MIT CEEPR

MIT Center for Energy and Environmental Policy Research

Working Paper Series

A Multi-control Climate Policy Process for a Designated Decision Maker

HENRI F. DRAKE, RONALD L. RIVEST, ALAN EDELMAN, AND JOHN DEUTCH

JULY 2020

CEEPR WP 2020-015

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

A multi-control climate policy process for a designated decision maker

Henri F. Drake*[1,2], Ronald L. Rivest[1], Alan Edelman[1], and John Deutch[1]

[1] *Massachusetts Institute of Technology, Cambridge, MA, USA*

[2] *MIT-WHOI Joint Program in Oceanography/Applied Ocean Science & Engineering, Cambridge and Woods Hole, MA, USA*

Abstract

Persistent greenhouse gas (GHG) emissions threaten global climate goals and have prompted consideration of climate controls supplementary to emissions mitigation. We present an idealized model of optimally-controlled climate change, which is complementary to simpler analytical models and more comprehensive Integrated Assessment Models. We show that the four methods of controlling climate damage– mitigation, carbon dioxide removal, adaptation, and solar radiation modification– are not interchangeable, as they enter at different stages of the causal chain that connects GHG emissions to climate damages. Early and aggressive mitigation is always necessary to stabilize GHG concentrations at a tolerable level. The most cost-effective way of keeping warming below 2° C is a combination of all four controls; omitting solar radiation modification– a particularly contentious climate control– increases net control costs by 31%. At low discount rates, near-term mitigation and carbon dioxide removal are used to permanently reduce the warming effect of GHGs. At high discount rates, however, GHGs concentrations increase rapidly and future generations are required to use solar radiation modification to offset a large greenhouse effect. We propose a policy response process wherein climate policy decision-makers re-adjust their policy prescriptions over time based on evolving climate outcomes and revised model assumptions. We demonstrate the utility of the process by applying it to three hypothetical scenarios in which model biases in 1) baseline emissions, 2) geoengineering (CDR and SRM) costs, and 3) climate feedbacks are revealed over time and control policies are re-adjusted accordingly.

Introduction

Climate change due to anthropogenic greenhouse gas (GHG) emissions poses an existential threat to society Steffen et al. (2018). Ever since the direct link between GHGs and global warming was established in climate models over fifty years ago (Manabe and Wetherald, 1967), scientists have advocated for substantial emissions mitigation to stabilize global GHG concentrations and temperatures (Revelle et al., 1965). The discovery that humans were unintentionally modifying the climate was unsurprisingly followed by speculation about intentional climate control Kellogg and Schneider (1974). With every year of increasing GHG emissions and climate goals slipping out of reach Peters et al. (2020), calls for serious consideration of climate controls beyond just mitigation–and their implications– grow louder Buck (2012); Council et al. (1991); Crutzen (2006); Parson (2017); Victor et al. (2009).

Four climate controls have emerged as plausible candidates for use in the near future: emissions **M**itigation, carbon dioxide **R**emoval (CDR), **G**eo-engineering by Solar Radiation Modification (SRM), and **A**daptation. The four controls are not directly interchangeable as they enter at different stages of the causal chain of climate damages (Figure 1; Deutch,

2019; Moreno-Cruz et al., 2018):

$$\text{Emissions} \xrightarrow{\mathbf{M}} \text{GHGs} \xrightarrow{\mathbf{R}} \text{Forcing} \xrightarrow{\mathbf{G}} \text{Warming} \xrightarrow{\mathbf{A}} \text{Damages}. \tag{1}$$

Controls further down the chain generally carry greater risks, since they require carefully off-setting the various downstream effects of GHG emissions, but also have advantages: CDR is the only control that decreases GHG concentrations; SRM is quick to deploy and has low direct costs McClellan et al. (2012); and adaptation allows for flexibility in the other controls as any residual climate damages can be reduced by adapting to the new climate, to some extent Dow et al. (2013).

Numerous social or geopolitical factors may substantially limit or block deployments of certain controls, with solar radiation modification standing out as particularly contentious Caldeira and Ricke (2013); Parson and Keith (2013); Schäfer et al. (2013). Problems related to inequity Flegal and Gupta (2018), distrust Haerlin and Parr (1999); Lacey et al. (2018), or lack of governance Flegal et al. (2019); Ricke et al. (2013) are just a handful of examples. Here, we ignore many of these complexities– except in as much as they are implicitly included in costs and socio-technological constraints– and focus on the "best-case" scenario where a globally-trusted decision-maker prescribes global control policies and their policy prescriptions are exactly realized.

Our hypothetical trusted decision-maker must follow some set of principles on which to base their control policies. Two commonly-studied approaches are 1) the cost-benefit approach (e.g. Nordhaus, 1992), in which control costs are balanced against the benefits of avoided damages, and 2) the cost-effectiveness approach (e.g. Luderer et al., 2013), in which control costs are minimized subject to a prescribed climate constraint. The cost-effectiveness approach underlies the Paris Climate Agreement United Nations Framework Convention on Climate Change (2015), which aims to keep global warming well below 2 °C above pre-industrial levels and currently organizes global climate policy[1].

The conventional tool for optimizing global climate control are Integrated Assessment Models (IAMs), which are the result of coupling simple climate system models to simple energy-economy models (see Weyant, 2017, for a general overview of IAMs and their utility to date). In this paper, we 1) present an idealized model of optimally-controlled climate change which is complementary to both simpler analytical models and more comprehensive IAMs and 2) we propose a sequential policy process for periodic and critical re-evaluation of inevitably biased forecasts, which we illustrate with three hypothetical examples.

MARGO: An idealized model of optimally-controlled climate change

The **MARGO** model consists of a physical energy balance model of Earth's climate coupled to an idealized socio-economic model of climate damages and controls (Figure 1):

Mitigation of greenhouse gas emissions,
Adaptation to climate impacts,
Removal of carbon dioxide (CDR),
Geoengineering by solar radiation modification (SRM), and
Optimal deployment of available controls.

[1]Intended nationally determined contributions to this effort imply 2.6–3.1 °C of warming and will need to be strengthened at upcoming re-negotiations (and realized) to have a reasonable chance of keeping warming below 2 °C Rogelj et al. (2016).

The model is modular, fast, and customizable and can be run with several options of objective functions and constraints.

Each of the climate controls acts, in its own distinct way, to reduce the damages caused by a changing climate but carry their own deployment costs (including direct costs, research and development costs, infrastructure costs, regulatory costs). The model is designed to include key features of climate physics, economics, and policy as concisely as possible and in ways consistent with both theory and more comprehensive General Circulation Models and IAMs. The shortcoming of the model's simplicity is that while its results provide qualitative insights, the quantitative results are unreliable.

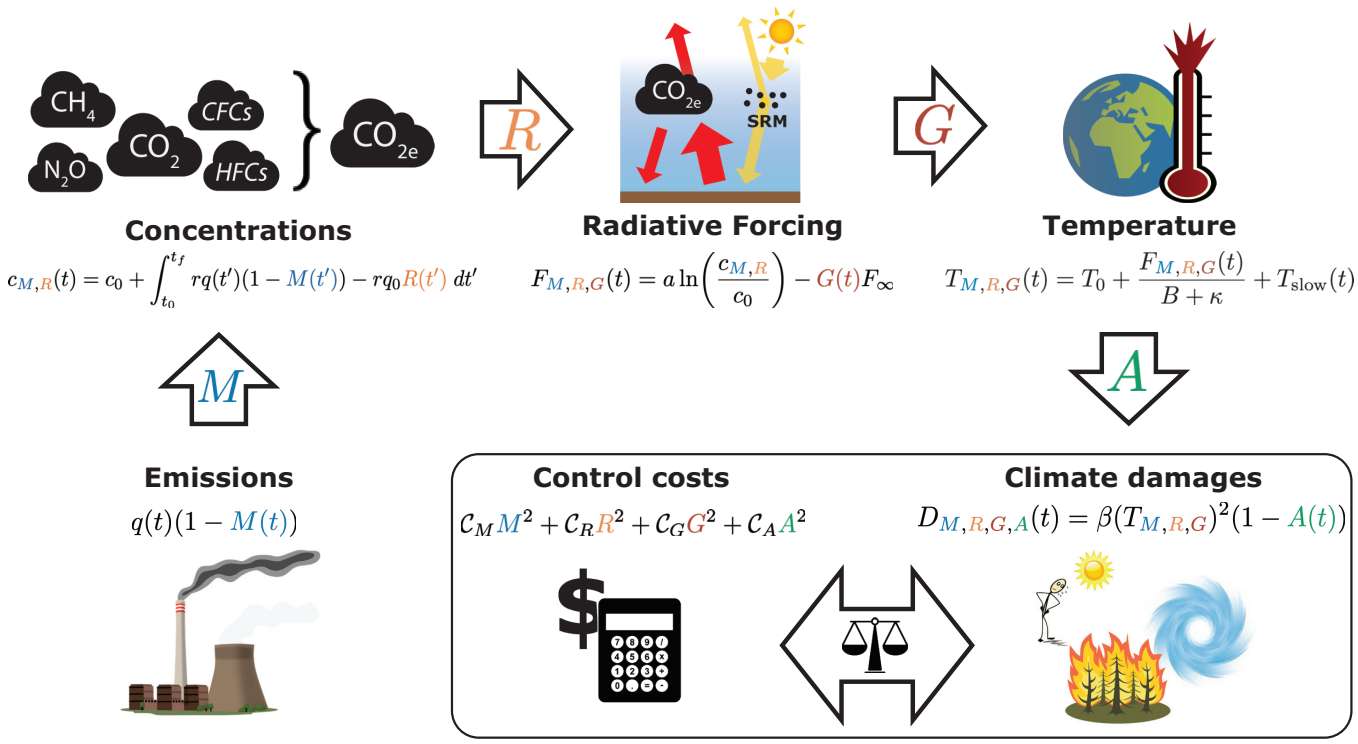

Figure 1: Schematic of the causal chain from greenhouse gas emissions to climate damages, including the unique effects of four climate controls: emissions **M**itigation, carbon dioxide **R**emoval, **G**eoengineering by Solar Radiation Management (SRM), and **A**daptation. Climate controls yield benefits in terms of avoided climate damages, which are balanced against control deployment costs.

The model is developed in open source using the Julia programming language Bezanson et al. (2017) at `github.com/hdrake/OptimizeClimate` (Drake et al., 2020). The model originated as an extension of a previous model Deutch (2019) to time-dependent control variables, although many improvements have been made since then. Each model component is expressed in closed form to facilitate analytical analysis and computation. Unlike most idealized climate-economic models, the entire MARGO framework can be explicitly written down in one or two expressions (Appendix A2). A derivation and interpretation of the two-box energy balance model– which has the same form as that of DICE Nordhaus and Sztorc (2013)– is included in the Methods. The parameter values used throughout the paper are set to the defaults mentioned in this section (and comprehensively listed in Appendix A2), except where explicitly stated otherwise. Validation experiments are summarized in the Methods and described in detail in the Appendix.

No-policy baseline scenario

Climate-controlled scenarios are considered relative to an exogenous no-policy baseline where carbon-dioxide equivalent (CO_{2e}) emissions $q(t)$ increase linearly four-fold by 2100 relative to 2020 and decrease linearly to zero by 2150, resulting in $7.3\,\mathrm{W/m^2}$ of radiative forcing by 2100 and $8.5\,\mathrm{W/m^2}$ by 2150, relative to preindustrial levels. As a result of this forcing, the global-mean temperature reaches $2\,^\circ\mathrm{C}$ by 2050 and soars to $T \approx 4.75\,^\circ\mathrm{C}$ by 2100, relative to preindustrial. We interpret this emission scenario as an idealized extension of the SSP3 baseline scenario, which is characterized by fossil-fueled growth Riahi et al. (2017).

There are five steps in the causal chain (eq. 1) between CO_{2e} emissions and climate damages.

1. CO_{2e} is emitted at a rate $q(t)$, with only a fraction $r = 50\%$ Solomon et al. (2009) remaining in the atmosphere after a few years, net of uptake by the ocean and terrestrial biosphere (Figure 2a).

2. CO_{2e} concentrations increase as long as the emissions $q(t)$ are non-zero, and are given by $c(t) = c_0 + \int_{t_0}^{t} rq(t)\,\mathrm{d}t$ (Figure 2b).

3. Increasing CO_{2e} concentrations strengthen the greenhouse effect, reducing outgoing longwave radiation and causing an increased radiative forcing of $F(t) = a\ln(c(t)/c_0)$, which exerts a warming effect on the surface.

4. Near-surface air temperatures eventually increase by $T(t) = F(t)/B$ to balance the reduced cooling to space, where $B/(\kappa + B) = 60\%$ of the warming occurs within a few years and the remaining $\kappa/(B + \kappa) = 40\%$ occurs over the course of several centuries due to ocean heat uptake Marshall and Zanna (2014). The feedback parameter B includes the effects of all climate feedbacks, except those involving the carbon cycle and the long-term ice sheet response (Figure 2c), and the ocean heat uptake rate κ parameterizes the combined effects of advection and diffusion of heat into the deep ocean.

5. Anthropogenic warming causes a myriad of climate impacts, which result in damages that increase non-linearly with temperature, $D = \beta T^2$.

Effects of climate controls

The four available climate controls enter as fractional controls at each link of the climate change causal chain (eq. 1).

Mitigation reduces emissions by a factor $M(t) \in [0, 1]$ such that the controlled emissions that remain in the atmosphere are $rq(t)\,(1 - M(t))$, where $M = 1$ corresponds to complete decarbonization of the economy.

Removal of CO_{2e}, $R(t) \in [0, 1]$, in contrast to mitigation, is de-coupled from instantaneous emissions and is expressed as the fraction of 2020 baseline emissions that are removed from the atmosphere in a given year, $q_0 R(t)$. A maximal value of $R = 1$ corresponds to removing $60\,\mathrm{GtCO_{2e}/year}$, which is more than twice a recent upper-bound estimate of the global potential for negative emission technologies Fuss et al. (2018).

