ENVIRONMENTAL RESEARCH
CLIMATE

PURPOSE-LED PUBLISHING™

Part of the ENVIRONMENTAL RESEARCH SERIES

PAPER • OPEN ACCESS

Economic assessment of SRM under socio-political and geophysical tipping dynamics

To cite this article: Francisco Estrada *et al* 2026 *Environ. Res.: Climate* **5** 015015

IOP Publishing Environ. Res.: Climate 5 (2026) 015015 https://doi.org/10.1088/2752-5295/ae33df

ENVIRONMENTAL RESEARCH
CLIMATE

OPEN ACCESS

RECEIVED
31 July 2025

REVISED
22 December 2025

ACCEPTED FOR PUBLICATION
6 January 2026

PUBLISHED
16 January 2026

PAPER

Economic assessment of SRM under socio-political and geophysical tipping dynamics

Francisco Estrada[1,2,3,*], Bernardo A Bastien-Olvera[1,3], Oscar Calderon-Bustamante[1], Miguel A Altamirano[2], Rodrigo Muñoz-Sánchez[1], Juan Moreno-Cruz[4] and Wouter Botzen[2]

[1] Instituto de Ciencias de la Atmósfera y Cambio Climático, Universidad Nacional Autónoma de México, CDMX, Mexico
[2] Institute for Environmental Studies, Vrije Universiteit, Amsterdam, The Netherlands
[3] Programa de Investigación en Cambio Climático, Universidad Nacional Autónoma de México, CDMX, Mexico
[4] School of Environment, Enterprise, and Development, University of Waterloo, Waterloo, ON, Canada
[*] Author to whom any correspondence should be addressed.

E-mail: feporrua@atmosfera.unam.mx

Keywords: SRM, termination shock, tipping points, IAM, economic impacts, abrupt climate change

Supplementary material for this article is available online

Abstract

Solar radiation modification (SRM) is increasingly discussed as a policy response to worsening climate impacts and stalled mitigation progress. Yet the viability of SRM hinges on its long-term governance, particularly the risk of abrupt and permanent termination, which could trigger rapid warming and catastrophic outcomes. Here we develop a coupled socio-political–geophysical tipping point (SPTP–GTP) framework to assess the economic conditions under which SRM might reduce rather than amplify global risk. We introduce a novel damage function sensitive to the rate of warming and accounts for both catastrophic and non-catastrophic damages. Using reduced-complexity climate projections and a probabilistic failure modeling, we estimate the expected present value of SRM deployment across a range of governance scenarios. Our findings show that SRM only proves beneficial under a narrow intersection of robust global mitigation, extremely low failure risk, and gradual phase-out. Paradoxically, the same governance failures that make SRM politically attractive undermine the very conditions needed for its safe operation. These findings provide a quantitative, risk–risk perspective on the governance debate, suggesting that the required conditions for SRM are at odds with current socio-political realities.

1. Introduction

Solar radiation modification (SRM) represents one of the few policy interventions capable of rapidly halting further warming-or even reducing current global and regional temperatures (UNEP 2023). Given persistent challenges in reducing greenhouse gas emissions and current political context, warming exceeding 2 °C this century, potentially by mid-century, appears increasingly likely (Moreno-Cruz *et al* 2024, Forster *et al* 2025, Gay *et al* 2025). As warming proceeds, SRM deployment and other geoengineering interventions may become an attractive supplement to mitigation efforts, offering temporary relief from increasing climate impacts on human and natural systems and extending the time available for adaptation and mitigation (Lawrence *et al* 2018, MacMartin *et al* 2018, Irvine *et al* 2019, UNEP 2023, Baum *et al* 2024).

However, SRM carries significant risks, including regional climate disruptions, additional menaces for some sectors and regions, and the threat of termination shock (Jones *et al* 2013, Irvine *et al* 2017, Trisos *et al* 2018, Fan *et al* 2021, Carlson *et al* 2022, Gay *et al* 2025), consisting in a rapid warming if deployment ceased abruptly, with significant additional damages to human and natural systems and potentially triggering large-scale, catastrophic responses in the climate system (e.g. Atlantic Meridional Overturning Circulation (AMOC) collapse, permafrost carbon feedbacks). There has been debate over the seriousness and manageability of a termination shock, ranging from global societal collapse to considering it as an ‘overstated risk’ (Halstead 2018). Parker and Irvine (2018) distinguish between forced and elective termination shock and argue that the risk is manageable through policy safeguards and, in case of occurrence, economically absorbable. We believe that the optimistic perspectives largely overlook two critical dimensions shared across geoengineering literature:

1) Warming rate driven catastrophes: termination shock’s uniquely rapid warming rate could trigger nonlinear geo/biophysical responses and irreversible Earth system shifts. Crucially, prior economic assessments lack formal frameworks to quantify damages from a termination shock and potential rate-driven catastrophes (Aaheim *et al* 2015, Harding *et al* 2020);
2) Governance dependency: SRM’s viability critically depends on successful mitigation governance. As a temporary supplement, not a solution to climate change, it requires eventual phase-out contingent on emission reductions. Critically, the governance failure enabling SRM adoption (via failed mitigation) undermines the very conditions needed for its safe management (Asayama and Hulme 2019). This problem adds to the moral-hazard problem discussed in the literature (Markusson *et al* 2018, Jebari *et al* 2021, UNEP 2023, Moreno-Cruz *et al* 2024).

Here, we reframe termination shock through an integrated climate catastrophe perspective. Traditionally, catastrophes are attributed to geophysical/biophysical tipping points (GTPs). We argue that human and climate systems constitute a coupled system: worsening warming and persistent global mitigation failures fuel societal despair, triggering socio-political tipping points (SPTPs). SPTPs occur when escalating observed and projected impacts override societal risk aversion, forcing risk–risk tradeoffs where high-stakes interventions like SRM become politically viable. This dynamic explains why the potential deployment of SRM is now seriously debated: it emerges directly from the despair induced by failing international mitigation. Under such conditions, SPTPs enable policies whose governance failure (e.g. abrupt SRM cessation) could cascade into broader catastrophe. Critically, SPTPs and GTPs interact: maladaptive interventions that could lead to a termination shock could trigger geophysical tipping points (Lenton *et al* 2019a, McKay *et al* 2022a), while pre-existing governance failures amplify both SPTP activation and SRM’s systemic risks. SPTPs manifest primarily through policy shifts, but may also include cultural or institutional tipping behaviors (Nijsse *et al* 2025). Note that SPTPs can be interpreted as part of the positive/negative social tipping points recently proposed, in which society could bring cascading changes that effectively address the climate crisis or that lead to catastrophic outcomes (Schimel and Miller 2023, Spaiser *et al* 2024, Nijsse *et al* 2025). SPTPs focus on the destabilizing effects that policy decisions to tackle climate change can have on the Earth system. This SPTP framework establishes the narrative context for our subsequent economic modeling in a world that has already crossed such a SPTP, characterized by a high-risk governance environment, weak institutions, and failing multilateral cooperation.

