ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

# A simple model for assessing climate control trade-offs and responding to unanticipated climate outcomes

To cite this article: Henri F Drake *et al* 2021 *Environ. Res. Lett.* **16** 104012





ENVIRONMENTAL RESEARCH
LETTERS

CrossMark

OPEN ACCESS

RECEIVED
9 January 2021

REVISED
26 August 2021

ACCEPTED FOR PUBLICATION
7 September 2021

PUBLISHED
21 September 2021



LETTER

# A simple model for assessing climate control trade-offs and responding to unanticipated climate outcomes

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

[1] Massachusetts Institute of Technology, Cambridge, MA, United States of America
[2] MIT-WHOI Joint Program in Oceanography/Applied Ocean Science & Engineering, Cambridge and Woods Hole, MA, United States of America
[*] Author to whom any correspondence should be addressed.

**E-mail:** henrifdrake@gmail.com

**Keywords:** climate policy, geoengineering, carbon dioxide removal, mitigation, adaptation, climate change

Supplementary material for this article is available online

## Abstract

Persistent greenhouse gas (GHG) emissions threaten global climate goals and have prompted consideration of climate controls supplementary to emissions mitigation. We present MARGO, an idealized model of optimally-controlled climate change, which is complementary to both simpler conceptual models and more complicated Integrated Assessment Models. The four methods of controlling climate damage—mitigation, carbon dioxide removal (CDR), adaptation, and solar radiation modification (SRM)—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 necessary to stabilize GHG concentrations below a tolerable level. While the most cost-beneficial and cost-effective pathways to reducing climate suffering include deployments of all four controls, the quantitative trade-offs between the different controls are sensitive to value-driven parameters and poorly-known future costs and damages.

Static policy optimization assumes perfect foresight and obscures the active role decision-makers have in shaping a climate trajectory. We propose an explicit policy response process wherein climate control policies are re-adjusted over time in response to unanticipated outcomes. We illustrate this process in two 'storyline' scenarios: (a) near-term increases in mitigation and CDR are deficient, such that climate goals are expected to slip out of reach; (b) SRM is abruptly terminated after 40 years of successful deployment, causing an extremely rapid warming which is amplified by an excess of GHGs due to deterred mitigation. In both cases, an optimized policy response yields substantial benefits relative to continuing the original policy.

The MARGO model is intentionally designed to be as simple, transparent, customizable, and accessible as possible, addressing concerns about previous climate-economic modelling approaches and enabling a more diverse set of stakeholders to engage with these essential and timely topics.

## 1. 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 50 years ago (Manabe and Wetherald 1967), scientists have advocated 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 GHG emissions continuing to increase and climate goals slipping out of reach (Peters *et al* 2020), the calls for both rapid emissions reductions (Forster *et al* 2020, Prakash and Girgenti 2020) and serious consideration of supplementary climate controls have grown louder (Council *et al* 1991, Crutzen 2006, Victor *et al* 2009, Buck 2012, Parson 2017).

Four climate controls have emerged as candidates for use in the 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; Caldeira *et al* 2013, Moreno-Cruz *et al* 2018, Deutch 2019):

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

Emissions mitigation is the only preventative solution which cuts off $CO_{2e}$ emissions at their source, but will become increasingly difficult (expensive) at high levels of penetration (Edenhofer *et al* 2014). In combination with mitigation, CDR can in principle be used to drive net-negative emissions which decrease GHG concentrations to compensate for historical emissions, but is unproven at scale (Fuss *et al* 2014). SRM is quick to deploy, immediately results in significant global cooling, and has low direct costs (McClellan *et al* 2012); however, SRM does not perfectly offset the impacts of $CO_{2e}$-induced warming (Irvine *et al* 2017) and more fundamentally does not solve the underlying problem of long-term anthropogenic $CO_{2e}$ accumulation (Pierrehumbert 2019). Finally, adaptation allows for flexibility in the other controls as remaining climate damages can be somewhat reduced by adapting to the new climate, but is subject to physical (Sherwood and Huber 2010) and social (Dow *et al* 2013) limits to adaptability.

Numerous social or geopolitical factors may substantially limit or block deployments of certain controls, with SRM standing out as particularly contentious (Caldeira and Ricke 2013, Parson and Keith 2013, Schäfer *et al* 2013, Talati and Higgins 2019). Problems related to inequity (Flegal and Gupta 2018), distrust (Haerlin and Parr 1999, Lacey *et al* 2018), or lack of governance (Ricke *et al* 2013, Flegal *et al* 2019) are just a handful of examples. While we do not explicitly represent these all of these complexities in our modeling, we explore them implicitly within parameter sensitivity experiments and in two 'storyline' scenarios (section 4), as recommended by Shepherd *et al* (2018). In these scenarios, we explore both: (a) the best (or 'optimized') case in which a unitary decision-maker, as a surrogate for the more complicated realistic international policy-making process, prescribes control trajectories and their prescriptions are exactly realized, and (b) more realistic cases (hereafter referred to as 'suboptimal') in which control deployments fall short of the prescribed optimal trajectory and it is thus beneficial for the decision-maker to readjust their policies.

Our hypothetical decision-maker must follow some set of principles on which to base their control policies. We explore two commonly-studied approaches: (a) the cost-benefit approach, in which control costs are balanced against the benefits of avoided damages, and (b) the cost-effectiveness approach, in which control costs are minimized subject to a prescribed climate constraint. The cost-effectiveness approach implicitly underlies the Paris Climate Agreement (United Nations Framework Convention on Climate Change 2015), which currently organizes global climate policy and aims to keep global warming *well below* 2 °C relative to preindustrial levels.

The conventional tool for optimizing global climate control is the integrated assessment model (IAM), the result of coupling a climate system model to an economic model see Weyant (2017), for a general overview of IAMs and their utility to date. In this paper, we (a) present a novel idealized climate-economic model of optimally-controlled climate change and (b) propose a sequential policy process for periodic policy re-evaluation, which we illustrate by two 'storyline' scenarios: (A) near-term mitigation and CDR shortfalls (i.e. the present policy landscape Rogelj *et al* 2016, Peters *et al* 2017, Olhoff and Christensen 2020) and (B) abrupt termination of SRM (e.g. Matthews and Caldeira 2007, Goes *et al* 2011).

## 2. 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):

- **M**itigation of GHG emissions,
- **A**daptation to climate impacts,
- **R**emoval of carbon dioxide (CDR),
- **G**eoengineering by SRM, and
- **O**ptimal deployment of available controls.

Each of the climate controls acts, in its own distinct way, to reduce the damages caused by a changing climate, but also carry their own deployment costs. The model is designed to include key features of climate physics, economics, and policy as concisely as possible. The shortcoming of the model's simplicity is that its quantitative results can not be relied on to quantitatively inform policy decisions, but instead provide intuition about system dynamics and linkages between variables and parameter values.

The model is developed openly using the Julia programming language (Bezanson *et al* 2017) at github.com/ClimateMARGO/ClimateMARGO.jl and includes comprehensive documentation. The model originated as an extension of a previous analytical model (Deutch 2019) to include time-dependent control variables. Each model component is expressed