A useful diagnostic quantity is the effective emissions

$$rq(t)(1 - M(t)) - rq_0 R(t), \tag{2}$$

which is the annual rate of CO_{2e} accumulation in the atmosphere (Figure 2a), with contributions from both emissions mitigation and CDR. The change in CO_{2e} concentrations is simply the integral of the effective emissions over time (Figure

2b),

$$c_{M,R}(t) = c_0 + \int_{t_0}^{t} rq(t')(1 - M(t'))\ \mathrm{d}t' - rq_0 \int_{t_0}^{t} R(t')\ \mathrm{d}t'. \tag{3}$$

Geoengineering by SRM, $G(t) \in [0, 1]$, acts to offset a fraction of the CO_{2e} forcing,

$$F_{M,R,G}(t) = F_{M,R}(t) - G(t)F_{\infty}, \tag{4}$$

where $F_{M,R} = a \ln(c_{M,R}(t)/c_0)$ is the controlled CO_{2e} forcing and $F_{\infty} = 8.5\,\mathrm{W/m^2}$ is the maximum baseline CO_{2e} forcing, which is attained starting in 2150, when baseline emissions are assumed to reach zero. A value of $G = 1$ thus corresponds to a complete cancellation between the equilibrium warming from baseline CO_{2e} increases and the cooling from a full deployment of SRM.

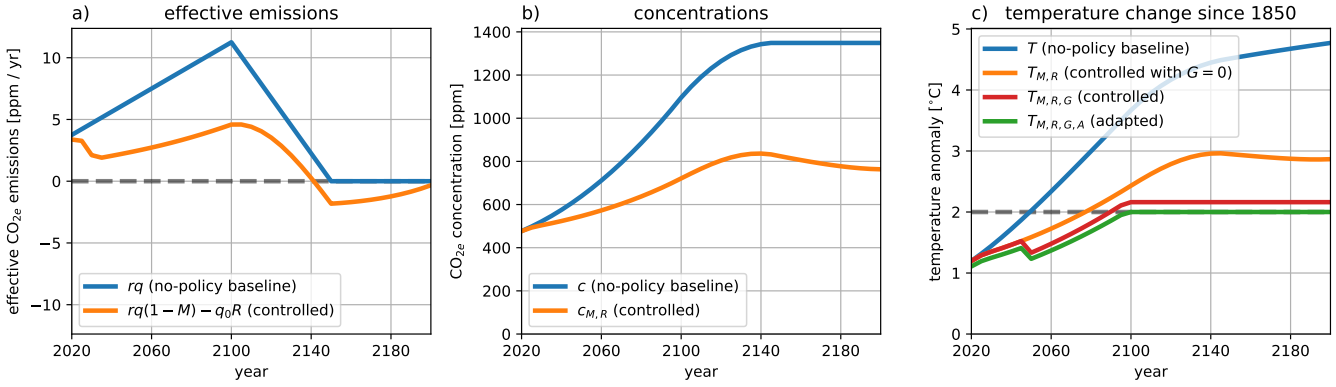

Figure 2: Baseline (blue) and optimally-controlled (orange) a) effective CO_{2e} emissions, b) CO_{2e} concentrations, and c) temperature anomaly relative to preindustrial from cost-effectiveness analysis. Panel c) shows the optimal temperature change that would occur: in a baseline scenario (blue); with just emissions **M**itigation and carbon dioxide **R**emoval (orange); with **M**itigation, **R**emoval, and solar-**G**eoengineering (red); and as an "adapted temperature" (eq. 7) with **A**daptation measures also taken into account. The dashed grey line marks the threshold adapted temperature of $T^{\star} = 2\,^{\circ}\mathrm{C}$ to be avoided. In (c), $T_{M,R,G}$ and $T_{M,R,G,A}$ decrease slightly in 2050 relative to $T_{M,R}$ as small but non-zero SRM deployment becomes permissible. Equivalent curves for cost-benefit analysis are shown in Figure S1.

The controlled near-surface air temperature (Figure 2c) evolves according to the total controlled forcing,

$$T_{M,R,G}(t) - T_0 = \frac{F_{M,R,G}(t)}{B + \kappa} + \frac{\kappa}{B} \int_{t_0}^{t} \frac{e^{\frac{t'-t}{\tau_D}}}{\tau_D} \frac{F_{M,R,G}(t')}{B + \kappa}\ \mathrm{d}t', \tag{5}$$

where $T_0 = 1.1\,^{\circ}\mathrm{C}$ is the present warming relative to preindustrial and $\tau_D = 240\,\mathrm{years}$ is the slow timescale of ocean heat uptake. The first term on the right-hand side of [20] represents a fast transient response while the second term represents a slow recalcitrant response due to the thermal inertia of the deep ocean (see Methods). Climate inertia decouples the temperature response from instantaneous forcing and implies that an additional fraction of short-term warming (or cooling) is locked in for the future, even if radiative forcing is stabilized (Lickley et al., 2019), as in the case of bringing emissions to zero in our model[2].

[2]In earth system models with a dynamic carbon cycle, the slow recalcitrant warming due to a reduction in ocean heat uptake happens to be roughly offset by the ocean carbon sink Solomon et al. (2009), such that bringing emissions to zero roughly stabilizes temperatures (Matthews and Caldeira, 2008). The model's realism would be improved by implementing a simple non-linear model of the ocean carbon cycle Glotter et al. (2014)

Adaptation to climate impacts acts to reduce damages by a fraction $A(t) \in [0, 40\%]$. Since some climate impacts are likely impossible to adapt to (Dow et al., 2013), we assume that adaptation can at most reduce climate damages by one-third. The controlled damages are thus given by

$$D_{M,R,G,A} = \beta(T_{M,R,G})^2(1 - A(t)), \tag{6}$$

where the damage parameter β is tuned such that a warming of 3 °C results in damages of the 2% of Gross World Product (GWP), consistent with DICE in the limit of non-catastrophic warming Nordhaus and Sztorc (2013). Although adaptation does not affect the planetary temperature directly, it is useful to consider an "adapted temperature" $T_{M,R,G,A}$ which yields controlled damages equivalent to the fully-controlled damages $\beta(T_{M,R,G,A})^2 = \beta(T_{M,R,G})^2(1 - A)$ and is defined

$$T_{M,R,G,A} \equiv T_{M,R,G}\sqrt{(1 - A)}. \tag{7}$$

Costs and benefits of controlling the climate

The costs of deploying climate controls are non-negligible and must be balanced with the benefits of controlling the climate to avoid climate impact damages. The costs of climate controls are parameterized as:

$$\mathcal{C} = \mathcal{C}_M M^2 + \mathcal{C}_R R^2 + \mathcal{C}_G G^2 + \mathcal{C}_A A^2, \tag{8}$$

where the C_* are the hypothetical annual costs of fully deploying that control (see Methods) and the cost functions are assumed to be convex functions of fractional deployment with zero initial marginal cost, as is customary Belaia (2019); Moreno-Cruz et al. (2018); Nordhaus (1992), and are here all taken to be quadratic for simplicity Deutch (2019); Moreno-Cruz et al. (2018). The benefits of deploying climate controls are the avoided climate damages relative to the no-policy baseline scenario,

$$\mathcal{B} = D - D_{M,R,G,A} = \beta(T^2 - (T_{M,R,G,A})^2). \tag{9}$$

Exogenous economic growth

In contrast to conventional IAMs, which follow classic economic theories of optimal economic growth and solve for the maximal welfare based on the discounted utility of consumption, we here treat economic growth as exogenous (as in Moreno-Cruz et al., 2018). The economy, represented by the GWP $E(t) = E_0(1 + \gamma)^{(t-t_0)}$, grows from its present value of $E_0 = 100$ trillion USD with a fixed growth rate $\gamma = 2\%$, consistent with DICE, expert opinion, and an econometric forecast model Christensen et al. (2018); Nordhaus and Sztorc (2013). We ignore feedbacks of climate abatement costs and climate damages on economic growth, since they are small variations relative to the exponential rate of economic growth in many IAM implementations Azar and Schneider (2002); Nordhaus and Sztorc (2013), but not all Glanemann et al. (2020).

Optimal deployments of climate controls

A trusted climate policy decision-maker specifies the objective function to maximize subject to additional policy constraints. The MARGO model is readily optimized in terms of the time-dependent climate control variables $M(t), R(t), G(t), A(t)$. The numerical implementation of the optimization, as well as additional socio-technological constraints on the permitted

timing and rates of deployments, are described in the Methods. Here, we describe the optimally-controlled results of two policy approaches, cost-benefit analysis and cost-effectiveness analysis, and explore their sensitivity to the discount rate ρ and possible limits to the fractional penetration of mitigation μ, respectively.

Cost-benefit analysis

A natural and widely-used approach is cost-benefit analysis, in which the cost $\mathcal{C}_{M,R,G,A}$ of deploying climate controls is balanced against the benefits $\mathcal{B}_{M,R,G,A}$ of the avoided climate damages. Formally, we aim to maximize the net present benefits:

$$\max\left\{\int_{t_0}^{t_f}\left(\mathcal{B}_{M,R,G,A}-\mathcal{C}_{M,R,G,A}\right)(1+\rho)^{-(t-t_0)}\,\mathrm{d}t\right\}, \tag{10}$$

where ρ is a social discount rate that determines the annual depreciation of future costs and benefits of climate control to society. There are different views about the appropriate non-zero discount rate to apply to multi-generational social utility Arrow et al. (2013); Ramsey (1928); Solow (1974); Stern et al. (2007). Here, we choose a discount rate of $\rho = 1\%$, on the low end of values used in the literature, motivated by our preference towards inter-generational equity Schneider (1989).

The results of maximizing net present benefits are shown in Figure 3. Early and aggressive emissions mitigation– and to a lesser extent CDR (Fig 3a)– drive net discounted costs of up to 1.5 trillion USD/year before 2075 relative to the no-policy baseline but deliver orders of magnitude more in net discounted benefits from 2075 to 2200 (Fig 3b). Effective CO_{2e} emissions reach net-zero by 2040 and concentrations stabilize at $c_{M,R} = 500\,\mathrm{ppm}$, slightly above present day $c_0 = 460\,\mathrm{ppm}$ (Figure S1a,b). In 2050, deployments of SRM become permissible and quickly scale up to a moderate level of $G = 15\%$, permanently bringing carbon-controlled temperatures from about $T_{M,R} \approx 1.5\,^{\circ}\mathrm{C}$ to $T_{M,R,G} \approx 0.75\,^{\circ}\mathrm{C}$ above preindustrial (Figure S1c). Deployments of adaptation are modest, reflecting its relatively high costs and its position at the end of the the causal chain of climate damage (eq. 1)

The preference for controls earlier in the causal chain, notably mitigation, is largely a result of the choice $\rho = 1\%$ for the discount rate (Figure 3c). In particular, if the discount rate increases above the economic growth rate Tol (2003), $\rho > \gamma = 2\%$, the time decay leads to a different regime of control preferences: the short-term fix offered by SRM overwhelmingly becomes the preferred control since the high future costs of its unintended climate damages are damped by the aggressive discounting of future costs. Adaptation emerges as the only control that peaks for intermediate values of the discount rate, since its benefits are experienced both in the short-term and long-term.

Cost-effectiveness of avoiding damage thresholds

The conventional cost-benefit approach to understanding climate change is limited by the poorly understood damage function Koomey (2013), which is likely to continue being revised as more is learned about its behavior at high levels of forcing (Alley et al., 2003; Burke et al., 2015). An alternative approach, which presently guides global climate policy negotiations, is to prescribe a threshold of climate damages– or temperatures, as in the Paris Climate Agreement United Nations Framework Convention on Climate Change (2015)– which is not to be surpassed.

In this implementation, we aim to find the lowest net present costs of control deployments

$$\min\left\{\int_{t_0}^{t_f}\mathcal{C}_{M,R,G,A}(1+\rho)^{-(t-t_0)}\,\mathrm{d}t\right\} \tag{11}$$

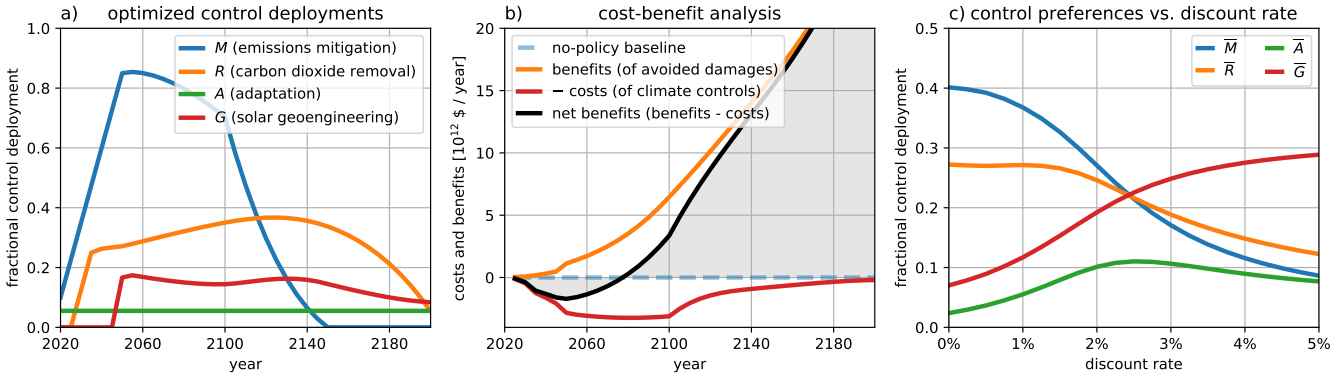

Figure 3: Results of cost-benefit analysis and sensitivity to the discount rate ρ. (a) Optimal control deployments and (b) corresponding discounted costs and benefits relative to the no climate-policy baseline scenario. The total positive area shaded in grey in (b) is the maximal net present benefits (eq. 10). (c) Time-mean control deployments as a function of the discount rate.

which keep controlled damages below the level corresponding to a chosen temperature threshold, $\beta(T_{M,R,G})^2(1-A(t)) < \beta(T^\star)^2$, which we rewrite

$$T_{M,R,G,A} < T^\star, \tag{12}$$

where $T_{M,R,G,A}$ is the "adapted temperature" (eq. 7).