Inspired by the socio-political–geophysical tipping point (SPTP-GTP) framework, we develop a novel damage function explicitly designed to address the identified limitations. This function quantifies catastrophe probability as a function of warming rate (relative to a ‘safe’ threshold), incorporates damage persistence to capture irreversible system shifts, and leverages recent econometric evidence to calibrate damages for both catastrophic and non-catastrophic outcomes.

We analyze SRM’s economic viability through a risk–risk lens. Specifically, we define critical viability boundaries, three of them that depend on operational and policy decisions: the tolerable termination shock probabilities, minimum safe phase-out durations for masked warming recovery, the ‘moral hazard’ and global mitigation policy. Two more boundaries that are studied are related to the climate-human system’s responses: the catastrophe-triggering likelihoods under abrupt warming, and the magnitude of catastrophic damages. These collectively determine the conditions where SRM reduces rather than amplifies expected climate risks.

There has been a longstanding debate about the role (if any) of SRM in tackling the worst impacts of climate change, and its governance including aspects like geopolitics, ethics, justice, and the global north and south divide (UNEP 2023, Jinnah and Dove 2025). Some authors argue that SRM inherently ungovernable and thus all SRM activities including small- and large-scale experimentation, any type of research and public funding should be banned (Stephens *et al* 2021, Biermann *et al* 2022). Others contend that governance is already emerging through ‘authoritative assessments’ that normalize and institutionalize the field (Gupta and Möller 2019). Middle-ground position favors moratorium on SRM use, but promote rigorous, responsible and ethical research and multilateral discussions about a global governance system (Climate Overshoot Commission 2023, UNEP 2023, Office of the European Union 2024, Parson *et al* 2024). Our analysis directly informs this debate by quantifying the conditions under which SRM reduces systemic risk. Our framework provides a way to test some core claims of these positions.

We show that the viability of SRM is inextricably linked to the success of global greenhouse gas mitigation governance, requiring essential prerequisites of robust international mitigation, very low failure risk, and gradual phase-out. Yet critically, these prerequisites are undermined by the very governance failures that make SRM politically attractive in the first place, revealing a fundamental paradox.

The paper proceeds as follows. Section 2 describes the methodology used, including the proposed damage function, parameter calibration and assumptions, as well as the framework of SPTP-GPT for stratospheric aerosol injection (SAI) interventions. The results are discussed in section 3, including the conditions that support the adoption of SRM interventions. This section also discusses some caveats of the analysis and outlines future research to expand current evaluations of the economic convenience SRM implementation. Section 4 concludes with policy implications.

2. Data and methods

2.1. Damage function

There are a limited number of assessments of the economic consequences of deploying SRM options and mostly focus on the SAI method. For instance, (Aaheim *et al* 2015) use the output of global climate models and a computable general equilibrium model to assess the impacts of implementing two SRM methods, SAI and marine cloud brightening. Their results suggest that under a moderate emissions baseline (RCP4.5), SRM deployment may bring benefits, but that these are small in general and negative for some regions. A more recent study (Harding *et al* 2020) reports contrasting results in which considerable income benefits at the global level would be obtained from deploying SAI under a high emissions scenario (RCP8.5). Moreover, income inequality would be significantly reduced not only in comparison with the RCP8.5 scenario, but with respect to 2010. The authors caution that their results are not consistent with general concerns about SRM increasing inequalities between global north and south, with the latter having lower benefits or higher residual damages. One of the reasons they offer is related to the econometric damage function used as it may not be adequately representing the effects of changes in climate under a geoengineering scenario.

However, none of such studies considered the potential catastrophic consequences of abrupt warming that can be produced by a sudden termination of SAI operation. In this paper we propose that the assessment of the economic benefits/costs of geoengineering should be analyzed through the lens of catastrophic changes in the climate system. Using standard damage functions to investigate the economic convenience of deploying geoengineering can severely underestimate the risks, and results can be more informative of the limitations of such damage functions than of the problem being studied. This has been previously discussed in the analysis of low-probability/high-impact events, geophysical tipping points and catastrophic climate change (Weitzman 2009, Pindyck 2013, Stern 2013, Anthoff *et al* 2016, Nordhaus 2019). While the probability of a termination shock that leads to catastrophic damages can be assumed small due to potential guardrails and security measures, it is not zero. As discussed in Weitzman (Weitzman 2011), fat-tails in welfare losses have large implications for assessing economic costs, income inequality effects, as well as for policy advice. This is the case of uncontrolled warming, but also SRM interventions. In the context of possible deployment of geoengineering interventions, the explicit inclusion of SPTPs and GTPs becomes increasingly relevant, as abrupt changes in the rate of warming can trigger tipping points and reorganization of synoptic to local processes, with large impacts on human and natural systems that are absent or poorly considered in standard damage functions. Risk–risk framework for SRM assessment should account for all the relevant benefits and risks from all the options considered (Felgenhauer *et al* 2025), and until now the risk of a termination shock in economic assessments of SRM has been largely absent.

To assess the damages of changes in the rate and level of warming, we propose a damage function that is the mixture of non-catastrophic and catastrophic components:

$$D_t = \alpha_1 \Delta T_t^2 + \omega_t \left[\alpha_2 \Delta T_t^{\beta}\right] \tag{1}$$

$$\omega_t = \frac{1}{\left(1 + e^{-a(\delta_t - c)}\right)} \tag{2}$$

$$\delta_t = \frac{(\Delta T_t - \Delta T_{t-k})}{R_k^*} = \frac{R_t}{R_k^*} \tag{3}$$

$$\delta_t = \max(\delta_t, \delta_{t-1}); \tag{4}$$

where α_1, α_2 are scaling parameters of the non-catastrophic and catastrophic components, respectively, of the damage function that convert global temperature change ΔT_t to percent losses in global

GDP (table S1 provides quick reference of the definitions of the parameters in equations (1)–(4). The catastrophic component represents the effects of abrupt climate change that could lead to severe economic losses. The difference with some of the previous work on catastrophic damage functions is that equations (1)–(4) are designed to account for catastrophic damages that depend on the level of warming but that are triggered by rapid changes in the warming rate (Mastrandrea and Schneider 2001, Weitzman 2012, Howard and Sterner 2017). As such, severe damages may occur even at moderate temperature levels if warming happens over a short period. Variable ω_t in equation (2) represents the probability of occurrence of catastrophic impacts, which is a function of the ratio of a time- and scenario-dependent rate of change in global temperature and a fixed change rate (equation (3)). δ_t is the ratio of the annual change in warming R_t of the scenario of interest and a safe rate of change R_k^*. This ratio is expressed as units of the safe rate of change R_k^*. A characteristic of catastrophic climate change is that it can have highly persistent or permanent consequences (Lenton *et al* 2008, Estrada *et al* 2015, Lenton *et al* 2019a, McKay *et al* 2022a, Tol 2024). Equation (4) accounts for this property by which the increase in risk of catastrophic climate change is permanent during this century.