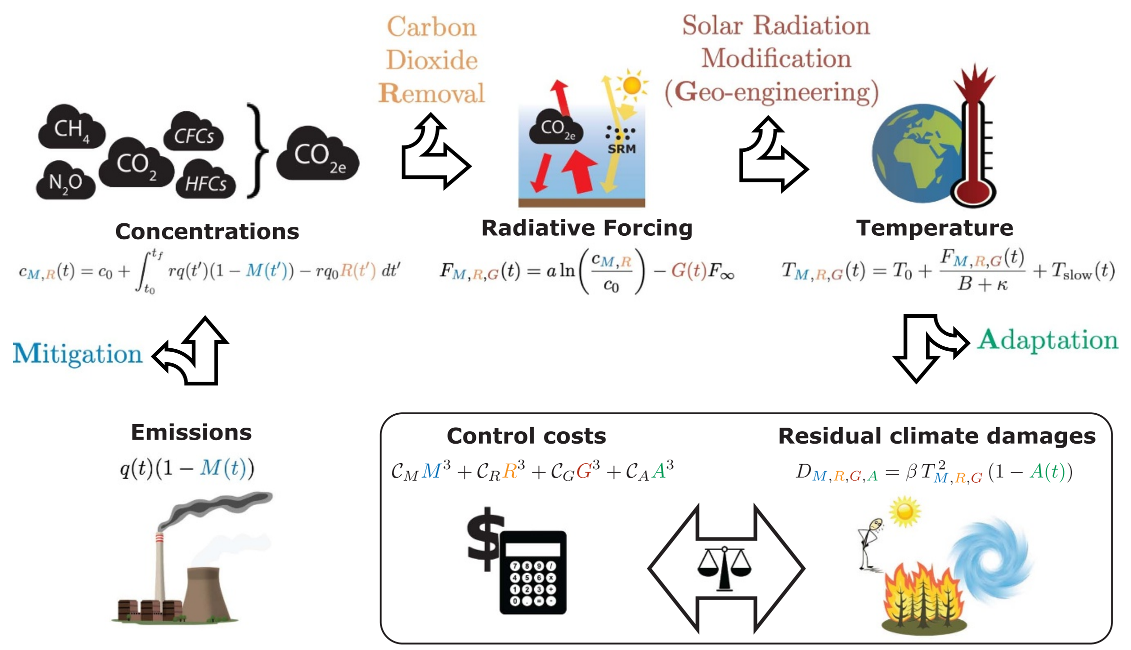


**Figure 1.** Schematic of the causal chain from GHG emissions to climate damages, including the unique effects of four climate controls: emissions **M**itigation, carbon dioxide **R**emoval (CDR), **G**eoengineering by solar radiation management (SRM), and **A**daptation. Climate controls yield benefits in terms of avoided climate damages, which are balanced against control deployment costs. Images (a)–(f) are all stated to be in the public domain. Reproduced from https://freesvg.org/. (g) Reproduced from Pixabay (2021).

in closed form to facilitate analytical analysis and computation, such that an entire MARGO optimization problem can be explicitly written down in one or two expressions, for cost-benefit and cost-effectiveness analyses, respectively (see section S3.1 (available online at stacks.iop.org/ERL/16/104012/mmedia)). The parameter values used throughout the paper are set to the defaults mentioned in this section (and comprehensively listed in table S1), except when exploring parameter sensitivities. Although we tune many of the model parameters based on existing literature and validated the model's behavior against another model (see sections S1 and S5), we caution that our 'optimized' results are not reliable as the basis for real-world policy recommendations because of the model's simplicity. In a few cases, where the choice of a parameter value is value-dependent or arbitrary, we perform sensitivity explorations to show how our results depend on these choices.

### 2.1. 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 and decrease linearly to zero by 2150 (for reasons independent of climate policy), resulting in 7.3 W m$^{-2}$ of radiative forcing by 2100 and 8.5 W m$^{-2}$ by 2150, relative to preindustrial levels. As a result of this forcing, global-mean warming reaches 2 °C by 2050 and soars to $T \approx 4.75$ °C by 2100, relative to preindustrial. We interpret this emission scenario as an idealized extension of the SSP3-7.0 baseline scenario, which is characterized by fossil-fueled growth (Riahi *et al* 2017).

There are five steps in the causal chain (equation (1) and figure 1) between $CO_{2e}$ emissions and climate damages:

(a) $CO_{2e}$ is emitted at a rate $q(t)$, with only a fraction $r \simeq 50\%$ (Solomon *et al* 2009, Joos *et al* 2013) remaining in the atmosphere after a few years, net of uptake by the ocean and terrestrial biosphere[3].
(b) Atmospheric $CO_{2e}$ increases as long as emissions $q(t)$ are greater than zero (Matthews and Caldeira 2008).
(c) Higher $CO_{2e}$ concentrations strengthen the greenhouse effect, reducing outgoing longwave radiation and causing an increased radiative forcing $F(t)$, which exerts a warming effect on the surface.
(d) Near-surface air temperatures increase by $\Delta T = T(t) - T(t_0)$ to balance the reduced cooling to space.
(e) Anthropogenic warming causes a myriad of climate impacts, resulting in gross economic damages $D(t) = \beta T^2$.

[3] The model's realism would be improved by replacing this crude carbon accounting model with multi-mode linear (Joos *et al* 2013) or non-linear (e.g. Glotter *et al* 2014) models of the ocean carbon cycle, but these would considerably complicate the exposition and interpretability of MARGO's equations.

### 2.2. Effects of climate controls

The four available climate controls enter as fractional controls at each link of the climate change causal chain (equation (1)). Each control is parameterized such that its full deployment would, in isolation, roughly remove or offset the climate damages due to the baseline emissions. This normalization allows a meaningful quantitative comparison of the control deployment variables.

**M**itigation reduces emissions by a factor $M(t) \in [0, 100\%]$ such that controlled emissions are given by $q(t)\,(1 - M(t))$, where $M = 100\%$ corresponds to complete decarbonization of the economy.

**R**emoval of $CO_{2e}$, $R(t) \in [0, 100\%]$, 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 $R = 100\%$ thus corresponds to removing 59 $GtCO_{2e}$/year, which is more than twice a recent upper-bound estimate of 24.5 $GtCO_{2e}$/year for 'feasible' potential deployments of negative emission technologies (Fuss *et al* 2018).

A useful diagnostic quantity is the effective emissions

$$r[q(t)(1 - M(t)) - q_0 R(t)], \tag{2}$$

which is the annual rate of $CO_{2e}$ accumulation in the atmosphere (figure 2(a)). The change in $CO_{2e}$ concentrations is the integral of the effective emissions over time (figure 2(b)),

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

**G**eoengineering by SRM, $G(t) \in [0, 100\%]$, acts to offset a fraction of the $CO_{2e}$ forcing (figure 2(c)),

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

where $F_{M,R} = a\ln(c_{M,R}(t)/c_0)$ is an empirically-determined $CO_{2e}$ forcing function (with $a = 5$ W m$^{-2}$) and $F_\infty = 8.5$ 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 $G = 100\%$ thus corresponds to a complete cancellation between the equilibrium warming from baseline $CO_{2e}$ increases and the cooling from a full deployment of SRM. Since the model timestep $\delta t = 5$ years is longer than the 1 year residence timescale of aerosols in the stratosphere (the most likely and persistent SRM technology candidate; Robock *et al* 2008), SRM forcing is treated as effectively instantaneous. Inefficiencies (Visioni *et al* 2017) or inefficacies (Modak *et al* 2016) in the SRM forcing mechanism do not appear explicitly and instead are factored into the SRMs deployment costs below.

The controlled warming (figure 2(d)), given by the deep-layer energy balance model solution

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

evolves in response to the total controlled forcing $F_{M,R,G}$, where $T_0 = 1.1$ °C is the present warming relative to preindustrial, $B = 1.13$ W m$^{-2}$ K$^{-1}$ is the climate feedback parameter, $\kappa = 0.73$ W m$^{-2}$ K$^{-1}$ is the ocean heat uptake rate, and $\tau_D = 240$ years is a slow deep ocean timescale (Geoffroy *et al* 2012). If forcing is stabilized for sufficiently long ($\Delta t \gg \tau_D$), warming asymptotes to an equilibrium response $T_{M,R,G} - T_0 = F_{M,R,G}/B$. Transiently, $B/(\kappa + B) = 60\%$ of the warming occurs effectively instantaneously (first term on right-hand side of equation (5)), while the remaining $\kappa/(B + \kappa) = 40\%$ is spread out over centuries due to the thermal inertia of the deep ocean (second term). This climate inertia decouples the temperature response from instantaneous forcing and implies that some 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[4]. We derive, interpret, and validate this energy balance model solution in greater detail in section S1.

**A**daptation to climate impacts acts to reduce gross controlled damages by a fraction $A(t) \in [0, 100\%]$ (following the AD-DICE model de Bruin *et al* 2009), resulting in the residual damages:

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

where we choose the damage parameter $\beta(t) = \tilde{\beta}E(t)$, where $\tilde{\beta} = 1\%/(°\mathrm{C})^2$ is a constant and $E(t)$ is the time varying gross world product (GWP), based on the preferences presented in a recent meta-analysis that accounts for various biases in previous estimates (Howard and Sterner 2017) and is, for example, 3–4 times larger than in the DICE model (de Bruin *et al* 2009). Although adaptation does not affect the planetary temperature directly, it is useful to consider an 'adaptive' pseudo-temperature $T_{M,R,G,A}$ (figure 2(d)) which yields controlled damages equivalent to the fully-controlled residual damages $\beta(T_{M,R,G,A})^2 = \beta T^2_{M,R,G}(1 - A(t))$ and is defined

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

[4] 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).

