parry, V. Pope, W. H. G. Roberts, E. J. Stone, G. J. L. Tourte, and J. H. T. Williams. The bridge hadcm3 family of climate models: Hadcm3@bristol v1.0. *Geoscientific Model Development*, 10(10):3715–3743, 2017.

[155] Long Cao, Lei Duan, Govindasamy Bala, and Ken Caldeira. Simulated long-term climate response to idealized solar geoengineering. *Geophysical Research Letters*, 43(5):2209–2217, 2016.

[156] Doug McNeall, Jonny Williams, Richard Betts, Ben Booth, Peter Challenor, Peter Good, and Andy Wiltshire. Correcting a bias in a climate model with an augmented emulator. *Geoscientific Model Development*, 13(5):2487–2509, 2020.

[157] Daniele Visioni, Douglas G. MacMartin, Ben Kravitz, Jadwiga H. Richter, Simone Tilmes, and Michael J. Mills. Seasonally modulated stratospheric aerosol geoengineering alters the climate outcomes. *Geophysical Research Letters*, 47(12):e2020GL088337, 2020. e2020GL088337 2020GL088337.

[158] Han-Li Liu, Charles G. Bardeen, Benjamin T. Foster, Peter Lauritzen, Jing Liu, Gang Lu, Daniel R. Marsh, Astrid Maute, Joseph M. McInerney, Nicholas M. Pedatella, Liying Qian, Arthur D. Richmond, Raymond G. Roble, Stanley C. Solomon, Francis M. Vitt, and Wenbin Wang. Development and validation of the whole atmosphere community climate model with thermosphere and

ionosphere extension (waccm-x 2.0). *Journal of Advances in Modeling Earth Systems*, 10(2):381–402, 2018.

[159] Ciyou Zhu, Richard H Byrd, Peihuang Lu, and Jorge Nocedal. Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on mathematical software (TOMS)*, 23(4):550–560, 1997.

[160] Matthieu Guimberteau, Philippe Ciais, Agnes Ducharne, Juan Pablo Boisier, Ana Paula Dutra Aguiar, Hester Biemans, Hannes De Deurwaerder, David Galbraith, Bart Kruijt, Fanny Langerwisch, German Poveda, Anja Rammig, Daniel Andres Rodriguez, Graciela Tejada, Kirsten Thonicke, Celso von Randow, Rita C. S. Von Randow, Ke Zhang, and Hans Verbeeck. Impacts of future deforestation and climate change on the hydrology of the Amazon Basin: a multi-model analysis with a new set of land-cover change scenarios. *Hydrology and Earth System Sciences*, 21(3):1455–1475, 2017.

[161] Tobias Bischoff and Tapio Schneider. The equatorial energy balance, itcz position, and double-itcz bifurcations. *Journal of Climate*, 29(8):2997–3013, 2016.

[162] J. E. Smyth, R. D. Russotto, and T. Storelvmo. Thermodynamic and dynamic responses of the hydrological cycle to solar dimming. *Atmospheric Chemistry and Physics*, 17(10):6439–6453, 2017.

[163] T. Andrews, P. Forster, O. Boucher, N. Bellouin, and A. Jones. Precipitation, radiative forcing and global temperature change. *Geophys. Res. Lett.*, 37:L14701, 2010.

[164] G. Bala, K. Caldeira, and R. Nemani. Fast versus slow response in climate change: Implications for the global hydrological cycle. *Climate Dyn.*, 35:423–434, 2010.

[165] MHJ Van Huijgevoort, P Hazenberg, HAJ Van Lanen, and R Uijlenhoet. A generic method for hydrological drought identification across different climate regions. *Hydrology and Earth System Sciences*, 16(8):2437–2451, 2012.

[166] Aiguo Dai. Precipitation characteristics in eighteen coupled climate models. *Journal of climate*, 19(18):4605–4630, 2006.

[167] Saralees Nadarajah. The exponentiated gumbel distribution with climate application. *Environmetrics: The official journal of the International Environmetrics Society*, 17(1):13–23, 2006.