Inductive knowledge obtained from observed data is always useful, but limited about what lies well beyond the range of experience (Weitzman 2011). This is why damage functions produced solely by econometric methods or CGEM are ill-posed to tackle scenarios that involve the possibility of catastrophic climate change. Such methods can be very useful to increase the confidence about central regions of the damage probability distribution, but not for its tails (Weitzman 2009, 2011). Analyzing catastrophic climate change requires accepting the burden of higher epistemic uncertainty and more speculative formulations. However, despite this, when combined with sensitivity analysis, relevant information can be extracted about the effects of low-probability, high impact outcomes. In this paper we make use of published studies and data to calibrate the parameters of the proposed damage function and present extensive sensitivity analyses to inform about the limits and conditions for SRM deployment.

The non-catastrophic component α_1 is calibrated using a recent literature review (Tol 2024) which suggest that a 2.5 increase in global temperatures would produce a 2.71% loss in GDP, of which 0.21% correspond to catastrophic damages. From the total loss estimate of 2.71% we subtract 0.21% to avoid double counting catastrophic damages (Tol 2024). The calibration of parameters α_2, β, and R_k^* is complex and speculative since little is known about the probability of occurrence of abrupt changes in the climate system, and what they could imply for human and natural systems. However, previous studies are helpful for proposing parameter values that reflect the current state of knowledge. Previous work on catastrophic economic consequences addresses mainly high levels of warming, regardless of the rate of increase in global temperature. Weitzman (2009) and Pindyck (2013) focus on the nonlinearity of damages at very high levels of warming. For example, Weitzman (2012) proposes a damage function in which an increase in global temperature of 6 °C produces a decrease of 50% in global GDP. The exponent in such damage function is $\beta = 6.754$ creating rapidly increasing damages for large temperature changes, but assumes that only catastrophic damages occur only under high levels of warming, ignoring the role abrupt changes in rate of warming can have. Other efforts have been centered on analyzing aspects like the probabilities, outcomes and levels of warming that would trigger specific tipping points (Vellinga and Wood 2002, Link and Tol 2011, Anthoff *et al* 2016, Cai *et al* 2016, Diaz and Moore 2017). The costs of the damages associated with the occurrence of different tipping points is estimated in the range from 2.5% to 20% of global GDP (table 1 in Cai *et al* 2016). However, these estimates seem conservative considering some recent studies for non-catastrophic climate change, particularly those based on econometric methods. Recent estimates of (non-catastrophic) economic losses due to climate change (Burke *et al* 2015, Kalkuhl and Wenz 2020, Kotz *et al* 2024) suggest decreases between 50%–75% for the end of the century (RCP8.5, about 4 °C increase in global temperature). Kotz *et al* (2024) project decreases of 19% (range: 11%–29%) of GDP by mid-century (about 2 °C) and about 60% by the end of the century (about 4 °C) with confidence intervals exceeding 70% and nearing 90% for the 1% significance level. Based on computable general equilibrium output and quantile regression, Van der Wijst *et al* (2023) estimates that 2 °C and 4 °C increases in global temperature (w.r.t. preindustrial times) would produce a reduction in global GDP of about 20% and 32%, respectively. Based on this information we propose as a calibration point for α_2 a 25% loss in global GDP for a severely abrupt (i.e. $\omega_t = 1$) increase in warming occurs and the level of warming is 3 °C. We select $\beta = 4$ which implies that the damages increase rapidly with warming, but not as much as in other catastrophic damage functions, such as Weitzman (2009). The maximum safe rate of change (R_k^*) in global temperatures is assumed 0.015 °C per year, for which the risks of nonlinearities in the climate system may already be significant. Note that even 0.15 °C/decade would lead to about 1 °C above current levels of warming ($\sim 1.4°C$) at the end of this century, providing sound reasons to argue for a much lower safe rate (IPCC 2018, McKay *et al* 2022a). This yearly change rate is within the warming rates in climate scenarios without

SRM intervention, while in the case of a termination shock such rates could be in the range of 0.13 °C to 0.76 °C (Ross and Damon Matthews 2009), with other estimates placing them between 10 to 20 times the warming rates in the 20th century (Matthews and Caldeira 2007). Studies suggest that the period for recovering the masked warming once an uncontrolled termination shock occurs is short, ranging for 5–8 yr (Irvine *et al* 2012, Jones *et al* 2013). The sigmoid function ω_t is parameterized with $c = 8$, implying that $\omega_t = 0.5$ when the observed rate of warming is eight times the safe rate of warming, and $a = 2$, such that $\omega_t \approx 1$ for $\delta_t = 11$. As mentioned above, given the high uncertainty in parameter calibrations, extensive sensitivity analyses are conducted and reported in the results section.

2.2. Estimating the expected present value of SRM deployment accounting for the risk of a termination shock event

In addition to the probability of GTPs, the deployment of SRM implies a non-zero risk of failure. The risk of deploying SRM and having to abruptly and permanently terminate its use is a low-probability/high-impact event that could be brought about by variety of reasons, such as unexpected outcomes of its implementation, technological issues, war, conflict, terrorism, pandemics, and breakdown of logistics due to lack of international cooperation, and perception of failure of SAI (Corry 2017, Avin *et al* 2018, Baum *et al* 2018, Tang and Kemp 2021, Keys *et al* 2022). This risk can be enhanced/reduced by the progress or lack of thereof in international mitigation policy, and by the type and strength of the governance constructed around SRM, which can also determine its duration and the steepness of masked warming recovery (Jones *et al* 2013, Irvine *et al* 2017).

Assume that the probability of such an abrupt and permanent termination of SRM (APTSRM) in any given year is P_{failure}. As discussed in more detail in the next subsection, this parameter is treated as exogenous, meaning that the specific socio-political processes that would lead to an APTSRM are not modeled. Instead, we explore a wide range of values to explore the bounds of tolerable probability of failure under which SRM remains a risk-reduction strategy. While this probability is unknown, it is expected to be very small but larger than zero. As such, if the SRM effort is sustained over a period n, the probability of not having any APTSRM during n is given by $(1 - P_{\text{failure}})^n$, which becomes increasingly smaller with n. For example, if $P_{\text{failure}} = 1\%$, then the probability of having at least one APTSRM during a century is $1 - (0.99)^{100} \approx 64\%$, while the complement is only about 36%. That is, with $P_{\text{failure}} = 1\%$, it is about twice as likely to have a APTSRM event than not. A $P_{\text{failure}} = 0.5\%$ implies an almost 40% chance of an APTSRM, and a 10% chance for $P_{\text{failure}} = 0.1\%$. Some studies claim SAI would imply a multi-century commitment if GHG emissions reductions are expected to come from scaling up carbon dioxide removal (CDR) technologies such as direct air capture (Baur *et al* 2023). As such, risk–risk analyses and the assessment of the expected economic costs/benefits of SRM interventions must include the possibility of a termination shock. Moreover, this type of analysis can help determine the range of probabilities of failure that may make the risks of SRM failure acceptable.