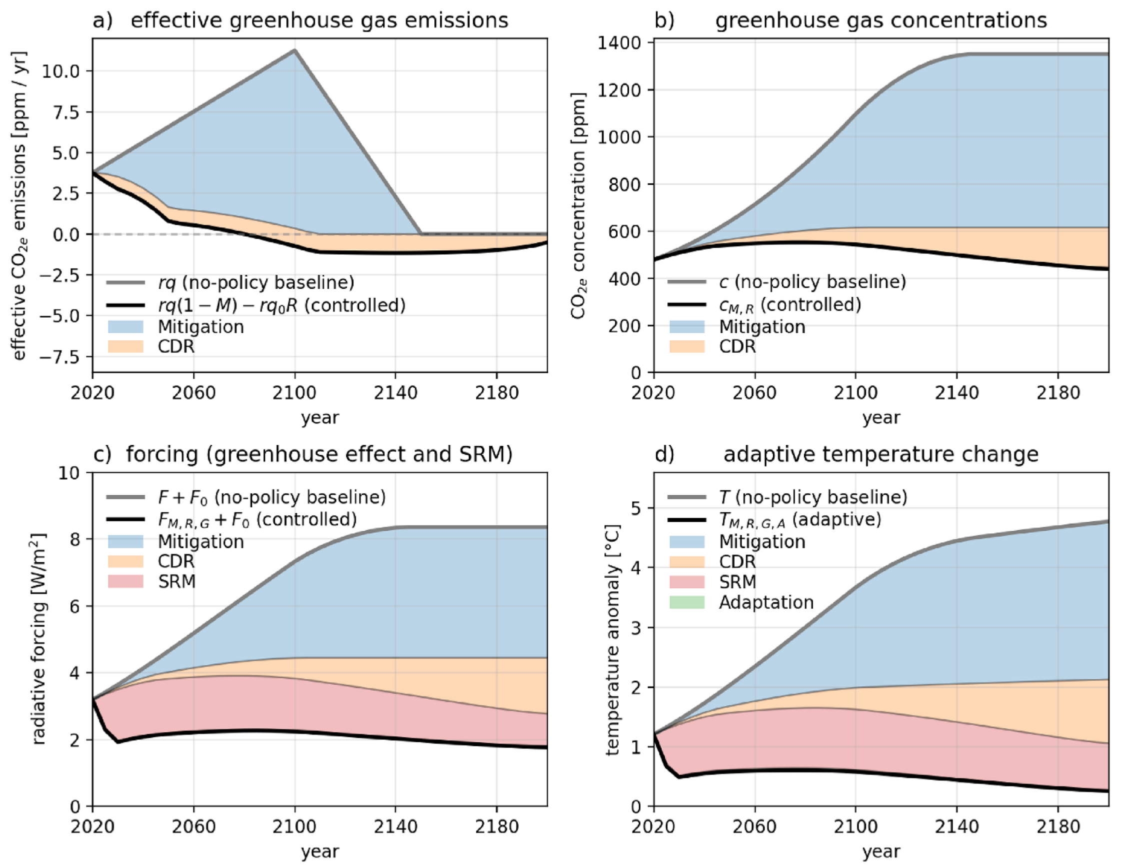


**Figure 2.** Baseline (thick grey line) and optimally cost-beneficial (thick black line) (a) effective $CO_{2e}$ emissions, (b) $CO_{2e}$ concentrations, (c) radiative forcing, and (d) temperature change (relative to preindustrial). Colored wedges show a natural decomposition of the effects of the four different climate controls, computed by setting downstream controls to zero in (5): **M**itigation (blue); carbon dioxide **R**emoval (CDR; orange); SRM or **G**eoengineering (red); and **A**daptation (green). Adaptation does not directly affect realized temperatures but is included in the plot using the adaptive pseudo-temperature construct (equation (7)); in this scenario, however, the optimal adaptation is so minor that its wedge is covered by the thick black line. Equivalent curves for the cost-effectiveness analysis are shown in figure S8.

### 2.3. Costs and benefits of controlling the climate

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

$$\mathcal{C} = \mathcal{C}_M M^3 + \mathcal{C}_R R^3 + \mathcal{C}_G G^3 + \mathcal{C}_A A^3. \tag{8}$$

The tuning of the control cost parameters, which represent the hypothetical annual costs of full deployment, is described in detail in section S2. The costs are summarized by: $\mathcal{C}_M = 55$ USD per $tCO_{2e}$ (marginal cost of 166 USD per $tCO_{2e}$ at $M = 100\%$) or 2% of the GWP in 2100 (based on Clarke *et al* 2014), $\mathcal{C}_R = 440$ USD per $tCO_{2e}$ for a sequestration rate of $q_0$ (based on Fuss *et al* 2018), $\mathcal{C}_G = 55\%$ of GWP per 8.5 W m$^{-2}$ of SRM (in the absence of better estimates, we conservatively assume the costs of side-effects are equal to the climate damages that would result from an equivalent magnitude of $CO_{2e}$ forcing), and $\mathcal{C}_A = 11.5\%$ GWP (de Bruin *et al* 2009). These cost functions are all convex functions of fractional deployment with zero initial marginal cost (as in Nordhaus 1992, Moreno-Cruz *et al* 2018) and are here all taken to be cubic for simplicity, such that marginal control costs increase quadratically with the deployment fraction. The benefits of deploying climate controls are the avoided residual climate damages relative to the no-policy baseline scenario,

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

### 2.4. Exogenous economic growth

We treat economic growth as exogenous: it is represented by the GWP, $E(t) = E_0(1+\gamma)^{(t-t_0)}$, which grows from its present value of $E_0 = 100$ trillion USD at a fixed growth rate $\gamma = 2\%$, consistent with expert opinion and an econometric forecast model (Christensen *et al* 2018). Several recent studies argue that the accumulated damages due to climate change slowing economic growth dwarf the direct damages on production, thus warranting more stringent mitigation (e.g. Moore and Diaz 2015, Glanemann *et al* 2020) As in DICE (Nordhaus and Sztorc 2013), we ignore such major feedbacks on economic growth; however, our substantially increased damage function (following Howard and Sterner 2017) results in similarly stringent climate control policies.

## 3. Optimizing a balanced climate policy portfolio

The surrogate climate policy decision-maker specifies an objective function to maximize, subject to additional policy constraints, and MARGO is readily optimized in terms of its time-dependent climate controls. The optimization is implemented in Julia using the Interior Point Optimizer (Wächter and Biegler 2006) and runs in a fraction of a second (figure S4); see details in section S4.1 (some additional policy constraints required for quasi-realism, such as maximum deployment rates, are described in section S3.2). Here, we describe the optimally-controlled results of two policy approaches, cost-benefit analysis and cost-effectiveness analysis, and explore their sensitivity to key value-driven or poorly-known parameters. See figures S6 and S7 for intuitive visualizations of the one- and two-dimensional versions of the optimization problem, respectively.

### 3.1. Cost-benefit analysis

A widely-used approach is cost-benefit analysis, in which costs (e.g. of deploying climate controls, $\mathcal{C}_{M,R,G,A}$) are balanced against benefits (e.g. of avoiding climate damage, $\mathcal{B}_{M,R,G,A}$). Formally, we aim to maximize the net present benefits[5]:

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

where $\rho$ 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 discount rate to apply to multi-generational social utility (Ramsey 1928, Solow 1974, Stern *et al* 2007, Arrow *et al* 2013). Here, we choose a default discount rate of $\rho=\gamma=2\%$ (corresponding to a pure time discount rate of zero), which is on the low end of values used in the literature due to our preference toward inter-generational equity (e.g. following Schneider 1989).

[5] This is equivalent to the conventional formulation of maximizing welfare changes $\Delta W=\int\lambda\Delta C\,\mathrm{d}t$, where $\Delta C=\mathcal{B}-\mathcal{C}$ is a change in consumption, $\lambda(t)$ is a discount factor. We assume a logarithmic utility function such that consumption changes are effectively discounted at a rate $-\dot{\lambda}/\lambda=\dot{C}/C+\delta\approx\gamma+\delta$ (with an implicit assumption of an iso-elasticity of marginal utility of consumption of unity, $\eta\simeq1$), where $\delta$ is the pure time discount rate and we assume similar growth rates for production and consumption $\dot{C}/C\approx\gamma$ (Stern *et al* 2007, Kelleher and Wagner 2019). We define $\rho\equiv\gamma+\delta$ to be the discount rate that describes exponential discounting $\lambda(t)\propto e^{-\rho t}$, which by expansion for $\rho\ll1$ gives us the more intuitive form $e^{-\rho t}\simeq(1+\rho)^{-t}$ used here.