The results of optimizing the cost-effectiveness of controls that keep adapted temperatures below $T^\star = 2\,^\circ\mathrm{C}$ are shown in Figures 2 and 4. Fractional emissions mitigation increases to a maximum of $M = 50\,\%$ decarbonization by 2035 and is maintained until emissions peak in 2100 (Figures 2a and 4a). Carbon dioxide is initially removed at rate of $Rq_0 \approx 15\%\, q_0 = 1.1\,\mathrm{ppm/year}$ starting in 2030, which ramps up to $Rq_0 \approx 30\%\, q_0 = 2.2\,\mathrm{ppm/year}$ by 2140. Since the optimally-controlled temperatures that result from the above cost-benefit analysis are already lower than $T^\star = 2\,^\circ\mathrm{C}$, the optimal controls from cost-effectiveness are less ambitious than for the cost-benefit analysis (Figures 3a, 4a), in contrast to some previous mitigation-only studies Hammitt (1999); Nordhaus (1992) but inline with recent analysis Glanemann et al. (2020) that uses an updated climate damage function Burke et al. (2015). As a consequence of relatively relaxed mitigation and CDR early on, a sizable deployment of SRM is used to shave off $1\,^\circ\mathrm{C}$ degree of warming at its peak in the mid-22nd Century in order to meet the temperature goal (Figure 4a and Figure 2c). Adaptation offsets $A = 15\,\%$ of damages and plays a moderate role in reducing damages to below the threshold. Even with discounting, annual costs of control deployments increase until 2100 and remain roughly constant in the 22nd Century (Figure 4b).

To explore the sensitivity of these results to our assumed mitigation costs $\mathcal{C}_M M^2$, which allow for up to 50% mitigation by 2035 at the relatively low cost of 700 billion USD/year, we compare the results against a re-optimization with steeper costs at high levels of mitigation. The mitigation cost function is modified to

$$C_M M^2 \left(1 - e^{-\left(\frac{1-M}{1-\mu}\right)}\right)^{-1}, \tag{13}$$

where we set the penetration limit of cheap mitigation to $\mu = 40\%$ and the function's structure is shown in Figure 4d. Mitigation costs are unchanged for $M \ll \mu$. Around $M \approx \mu$, low-hanging mitigation options are increasingly exhausted and costs begin to increase much more rapidly than the default assumption M^2. The high costs of deep decarbonization drive a reduction in the peak mitigation from $M = 50\%$ to nearly $M = 30\%$ in 2060, with the decreased mitigation being

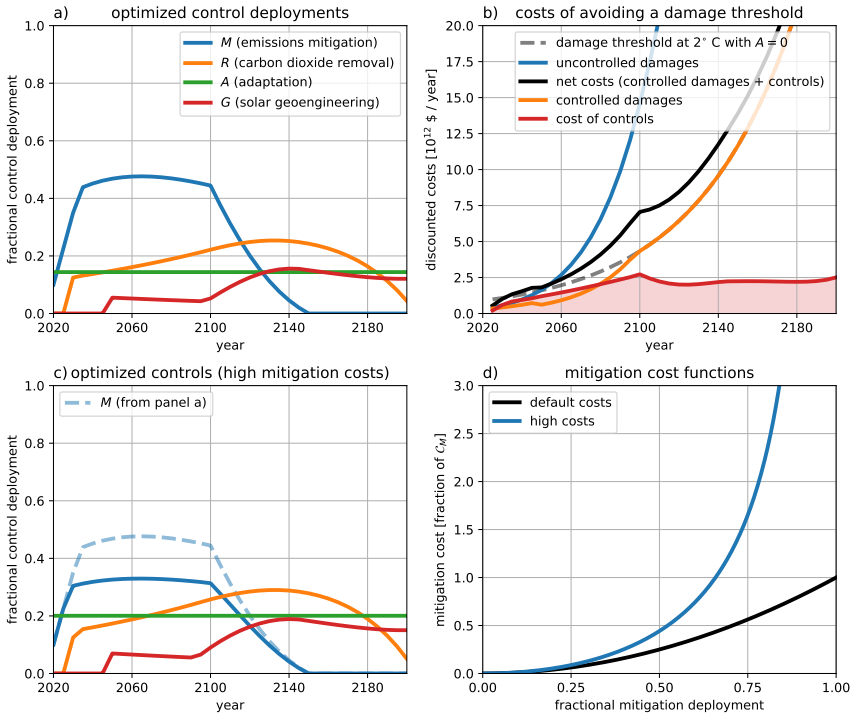

Figure 4: Results of cost-effectiveness analysis and sensitivity to potential limits μ to mitigation. (a) Optimal control deployments and (b) corresponding costs and damages. In panel (b), the blue line shows the discounted baseline uncontrolled damages; the dashed grey line shows the discounted damages associated with 2° of warming, which are to be avoided at all costs; the orange line shows the discounted damages in the optimally-controlled solution; and the red line shows the optimal discounted costs of controls such that the shaded area below is the minimal net present costs of controls (eq. 11). (c) Control deployments, as in (a), but re-optimized with high costs of deep decarbonization (blue line in d, eq. 13) relative to the default mitigation costs (black line in d). Mitigation in the default scenario (a) is reproduced as a dashed line in (c) for ease of comparison.

compensated by increases in the other three controls (Figure 4c).

Benefits of a complete portfolio of climate controls

To quantify the benefits of considering a complete portfolio of climate controls, as opposed to considering control technologies in isolation, we compute optimal control trajectories with all 15 combinations of the controls $\alpha \in \{M, A, R, G\}$, setting $\alpha \equiv 0$ for omitted technologies. The most cost-effective strategy includes all four controls and has a net present cost of 136 trillion USD (discount rate of $\rho = 1\%$). Since mitigation is the dominant control in the $\{MARG\}$ scenario (Figure 4a), the six most cost-effective portfolios include mitigation, with the no-SRM $\{MAR\}$ and mitigation-plus-CDR $\{MR\}$ scenarios costing only 31% and 38% more than the $\{MARG\}$ scenario, respectively (Table 1). The costs in single-control scenarios are much larger, with additional costs of 136% for the mitigation-only scenario $\{M\}$ to 201% for the SRM-only scenario $\{G\}$. In the adaptation-only $\{A\}$ and CDR-only $\{R\}$ scenarios, there is no solution that avoids an adapted temperature of $T^\star = 2\,°\mathrm{C}$, because we have imposed an adaptability limit $A < 40\%$ Dow et al. (2013) and limits to plausible levels of CDR $q_0 R < q_0 = 60\,\mathrm{GtCO_{2e}/year}$ (see Methods).

Table 1: Additional net present cost of avoiding an adapted temperature of $T^\star = 2\,^\circ\mathrm{C}$, relative to the 136 trillion USD net present cost of controls in the $\{MARG\}$ reference scenario with all four controls available: mitigation (M), adaptation (A), CDR (R) and SRM (G). Since we have imposed upper bounds $A < 40\%$ and $q_0 R < q_0 = 60\,\mathrm{GtCO_{2e}}$/year on adaptation and CDR, there is no scenario in which they can, in isolation, keep damages below those associated with $T^\star = 2\,^\circ\mathrm{C}$ of warming.

MARG	MRG	MAR	MAG	MR	MG	ARG	RG
0%	5%	31%	34%	38%	46%	63%	96%

MA	AG	M	G	AR	R	A	
105%	109%	136%	201%	216%	N/A	N/A	

A policy process for responding to uncertain future outcomes

Integrated Assessment Modelling (IAM) approaches assume perfect foreknowledge of model dynamics, parameters (or parameter distributions), and inputs. Future outcomes will differ from projections because the models are imperfect approximations of the socio-economic and physical climate systems they represent. For example, socio-economic models may assume erroneous future costs of climate controls Weyant (2008) and physical climate models may omit tipping elements (Steffen et al., 2018), both of which would lead to biases in model projections with respect to actual outcomes. Furthermore, the assumption of perfect foreknowledge degrades the active roles of policy decision-makers in determining baselines and control cost functions, and of climate researchers in refining estimates of physical model parameters.

A hypothetical trusted climate policy decision-maker must be in a position to respond to the inevitable differences that arise between model projections and actual outcomes and to revise their system understanding based on the newest developments in research. We show how our model equips climate policy decision-makers with the ability to periodically re-evaluate policy prescriptions by revising the underlying model structure and parameter values to correct for revealed biases.

The responsive control strategy process we propose is as follows:

1. Initial future trajectories of optimal control deployments are computed from the vantage point of $t = t_0$;
2. Model projections and control deployments are integrated forward one policy-making period to $t_1 = t_0 + \Delta t$;
3. Model structure and parameter values are revised, owing to new information obtained from observed outcomes and research developments;
4. Future trajectories of control deployments are re-optimized, now from the vantage point of $t_1 = t_0 + \Delta t$ and with revised model parameters;
5. Return to step 2, replacing $t_1 = t_0 + \Delta t$ with $t_n = t_{n-1} + \Delta t$ for period n, and repeat the process for the desired number of periods.

To illustrate the utility of the policy response process, we apply it to three hypothetical future scenarios, in which the most cost-effective controls for keeping adapted temperatures below $T^\star = 2\,^\circ\mathrm{C}$ are sequentially re-optimized in response to changes in model inputs and parameters. As a point of reference, we note that the passage of time itself leads to minor

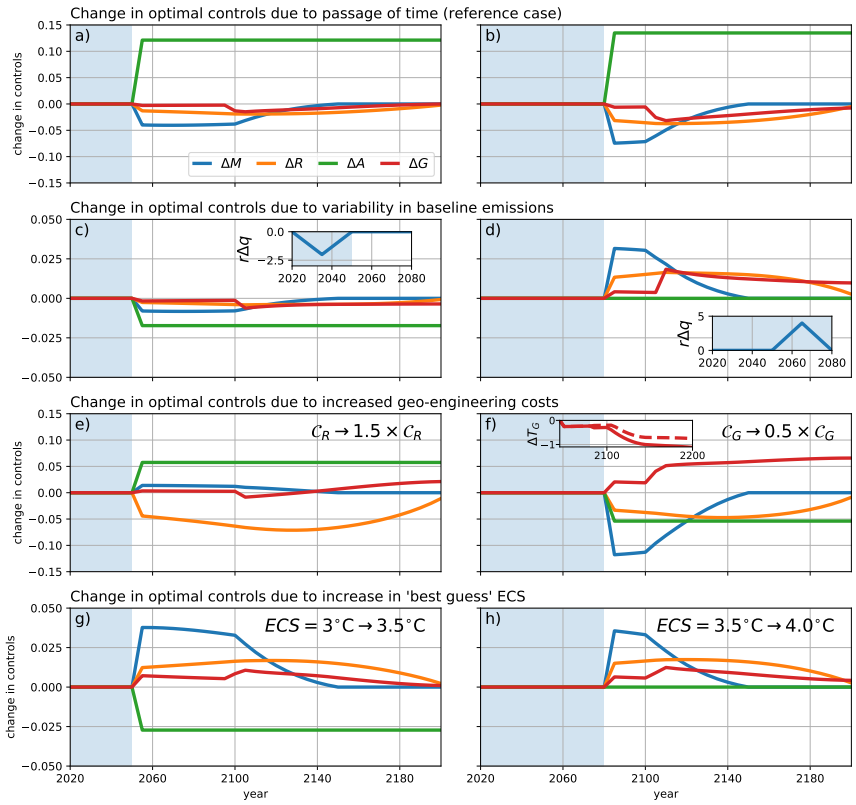

Figure 5: Illustration of the proposed policy process in which the optimally cost-effective control policies are periodically re-adjusted, relative to the original policies prescribed in 2020 (Figure 4a). In a reference case (a,b), time advances sequentially to 2050 (a) and 2080 (b) and policies are re-adjusted to reflect the new timelines. The blue shading shows the passage of time. The changes in control deployments shown in (c-h) are due to sequential re-optimization at 2050 (left) and 2080 (right), relative to the reference case (a,b), but now with revised model parameters: (c,d) where historical effective emissions $rq(t)$ are sequentially decreased and then increased (see insets); (e,f) where the costs of CDR and SRM are sequentially increased and decreased, respectively; and (g,h) where the best guess of the Equilibrium Climate Sensitivity (ECS) is revised upwards in 2050 and again in 2080. The inset in (f) shows the cooling due to SRM $\Delta T_G = T_{M,R,G} - T_{M,R}$ in the default scenario (dashed) and after the re-evaluation in 2080 shown in panel (d) (solid).

adjustments in the optimal combination of control deployments. As each successive generation is exposed to increasingly damaging temperatures, their most cost-effective solution is to increase adaptation measures, which past generations did not yet need, and save costs by slightly decreasing all other controls in the near future (Figure 5a,b). The control adjustments in the three scenarios below (Figure 5c-h) are shown relative to those in the reference case (Figure 5a,b).

Scenario 1: revealed bias in projected near-term baseline emissions

Suppose in $t_0 = 2020$ that the policy decision-maker prescribes aggressive climate control policies based on their cost-effectiveness at keeping warming below $T^\star = 2\,^\circ\mathrm{C}$ (step 1; Figure 4a) and that these optimal climate controls are perfectly implemented over the following $\Delta t = 30$ years (step 2).

The policy decision-maker directs a re-evaluation of the optimal control strategy at $t_1 = 2050$. The actual baseline emission trajectory between $t_0 = 2020$ and $t_1 = 2050$ is found to be $r\Delta q = 1\,\mathrm{ppm/year}$ lower than projected on average (Figure 5c,

inset), resulting in lower CO_{2e} concentrations than anticipated and a projected maximum warming of $\max(T_{M,R,G,A}) = 1.9\,^{\circ}\text{C}$, well below the $T^{\star} = 2\,^{\circ}\text{C}$ goal. The model inputs are thus revised to account for these lower-than-expected historical baseline emissions (step 3) and the optimal future control trajectories are re-computed (step 4). Reduced historical emissions imply a larger remaining carbon budget (Millar et al., 2016) and allow the policy decision-maker to slightly relax control deployments while still remaining below $T^{\star} = 2\,^{\circ}\text{C}$ of warming (Figure 5c), resulting in 12 trillion USD of avoided net present control costs. At this point, the policy decision maker must decide whether to continue existing policies that lead to $1.9\,^{\circ}\text{C}$ of warming or to reduce future controls deployments (and costs) at the risk of increased climate impacts due to an additional $0.1\,^{\circ}\text{C}$ of warming.

Suppose that, after following the re-optimized control trajectories for another $\Delta t = 30$ years (step 5), the historical effective baseline emissions must now be revised upwards by 2 ppm/year on average (Figure 5d, inset). With existing policies, the increased historical emissions would result in a $0.13\,^{\circ}\text{C}$ overshoot of the $T^{\star} = 2\,^{\circ}\text{C}$ degree goal. The most cost-effective adjustment to existing control policies that is consistent with the temperature goal is to increase mitigation, CDR, and SRM efforts by an additional $\Delta M = 3\%$, $\Delta R = 2\%$, and $\Delta G = 2\%$ (Figure 5d), at a net-present cost of 10 trillion USD.