The calculation of the expected present value of damages under an SRM intervention requires evaluating the present value of damages for each year in which APTSRM event can happen, weighted by their probability of occurrence. These probabilities can be calculated using a geometric probability distribution:

$$P(\text{First failure in year } t) = P(t) = (1 - P_{\text{failure}})^{t-1} P_{\text{failure}}. \tag{5}$$

The expected present value of damages under an SRM intervention is calculated as follows:

$$\text{EPV}_{\text{SRM}} = \sum_{t=1}^{n} P(t)\, PV_{\text{failure at t}} + (1 - P_{\text{failure}})^n PV_{\text{no failure}}. \tag{6}$$

The EPV_{SRM} value can be compared with the present value of a no intervention scenario ($PV_{\text{no intervention}}$) to evaluate the economic convenience of deploying SRM.

2.3. Why exogenous failure probabilities: the governance paradox

History shows that societies tend to invest significantly in reducing probabilities of failure in high-risk domains, such as nuclear power generation and aviation, precisely because the consequences of failure are catastrophic. This could suggest that a model that includes an endogenous failure probability would be more appropriate, where higher potential damages would incentivize better governance and lower probabilities of failure. However, this logic holds for mature technologies within stable and effective governance systems, which we argue is not consistent with the characteristics of SRM deployment under the SPTP scenario analyzed here. SRM is considered as a response to the breakdown of robust governance conditions. The core paradox we model is that the conditions making SRM politically viable are the same that make low failure probabilities exceptionally hard to achieve. In this context, treating the probability of failure as an endogenous, optimizable variable would imply an unrealistic capacity for self-correcting governance within a system defined by its fragility.

Modeling P_{failure} as an exogenous parameter is based in the governance paradox described by the SPTP framework which is supported by previous work (Clark 2023). SPTP drivers, such as eroded coordination capacity, contested legitimacy, and multilateral instability, operate upstream. They constitute the antecedent socio-political conditions that define a fragile governance environment, shaping both the feasibility of managed phase-out and the entire distribution of potential failure risks. This context frames SRM as a high-stakes response to governance failure, not a technically managed option within a stable system.

P_{failure} in our model is a downstream, reduced-form representation of the aggregate systemic risk emerging from this fragile context. The probability of failure is conditional on the upstream conditions created by the occurrence of SPTP (i.e. SRM has been deployed as a response to governance failure). Our quantitative framework is structurally agnostic to the specific micro-foundations of failure, whether technical or socio-political in origin. It is designed as a governance-conditional risk screening to capture the economic consequences of operating under a given level of systemic risk, whatever its underlying drivers.

There are three reasons why we treat P_{failure} as an exogenous parameter:

1) The high-risk governance context of SRM deployment: SRM becomes a viable policy option only after the governance of international greenhouse gas mitigation has failed, with institutions being unsuccessful at simpler coordination tasks. Moreover, the pathway to SRM involves escalating crises, which typically weaken rather than strengthen governance. We model a world where crises are more likely to lead to desperation, SPTP and institutional fragility than to a coordinated, pre-emptive bolstering of global governance. These conditions are at odds with the highly reliable, long-term coordination required for safe SRM management.
2) Framing the governance precondition: by treating P_{failure} as exogenous we frame the problem as a function of governance quality such that the preconditions that governance must meet can be identified.
3) Sensitivity analysis: rather than presupposing how governance responds to stress, our framework allows the critical question 'What level of governance is required for SRM to be risk-reducing?' to be answered quantitatively. Readers who are optimistic about the global capacity to build robust SRM governance can focus on the low-probability results, while those who are skeptical, based on the historical precedent of failed climate cooperation, can see their concerns reflected in the high-probability scenarios.

The SPTP perspective motivates the consideration of a wide range for key parameters like P_{failure}, provides the narrative for our scenario design, and, most importantly, guides the policy interpretation of our economic decision surfaces, highlighting the governance paradox at the heart of the SRM dilemma.

2.4. SPTP framework for SRM deployment analysis

The consideration of SRM as an option to reduce the damages and risks of unabated or partially abated anthropogenic climate change results from the governance failure of global climate policy, particularly with respect to mitigation goals. As has been discussed in the literature, the SRM deployment decision can be analyzed and understood through a risk–risk analysis framework, in which the risks of anthropogenic warming without SRM are compared to those of SRM deployment plus residual risk. Under such framework, the decision to consider and eventually deploying SRM is a function of the severity of the (present and future) risks perceived by global society (or a part of it) (Aldy *et al* 2021). This risk perception is in turn influenced by the lack of past progress in the global mitigation governance and to physical, political and/or socioeconomical constraints to reduce warming in the short- to mid-term. Risk–risk analysis entails society will consider and implement increasingly riskier options only as climate change becomes increasingly dire. In an integrated human and climate system, some socio-political decisions can cascade into triggering drastic and possibly irreversible changes in the climate system such as abrupt warming and catastrophic events like the collapse of the AMOC, and the disintegration of ice sheets, as well as the possible destabilization of biological and human systems. We refer to the conditions in which societal risk aversion is overridden into accepting high-stakes risk–risk tradeoffs as sociopolitical tipping points. These are illustrated by the implementation of some geoengineering options like SAI and large-scale CDR, but may not be limited to them.

Once the SPTP is triggered and SRM is deployed, its success critically depends on both the governance of SRM itself and that of global mitigation of GHG (Rabitz 2019). Regarding the SRM governance, there are two essential aspects considered in this paper that are related to maintaining risks from materializing and to minimize the negative impacts of SRM (Brent *et al* 2024): 1) the safe and continuous operation of the SRM once it is implemented, which we will investigate through its probability of failure; 2) how phase-out is managed in case of having to terminate SRM deployment, that is represented here by the number of years to recover the masked warming. With respect to the governance of global mitigation, our analysis considers various global emissions trajectories that represent different international mitigation efforts ranging from a prompt start of significant mitigation to severely delayed mitigation or no mitigation. These scenarios can be interpreted as possible realizations of what has been called the moral hazard effect of SRM that could deter mitigation (long delays in starting mitigation) or that could even promote it (mitigation soon after SRM starts). The scenarios used to illustrate the framework described in the following subsection.

2.5. Climate change scenarios for a risk–risk economic assessment of SRM including different global GHG mitigation governance

The global temperature change scenarios in this paper were generated using the MAGICC7 reduced complexity climate model available at https://live.magicc.org/ (Meinshausen *et al* 2011, 2020, Nicholls *et al* 2021).