The results of maximizing net present benefits are shown in figure 3. Early and aggressive emissions mitigation—and to a lesser extent CDR (figure 3(a))—drive discounted costs of up to 2 trillion USD/year (or 2% of GWP) relative to the no-policy baseline, but immediately deliver even larger benefits; after 2100, control costs begin decreasing toward zero while the benefits of avoided damages continue for as long as the time horizon allows (figure 3(b)). Effective $CO_{2e}$ emissions reach net-zero by 2080, such that concentrations stabilize below $c_{M,R}<550$ ppm and are brought back down below present-day values by 2200 (figures 2(a) and (b)) as CDR is ramped up to $R=30\%$ (or $q_0R\approx$ 9 $\mathrm{GtCO_{2e}}$/year). Significant SRM deployments of about $G=20\%$ (cooling of 1.8 W m$^{-2}$; figure 2(c)) results in a net decrease in radiative forcing, causing the maximum controlled temperature $T_{M,R}\approx1.6\,^\circ$C to plummet to just $T_{M,R,G}\approx0.5\,^\circ$C above preindustrial (figure 2(d)). Modest deployments of adaptation additionally offset about 12% of gross climate damages, but this effect is weak since the other controls have already reduced damages by orders of magnitude relative to the baseline.

The adaptive temperatures that result from cost-benefit analysis are only modestly sensitive to the choice of the discount rate $\rho$ (figure 3(c)): as the discount rate is increased well above the economic growth rate (Tol 2003), $\rho>\gamma=2\%$, for example, the optimized adaptive warming in 2100 only increases from $T_{M,R,G,A}\approx0.5\,^\circ$C to 1.3 °C. However, these relatively modest changes in residual climate damages obscure a fundamental shift in control preferences away from preventative mitigation toward unproven restorative (CDR) and reactive controls (SRM and adaptation).

### 3.2. 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), especially at high levels of forcing (Alley *et al* 2003, Burke *et al* 2015, Howard and Sterner 2017), and is sensitive to value judgements about appropriate discounting. An alternative approach, which presently guides global climate policy negotiations, is to prescribe a threshold of global warming (or related climate damages), which is not to be surpassed.

In MARGO's cost-effectiveness formulation, we aim to find the lowest net present costs of control deployments

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

which keep global warming below a chosen temperature (or climate damage) threshold,

$$T_{M,R,G}<T^{\star}=1.5\,^\circ\mathrm{C}. \tag{12}$$

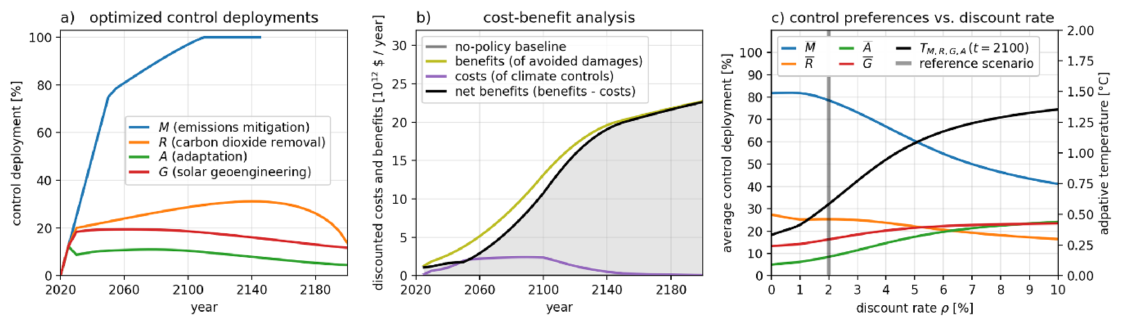


**Figure 3.** Results of cost-benefit analysis and their sensitivity to the discount rate $\rho$. (a) Optimized fractional control deployments and (b) corresponding discounted costs and benefits relative to the no-climate-policy baseline scenario. The mitigation curve in (a) ends in 2150 because baseline emissions go to zero and there is nothing to mitigate thereafter. The total positive area shaded in grey in (b) is the maximum net present benefit (equation (10)), defined as the benefits (green) minus the costs (purple). Since the discount rate is equal to the economic growth rate and $E_0 = 100 \times 10^{12}$ USD yr$^{-1}$, the $y$-axis in (b) can also be read as a percentage of GWP in a given year. (c) Time-mean control deployments (colors; left axis) and adaptive temperatures in 2100 (black; right axis), as a function of the discount rate.

As in the Paris Climate Agreement (United Nations Framework Convention on Climate Change 2015), we leave adaptation to be specified in a separate decision stage that is not as easily aggregated to the global level. While the Paris Agreement only explicitly considers mitigation and CDR as controls, we include SRM for illustrative purposes (see also MacMartin *et al* 2018).

The results of this cost-effectiveness optimization (SI figures 1 and 2) are qualitatively similar to the cost-benefit analysis above and are summarized as follows: substantial emissions mitigation keep $T_M < 3\,°\mathrm{C}$, sustained CDR brings warming down to $T_{M,R} = 2.2\,°\mathrm{C}$ by 2200, and SRM shaves off the remaining $\Delta T_G = T_{M,R} - T_{M,R,G} \approx 0.7\,°\mathrm{C}$ of warming to keep $T_{M,R,G} < T^\star = 1.5\,°\mathrm{C}$ at all times.

The relative importance of SRM in this cost-effectiveness optimization is sensitive to its poorly-known cost (assumed to be dominated by the damages due to side-effects of offsetting $CO_{2e}$ forcing with SRM) and the value-dependent discount rate (figure 4). For the conservative reference SRM cost used here, less than 20% of the optimized cooling is achieved with SRM at low discount rates ($\rho < 1\%$), while more than 40% is achieved with SRM at high discount rates ($\rho > 5\%$). Similarly, if SRM costs are reduced by an order of magnitude (e.g. if side-effects are found to be insignificant), more than 50% of the cooling (relative to the baseline warming) is achieved with SRM even under the low default discount rate $\rho = 2\%$.

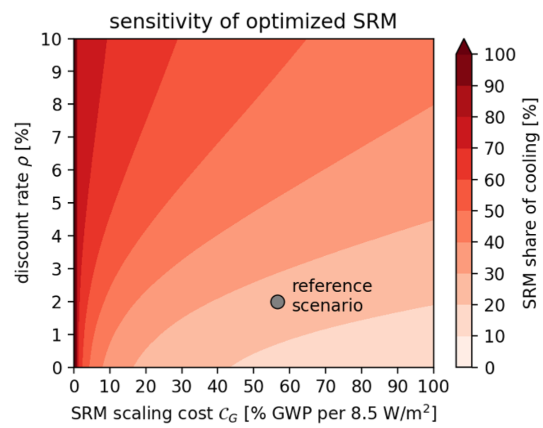


**Figure 4.** Sensitivity of optimally cost-effective SRM deployments to the discount rate $\rho$ and SRM scaling cost $C_G$. Colored contours show SRM's average share of the cooling required to keep global warming $T_{M,R,G}$ below the threshold $T^\star$, normalized by the baseline temperature deficit $T^\star - T$. A value of 100% corresponds to an SRM-only control portfolio, whereas a value of 0% means only other controls are used. The grey dot shows the location of the relatively conservative (risk-averse) default values used here, where SRM provides roughly 25% of the required cooling.

## 4. A policy process for responding to policy shortfalls and climate surprises

The cost-benefit and cost-effectiveness calculations presented above assume the surrogate policy decision-maker has perfect foresight and that their prescribed optimal control policies are perfectly implemented, such that the anticipated climate outcomes are exactly realized. Here, we present a policy process that allows the decision-maker to respond to suboptimal outcomes, such as policy shortfalls or climate surprises.

### 4.1. The policy response process

Step 1: The process begins with a single optimization, which produces optimized climate control trajectories and corresponding projections of climate outcomes, from an initial vantage point of $t_0$ (e.g. present-day).