Scenario 2: revealed bias in projected geoengineering (CDR and SRM) costs

Suppose that at a re-evaluation in 2050, CDR is found to be 50 % more expensive than projected. The climate policy-maker directs deployment of the most cost-effective control trajectories which keep warming below $T^{\star} = 2\,^{\circ}\text{C}$, which are re-optimized with the revised cost of CDR. The result is to decrease CDR by $\Delta R = -5\,\%$ and instead increase adaptation by $\Delta A = 5\%$ (Figure 5e). The shift away from expensive CDR towards adaptation results in 11.5 trillion USD of avoided net present costs of control deployments, with little difference in climate damage outcomes.

Suppose that after an additional 30 years, during which SRM is ramped up to a modest but non-zero level $G = 5\,\%$ (Figure 4a), it becomes clear that the costs of unintended side-effect damages of SRM are less than half as large as expected. In this scenario, the optimal future trajectory is to expand SRM deployments in the 22nd Century to $G \approx 20\,\%$ (resulting in $\Delta T_G = T_{M,R,G} - T_{M,R} \approx -1.0\,^{\circ}\text{C}$ of cooling, up from $-0.6\,^{\circ}\text{C}$; Figure 5f, inset) and reduce future mitigation levels by $\Delta M = -10\,\%$ (Figure 5f), resulting in another 12.6 trillion USD of avoided net present control costs.

Scenario 3: revealed bias in estimates of climate sensitivity

Suppose that by 2050, a dramatically improved suite of general circulation climate models robustly exhibits Equilibrium Climate Sensitivities of $ECS = 3.5\,^{\circ}\text{C}$, up from $3\,^{\circ}\text{C}$ in recent years (Geoffroy et al., 2012), and further improvements result in $ECS = 4\,^{\circ}\text{C}$ by 2080. Each of these revisions effectively shrinks the remaining cumulative carbon budget and thus requires sequentially increased deployments of mitigation, CDR, and SRM in order to keep warming below $T^{\star} = 2\,^{\circ}\text{C}$ (Figures 5g, h).

This responsive policy process only works if adjustments are made sufficiently frequently. If the policy decision-maker had waited from 2020 until 2100 before re-adjusting their course for a higher climate sensitivity of $ECS = 4\,^{\circ}\text{C}$, there would already be enough warming baked into the system that $T_{M,R,G,A} = 2.2\,^{\circ}\text{C} > T^{\star}$ of warming would be inevitable– even if the optimal policy from 2020 (Figure 4a,b) had been perfectly implemented.

Discussion

Few studies have considered the combined use of mitigation, carbon dioxide removal (CDR), solar radiation modification (SRM), and adaptation for controlling climate damages. We have developed a multi-control, time-dependent model of optimally cost-beneficial or cost-effective climate policies, which extends and improves upon previous work Deutch (2019). Another recent study Moreno-Cruz et al. (2018) uses a similar conceptual model with time-dependent controls to analytically investigate the differences between different climate controls; however, this model's climate physics are reduced to a simple empirical relationship that is not as clearly applicable to the case of significant SRM, where the direct link between cumulative emissions and temperature falls apart. Despite these differences, our study reproduces two key conceptual results of both earlier studies: 1) the four different climate controls are not interchangeable, as they enter at different stages of the causal chain between emissions and damages, and 2) the most cost-effective solution to limiting climate damages is to use all four controls at our disposal. The first result emerges from the role of each control in modifying the basic stock-flow properties of the carbon and heat budgets in the climate system. The second result is a direct consequence of marginal control costs which 1) begin at zero and 2) are concave, and is not guaranteed to hold if either assumption fails. For example, if learning effects are strong enough to cause fractional deployments costs to become convex, then a single-control strategy could be more appealing. Alternatively, if substantial R&D investments are necessary before a control is deployed, the large up-front marginal cost may be disqualifying.

We have proposed a policy response process which highlights the iterative nature of climate policy decision-making. We show that this process can be used to periodically correct for revealed biases in our understanding of the climate-economic system, in order to avoid unanticipated climate damages or "excessive" spending on climate controls. We view our proposed policy response process as an improvement over previously proposed "sequential" and "adaptive" strategies, in which policies are periodically re-evaluated by following instructions from a subjectively-defined decision flow chart (e.g. Lempert et al., 1996). In our process, policy re-evaluations are always optimally cost-beneficial or cost-effective, although the parameters that govern this optimization can be periodically re-adjusted. We argue that our policy process based on re-optimization is more defensible than previous approaches but retains the benefits of the process being "adaptive".

For clarity of exposition, we have presented a fully deterministic version of the MARGO model. In actuality, key inputs such as the climate feedback parameter B (and the related climate sensitivity ECS) and the damage function $D(T)$ are extremely uncertain. Propagation of these uncertainties through a convex damage function typically increases expected climate damages and strengthens the case for early and aggressive climate control Wagner and Zeckhauser (2016). Future work includes 1) extending MARGO to a stochastic programming approach that accounts for uncertainty in the various input parameters (see Methods) and 2) implementing a Bayesian policy response process where prior parameter distributions can be updated based on observed outcomes Shayegh and Thomas (2015) or improved parameter estimates from research developments. Stochastic programming of IAMs is significantly complicated by their endogenous economic models Crost and Traeger (2013); the model presented here is significantly more endogenous and may prove to be a useful framework for straight-forward multi-stage stochastic programming Webster et al. (2012).

The greatest caveat of the present study may be the assumption of a single trusted decision-maker. This device evidently avoids the complexities of a realistic decision making process that involve multiple stake holders with conflicting interests. The costs and benefits defined here are globally-aggregated; asymmetric costs and benefits between different regions lead to diverging incentives, which are further complicated as the number of unique climate controls increases. Asymmetric

multi-control incentives can be counter-intuitive: for example, one study suggests that high asymmetry in SRM damages drives even higher levels of mitigation because of the risk of SRM "free-drivers" Moreno-Cruz (2015).

Even in the case where climate control policies are prescribed by a single hypothetical decision-maker, there are sure to be inefficiencies in their implementation which we argue are more likely to result in under-deployment of controls than over-deployment. Considerable caution must be taken whenever relying on substantial CDR or SRM since neither of these controls exist as socio-technological systems capable of influencing climate, resulting in a "moral hazard" that shifts the burden to unconsenting future generations Flegal et al. (2019); Fuss et al. (2014).

The MARGO model is an idealized model which highlights the qualitatively different roles of mitigation, CDR, adaptation, and SRM in climate control. Both economic and physical components of the model have been abstracted as much as possible to highlight a small number ($N \approx 9$) of key parameters that govern the leading order behavior of the system (as compared to widely-used IAMs: Hope, 2006; Nordhaus, 1992; Tol, 1997): the climate feedback parameter B (related to the equilibrium climate sensitivtiy $ECS = F_{2\times \mathrm{CO}_2}$), the ocean heat uptake rate κ, the exogenous economic growth rate γ, the discount rate ρ, the climate damage parameter β, and the controls costs $\mathcal{C}_M$, $\mathcal{C}_R$, $\mathcal{C}_A$, $\mathcal{C}_G$ (Appendix A2 and Table 2). We show how the model can be used to investigate the sensitivity of "optimal" climate control policies to poorly constrained parameters, such as future control costs, and value-dependent parameters, such as the discount rate. We believe that our model resides in a sweet spot of being more realistic than semi-analytic models and easier to understand than conventional IAMs. We demonstrate that our model can be easily modified to reproduce the qualitative results of other studies (e.g. Belaia, 2019; Soldatenko and Yusupov, 2018, Appendix A3) and hope that it will be a useful community tool for extending simpler models, interpreting more comprehensive models, and bridging the gaps between climate economists, scientists, policy decision-makers, and the public Buck (2010); Pindyck (2017); Schneider (1997).

Methods

All data and figures used in the study can be found at `github.com/hdrake/OptimizeClimate` and are readily reproduced or modified by the Jupyter notebooks therein.

Control costs

The scaling costs for the four controls used in the present study are subjectively tuned; we here describe our rationale for choosing the parameter values. We remind the reader that the purpose of the MARGO model is to reveal insights about trade offs between the multiple controls and the dependence of model results on structural and parameteric choices. The interested reader can choose their own parameter values and see how the results change by visiting our web-browser application at `github.com/hdrake/OptimizeClimate` (placeholder until we have a better webapp).

The costs of mitigation are set according to the Working Group III contribution to Intergovernmental Panel on Climate Change's Fifth Assessment Report Clarke et al. (2014). In aggressive mitigation scenarios where CO_{2e} emissions decrease 78% to 118% by 2100, they estimate abatement costs of about 2% of GWP (see their Figure 6.21, panel f). Thus, we set the scaling cost of mitigation controls to $\mathcal{C}_M = \tilde{\mathcal{C}}_M E(t)$, where the cost of mitigating all emissions is $\tilde{\mathcal{C}}_M = 2\%$ of the GWP $E(t)$.

The costs of CDR are set according to bottom-up cost estimates from (Fuss et al., 2018, their Table 2). We compute the

mean cost of negative-emissions technologies, where we weight the median cost of each negative-emissions technology (in $\mathrm{USD/tCO_2}$) by its upper-bound potential for carbon-dioxide removal (in $\mathrm{GtCO_2/year}$). This leads to a total potential of roughly $q_0/2 \approx 26\,\mathrm{GtCO_2/year}$ at an average cost of $\overline{C}_R = 110\,\mathrm{USD/tCO_2}$. The scaling cost is thus set based on an estimate for $R = 50\%$, i.e. $\mathcal{C}_R \left(\frac{1}{2}\right)^2 = \overline{C}_R\, q_0/2$ or $\mathcal{C}_R = 2\overline{C}_R\, q_0 = 13$ trillion USD/year.

The costs of SRM largely reflect the costs of unintended climate damages that result due to their imperfect compensation for GHG forcing Irvine et al. (2017). Relative to both the costs of unintended damages and the costs of other climate controls, the direct costs of SRM measures are thought to be small (McClellan et al., 2012), as in the most commonly studied proposal of releasing gaseous sulfate aerosol precursors into the stratosphere to reflect sunlight back to space. The reference cost of SRM is thus given by $\mathcal{C}_G(t) = \tilde{\mathcal{C}}_G E(t)$, where $\tilde{\mathcal{C}}_G$ is the damage due to deploying $-F_\infty \equiv -F(t \to \infty) = -8.5\,\mathrm{Wm^{-2}}$ worth of SRM, as a fraction of the exogenous GWP $E(t)$. In the face of considerable uncertainties about the climate impacts of large-scale SRM (Irvine et al., 2017), we make the conservative assumption that the unintended damages of SRM are as large as the uncontrolled damages due to an equivalent amount of $\mathrm{CO_{2e}}$ forcing (as in Belaia, 2019; Goes et al., 2011), i.e. $\tilde{\mathcal{C}}_G \equiv \tilde{\beta}(F_\infty/B)^2 \approx 4.6\,\%$, where F_∞/B is the equilibrium temperature response to a fixed radiative forcing of $F_\infty = 8.5\,\mathrm{Wm^{-2}}$.

The costs of adaptation are estimated based on a recent joint report from the United Nations, the Bill and Melinda Gates Foundation, and the World Bank. They estimate that adaptation measures costing 1.8 trillion USD from 2020 to 2030 generate more than five times as much in total net benefits. Here, we make the crude assumption that this level of spending (180 billion USD / year) reduces climate damages by $A = 20\%$, i.e. $\mathcal{C}_A \left(\frac{1}{5}\right)^2 = 180$ billion USD / year, or $\mathcal{C}_A = 4.5$ trillion USD / year. We additionally cap adaptation at $A < 1/2$, recognizing that adaptation to all climate impacts is impossible: there will always be residual damages that can not be adapted to Dow et al. (2013).

Optimization method

We use the Interior Point Optimizer Wächter and Biegler (2006) (`https://github.com/coin-or/Ipopt`), an open source software package for large-scale nonlinear optimization, to minimize objective functions representing benefits and costs to society subject to assumed policy constraints. In practice, the control variables $\alpha \in \mathcal{A} = \{M, R, G, A\}$ are discretized into $N = (t_f - t_0)/\delta t$ timesteps (default $\delta t = 5\,\mathrm{years}$, $N = 36$) resulting in a $4N$-dimensional optimization problem. In the default (deterministic and convex) configuration, the model takes only $\mathcal{O}(10\,\mathrm{ms})$ to solve after just-in-time compiling and effectively provides user feedback in real time. This makes the model amenable to our forthcoming interactive web application, which is inspired by the impactful En-ROADS model web application Siegel et al. (2018).

The model was designed from the beginning with the goal of eventual use in stochastic simulations where 1) the determinstic scalar objective function can be generalized to an expected value of a probabilistic ensemble of simulations that sample an uncertain parameter space, and 2) determinstic constraints can be generalized to probablistic constraints (e.g. having a two-thirds chance of keeping temperatures below a goal $T^\star$), although these features are still under active development.

Social, technological, and economic inertia

For each control $\alpha \in \mathcal{A} = \{M, R, G, A\}$, we assert a maximum deployment rate

$$\left|\frac{\mathrm{d}\alpha}{\mathrm{d}t}\right| \leq \dot{\alpha}, \tag{14}$$

as a crude parameterization of social, technological, and economic inertia Ha-Duong et al. (1997), which acts to forbid implausibly aggressive deployment Buck (2016) and phase-out scenarios (see Appendix A2 for more discussion). We set $\dot{M} \equiv \dot{R} \equiv 1/40\,\mathrm{years}^{-1}$ in line with the most ambitious climate goals Intergovernmental Panel on Climate Change (2018) and $\dot{G} = 1/20\,\mathrm{years}^{-1}$ to reflect the technological simplicity of attaining a large SRM forcing relative to mitigation and CDR. We interpret adaptation deployment costs as buying insurance against future damages at a fixed annual rate $\mathcal{C}_A A^2$, with $\dot{A} = 0$, which can be increased or decreased upon re-evaluation at a later date.

We also set a control readiness condition which optionally limits how soon each control is "ready" to be deployed. In particular, in the default configuration we set $t_R = 2030$ and $t_G = 2050$ because CDR has not yet been deployed at a climatically significant scale Minx et al. (2018) and SRM does not yet exist as a socio-technological system (Flegal et al., 2019).