The SSP585 was chosen as the inaction scenario in which no SRM and no mitigation take place. This selection responds not to how likely the world is to follow that trajectory, but to allow the possibility of reaching tipping points under both the inaction and SRM intervention scenarios. Allowing for catastrophic outcomes in the inaction and SRM scenarios can be more informative for risk–risk analysis. The baseline SRM (SRM20) scenario is represented by a successful deployment of SAI in the year 2020 that results in a SSP245 global temperature change trajectory, absent of termination shock failures. By 2100, the reduction in global temperatures from deploying SRM is 2.13 °C (i.e. the difference between SSP585 and the SRM20 scenarios). To represent possible results of global GHG mitigation while SRM is deployed, we explore four scenarios that can be interpreted as illustrations of the deterrence or enhancing effects of SRM (figure 1):

1) GHG governance scenario M30: SRM deployment is started in 2020 and global mitigation actions start 10 years after, following the emission path of the SSP245. This results in an overshoot in 2100 of 0.39 °C over the target SSP245 scenario. Conversely, the amount of SRM cooling needed to match the SSP245 target scenario is 0.39 °C by year 2100.
2) GHG governance scenario M40: as in the M30 scenario but global mitigation actions start 20 years after SRM deployment. With respect to the SSP245 scenario, in 2100 the M40 has an overshoot of 1.05 °C. The amount of SRM cooling that would be required to match the SSP245 target scenario is 1.05 °C by year 2100.
3) GHG governance M50: as in the M30 scenario but global mitigation actions start 30 years after SRM deployment. The M50 represents an overshoot of 1.6 °C by 2100. This implies that the amount of SRM cooling that would be required to match the SSP245 target scenario is 1.6 °C by year 2100.
4) GHG governance MF: this scenario represents global mitigation failure and follows the SSP585 trajectory. The amount of SRM cooling that would be required to match the SSP245 target scenario is 2.13 °C by year 2100.

A fifth scenario (M60) was considered in which mitigation actions start in 2060 but the difference between the MF and this late mitigation scenario is only 0.16 °C, which is negligible for our analysis and thus is omitted. Note that these scenarios do not include the effects of SRM deployment as they are meant to represent the baseline temperature change without SRM.

In addition to these scenarios, we also include a comparison of different levels of SRM intervention leading to 0.3 °C, 1.0 °C and 2.0 °C for each scenario discussed above. This allows us to investigate which levels of SRM intervention would lead to the largest risk-reduction for each mitigation scenario considered.

2.6. GDP projection

For the analysis presented in this paper, the SSP5 GDP projection was used. This socioeconomic scenario can be combined with radiative forcing levels of 8.5 W m^2 and also of 4.5 W m^2 (Riahi *et al* 2017). This socioeconomic scenario depicts a world with fast economic growth based on fossil fuels and technological progress. The OECD Env-growth quantification of the SSP5 narrative is used and was obtained

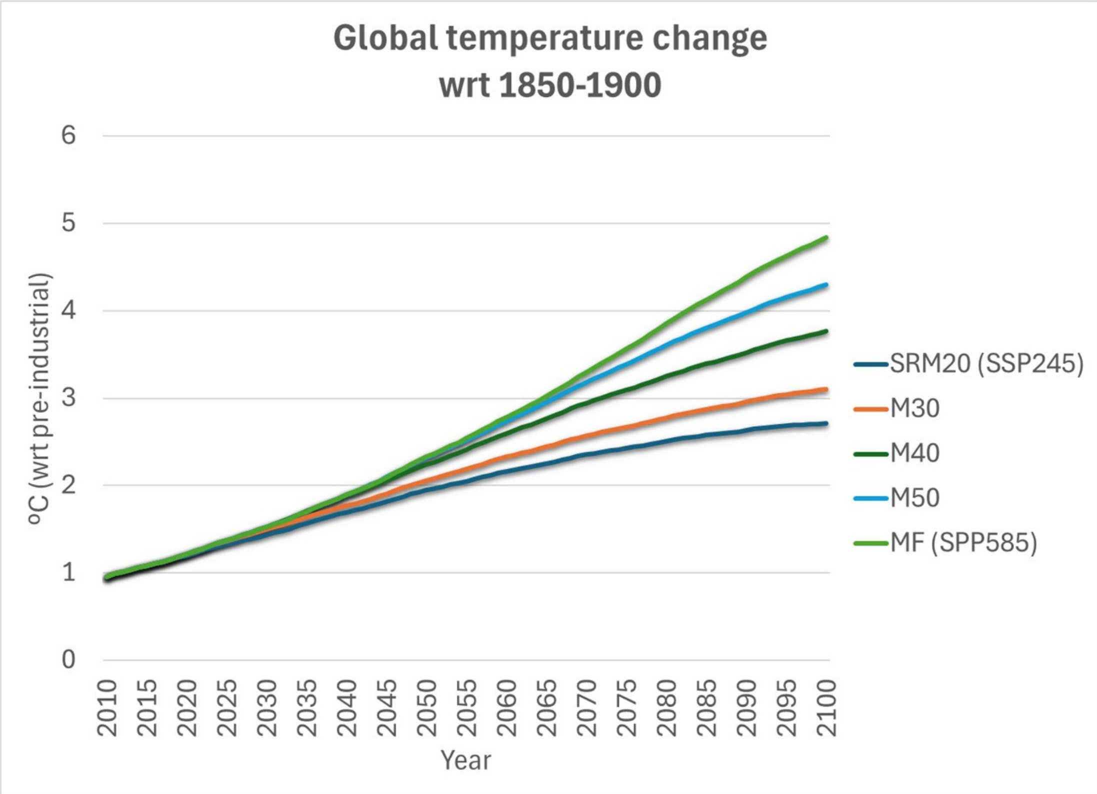

Figure 1. Global temperature projections depicting different levels of GHG mitigation and the SRM20 target scenario. M30, M40 and M50 represent the global temperature change trajectory for GHG mitigation efforts starting in 2030, 2040 and 2050, respectively. MF shows global temperature change for a mitigation failure scenario, while the SRM20 scenario describes temperature change achieved through stratospheric aerosol injection.

from the SSP Database (https://tntcat.iiasa.ac.at/SspDb/dsd?Action=htmlpage&page=30). The data was linearly interpolated from 5 year frequency to annual.

For estimating the impact of a termination shock various sets of climate scenarios were developed. Each of these sets is based on two scenarios: the target SRM scenario in which no termination shock occurs, and one of the four GHG governance scenarios described above. For each pair of these 'parent' scenarios, individual scenarios are generated in which a single and permanent termination shock is imposed in each year t in the period of analysis. The termination shock leads to a transition from the successful SRM to the selected GHG governance scenario, and the transition is characterized by an n-year time to recover the masked warming. Figure 2 illustrates the individual SRM failure scenarios generated for the SRM and MF scenarios, using a 5 year transition window. To investigate the governance of safe implementation of SRM, our analysis explores two main controllers: (1) the probability of failure, described above, and, (2) the phase-out management in case of unexpected termination of SRM deployment, which is represented by the window for recovering the masked warming.