Step 2: Time advances, $t \rightarrow t_1 = t_0 + \Delta t$, such that climate control deployments and outcomes take on their realized values over this period. From the vantage point of the decision-maker in $t_0$, these outcomes will be suboptimal by definition

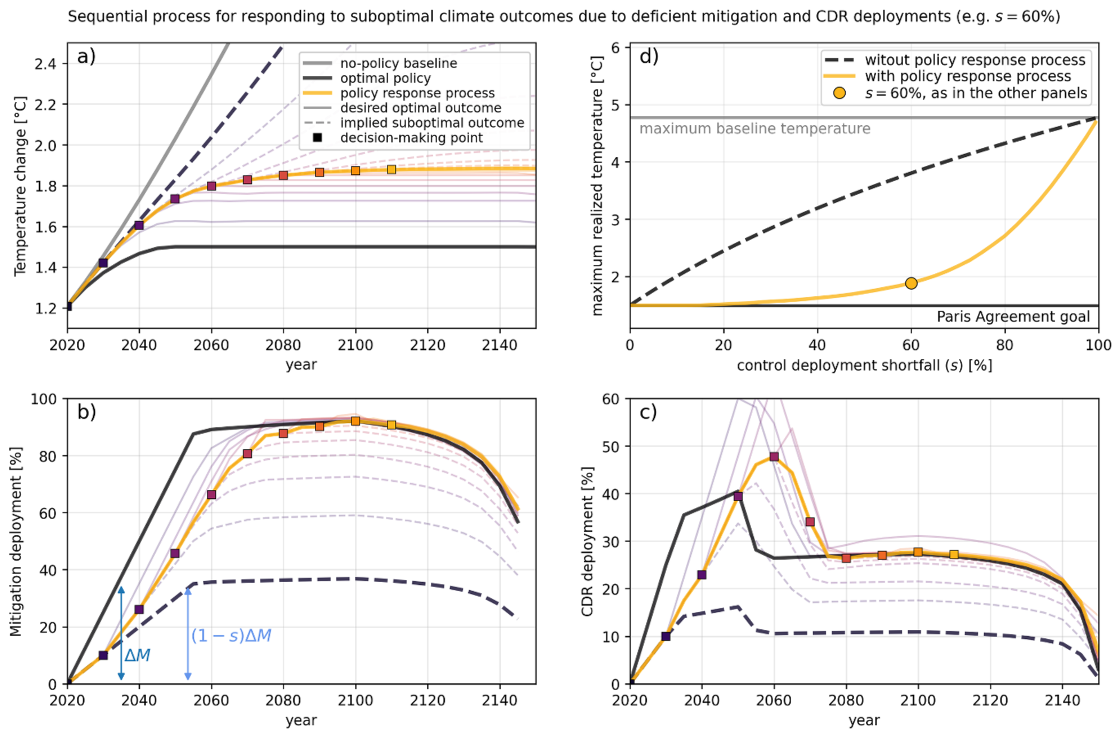


**Figure 5.** Storyline A: a surrogate policy decision-maker prescribes optimized mitigation and CDR trajectories which would limit warming below 1.5 °C (thick black lines), but realized deployments repeatedly fall short of these control targets by a fraction $s$, the *shortfall*. In the absence of a policy response, these control shortfalls lead to a substantial overshoot of the $T^{\star} = 1.5$ °C goal (thick dashed lines). Panels (a)–(c) illustrate a policy process for sequentially responding to these control shortfalls every $\Delta t = 10$ years, for an arbitrary example value $s = 60\%$. After computing an optimized future projection, $M \to M + \Delta M$ (thin solid lines), realized climate controls are incremented suboptimally, $M \to M + (1-s)\Delta M$ (thin dashed lines; see also annotated arrows). After $\Delta t = 10$ years of realized shortfalls (gold line from one square to the next), the decision-maker re-optimizes their prescription of future deployments, and the process repeats. For $s = 60\%$, a temporary shortfall in mitigation (b) and CDR (c) results in an overshoot of the temperature goal by 0.4 °C (a), although temperatures are eventually stabilized at this level by a delayed decarbonization of $M \approx 100\%$ (b) and intensified CDR deployments (c), which partially compensates for earlier control shortfalls. Panel (d) shows how the maximum realized temperature asymptotes to the temperature goal of 1.5 °C for $s < 30\%$ in the optimal limit and to the catastrophic baseline warming of 4.75 °C in the sub-optimal limit of no climate control ($s \to 100\%$).

if they differ at all from the original optimized projections.

Step 3: To account for policy changes in response to realized climate outcomes, climate control deployments are re-optimized, now from the vantage point of $t_1$ and with modified policy constraints or parameter values.

Step 4: Repeat, as desired, starting from Step 2, advancing in time to $t_{n+1} = t_n + \Delta t$.

This policy response process is illustrated below via two 'storyline' scenarios (Shepherd *et al* 2018).

### 4.2. Storyline A: mitigation and CDR shortfalls

A surrogate decision-maker prescribes the most cost-effective control trajectories for keeping $T_{M,R} < T^{*} = 1.5$ °C (figures 5(a)–(c), solid black lines), omitting SRM and adaptation, and thus calls for a rapid ramp-up of emissions mitigation ($M = 90\%$ by 2055) and CDR ($R = 40\%$ by 2050). Suppose that, by $t_1 = 2030$, $\Delta t = 10$ years later, it becomes apparent that these anticipated increases in mitigation and CDR have not been met, and realized control deployments instead fall short by $s = 60\%$ (the *shortfall* parameter) between $t_0 = 2020$ and $t_1 = 2030$. If this trend were to continue, and only $1 - s = 40\%$ of mitigation and CDR were to be deployed between 2020 and 2150, then temperatures would skyrocket, eventually reaching 3.8 °C by 2200 (black dashed lines).

Instead, the decision-maker responds by re-optimizing the model, now from the vantage point of $t_1 = 2030$, prescribing a larger ramp-up of CDR to compensate for the previous control shortfalls (figure 5(c)). However, short-term mitigation and CDR increases are constrained by the maximum deployment rate (100% over 40 years, see section S3.2), such that even the most ambitious deployments allowed by the model result in warming that overshoots the target $T^{\star} = 1.5$ °C by 0.1 °C (figure 5(a), lowest transparent line). The decision-maker is thus forced to relax their temperature goal, i.e. increase $T^{\star}$, if the optimal solution is to satisfy the temperature constraint (details in section S4.2). Suppose this sequential process repeats every $\Delta t = 10$ years, such that incremental deployments of mitigation and CDR always fall short of the decision-maker's control prescriptions by 60% (thin dashed lines) and the decision-maker responds by prescribing ever more

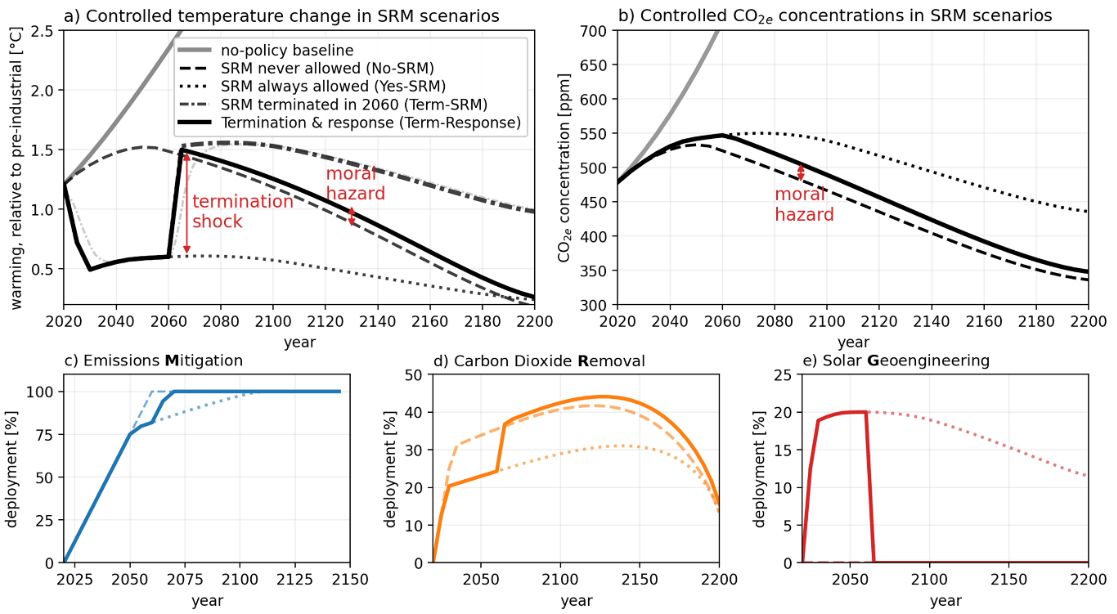


**Figure 6.** Storyline B: after $\Delta t = 40$ years of stabilizing global temperatures near present-day levels due in part to cooling by SRM (Yes-SRM; dotted lines), a surrogate policy decision-maker decides to abruptly terminate all SRM deployments (Term-SRM; dash-dotted lines). As SRM is terminated, temperatures rapidly rebound 1 °C (a 'termination shock'), which is 0.05 °C warmer than in the counterfactual optimal trajectory in which SRM was never allowed (No-SRM, dashed lines) due to a 25 ppm increase in $CO_{2e}$ attributed to deterred mitigation (a 'moral hazard'). If the decision-maker responds to the termination by re-optimizing future control deployments (Term-Response; solid lines), however, they are able to eventually reduce $CO_{2e}$ concentrations and temperatures back below pre-termination levels.