Two-box energy balance model

The evolution of the global-mean near-surface temperature anomaly (relative to the initial time $t_0 = 2020$) is determined by the two-box linear energy balance model (Held et al., 2010):

$$C_U \frac{\mathrm{d}T}{\mathrm{d}t} = -BT - \kappa(T - T_D) + F(t), \tag{15}$$

$$C_D \frac{\mathrm{d}T_D}{\mathrm{d}t} = \kappa(T - T_D), \tag{16}$$

where eq. 15 represents the upper ocean with average temperature anomaly T, and eq. 16 represents the deep ocean with an average temperature T_D. The near-surface atmosphere exchanges heat rapidly with the upper ocean and thus the global-mean near-surface air temperature is also given by T. The physical model parameters are: the upper ocean heat capacity $C_U = 7.3\,\mathrm{W\ yr\ m^{-2}\ K^{-1}}$ (including a negligible contribution $C_A \ll C_U$ from the atmosphere); the deep ocean heat capacity $C_D = 106\,\mathrm{W\ yr\ m^{-2}\ K^{-1}}$; the climate feedback parameter $B = 1.13\,\mathrm{W\ m^{-2}\ K^{-1}}$; and the ocean mixing rate $\kappa = 0.73\,\mathrm{W\ m^{-2}\ K^{-1}}$. The parameter values are taken from the multi-model mean of values diagnosed from 16 CMIP5 models Geoffroy et al. (2012). The radiative forcing and temperature anomalies at $t_0 = 2020$ relative to preindustrial are $F(t_0) - F(t_{\mathrm{pre}}) = 2.5\,\mathrm{W\,m^{-2}}$ and $T_0 \equiv T(t_0) - T(t_{\mathrm{pre}}) = 1.1\,\mathrm{K}$, where we set $F_0 \equiv F(t_0) = 0\,\mathrm{W\,m^{-2}}$ and $T(t_{\mathrm{pre}}) = 0\,\mathrm{K}$ for convenience.

Since, by construction, the anthropogenic forcing $F(t)$ varies on timescales longer than the fast relaxation timescale $\tau_U = C_U/(B + \kappa) = 4\,\mathrm{years}$, we can ignore the time-dependence in the upper ocean and approximate

$$T \approx \frac{F + \kappa T_D}{B + \kappa}, \tag{17}$$

where the evolution of the deep ocean

$$C_D \frac{\mathrm{d}T_D}{\mathrm{d}t} \approx -\frac{B\kappa}{B + \kappa} T_D + \frac{\kappa}{B + \kappa} F \tag{18}$$

occurs on a slower timescale $\tau_D \equiv \dfrac{C_D}{B}\dfrac{B + \kappa}{\kappa} = 240\,\mathrm{years}$ (Held et al., 2010). This approximation is convenient because it permits a simple closed form solution, but should be avoided if the model is applied to scenarios with rapidly changing

forcing, such as studies of the transient response to an instantaneous doubling of CO_2 or the SRM "termination effect" (see Appendix A1 for validation of the approximation). Plugging the exact solution to eq. 18 into eq. 17 gives the closed-form solution

$$T(t) - T_0 = \frac{F(t)}{B+\kappa} + \frac{\kappa}{B}\frac{1}{(B+\kappa)}\int_{t_0}^{t} \frac{e^{-(t-t')/\tau_D}}{\tau_D} F(t')\,\mathrm{d}t'. \tag{19}$$

The evolution of the controlled temperature anomaly (eq. 20; Figure 2c) has the same form but is instead driven by the controlled net radiative forcing $F_{M,R,G}$.

We identify the first term on the right hand side of eq. 19 and eq. 20 as the transient climate response (Gregory and Forster, 2008), which dominates for $t - t_0 \ll \tau_D$, while the second term is a slower "recalcitrant" response due to a weakening of ocean heat uptake as the deep ocean comes to equilibrium with the upper ocean (Held et al., 2010). While the contribution of the recalcitrant component to historical warming is thought to be small, it contributes significantly to 21st century and future warming (Gregory and Forster, 2008; Held et al., 2010).

The behavior of the model on short and long timescales is illustrated by applying it to the canonical climate change experiment in which CO_2 concentrations increase at 1% per year until doubling. The temperature anomaly first rapidly increases until it reaches the Transient Climate Sensitivity $TCS = \frac{F_{2\times}}{B+\kappa} = 1.9\,^{\circ}\mathrm{C}$ around the time of doubling $t = t_{2\times}$, with $t_{2\times} - t_0 \ll \tau_D$ and $F_{2\times} = \alpha \ln(2)$, and then gradually asymptotes to the Equilibrium Climate Sensitivity $ECS = \frac{F_{2\times}}{B} = 3.1\,^{\circ}\mathrm{C} > TCS$ on a much longer timescale $t - t_0 \gg \tau_D$.

Model validation

In Section 1 of the SI, we show that subjecting the MARGO energy balance model to a stylized RCP8.5-like forcing accurately reproduces the multi-model mean response from an ensemble of 35 comprehensive general circulation climate models from the CMIP5 ensemble (Figure S2). In Appendix A3, we show that by tweaking just a few of these default parameter values (Tables 2 and 3), the model replicates the qualitative results of studies ranging from analytical control theory analysis of SRM deployments Soldatenko and Yusupov (2018) to numerical optimizations of mitigation, CDR, and SRM deployments in a recent application of DICE Belaia (2019), a commonly used Integrated Assessment Model Nordhaus (1992).

Acknowledgements

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 174530. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

A Validation of MARGO's approximate two-box Energy Balance Model

A.1 Comparison with CMIP5 simulations under RCP8.5

The two-box Energy Balance Model (EBM) used in the **MARGO** model is described in the main text Methods. Here, we validate the MARGO-EBM by comparing it to an ensemble of 35 CMIP5 models under the RCP8.5 forcing scenario. We further validate the MARGO-EBM's approximation to the two-layer box model (in the equilibrated-thermocline limit $C_U \ll C_D$). We validate the approximation in three different high-forcing regimes: 1) the RCP8.5 scenario with large but gradual changes in forcing over the 21st Century; 2) the long-term (800 year) approach to equilibrium in an extended RCP8.5 scenario (ECP8.5); and 3) the short-term response to deployment and termination of large-amplitude solar radiation modification (SRM).

First, we construct an idealized forcing scenario that is meant to appromximate RCP8.5 Riahi et al. (2007) and its extension beyond 2100, ECP8.5 Meinshausen et al. (2011). In our scenario, baseline CO_{2e} emissions: 1) increase exponentially with a growth rate of $1/37\,\text{years}^{-1}$ to reach a maximum of $410\,\text{GtCO}_{2e}$/year in 2100, approximately 7 times present-day emissions; 2) plateau between 2100 and 2120; and 3) decrease linearly to zero between 2120 and 2200 (Figure 7a). As a result, CO_{2e} concentrations increase exponentially from the preindustrial value $c_0 = 280\,\text{ppm}$ in 1850 to 1400 ppm in 2100. In the extended scenario ECP8.5, CO_{2e} concentrations continue to grow until stabilizing at 3000 ppm in 2200[3] (Figure 7b). These increases in CO_{2e} drive a radiative forcing which increases to $F = 8.5\,\text{W/m}^2$ by 2100 and stabilizes at $F = 12\,\text{W/m}^2$ by 2200 (Figure 7c). The forcing timeseries constructed here approximates the RCP8.5 and ECP8.5 scenarios reasonably well– compare our Figure 7c with Figure 4 of Meinshausen et al (2011; Meinshausen et al., 2011).

When subjecting the MARGO-EBM to the RCP8.5-like scenario introduced above, we almost exactly recover the multi-model-mean warming from the CMIP5 ensemble under RCP8.5 (Figure 7d, solid black and blue lines). The excellent agreement is not surprising, given that we have tuned our MARGO-EBM with parameter values calibrated to the CMIP5 models Geoffroy et al. (2012). The climate physics-based calibration used here Geoffroy et al. (2012) is more realistic than the calibrations of commonly-used IAMs Calel and Stainforth (2016) and more robust to out-of-sample climate forcings.

A.2 Evaluation of the equilibrated-thermocline approximation

The MARGO-EBM uses the equilibrated-thermocline approximation,

$$T_{M,R,G}(t) - T_0 = \frac{F_{M,R,G}(t)}{B+\kappa} + \frac{\kappa}{B}\int_{t_0}^{t} \frac{e^{\frac{t'-t}{\tau_D}}}{\tau_D}\frac{F_{M,R,G}(t')}{B+\kappa}\,\mathrm{d}t', \tag{20}$$

which is a valid solution of the two-layer equations (15-16) in the limit $C_U \ll C_D$. In Figure 7e we show that this approximation (dashed black line) introduces only very small errors relative to the full solution under the ECP8.5 forcing scenario (solid black line). The full solution is computed numerically by solving the two-layer EBM equations 15 and 16 using forward finite differences. If we dramatically reduce either the deep ocean heat uptake rate κ or the deep ocean heat capacity C_D, as is customary in IAMs Calel and Stainforth (2016), then the model 1) equilibrates much too quickly with the instantaneous forcing and 2) underestimates recalcitrant changes that occurs long after the radiative forcing is stabilized (Figure 7e, dotted black line).

[3] In the original definition of the ECP8.5 scenario Meinshausen et al. (2011), much of these CO_{2e} increases are the result of increases in other gases such as Methane, Nitrous Oxide, and Hydrofluorocarbons.

Since we are interested in the response of the MARGO-EBM to climate controls which may cause the controlled radiative forcing $F_{M,R,G}$ to deviate substantially from a high-emissions baseline scenario, we here validate the MARGO-EBM's response to a short-term impulse of radiative forcing. In Figure 8, we modify the above ECP8.5 scenario by adding a Gaussian negative radiative forcing anomaly due to short-term SRM. The negative forcing impulse is centered around 2075, has a magnitude of $F_G = -GF(t \to \infty) = -3.4\,\mathrm{Wm}^{-2}$ (for $G = 40\%$), and a timescale of $\sigma = 20\,\mathrm{years}$ (Figure 8a). This negative forcing results in a pronounced short-term *net* cooling between 2050-2070, followed by an extremely rapid warming from 2070 to 2080 as the SRM program terminates (Figure 8b,c). A weak residual cooling of $0.1\,^{\circ}\mathrm{C}$ propagates into the deep ocean and lingers for centuries (Figure 8c). Despite the neglect of upper-ocean thermal inertia in the equilibrated-thermocline approximation, the MARGO-EBM agrees well with the full solution of the two-box equations, the approximation lagging behind the full solution by roughly $\tau_U = 5\,\mathrm{years}$ (Figure 8c).

B Comprehensive model equations and parameter values

In the cost-effectiveness framing, the full formulation of the problem

$$\min \{ \text{ discounted costs } \} \quad \text{subject to} \quad T_{M,R,G,A} < T^{\star}$$

is given, in closed form, by:

$$\min \left\{ \left[E_0 (1+\gamma)^{(t-t_0)} \left(\tilde{\mathcal{C}}_M M^2 + \tilde{\mathcal{C}_G} G^2 \right) + \mathcal{C}_R R^2 + \mathcal{C}_A A^2 \right] (1+\rho)^{-(t-t_0)} \right\} \tag{21}$$

Subject to

$$\sqrt{1-A} \left[T_0 + \frac{a \ln \left(\frac{c_0 + \int_{t_0}^{t} rq(1-M)\mathrm{d}t' - q_0 \int R \mathrm{d}t'}{c_0} \right) - F_{\infty} G}{B + \kappa} + \frac{\kappa}{B} \int_{t_0}^{t} \frac{e^{\frac{t'-t}{\tau_D}}}{\tau_D} \left\{ \frac{a \ln \left(\frac{c_0 + \int_{t_0}^{t'} rq(1-M)\mathrm{d}t'' - q_0 \int_{t_0}^{t'} R \mathrm{d}t''}{c_0} \right) - F_{\infty} G}{B + \kappa} \right\} \mathrm{d}t' \right] < T^{\star}, \tag{22}$$

where $\tau_D = \dfrac{C_D}{B} \dfrac{B+\kappa}{\kappa}$ is a timescale specified by the physical parameter C_D. The cost-benefit equation can similarly be derived based on the equations in the main text.

The problem is fully characterized by the 19 "free" parameters in equations 21 and 22, the default values of which are reported in Table 2 (18 in the case of cost-effectiveness, which avoids the use of a poorly-constrained damage coefficient β). The 19 parameters are: 3 grid parameters t_0, t_f, δt; the 3 initial conditions T_0, c_0, E_0; the 1 carbon cycle parameter r; the 4 physical parameters a, B, κ, and C_D; the 3 economic parameters β, ρ, γ; and the 5 control cost parameters $\mathcal{C}_A$, $\mathcal{C}_R$, $\tilde{\mathcal{C}}_M$, $\tilde{\mathcal{C}}_G$, F_{∞}. The baseline emissions timeseries $q(t)$ is treated as exogenous and must be prescribed as an input. In the cost-effectiveness framework, the poorly-constrained damage parameter β is replaced by a prescribed temperature goal $T^{\star}$. The grid, initial condition, and physical parameters are well constrained, while the economic and cost parameters are heuristic interpretations of the wider climate and economic literature.

The control variables $\alpha \in \mathcal{A} = \{M, R, G, A\}$ satisfy several additional constraints, which could be thought of as an additional 20 parameters, at most, although many end up being unimportant or redundant across several parameters (1 and 2 are necessary physical constraints on the controls whereas 3, 4, and 5 simply make the model's behavior more realistic):

1. The controls must be positive, $\alpha \geq 0$;

2. They have an upper bound: $\alpha < \alpha_{\max}$. $M_{\max} = 1$ is by set by the definition of mitigation. $G_{\max} = 1$ is chosen because it results in a negative radiative forcing that exactly offsets the maximum GHG forcing of $8.5\,\mathrm{W/m^2}$. We set $A_{\max} = 40\%$ in acknowledgement of practical Dow et al. (2013) and theoretical Sherwood and Huber (2010) limits to adaptability (this is meant as more of a symbolic gesture rather than an estimate of how much climate damage might be adaptable). Finally, $R = 50\%$ is set based on a recent bottom-up estimate of the potential for carbon dioxide removal of existing (but not necessarily scalable) negative emissions technologies.

3. They have an initial condition $\alpha(t_0) = \alpha_0$, which are all set to zero except for $M_0 = 10\%$, since none of the other controls have yet been deployed at scale.