3. Results and discussion

3.1. Assessing the economic damages from two illustrative termination shock scenarios

Supplementary figures S1–S2 illustrate the behavior of the proposed damage function for the combination of the SRM20 and MF scenarios and two different dates of SRM failure and two windows for recovering the masked warming. In figure S1, the termination shock occurs in 2030 and the transition years for recovering the masked warming is 8 years, which is in the upper bound for an uncontrolled termination of SRM. Since the termination occurs shortly (10 years) after the start of SRM deployment the amount of masked warming is relatively small, about 0.4 °C in 2030 and 0.6 °C in 2038. In all three scenarios (MF, SRM20 and SRMfailure) δ values increase as the warming progresses, leading to about 4 times the safe rate of warming in the MF baseline scenario and remaining limited to about 2 when the SRM is successfully deployed. This illustrates the value of SRM to limit the potential of catastrophic outcomes in comparison to the baseline scenario, if the SRM is managed to be safely maintained. However,

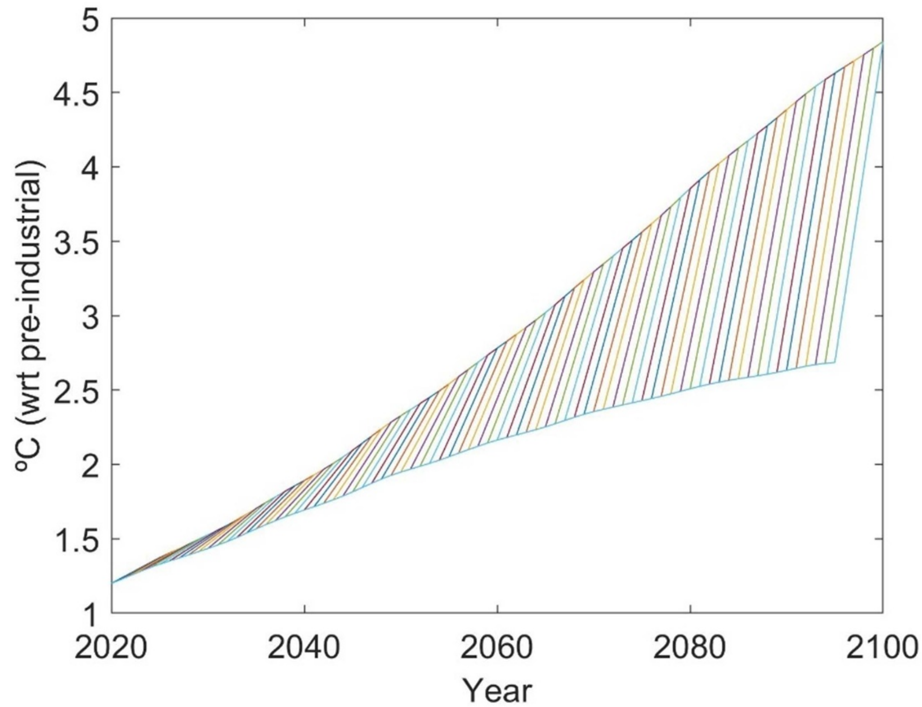

Figure 2. Illustration of a set of potential SRM failure scenarios. The individual scenarios result from imposing a single and permanent termination shock in every possible year, using the SRM20 as the intervention scenario and the MF as the unmasked warming scenario global temperature would revert to in case of a termination shock. The window for recovering the masked warming is set to 5 years.

if the termination shock occurs, δ increases to values of about 6 times the chosen safe rate of warming, leading to a probability of catastrophic damages of about 2%. These differences in the amount and rate of warming among scenarios lead to notable dissimilar damages: the successful implementation of SRM20 can lead to a reduction of about 70% of economic losses in 2100 in comparison to the MF baseline scenario; however, if a termination shock occurs, damages in 2100 are 30% larger than in the MF scenario. Using a 1.5% discount rate, the present value of MF baseline, SRM20 and SRMfailure are US$868 387.66 billion, US$392 872.81 billion, and US$1000 683.13 billion, respectively. That is, a successful SRM20 would reduce the present value of damages by about 55% in comparison to the MF baseline, while the present value in the case of an early uncontrolled termination of SRM would be 15% larger.

Figure S2 shows the case of a more severe termination shock later in the century with a much larger differential warming (1 °C in 2050 and 1.5 °C at the time of full recovery of masked warming) between the MF baseline and the SRM20 scenarios. For this example, a 15 year window for recovering the masked warming was chosen to represent an effort to manage phase-out instead of a more abrupt termination. Under this termination shock scenario, the values of δ exceed 7 times the safe rate of warming and drive the probability of catastrophic damages to about 40%, with damages steeply increasing during the second part of the century. Under the SRMfailure scenario, losses in 2100 are about 6.5 times larger than those of the MF baseline scenario. The present value of this controlled termination shock is US$3697 948.90 billion, about 3 times larger than that of the MF baseline.

Since results vary widely depending on the date of failure and the recovery window, as well as other parameters, the next subsection presents decision surfaces with extensive variations of the main parameters and GHG and SRM governance levels. Moreover, the date of failure of SRM is treated as unknown and the expected present value is used.

3.2. Constructing economic decision surfaces for assessing SRM deployment

This section uses a risk–risk assessment approach to underline the SRM and GHG governance conditions under which an SRM intervention would be economically desirable. For this purpose, the expected present value of SRM deployment is compared with alternative baseline scenarios representing different levels of global GHG mitigation efforts. The results presented below are derived from extensive sensitivity analyses of key parameter values with the objective of producing economic decision surfaces that