ambitious future control deployments in an attempt to compensate for their earlier shortfalls (thin solid lines). As time passes, realized control deployments converge toward their optimal response trajectories, and temperatures eventually stabilize, although they overshoot the original temperature goal $T^\star$ by 0.4 °C (gold lines).

A shortfall of $s = 60\%$ was chosen arbitrarily because it yields moderate but visually distinguishable results (figures 5(a)–(c)). Figure 5(b) shows the sensitivity of the maximum realized warming as a function of the control deployment shortfall $s$, which is varied from 0% (optimal) to 100% (zero controls). In the absence of a policy response, the realized warming increases roughly linearly with the shortfall, such that even a small shortfall of 10% results in a 0.5 °C overshoot of the $T^\star = 1.5$ °C goal (figure 5(d), dashed line). The policy response process (gold line), however, yields dramatically better outcomes and allows some room for error. Moderate shortfalls ($s \lesssim 30\%$) yield higher control costs (and are thus sub-optimal), but the decision-maker still has enough room to compensate for earlier shortfalls and keep warming below the goal of $T^\star = 1.5$ °C. It is only for large shortfalls ($s > 50\%$) that the maximum realized warming increases rapidly with $s$, reaching 2 °C at $s = 65\%$ and asymptopting to the baseline warming of 4.75 °C in the limit of $s \to 100\%$ (zero control deployments).

Alternatively, the decision-maker may first attempt to only relax the temperature constraint for the short term ($t < 2100$), allowing a temporary overshoot of the temperature goal to buy time for CDR to compensate for excess emissions, before resorting to relaxing the long-term ($t \geqslant 2100$) temperature constraint if required (figure S10). While allowing a temporary temperature overshoot allows temperature goals to be met for twice as large a shortfall as the more rigid process above, it requires betting on large CDR deployments ($R = 90\%$ by 2090) which push the limits of feasibility (Fuss *et al* 2014, 2018).

### 4.3. Storyline B: abrupt termination of SRM

In $t_0 = 2020$, a surrogate decision-maker prescribes the most cost-beneficial combination of emissions mitigations, CDR, and SRM (hereafter Yes-SRM; figure 6). Suppose that after a perfect deployment of these control trajectories for $\Delta t = 40$ years, yielding optimal climate outcomes, the policy decision-maker decides to abruptly cancel SRM deployments (within one model timestep $\delta t = 5$ years) and forbid their future use (figure 6(e); see Parker and Irvine 2018 for a discussion of this storyline's plausibility). The abrupt termination of SRM forcing results in an abrupt warming of 1 °C over a decade (Term-SRM; figure 6), known as a 'termination shock'. To counteract additional future damages due to this unanticipated warming, the policy decision-maker responds by re-optimizing their portfolio of future climate deployments from the vantage point of $t_1 = 2060$ (Term-Response), prescribing enhancements to mitigation and CDR which accelerate the approach to net-zero emissions by 10 years and result in 0.75 °C less warming by 2200 than in Term-SRM (figure 6). Despite the increases in controls after $t_1 = 2060$ due to this policy response, $CO_{2e}$ concentrations (figure 6(b)) and warming (figure 6(a)) remain higher than in a counterfactual world in which SRM was never allowed (No-SRM; figure 6). This 'moral hazard' (Lin 2013, McLaren 2016), whereby investment in SRM deters emissions mitigation and CDR, amplifies the termination shock by an additional 0.05 °C and results in a persistent increase in $CO_{2e}$ levels of 25 ppm relative to No-SRM. A rapid warming of 1 °C would likely result in substantial damages and pose challenges to adaptation. However, the termination shock considered here, which occurs within a balanced control portfolio including substantial mitigation and CDR, is small compared to the >1.5 °C/decade termination shock that would arise in a world controlled by SRM alone (figure S11), which is the worst-case scenario most often discussed in the literature (Parker and Irvine 2018).

From the vantage point of $t_0 = 2020$, we can order the four scenarios described above based on their net present benefits (from most to least): Yes-SRM > Term-Response > Term-SRM > No-SRM. Since SRM decouples temperatures from $CO_{2e}$ and can result in rapid temperature changes, our damage function based on temperature alone may underestimate damages. For example, adding damages due to $CO_{2e}$ concentrations (e.g. due to ocean acidification) disadvantages Yes-SRM and Term-SRM relative to No-SRM and Term-Response (figure S12(b)). On the other hand, adding damages due to the *rate* of warming disadvantages the termination scenarios Term-SRM and Term-Response (figure S12(a)). If these costs are both sufficiently high, the ordering instead becomes: No-SRM > Yes-SRM > Term-Response > Term-SRM (see figure S12 for the full sensitivity curves).

## 5. Discussion

Optimization of climate control in cost-benefit IAMs typically focuses on trade-offs between emissions mitigation and climate suffering (e.g. Nordhaus 1992, Tol 1997), although numerous studies have also considered trade-offs between mitigation and alternative climate control strategies: adaptation (de Bruin *et al* 2009), CDR (CDR; Kriegler *et al* 2013), and SRM (SRM; Goes *et al* 2011). Here, we explore trade-offs between all four of these approaches to climate control simultaneously. The optimized deployment levels of the climate controls depend upon their respective marginal costs per marginal benefit, which themselves are a complicated function of: their deployment cost curves, the causal chain of processes by which they affect downstream climate damages, and the choice of value-dependent parameters. We developed MARGO, a multi-control and time-dependent numerical model of optimized climate policies, to quantitatively explore these trade-offs. In our optimized simulations, emissions mitigation emerges as the preferred climate control, although a time-dependent combination of non-negligible deployments of all four controls yields the most cost-beneficial and cost-effective climate outcomes.

For clarity of exposition, we present a fully deterministic version of the MARGO model. In actuality, key inputs such as the climate feedback parameter $B$ 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). Similarly, economic models with formulations of preferences that naturally incorporate uncertainty and risk yield more stringent optimal controls (Cai and Lontzek 2018, Daniel *et al* 2019). Future work includes (a) extending MARGO to a stochastic programming approach that accounts for uncertainty in input parameters or stochastic climate dynamics (see section S4.3) and (b) implementing a Bayesian policy response process where prior parameter distributions can be updated based on observed stochastic outcomes (e.g. Shayegh and Thomas 2015) or improved parameter estimates from research developments (e.g. Hope 2015).

Climate outcomes will inevitably differ from the anticipated outcomes of prescribed control policies, whether because of imperfect control deployments, inherent variability, or structural uncertainty and bias in projected climate outcomes. We propose a policy process by which a surrogate decision-maker responds to undesirable real world outcomes by sequentially re-optimizing prescriptions of future climate control deployments, as an improvement over previously-proposed strategies based on arbitrary decision trees (e.g. Hammitt *et al* 1992, Lempert *et al* 1996, Goes *et al* 2011). We demonstrate the utility of this policy response process by quantifying its beneficial outcomes compared to alternative 'static' policies in which the decision-maker adheres to their original strategy despite control shortfalls or changing policy constraints.