4. We set maximum deployment and termination rates $\left|\frac{\mathrm{d}\alpha}{\mathrm{d}t}\right| < \dot{\alpha}$, which represent economic, technological, and social inertia. We set $\dot{M} = \dot{R} = 1/40\,\mathrm{years}^{-1}$ as an upper limit on plausible timescales of global energy transition. On the other hand, we set $\dot{G} = 1/20\,\mathrm{years}^{-1}$ to reflect the fact that solar geo-engineering deployment capacity could in principle be ramped-up very quickly, possibly even in the absence of global governance or regulation. We interpret adaptation costs as buying insurance against future damages up-front, with both benefits and costs spread evenly in the future. Thus, we set $\dot{A} = 0$. The caveat is that we allow control policy re-evaluations, at which point the value of adaptation can in that timestep be increased or decreased to a new level (see Figure 5 of main text), without a limit on the rate of increase.

5. We implement "readiness" constraints, $\alpha(t) = 0$ for all $t < t_\alpha$, to reflect the fact that some controls, such as geoengineering (both carbon and solar), do not yet exist as climate-relevant socio-technological systems Flegal et al. (2019). In particular, we set $t_R = 2030$ and $t_G = 2050$.

C Qualitative replications of other climate control model analysis

To illustrative the potential utility of MARGO as a community tool, we show how run-time parameter values in MARGO can be tweaked to match the model configurations and results of other studies of climate control policies. One the one hand, MARGO can be tuned to the inputs and outputs of a comprehensive multi-control IAM configuration to reproduce its qualitative results (Section A; Belaia, 2019); on the other hand, MARGO can be simplified by setting many of the parameters to zero to emulate an analytical model of climate control by solar radiation modification (SRM) only (Section B; Soldatenko and Yusupov, 2018). The goal of this section is to show how with minimal modifications to the default MARGO model, we are able to replicate key figures from two very different studies. For discussion of the figures we attempt to replicate, we refer readers to the original studies Belaia (2019); Soldatenko and Yusupov (2018).

C.1 Belaia (2019): A multi-control extension of DICE with Mitigation, Carbon Dioxide Removal, and Solar Geo-engineering

Belaia (2019) extend DICE, a commonly-used globally-aggregated general equilibrium IAM, to include carbon dioxide removal (CDR) and solar radiative modification– which they refer to as solar geoengineering (SG)– to supplement DICE's emissions mitigation in controlling climate damages Belaia (2019).

Parameter	Default Configuration
t_0	2020
t_f	2200
δt	$5\,\mathrm{yr}$
c_0	$460\,\mathrm{ppm}$
T_0	$1.1\,\mathrm{K}$
a	$4.97\,\mathrm{Wm}^{-2}$
r	50%
B	$1.13\,\mathrm{W\,m^{-2}\,K^{-1}}$
κ	$0.72\,\mathrm{W\,m^{-2}\,K^{-1}}$
C_D	$106\,\mathrm{W\,yr\,m^{-2}\,K^{-1}}$
β	$0.22 \times 10^{12}\,\$\,\mathrm{yr^{-1}\,K^{-2}}$
ρ	1%
E_0	$100 \times 10^{12}\,\$\,\mathrm{yr^{-1}}$
γ	2%
$\mathcal{C}_A$	$4.5 \times 10^{12}\,\$\,\mathrm{yr^{-1}}$
$\mathcal{C}_R$	$13 \times 10^{12}\,\$\,\mathrm{yr^{-1}}$
$\tilde{\mathcal{C}}_M$	2 % (of GWP)
$\tilde{\mathcal{C}}_G$	4.6 % (of GWP)
F_∞	$8.5\,\mathrm{Wm}^{-2}$

Table 2: Values of the 19 free parameters that characterize the MARGO model.

To implement CDR and SRM, Belaia (2019) make two fundamental changes to DICE. Their modelling of SRM forcing is identical to ours. In terms of costs, they similarly make the conservative assumption that SRM costs are dominanted by unintended side effects and scale with the damage of an equivalent amount of GHG forcing, but they include this damage cost as an additive term to the climate damages rather than the control costs. Their approach is thus similar to ours in the case of cost-benefit analysis, but in the cost-effectiveness case they effectively ignore indirect SRM damages while reaping the benefits of its low direct costs. The version of DICE they use already permits moderate negative emissions, as an extension of the emissions mitigation curve to 120%, i.e. 100% mitigation of baseline emissions mitigation plus removal of an addition 20% of baseline emissions). To extend this further, Belaia (2019) allow for substantial CDR by extending the mitigation curve indefinitely, although the cost curves are convex such that CDR becomes increasingly expensive. They also appear to have modified the functional form of emissions mitigation to keep CDR costs relatively low. The rationale for modelling CDR as an extension of mitigation is unclear, since 1) emissions mitigation and carbon dioxide removal are distinct physical, industrial, and economic processes and 2) marginal CDR costs today are already lower than the backstop mitigation costs assumed in their scenarios.

To approximate the DICE configuration used by Belaia (2019), we make the changes to MARGO's default parameter values reported in Table 3. Notably, we extended the time from 2200 to 2500, increased the reference costs for mitigation by about 75%, and increased the reference costs for SRM by about 175%. We found it necessary to modify the physical climate

Parameter	Belaia (2019)	Soldatenko and Yusupov (2018)
t_0		
t_f	2500	2100
δt	1 year	1 year
c_0		
T_0		
a		
r	75%	
B	$0.8\times$ 1.13 W m^{-2} K^{-1}	
κ	$0.75\times$ 0.72 W m^{-2} K^{-1}	
C_D	$0.75\times$ 106 W yr m^{-2} K^{-1}	
β		
ρ	1.5%	
E_0		
γ		
$\mathcal{C}_A$		
$\mathcal{C}_R$		
$\tilde{\mathcal{C}}_M$	3.6 % (of GWP)	
$\tilde{\mathcal{C}}_G$	12.5 % (of GWP)	
F_∞	7.5 Wm^{-2}	

Table 3: Values of the 19 free parameters that characterize the MARGO model, modified to replicate results from other models. Blank cells denote parameters that are not changed from the default values in Table 2.

parameters in order to match their CO_{2e} concentrations, radiative forcing, and temperatures based on their baseline emissions scenario $q(t)$, which we approximated with piece-wise quadratic functions (Figure 9a, blue line). Additionally, we omit adaptation and carbon dioxide removal, $A_{\max} \equiv R_{\max} \equiv 0$; we effectively remove the upper limit on mitigation $M_{\max} = 10$; we increase socio-technological intertia for all controls to $\dot{\alpha} = 1/90\,\text{years}^{-1}$; we set initial mitigation to $M_0 = 3\%$; and we remove all "readiness" constaints, $t_\alpha = 2020$. Additionally, in order to match the mitigation cost curves in their Figure 1 9, we found it necessary to decrease the mitigation cost exponent from 2 to 1.8, as compared to 2.8 in DICE-2013 Nordhaus and Sztorc (2013) or 2.6 in DICE-2016 Nordhaus (2017).

Figure 9 shows the results of cost-benefit analysis for: a baseline scenario, a mitigation only scenario, a mitigation and CDR scenario, and a scenario with mitigation, CDR, and SRM. Figure 9 has been formatted exactly as Figure 4 of Belaia (2019; Belaia, 2019), which presents the results from equivalent simulations in their extension of DICE, for convenient side-by-side comparison.

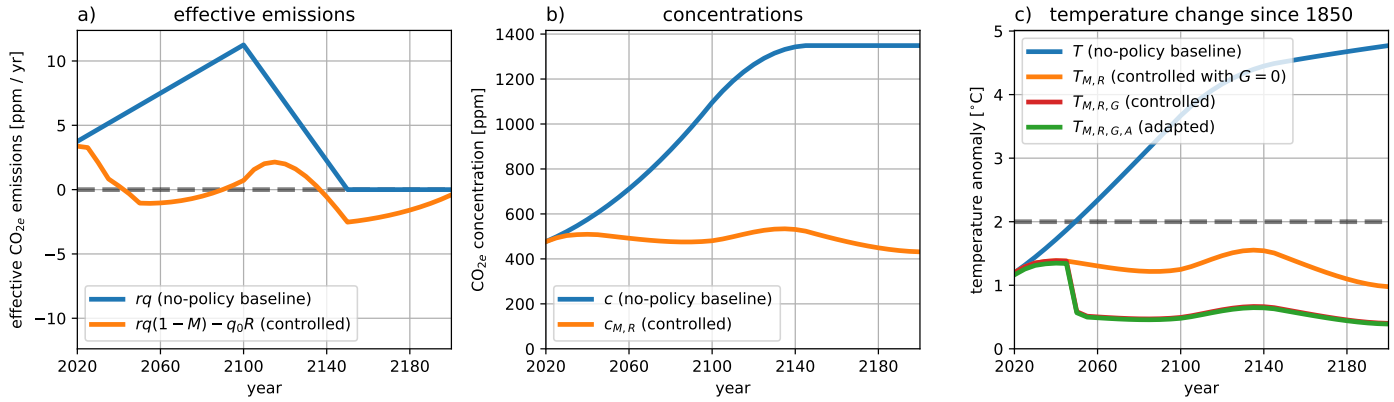

Figure 6: Baseline (blue) and optimally-controlled (orange) a) effective CO_{2e} emissions, b) CO_{2e} concentrations, and c) temperature anomaly relative to preindustrial from cost-benefit analysis. Panel c) shows the optimal temperature change that would occur: in a baseline scenario (blue); with just emissions **M**itigation and carbon dioxide **R**emoval (orange); with **M**itigation, **R**emoval, and solar-**G**eoengineering (red); and as an "adapted temperature" with **A**daptation measures also taken into account. The dashed grey line marks 2 °C for context. In (c), $T_{M,R,G}$ and $T_{M,R,G,A}$ decrease dramatically in 2050 relative to $T_{M,R}$ as moderate levels of SRM become permissible.

C.2 Soldatenko and Yusupov (2018): Analytical control theory applied to solar radiation modification

Soldatenko and Yusupov (2018; Soldatenko and Yusupov, 2018) develop an analytical model for the optimally cost-effective time-dependent deployment of solar radiation modification (SRM) which keeps temperatures in all years below $T^\star = T_0 + 1\,°\text{C}$ and keeps temperatures at their end date of 2100 below T_0. Although their representation of SRM forcing is more involved then ours and depends on the mass of sulfate aerosol injected, the resulting optimization problem is remarkably similar to an SRM-only configuration of the default MARGO model.

To approximate Soldatenko and Yusupov (2018)'s analytical model Soldatenko and Yusupov (2018), we make the changes to MARGO's default parameter values reported in Table 3. Additionally, we omit adaptation, carbon dioxide removal, and mitigation, $A_{\max} \equiv R_{\max} \equiv M_{\max} \equiv 0$; we remove all "readiness" constaints, $t_\alpha = 2020$, we set $T^\star = 2.1\,°\text{C}$ (1 °C above T_0) and add an additional constraint $T_{M,R,G} < T_0$ on the final timestep at $t_f = 2100$ (the latter is the only modification that required modifying compiled model source code).

Figure 10 shows the result of cost-effectiveness optimization for an SRM-only scenario, which is formatted to be directly comparable to Figure 3 of Soldatenko and Yusupov (2018).

References

Alley, R. B., Marotzke, J., Nordhaus, W. D., Overpeck, J. T., Peteet, D. M., Pielke, R. A., Pierrehumbert, R. T., Rhines, P. B., Stocker, T. F., Talley, L. D., and Wallace, J. M. (2003). Abrupt Climate Change. *Science*, 299(5615):2005–2010.

Arrow, K., Cropper, M., Gollier, C., Groom, B., Heal, G., Newell, R., Nordhaus, W., Pindyck, R., Pizer, W., Portney, P.,

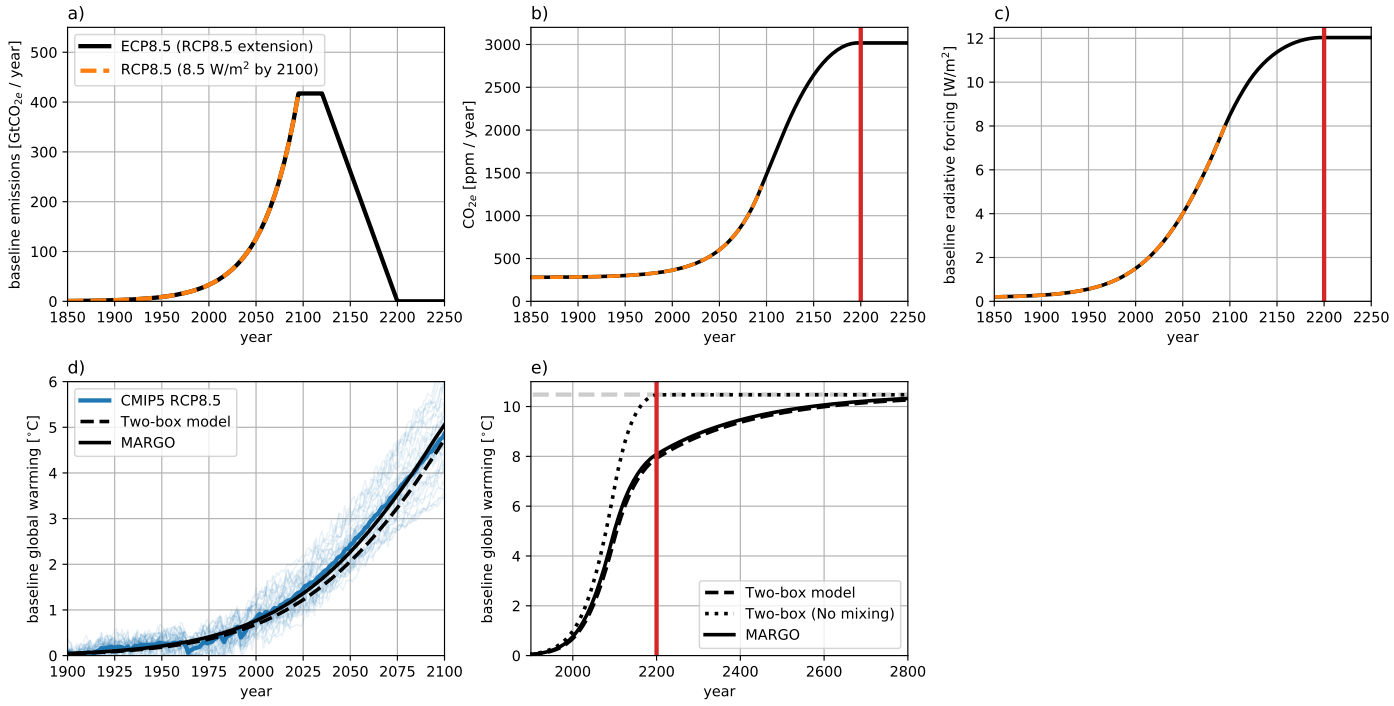

Figure 7: Validation of the 21st Century and equilibrium responses of the MARGO Energy Balance Model (EBM). a) Baseline CO_{2e} emissions, b) concentrations, and c) radiative forcing in an RCP8.5-like scenario (dashed orange line) and its extension beyond 2100 (ECP8.5; solid black line). d) The temperature response of CMIP5 models to the RCP8.5 forcing scenario (thin blue lines for individual models; thick blue line for multi-model mean) and of the MARGO-EBM to the RCP8.5-like scenario. The dashed black line shows the full solution to the two-layer equations 15 and 16 with the same parameter values (Geoffroy 2013; Geoffroy et al., 2012) as the approximate solution 20 used in the MARGO-EBM. e) The temperature response to the ECP8.5 scenario for: the MARGO-EBM (solid), the full two-box model (dashed black line) and the full two-box model with $\kappa = 0$ (dotted line). The vertical red lines delineate 2200, the year in which the ECP8.5 emissions reach net zero and concentrations are stabilized.