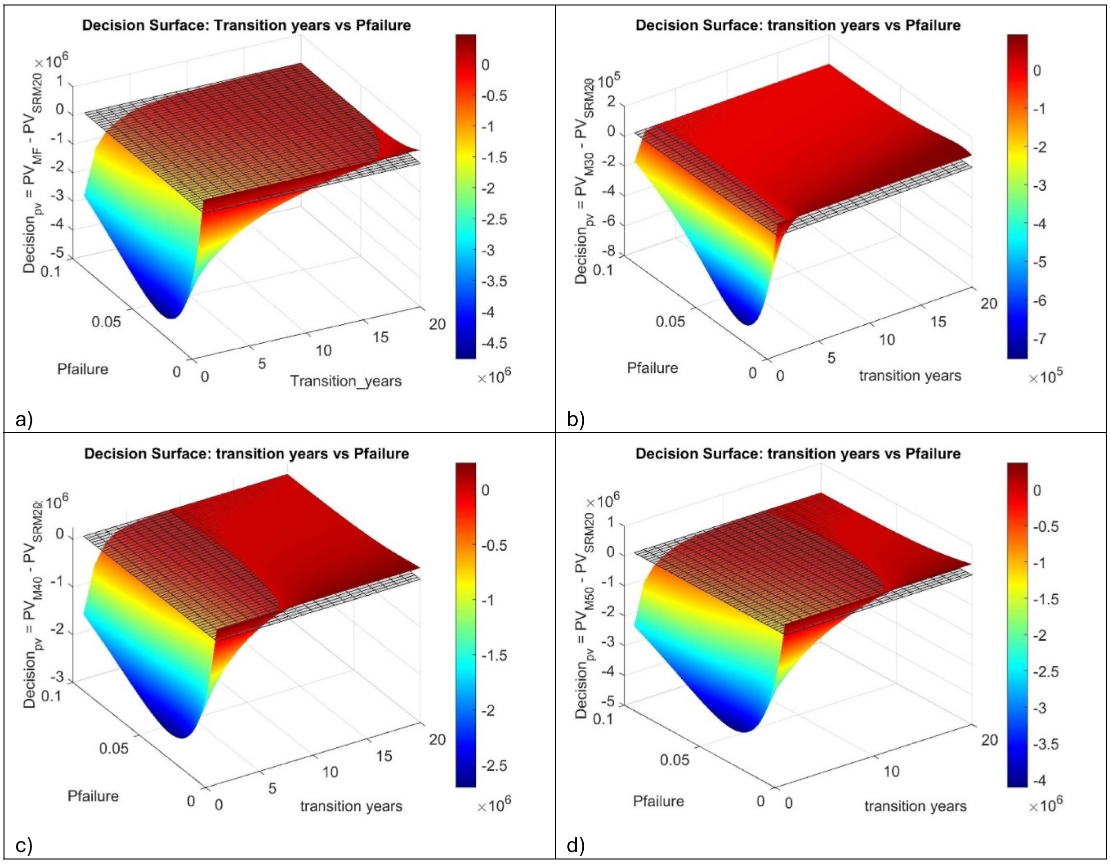

Figure 3. Decision surfaces comparing SRM20 and four different scenarios of global GHG mitigation. Panel (a) shows the decision surface based on comparing the present value of MF and the expected present value of SRM20 as a function of the probability of failure of SRM (from 0% to 10%) and the length of the window for recovering the masked warming. Panels (b)–(d) are as in panel (a) but for the M30, M40 and M50 scenarios, respectively. Positive values in the *z*-axis indicate regions of the decision surface in which SRM20 is preferred (units in billion dollars).

define viability boundaries for SRM from an economic perspective. As shown in table S1, the default discount rate used for the calculations below is 1.5%, unless otherwise stated for sensitivity analyses.

3.3. Decision surfaces for SRM deployment under different global mitigation efforts

The probability of SRM failure and the window to recover the masked warming are the main SRM governance controllers in our framework to maintain the safe operation of this intervention and are the focus of this subsection. A third controller is to influence GHG governance to reduce the risks in case of termination of SRM. For this purpose, the four mitigation scenarios (MF, M30, M40 and M50) in figure 1 are used to define four cases of combined SRM and GHG governance levels of success. Case 1 compares the expected present value of SRM deployment (SRM20) to the present value that would be obtained under no global mitigation and no SRM (MF). This combination of scenarios can be also interpreted as the case in which the SRM 'moral hazard' materializes and inhibits international mitigation efforts.

Figure 3(a) shows the differences between the present value of the MF scenario and expected present value of SRM20, considering the possibility of catastrophic outcomes, for a range of possible values of the probability of failure of SRM and the window to recover the masked warming. In this decision surface, positive values denote combinations of the parameters space for which the implementation of SRM20 would provide larger economic benefits, while negative values denote those for which the MF scenario is preferrable. In case 1, the viability boundary for SRM intervention is strictly limited to a thin strip defined by failure probabilities between 0% and 0.5% that becomes wider for longer warming recovery windows (figure 3(a)). The tolerable failure probability goes up to 4.5% if the SRM governance provides a transition year window of 20 years. The highest positive difference between the expected present values of SRM20 and the present value of MF occurs for the long transition windows and the lowest failure probabilities, and represents about a benefit of about US$410 000 billion. The case of a zero probability of failure is of special interest as it illustrates the limitations of previous studies that do not consider the possibility of catastrophic outcomes. Under such a stringent scenario, SRM always results in benefits when compared to the no intervention scenario and these benefits increase for high α_1 and low discount rates (figure S3). The omission of a catastrophic component in the analysis can explain more positive evaluations of SRM suggested in previous studies (Aaheim *et al* 2015, Harding *et al* 2020). Moreover, considering that *ceteris paribus* low-income countries are generally associated with higher damages (i.e. higher α_1), this finding also could partially explain why previous assessments suggest SRM could reduce inequality, as reduced warming would bring larger benefits to them (Harding *et al* 2020).

Figure 3 also provides relevant insights about the probability of failure for the governance objective of safe SRM operation. Critically, low probabilities of failure can produce much higher risks. In Case 1 (i.e. MF-SRM20 combination of scenarios), the minimum of the decision surface occurs for a probability of about 3%. This occurs because when probabilities of failure are very close to zero, it is likely the termination shock and the corresponding catastrophic damages do not occur; conversely, for high probabilities of failure, the termination shock is likely to occur earlier in the century, where the distance between MF and SRM20 is still small, and the damages are more limited. However, for probabilities larger than $\sim$ 0.5% and smaller than $\sim$ 4% the damages of a termination shock are maximized. This illustrates the tight safe operational space that SRM governance needs to commit to in case of deployment. Moreover, this result is similar across all mitigation scenarios (M30, M40, M50) although the magnitude of the losses is about half an order of magnitude smaller for early GHG mitigation. Figure S4 shows that higher discount rates, which impose lower weights to longer term outcomes, also increase the area of the decision surface for which combinations of probability of failure and the number of transition years lead to benefits from SRM20 in comparison to MF. The result of risk maximization for probabilities of failure in the range $\sim$ 0.5% to $\sim$ 4% is robust to the discount rate selection.

Cases 2–4 replace the MF scenario with active global mitigation policy starting at different dates. Note that for the analysis presented here, these GHG reduction scenarios imply that, in case of an active global mitigation world, the amount of SAI would be adjusted to maintain the SSP245 target. Case 2 uses the M30 scenario to represent the instance in which SRM has a positive influence on global mitigation, which starts in 2030 with a reduction effort that would lead to a warming very similar to the SSP245 at the end of the century. The M30 mitigation scenario extends the viability boundaries in the decision surface for a much wider range of the parameter values of failure probability and of length of the window for recovering the masked warming. If SRM deployment is accompanied by decided international mitigation action to rapidly eliminate the masked warming, the operational space for SRM benefits is less limited. Under the M30 scenario, a large fraction of warming reduction is done by GHG mitigation and SRM is used to eliminate a small part of the warming, and the associated impacts, that would happen otherwise. Benefits from SRM20 occur for most of the combinations of values of the probability of failure and length of window for recovering the masked warming, except for cases in which the warming is recovered in less than 5 years. As noted before, even under the M30 scenario, SRM deployment can lead to severe losses for cases in which the probability of failure is larger than 0.5%. Cases 3 and 4 show that the delay in strong global mitigation efforts rapidly erode the potential for benefits from SRM interventions. In Case 3 (M40), SRM20 is economically preferrable mostly for windows for recovering the masked warming at least larger than 8–10 years, while for Case 4 (M50) the window length in most cases must be over 13 years. However, for these late and weak mitigation scenarios, the magnitude and the range of parameter values that lead to catastrophic economic losses increases rapidly.