Presently, intended nationally determined contributions (INDCs) imply warming of 2.6 °C–3.1 °C and thus will need to be strengthened at upcoming re-negotiations—and then actualized—to have a reasonable chance of meeting the Paris Agreement's goal of well below 2 °C of warming (United Nations Framework Convention on Climate Change 2015, Rogelj *et al* 2016, Olhoff and Christensen 2020). Holz *et al* (2018) explore plausible 'ratcheting' scenarios in which mitigation efforts are iteratively increased relative to the INDCs, and optionally supplemented by varying levels of CDR, until expected warming remains below 1.5 °C in 2100. In Storyline A (section 4.2), we introduce an alternative ratcheting process through which a policy decision-maker sequentially updates their climate control prescriptions to compensate for realized climate control shortfalls, in order to salvage climate targets in an optimally cost-effective way.

Rapid warming due to an abrupt termination of SRM deployments is commonly considered one of the greatest risks of SRM (Parker and Irvine 2018). Goes *et al* (2011) show that substituting SRM for mitigation fails a cost-benefit test, especially when accounting for the risk of a termination shock. However, Bickel and Agrawal (2013) extend their analysis and argue that if decision-makers respond to termination by mitigating emissions (using a fixed decision-tree response), then SRM passes the cost-benefit over a much larger range of termination probabilities than in Goes *et al* (2011). Helwegen *et al* (2019) perform a similar analysis, but allow the decision-maker to respond with optimal changes in mitigation, and show that SRM deployments robustly enhance welfare, even when taking into account the risk of SRM termination triggering climate 'tipping points'. In Storyline B (section 4.3), we extend this analysis by allowing decision-makers to optimally respond to SRM termination by also increasing CDR deployments, which yields substantial benefits relative to the scenario without a policy response. This storyline also provides quantitative evidence for a novel interaction between two processes that are independently considered major risks of SRM: the 'termination shock' due to abrupt SRM termination is amplified by about 5% due to the 'moral hazard' of SRM costs deterring mitigation deployments.

Our assumption of a unitary surrogate decision-maker evidently avoids the complexities of a realistic decision-making process that involves multiple stake holders with conflicting interests. While some of these interactions are implicitly embedded in the two storylines described above, they could instead be explicitly included in a multi-agent extension of the MARGO model, in which the global climate response is the aggregated result of multiple agents exerting controls on the climate, according to their own diverging incentives (Ricke *et al* 2013, Heyen *et al* 2019, Emmerling *et al* 2020).

The MARGO model fills the complexity gap between semi-analytic theoretical models (e.g. Moreno-Cruz *et al* 2018, Deutch 2019) and simple—but relatively opaque—IAMs (e.g. Nordhaus 1992, Tol 1997, Hope 2006): its dynamics are governed by only $N = 12$ intuitive free parameters but it still produces quasi-realistic climate trajectories (see section S3.2 and table S1). We show how MARGO can be used to investigate the sensitivity of optimized climate control policies to poorly-known parameters, such as future control costs, and value-dependent parameters, such as the discount rate. We also demonstrate that MARGO can be modified to reproduce the qualitative results of other multi-control studies (e.g. Belaia *et al* 2020; see section S5) and hope that it will be a useful community tool for climate policy research, interactive teaching, and public outreach, and will help bridge the gaps between climate economists, scientists, policy decision-makers, and the public (Schneider 1997, Buck 2010, Pindyck 2017, Stainforth and Calel 2020). We encourage readers to interactively run the MARGO model themselves by visiting any of our web-apps at https://github.com/ClimateMARGO/ClimateMARGO.jl/blob/master/Gallery.md.

## Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/ClimateMARGO/MARGO-paper; https://doi.org/10.5281/zenodo.5503978.

## Acknowledgment

We thank Lyssa Freese and four anonymous reviewers for comments on earlier versions of the manuscript. We thank Fons van der Plas for leading the development of our online web-apps. 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.

## ORCID iD

Henri F Drake https://orcid.org/0000-0003-0135-0814

## References

Alley R B *et al* 2003 Abrupt climate change *Science* **299** 2005–10
Arrow K *et al* 2013 Determining benefits and costs for future generations *Science* **341** 349–50
Belaia M, Moreno-Cruz J B and Keith D W 2020 Optimal climate policy in three dimensions *Preprint*
Bezanson J, Edelman A, Karpinski S and Shah V 2017 Julia: a fresh approach to numerical computing *SIAM Rev.* **59** 65–98
Bickel J E and Agrawal S 2013 Reexamining the economics of aerosol geoengineering *Clim. Change* **119** 993–1006
Buck H J 2010 What can geoengineering do for us? Public participation and the new media landscape *Paper for Workshop: The Ethics of Solar Radiation Management (18 October 2010)* (University of Montana)
Buck H J 2012 Geoengineering: re-making climate for profit or humanitarian intervention? *Dev. Change* **43** 253–70
Burke M, Hsiang S M and Miguel E 2015 Global non-linear effect of temperature on economic production *Nature* **527** 235–9
Cai Y and Lontzek T S 2018 The social cost of carbon with economic and climate risks *J. Polit. Econ.* **127** 2684–734
Caldeira K, Bala G and Cao L 2013 The science of geoengineering *Annu. Rev. Earth Planet. Sci.* **41** 231–56
Caldeira K and Ricke K L 2013 Prudence on solar climate engineering *Nat. Clim. Change* **3** 941