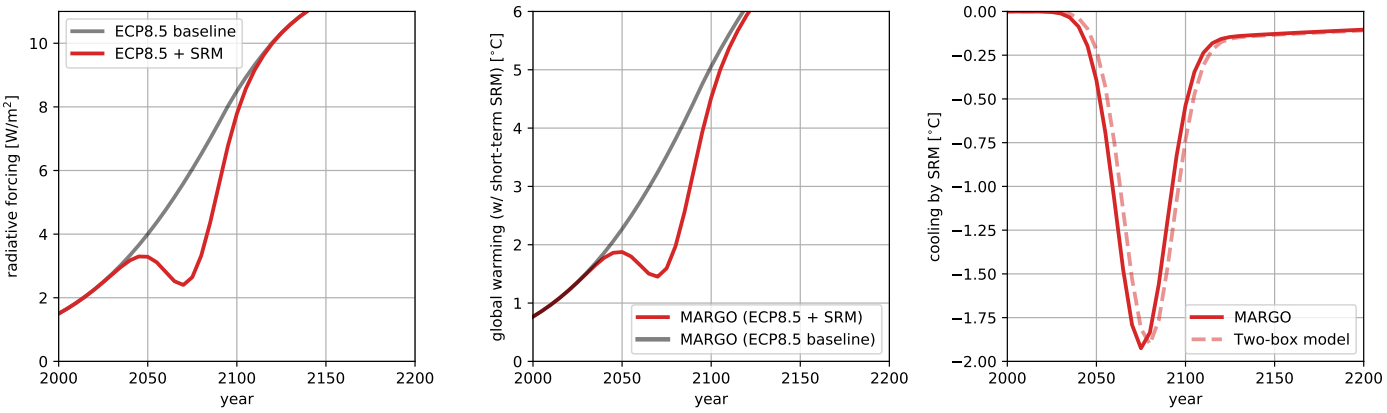

Figure 8: Response of the MARGO-EBM to the ECP8.5 scenario (grey) and to an additional short-term variation in forcing caused by a Gaussian deployment of SRM (red). a) Radiative forcing; b) Temperature response; c) Anomalous cooling in SRM scenario relative to the ECP8.5 baseline in MARGO (solid line) and the full solution to the two-box model (dashed line).

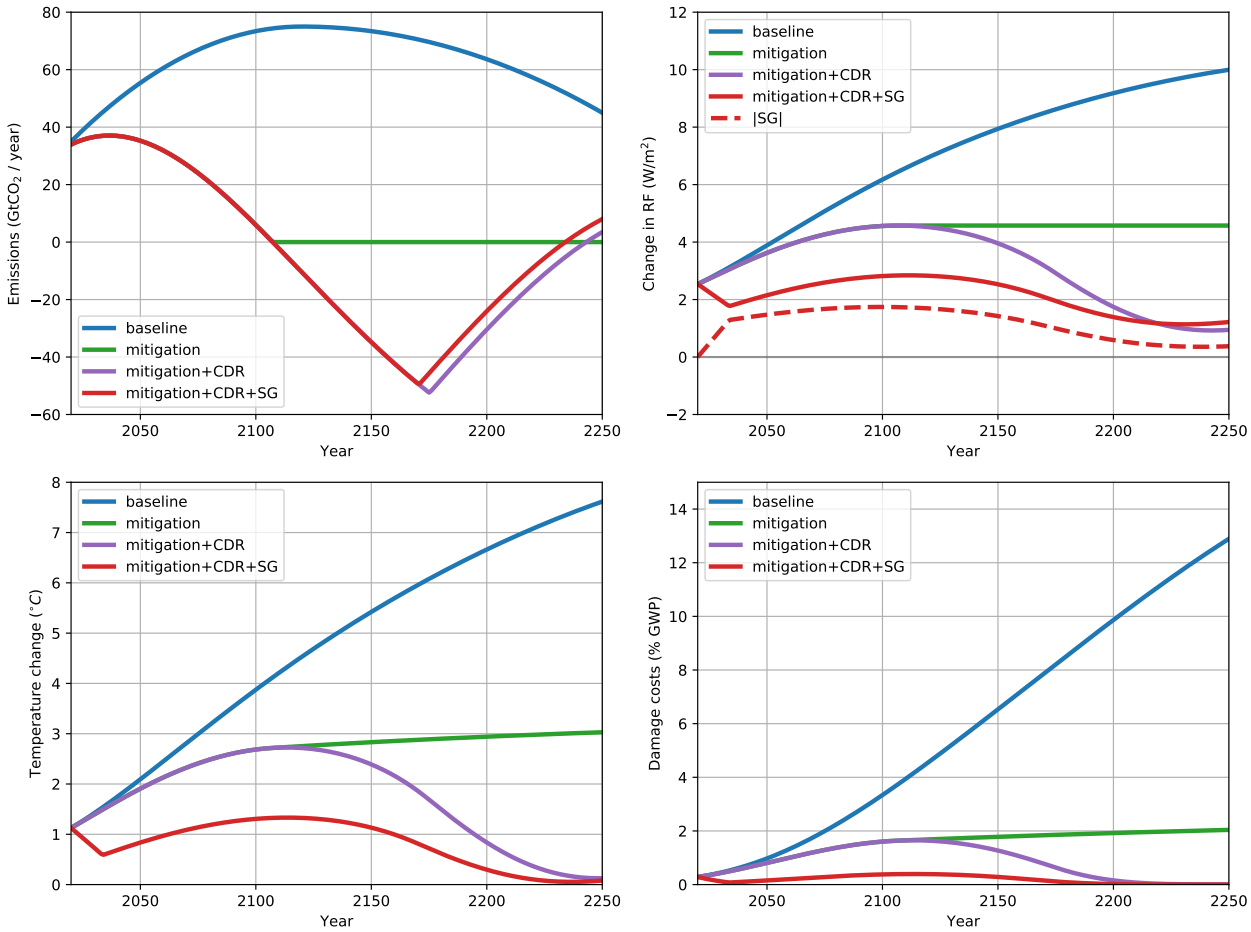

Figure 9: A qualitative replication of Figure 4 of Belaia (2019; Belaia, 2019); see their figure caption and accompanying discussion of the results.

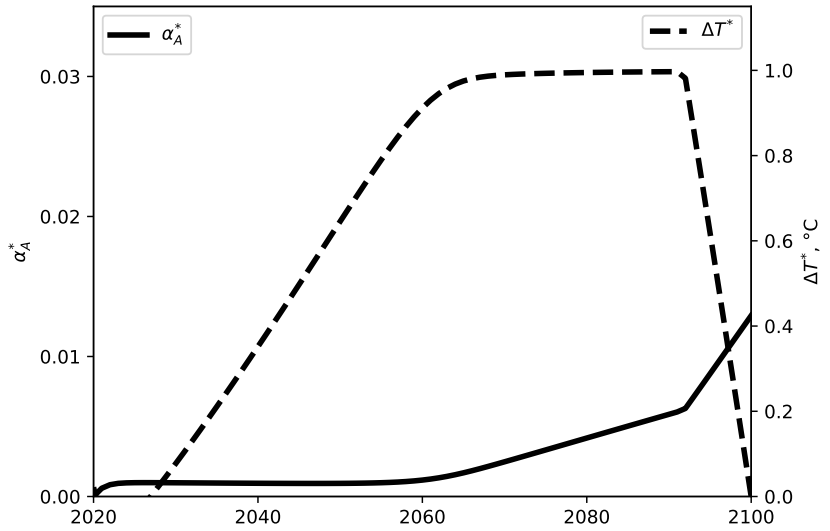

Figure 10: A qualitative replication of Figure 3 of Soldatenko and Yusupov (2018; Soldatenko and Yusupov, 2018), who consider the optimally cost-effective deployments of SRM which satisfy the following temperature constraints: $\Delta T^*(t) \leq 1\,°\text{C}$ and $\Delta T^*(t_f) \leq 0\,°\text{C}$, where $\Delta T^* \equiv T_{M,R,G} - T_0$ is the temperature anomaly relative to 2020 (ignoring mitigation and CDR, $M \equiv R \equiv 0$) and $t_f = 2100$ is the final date. The dashed curve shows the optimal SRM albedo $\alpha_A^* \equiv \frac{G(t)F_\infty}{S_0/4}$ and the solid black line shows the temperature anomaly ΔT^*.

Sterner, T., Tol, R. S. J., and Weitzman, M. (2013). Determining Benefits and Costs for Future Generations. *Science*, 341(6144):349–350. Publisher: American Association for the Advancement of Science Section: Policy Forum.

Azar, C. and Schneider, S. H. (2002). Are the economic costs of stabilising the atmosphere prohibitive? *Ecological Economics*, 42(1):73–80.

Belaia, M. (2019). Optimal Climate Strategy with Mitigation, Carbon Removal, and Solar Geoengineering. *arXiv:1903.02043 [econ, q-fin]*. arXiv: 1903.02043.

Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. (2017). Julia: A Fresh Approach to Numerical Computing. *SIAM Review*, 59(1):65–98.

Buck, H. J. (2010). What can geoengineering do for us? public participation and the new media landscape. In *Paper for workshop: The ethics of solar radiation management, 18 Oct 2010, University of Montana*.

Buck, H. J. (2012). Geoengineering: Re-making Climate for Profit or Humanitarian Intervention? *Development and Change*, 43(1):253–270. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-7660.2011.01744.x.

Buck, H. J. (2016). Rapid scale-up of negative emissions technologies: social barriers and social implications. *Climatic Change*, 139(2):155–167.

Burke, M., Hsiang, S. M., and Miguel, E. (2015). Global non-linear effect of temperature on economic production. *Nature*, 527(7577):235–239. Number: 7577 Publisher: Nature Publishing Group.

Caldeira, K. and Ricke, K. L. (2013). Prudence on solar climate engineering. *Nature Climate Change*, 3(11):941–941. Number: 11 Publisher: Nature Publishing Group.

Calel, R. and Stainforth, D. A. (2016). On the Physics of Three Integrated Assessment Models. *Bulletin of the American Meteorological Society*, 98(6):1199–1216. Publisher: American Meteorological Society.

Christensen, P., Gillingham, K., and Nordhaus, W. (2018). Uncertainty in forecasts of long-run economic growth. *Proceedings of the National Academy of Sciences*, 115(21):5409–5414.

Clarke, L. E., Jiang, K., Akimoto, K., Babiker, M., Blanford, G. J., Fisher-Vanden, K., Hourcade, J.-C., Krey, V., Kriegler, E., Loschel, A., et al. (2014). Assessing Transformation Pathways. In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Technical report.

Council, N. R. et al. (1991). Policy implications of greenhouse warming. In *Report of the Committee on Science, Engineering and Public Policy*, page 127. National Academy Press Washington, DC.

Crost, B. and Traeger, C. P. (2013). Optimal climate policy: Uncertainty versus Monte Carlo. *Economics Letters*, 120(3):552–558.

Crutzen, P. J. (2006). Albedo Enhancement by Stratospheric Sulfur Injections: A Contribution to Resolve a Policy Dilemma? *Climatic Change*, 77(3):211.

Deutch, J. M. (2019). Joint allocation of climate control mechanisms is the cheapest way to reduce global climate damage. *MIT Center for Energy and Environmental Policy Research Working Paper Series*.

Dow, K., Berkhout, F., Preston, B. L., Klein, R. J. T., Midgley, G., and Shaw, M. R. (2013). Limits to adaptation. *Nature Climate Change*, 3(4):305–307. Number: 4 Publisher: Nature Publishing Group.

Flegal, J. A. and Gupta, A. (2018). Evoking equity as a rationale for solar geoengineering research? Scrutinizing emerging expert visions of equity. *International Environmental Agreements: Politics, Law and Economics*, 18(1):45–61.

Flegal, J. A., Hubert, A.-M., Morrow, D. R., and Moreno-Cruz, J. B. (2019). Solar Geoengineering: Social Science, Legal, Ethical, and Economic Frameworks. *Annual Review of Environment and Resources*, 44(1):399–423. _eprint: https://doi.org/10.1146/annurev-environ-102017-030032.

Fuss, S., Canadell, J. G., Peters, G. P., Tavoni, M., Andrew, R. M., Ciais, P., Jackson, R. B., Jones, C. D., Kraxner, F., Nakicenovic, N., Quéré, C. L., Raupach, M. R., Sharifi, A., Smith, P., and Yamagata, Y. (2014). Betting on negative emissions. *Nature Climate Change*, 4(10):850–853. Number: 10 Publisher: Nature Publishing Group.

Fuss, S., Lamb, W. F., Callaghan, M. W., Hilaire, J., Creutzig, F., Amann, T., Beringer, T., Garcia, W. d. O., Hartmann, J., Khanna, T., Luderer, G., Nemet, G. F., Rogelj, J., Smith, P., Vicente, J. L. V., Wilcox, J., Dominguez, M. d. M. Z., and Minx, J. C. (2018). Negative emissions—Part 2: Costs, potentials and side effects. *Environmental Research Letters*, 13(6):063002. Publisher: IOP Publishing.

Geoffroy, O., Saint-Martin, D., Olivié, D. J. L., Voldoire, A., Bellon, G., and Tytéca, S. (2012). Transient Climate Response in a Two-Layer Energy-Balance Model. Part I: Analytical Solution and Parameter Calibration Using CMIP5 AOGCM Experiments. *Journal of Climate*, 26(6):1841–1857. Publisher: American Meteorological Society.

Glanemann, N., Willner, S. N., and Levermann, A. (2020). Paris Climate Agreement passes the cost-benefit test. *Nature Communications*, 11(1):1–11. Number: 1 Publisher: Nature Publishing Group.

Glotter, M. J., Pierrehumbert, R. T., Elliott, J. W., Matteson, N. J., and Moyer, E. J. (2014). A simple carbon cycle representation for economic and policy analyses. *Climatic Change*, 126(3):319–335.

Goes, M., Tuana, N., and Keller, K. (2011). The economics (or lack thereof) of aerosol geoengineering. *Climatic Change*, 109(3):719–744.

Gregory, J. M. and Forster, P. M. (2008). Transient climate response estimated from radiative forcing and observed temperature change. *Journal of Geophysical Research: Atmospheres*, 113(D23).

Ha-Duong, M., Grubb, M. J., and Hourcade, J.-C. (1997). Influence of socioeconomic inertia and uncertainty on optimal CO 2 -emission abatement. *Nature*, 390(6657):270–273. Number: 6657 Publisher: Nature Publishing Group.

Haerlin, B. and Parr, D. (1999). How to restore public trust in science. *Nature*, 400(6744):499–499. Number: 6744 Publisher: Nature Publishing Group.

Hammitt, J. K. (1999). Evaluation Endpoints and Climate Policy: Atmospheric Stabilization, Benefit-Cost Analysis, and Near-Term Greenhouse-Gas Emissions. *Climatic Change*, 41(3):447–468.

Held, I. M., Winton, M., Takahashi, K., Delworth, T., Zeng, F., and Vallis, G. K. (2010). Probing the Fast and Slow Components of Global Warming by Returning Abruptly to Preindustrial Forcing. *Journal of Climate*, 23(9):2418–2427.

Hope, C. (2006). The marginal impact of co2 from page2002: an integrated assessment model incorporating the ipcc's five reasons for concern. *Integrated assessment*, 6(1).

Intergovernmental Panel on Climate Change (2018). *Global warming of 1.5°C*. OCLC: 1056192590.

Irvine, P. J., Kravitz, B., Lawrence, M. G., Gerten, D., Caminade, C., Gosling, S. N., Hendy, E. J., Kassie, B. T., Kissling, W. D., Muri, H., Oschlies, A., and Smith, S. J. (2017). Towards a comprehensive climate impacts assessment of solar geoengineering. *Earth's Future*, 5(1):93–106. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016EF000389.

Kellogg, W. W. and Schneider, S. H. (1974). Climate Stabilization: For Better or for Worse? *Science*, 186(4170):1163–1172. Publisher: American Association for the Advancement of Science Section: Articles.

Koomey, J. (2013). Moving beyond benefit–cost analysis of climate change. *Environmental Research Letters*, 8(4):041005. Publisher: IOP Publishing.

Lacey, J., Howden, M., Cvitanovic, C., and Colvin, R. M. (2018). Understanding and managing trust at the climate science–policy interface. *Nature Climate Change*, 8(1):22–28. Number: 1 Publisher: Nature Publishing Group.

Lempert, R. J., Schlesinger, M. E., and Bankes, S. C. (1996). When we don't know the costs or the benefits: Adaptive strategies for abating climate change. *Climatic Change*, 33(2):235–274.

Lickley, M., Cael, B. B., and Solomon, S. (2019). Time of Steady Climate Change. *Geophysical Research Letters*, 46(10):5445–5451. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018GL081704.

Luderer, G., Pietzcker, R. C., Bertram, C., Kriegler, E., Meinshausen, M., and Edenhofer, O. (2013). Economic mitigation challenges: how further delay closes the door for achieving climate targets. *Environmental Research Letters*, 8(3):034033. Publisher: IOP Publishing.

Manabe, S. and Wetherald, R. T. (1967). Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *Journal of the Atmospheric Sciences*, 24(3):241–259.

Marshall, D. P. and Zanna, L. (2014). A Conceptual Model of Ocean Heat Uptake under Climate Change. *Journal of Climate*, 27(22):8444–8465. Publisher: American Meteorological Society.

Matthews, H. D. and Caldeira, K. (2008). Stabilizing climate requires near-zero emissions. *Geophysical Research Letters*, 35(4). _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2007GL032388.

McClellan, J., Keith, D. W., and Apt, J. (2012). Cost analysis of stratospheric albedo modification delivery systems. *Environmental Research Letters*, 7(3):034019. Publisher: IOP Publishing.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and van Vuuren, D. P. (2011). The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, 109(1):213.

Millar, R., Allen, M., Rogelj, J., and Friedlingstein, P. (2016). The cumulative carbon budget and its implications. *Oxford Review of Economic Policy*, 32(2):323–342. Publisher: Oxford Academic.

Minx, J. C., Lamb, W. F., Callaghan, M. W., Fuss, S., Hilaire, J., Creutzig, F., Amann, T., Beringer, T., Garcia, W. d. O., Hartmann, J., Khanna, T., Lenzi, D., Luderer, G., Nemet, G. F., Rogelj, J., Smith, P., Vicente, J. L. V., Wilcox, J., and Dominguez, M. d. M. Z. (2018). Negative emissions—Part 1: Research landscape and synthesis. *Environmental Research Letters*, 13(6):063001. Publisher: IOP Publishing.

Moreno-Cruz, J., Wagner, G., and Keith, D. (2018). An Economic Anatomy of Optimal Climate Policy. SSRN Scholarly Paper ID 3001221, Social Science Research Network, Rochester, NY.

Moreno-Cruz, J. B. (2015). Mitigation and the geoengineering threat. *Resource and Energy Economics*, 41:248–263.

Nordhaus, W. and Sztorc, P. (2013). Dice 2013r: Introduction and user's manual. *Yale University and the National Bureau of Economic Research, USA*.

Nordhaus, W. D. (1992). An Optimal Transition Path for Controlling Greenhouse Gases. *Science*, 258(5086):1315–1319. Publisher: American Association for the Advancement of Science Section: Articles.

Nordhaus, W. D. (2017). Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences*, 114(7):1518–1523. Publisher: National Academy of Sciences Section: Social Sciences.

Parson, E. A. (2017). Opinion: Climate policymakers and assessments must get serious about climate engineering. *Proceedings of the National Academy of Sciences*, 114(35):9227–9230. Publisher: National Academy of Sciences Section: Opinion.

Parson, E. A. and Keith, D. W. (2013). End the Deadlock on Governance of Geoengineering Research. *Science*, 339(6125):1278–1279. Publisher: American Association for the Advancement of Science Section: Policy Forum.

Peters, G. P., Andrew, R. M., Canadell, J. G., Friedlingstein, P., Jackson, R. B., Korsbakken, J. I., Quéré, C. L., and Peregon, A. (2020). Carbon dioxide emissions continue to grow amidst slowly emerging climate policies. *Nature Climate Change*, 10(1):3–6. Number: 1 Publisher: Nature Publishing Group.

Pindyck, R. S. (2017). The Use and Misuse of Models for Climate Policy. *Review of Environmental Economics and Policy*, 11(1):100–114. Publisher: Oxford Academic.

Ramsey, F. P. (1928). A Mathematical Theory of Saving. *The Economic Journal*, 38(152):543–559. Publisher: [Royal Economic Society, Wiley].

Revelle, R., Broecker, W., Craig, H., Kneeling, C., and Smagorinsky, J. (1965). Restoring the quality of our environment: report of the environmental pollution panel. Atmospheric carbon dioxide. *President's Science Advisory Committee, United States, US Government Printing Office: Washington, DC*.

Riahi, K., Grübler, A., and Nakicenovic, N. (2007). Scenarios of long-term socio-economic and environmental development under climate stabilization. *Technological Forecasting and Social Change*, 74(7):887–935.

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

Ricke, K. L., Moreno-Cruz, J. B., and Caldeira, K. (2013). Strategic incentives for climate geoengineering coalitions to exclude broad participation. *Environmental Research Letters*, 8(1):014021. Publisher: IOP Publishing.

Rogelj, J., Elzen, M. d., Höhne, N., Fransen, T., Fekete, H., Winkler, H., Schaeffer, R., Sha, F., Riahi, K., and Meinshausen, M. (2016). Paris Agreement climate proposals need a boost to keep warming well below 2 °C. *Nature*, 534(7609):631–639. Number: 7609 Publisher: Nature Publishing Group.

Schneider, S. H. (1989). The Greenhouse Effect: Science and Policy. *Science*, 243(4892):771–781. Publisher: American Association for the Advancement of Science Section: Articles.

Schneider, S. H. (1997). Integrated assessment modeling of global climate change: Transparent rational tool for policy making or opaque screen hiding value-laden assumptions? *Environmental Modeling & Assessment*, 2(4):229–249.

Schäfer, S., Irvine, P. J., Hubert, A.-M., Reichwein, D., Low, S., Stelzer, H., Maas, A., and Lawrence, M. G. (2013). Field tests of solar climate engineering. *Nature Climate Change*, 3(9):766–766. Number: 9 Publisher: Nature Publishing Group.

Shayegh, S. and Thomas, V. M. (2015). Adaptive stochastic integrated assessment modeling of optimal greenhouse gas emission reductions. *Climatic Change*, 128(1):1–15.

Sherwood, S. C. and Huber, M. (2010). An adaptability limit to climate change due to heat stress. *Proceedings of the National Academy of Sciences*, 107(21):9552–9555. Publisher: National Academy of Sciences _eprint: https://www.pnas.org/content/107/21/9552.full.pdf.

Siegel, L. S., Homer, J., Fiddaman, T., McCauley, S., Franck, T., Sawin, E., Jones, A. P., Sterman, J., and Interactive, C. (2018). En-roads simulator reference guide. Technical report, Technical Report.

Soldatenko, S. A. and Yusupov, R. M. (2018). Optimal Control of Aerosol Emissions into the Stratosphere to Stabilize the Earth's Climate. *Izvestiya, Atmospheric and Oceanic Physics*, 54(5):480–486.

Solomon, S., Plattner, G.-K., Knutti, R., and Friedlingstein, P. (2009). Irreversible climate change due to carbon dioxide emissions. *Proceedings of the National Academy of Sciences*, 106(6):1704–1709. Publisher: National Academy of Sciences Section: Physical Sciences.

Solow, R. M. (1974). The Economics of Resources or the Resources of Economics. *The American Economic Review*, 64(2):1–14. Publisher: American Economic Association.

Steffen, W., Rockström, J., Richardson, K., Lenton, T. M., Folke, C., Liverman, D., Summerhayes, C. P., Barnosky, A. D., Cornell, S. E., Crucifix, M., Donges, J. F., Fetzer, I., Lade, S. J., Scheffer, M., Winkelmann, R., and Schellnhuber, H. J. (2018). Trajectories of the Earth System in the Anthropocene. *Proceedings of the National Academy of Sciences of the United States of America*, 115(33):8252–8259.

Stern, N., Stern, N. H., and Treasury, G. B. (2007). *The Economics of Climate Change: The Stern Review*. Cambridge University Press.

Tol, R. S. (1997). On the optimal control of carbon dioxide emissions: an application of FUND. *Environmental Modeling & Assessment*, 2(3):151–163.

Tol, R. S. J. (2003). Is the Uncertainty about Climate Change too Large for Expected Cost-Benefit Analysis? *Climatic Change*, 56(3):265–289.

United Nations Framework Convention on Climate Change (2015). Paris agreement. Article 2(a). `https://unfccc.int/process-and-meetings/the-paris-agreement/theparis-agreement`.

Victor, D. G., Morgan, M. G., Apt, F., and Steinbruner, J. (2009). The Geoengineering Option - A Last Resort against Global Warming Essay. *Foreign Affairs*, 88(2):64–76.

Wagner, G. and Zeckhauser, R. J. (2016). Confronting Deep and Persistent Climate Uncertainty. SSRN Scholarly Paper ID 2818035, Social Science Research Network, Rochester, NY.

Webster, M., Santen, N., and Parpas, P. (2012). An approximate dynamic programming framework for modeling global climate policy under decision-dependent uncertainty. *Computational Management Science*, 9(3):339–362.

Weyant, J. (2017). Some Contributions of Integrated Assessment Models of Global Climate Change. *Review of Environmental Economics and Policy*, 11(1):115–137. Publisher: Oxford Academic.

Weyant, J. P. (2008). A Critique of the Stern Review's Mitigation Cost Analyses and Integrated Assessment. *Review of Environmental Economics and Policy*, 2(1):77–93. Publisher: Oxford Academic.

Wächter, A. and Biegler, L. T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1):25–57.

MIT CEEPR

MIT Center for Energy and Environmental Policy Research

Since 1977, the Center for Energy and Environmental Policy Research (CEEPR) has been a focal point for research on energy and environmental policy at MIT. CEEPR promotes rigorous, objective research for improved decision making in government and the private sector, and secures the relevance of its work through close cooperation with industry partners from around the globe. Drawing on the unparalleled resources available at MIT, affiliated faculty and research staff as well as international research associates contribute to the empirical study of a wide range of policy issues related to energy supply, energy demand, and the environment.

An important dissemination channel for these research efforts is the MIT CEEPR Working Paper series. CEEPR releases Working Papers written by researchers from MIT and other academic institutions in order to enable timely consideration and reaction to energy and environmental policy research, but does not conduct a selection process or peer review prior to posting. CEEPR's posting of a Working Paper, therefore, does not constitute an endorsement of the accuracy or merit of the Working Paper. If you have questions about a particular Working Paper, please contact the authors or their home institutions.

MIT Center for Energy and Environmental Policy Research
77 Massachusetts Avenue, E19-411
Cambridge, MA 02139
USA

Website: ceepr.mit.edu

MIT CEEPR Working Paper Series is published by the MIT Center for Energy and Environmental Policy Research from submissions by affiliated researchers.

For inquiries and/or for permission to reproduce material in this working paper, please contact:

Email ceepr@mit.edu
Phone (617) 253-3551
Fax (617) 253-9845