Christensen P, Gillingham K and Nordhaus W 2018 Uncertainty in forecasts of long-run economic growth *Proc. Natl Acad. Sci.* **115** 5409–14
Clarke L E *et al* 2014 Assessing transformation pathways *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge: Cambridge University Press)
Council N R *et al* 1991 Policy implications of greenhouse warming *Report of the Committee on Science, Engineering and Public Policy* (Washington, DC: National Academy Press) p 127
Crutzen P J 2006 Albedo enhancement by stratospheric sulfur injections: a contribution to resolve a policy dilemma? *Clim. Change* **77** 211
Daniel K D, Litterman R B and Wagner G 2019 Declining $CO_2$ price paths *Proc. Natl Acad. Sci.* **116** 20886–91
de Bruin K C, Dellink R B and Tol R S J 2009 AD-DICE: an implementation of adaptation in the DICE model *Clim. Change* **95** 63–81
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 *Nat. Clim. Change* **3** 305–7
Drake H F 2021 ClimateMARGO/MARGO-paper: Publication of Drake et al. (2021) ClimateMARGO.jl paper (https://doi.org/10.5281/zenodo.5503978)
Edenhofer O *et al* 2014 Technical summary *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge: Cambridge University Press) pp 33–107
Emmerling J, Kornek U, Bosetti V and Lessmann K 2020 Climate thresholds and heterogeneous regions: implications for coalition formation *Rev. Int. Organ.* **16** 293–316
Flegal J A and Gupta A 2018 Evoking equity as a rationale for solar geoengineering research? Scrutinizing emerging expert visions of equity *Int. Environ. Agreem.: Polit. Law Econ.* **18** 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 *Annu. Rev. Environ. Resour.* **44** 399–423
Forster P M, Maycock A C, McKenna C M and Smith C J 2020 Latest climate models confirm need for urgent mitigation *Nat. Clim. Change* **10** 7–10
Fuss S *et al* 2014 Betting on negative emissions *Nat. Clim. Change* **4** 850–3
Fuss S *et al* 2018 Negative emissions—part 2: costs, potentials and side effects *Environ. Res. Lett.* **13** 063002
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 *J. Clim.* **26** 1841–57
Glanemann N, Willner S N and Levermann A 2020 Paris climate agreement passes the cost-benefit test *Nat. Commun.* **11** 1–11
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 *Clim. Change* **126** 319–35
Goes M, Tuana N and Keller K 2011 The economics (or lack thereof) of aerosol geoengineering *Clim. Change* **109** 719–44
Haerlin B and Parr D 1999 How to restore public trust in science *Nature* **400** 499
Hammitt J K, Lempert R J and Schlesinger M E 1992 A sequential-decision strategy for abating climate change *Nature* **357** 315–18
Helwegen K G, Wieners C E, Frank J E and Dijkstra H A 2019 Complementing $CO_2$ emission reduction by solar radiation management might strongly enhance future welfare *Earth Syst. Dyn.* **10** 453–72
Heyen D, Horton J and Moreno-Cruz J 2019 Strategic implications of counter-geoengineering: clash or cooperation? *J. Environ. Econ. Manage.* **95** 153–77
Holz C, Siegel L S, Johnston E, Jones A P and Sterman J 2018 Ratcheting ambition to limit warming to 1.5 °C—trade-offs between emission reductions and carbon dioxide removal *Environ. Res. Lett.* **13** 064028
Hope C 2006 The marginal impact of $CO_2$ from PAGE2002: an integrated assessment model incorporating the IPCC's five reasons for concern *Integr. Assess.* **6** 19–56 (available at: http://116.203.146.222:8080/index.php/iaj/article/viewArticle/227)
Hope C 2015 The $10 trillion value of better information about the transient climate response *Phil. Trans. R. Soc.* A **373** 1–21
Howard P H and Sterner T 2017 Few and not so far between: a meta-analysis of climate damage estimates *Environ. Resour. Econ.* **68** 197–225
Irvine P J *et al* 2017 Towards a comprehensive climate impacts assessment of solar geoengineering *Earth's Future* **5** 93–106
Joos F *et al* 2013 Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis *Atmos. Chem. Phys.* **13** 2793–825
Kelleher J P and Wagner G 2019 Ramsey discounting calls for subtracting climate damages from economic growth rates *Appl. Econ. Lett.* **26** 79–82
Kellogg W W and Schneider S H 1974 Climate stabilization: for better or for worse? *Science* **186** 1163–72
Koomey J 2013 Moving beyond benefit–cost analysis of climate change *Environ. Res. Lett.* **8** 041005
Kriegler E, Edenhofer O, Reuster L, Luderer G and Klein D 2013 Is atmospheric carbon dioxide removal a game changer for climate change mitigation? *Clim. Change* **118** 45–57
Lacey J, Howden M, Cvitanovic C and Colvin R M 2018 Understanding and managing trust at the climate science–policy interface *Nat. Clim. Change* **8** 22–28
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 *Clim. Change* **33** 235–74
Lickley M, Cael B B and Solomon S 2019 Time of steady climate change *Geophys. Res. Lett.* **46** 5445–51
Lin A C 2013 Does geoengineering present a moral hazard *Ecol. Law Q.* **40** 673–712
MacMartin D G, Ricke K L and Keith D W 2018 Solar geoengineering as part of an overall strategy for meeting the 1.5 °C Paris target *Phil. Trans. R. Soc.* A **376** 20160454
Manabe S and Wetherald R T 1967 Thermal equilibrium of the atmosphere with a given distribution of relative humidity *J. Atmos. Sci.* **24** 241–59
Matthews H D and Caldeira K 2007 Transient climate–carbon simulations of planetary geoengineering *Proc. Natl Acad. Sci.* **104** 9949–54
Matthews H D and Caldeira K 2008 Stabilizing climate requires near-zero emissions *Geophys. Res. Lett.* **35** L04705
McClellan J, Keith D W and Apt J 2012 Cost analysis of stratospheric albedo modification delivery systems *Environ. Res. Lett.* **7** 034019
McLaren D 2016 Mitigation deterrence and the "moral hazard" of solar radiation management *Earth's Future* **4** 596–602
Modak A, Bala G, Cao L and Caldeira K 2016 Why must a solar forcing be larger than a $CO_2$ forcing to cause the same global mean surface temperature change? *Environ. Res. Lett.* **11** 044013
Moore F C and Diaz D B 2015 Temperature impacts on economic growth warrant stringent mitigation policy *Nat. Clim. Change* **5** 127–31
Moreno-Cruz J, Wagner G and Keith D 2018 An economic anatomy of optimal climate policy *SSRN Scholarly Paper ID* 3001221 (Rochester, NY: Social Science Research Network)
Nordhaus W D 1992 An optimal transition path for controlling greenhouse gases *Science* **258** 1315–19

Nordhaus W and Sztorc P 2013 DICE 2013R: introduction and user's manual (Yale University and the National Bureau of Economic Research, USA)
Olhoff A and Christensen J M 2020 *Emissions Gap Report 2020* (UNEP DTU Partnership)
Parker A and Irvine P J 2018 The risk of termination shock from solar geoengineering *Earth's Future* **6** 456–67
Parson E A 2017 Opinion: climate policymakers and assessments must get serious about climate engineering *Proc. Natl Acad. Sci.* **114** 9227–30
Parson E A and Keith D W 2013 End the deadlock on governance of geoengineering research *Science* **339** 1278–9
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 *Nat. Clim. Change* **10** 3–6
Peters G P, Andrew R M, Canadell J G, Fuss S, Jackson R B, Korsbakken J I, Le Quéré C and Nakicenovic N 2017 Key indicators to track current progress and future ambition of the Paris agreement *Nat. Clim. Change* **7** 118–22
Pierrehumbert R 2019 There is no plan B for dealing with the climate crisis *Bull. At. Sci.* **75** 215–21
Pindyck R S 2017 The use and misuse of models for climate policy *Rev. Environ. Econ. Policy* **11** 100–14
Prakash V and Girgenti G 2020 *Winning the Green New Deal: Why We Must, How We Can* (New York: Simon and Schuster)
Ramsey F P 1928 A mathematical theory of saving *Econ. J.* **38** 543–59
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*) (Washington, DC: President's Science Advisory Committee, United States, US Government Printing Office)
Riahi K *et al* 2017 The shared socioeconomic pathways and their energy, land use and greenhouse gas emissions implications: an overview *Glob. Environ. Change* **42** 153–68
Ricke K L, Moreno-Cruz J B and Caldeira K 2013 Strategic incentives for climate geoengineering coalitions to exclude broad participation *Environ. Res. Lett.* **8** 014021
Robock A, Oman L and Stenchikov G L 2008 Regional climate responses to geoengineering with tropical and Arctic $SO_2$ injections *J. Geophys. Res.: Atmos.* **113** D16
Rogelj J *et al* 2016 Paris agreement climate proposals need a boost to keep warming well below 2 °C *Nature* **534** 631–9
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 *Nat. Clim. Change* **3** 766
Schneider S H 1989 The greenhouse effect: science and policy *Science* **243** 771–81
Schneider S H 1997 Integrated assessment modeling of global climate change: transparent rational tool for policy making or opaque screen hiding value-laden assumptions? *Environ. Model. Assess.* **2** 229–49
Shayegh S and Thomas V M 2015 Adaptive stochastic integrated assessment modeling of optimal greenhouse gas emission reductions *Clim. Change* **128** 1–15
Shepherd T G *et al* 2018 Storylines: an alternative approach to representing uncertainty in physical aspects of climate change *Clim. Change* **151** 555–71
Sherwood S C and Huber M 2010 An adaptability limit to climate change due to heat stress *Proc. Natl Acad. Sci.* **107** 9552–5
Solomon S, Plattner G-K, Knutti R and Friedlingstein P 2009 Irreversible climate change due to carbon dioxide emissions *Proc. Natl Acad. Sci.* **106** 1704–9
Solow R M 1974 The economics of resources or the resources of economics *Am. Econ. Rev.* **64** 1–14
Stainforth D A and Calel R 2020 New priorities for climate science and climate economics in the 2020s *Nat. Commun.* **11** 3864
Steffen W *et al* 2018 Trajectories of the earth system in the anthropocene *Proc. Natl Acad. Sci. USA* **115** 8252–9
Stern N, Stern N H and Treasury G B 2007 *The Economics of Climate Change: The Stern Review* (Cambridge: Cambridge University Press)
Talati S and Higgins P 2019 Policy sector perspectives on geoengineering risk and governance *J. Sci. Policy Gov.* **14** (available at: https://www.sciencepolicyjournal.org/uploads/5/4/3/4/5434385/jpsg_talati_and_higgins_final.pdf)
Tol R S J 2003 Is the uncertainty about climate change too large for expected cost-benefit analysis? *Clim. Change* **56** 265–89
Tol R S 1997 On the optimal control of carbon dioxide emissions: an application of FUND *Environ. Model. Assess.* **2** 151–63
United Nations Framework Convention on Climate Change 2015 Paris agreement *Article 2(a)* (available at: 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 Aff.* **88** 64–76
Visioni D, Pitari G and Aquila V 2017 Sulfate geoengineering: a review of the factors controlling the needed injection of sulfur dioxide *Atmos. Chem. Phys.* **17** 3879–89
Wächter A and Biegler L T 2006 On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming *Math. Program.* **106** 25–57
Wagner G and Zeckhauser R J 2016 Confronting deep and persistent climate uncertainty *SSRN Scholarly Paper ID* 2818035 (Rochester, NY: Social Science Research Network)
Weyant J 2017 Some contributions of integrated assessment models of global climate change *Rev. Environ. Econ. Policy* **11** 115–37