3.4. Sensitivity analysis of the probability of occurrence of catastrophic economic losses from exceeding geophysical tipping points

Several publications suggest that geophysical tipping elements may be activated at relatively low levels of global warming (Drijfhout *et al* 2015, Lenton *et al* 2019a, Bathiany *et al* 2020, McKay *et al* 2022a). Using 57 climate models from the Coupled Model Intercomparison Project 6 (CMIP6), a recent study detects abrupt shifts in 6 of 10 subsystems already at 1.5 °C, and finds that models with faster warming rates exhibit more abrupt shifts. (Terpstra *et al* 2025). Motivated by this evidence, we examine how uncertainty in tipping-point trigger thresholds and damage severity alters the viability of SRM.

This subsection presents a sensitivity analysis of the viability boundaries across three GHG governance scenarios (MF, M30, and M40) varying (i) the threshold/centering parameter c of the sigmoid that maps the normalized warming rate to catastrophe probability ω and (ii) the catastrophic loss scale α_2. In our setup, is the excess-rate factor at which $\omega = 50\%$, and a range of $c \in [4, 12]$ is used for the analysis, meaning a 50% catastrophe probability when the warming rate is 4–12 times the chosen 'safe' rate. The range for α_2 spans losses of 10% to 50% of global GDP in the catastrophic component.

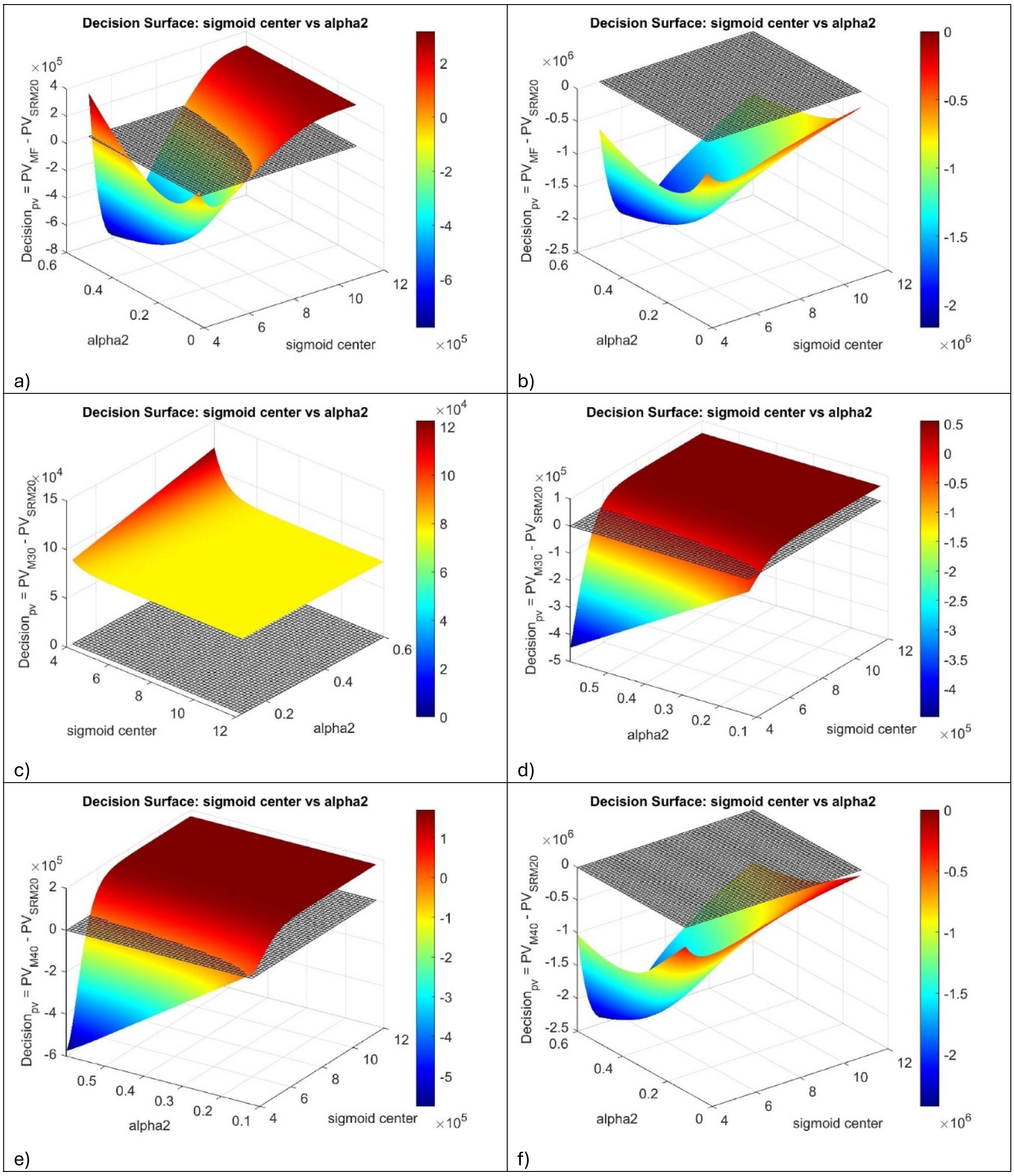

Figure 4. Decision surfaces comparing the SRM20 with the MF, M30 and M40 scenarios a range of values of the sigmoid center (c) and catastrophic losses (α_2) parameters. The upper panel presents the decision surfaces based on comparing the present value of MF and the expected present value of SRM20, while the middle and lower panels show the decision surfaces for comparisons between SRM20 and M30, and SRM20 and M40, respectively. Figures in the left and right panels are calculated for a 15 year and 5 year window for recovering the masked warming, respectively. Positive values in the z-axis indicate regions of the decision surface in which SRM20 is preferred (units in billion dollars).

The upper panel of figure 4 shows decision surfaces in the space. When SRM governance can ensure a backstop to slow the recovery of the masked warming (15 year ramp-down; figure 4(a)), lower tipping thresholds (small c) and larger catastrophic losses (high α_2) expand the region where SRM20 dominates MF: catastrophic outcomes become more probable under MF (figure S5), so moderating the peak and rate of warming via SRM yields sizable net benefits. In contrast, if SRM is unable to manage a slow phase-out (figure 4(b)), most parameter combinations lead to large global losses, making MF preferable.

Under strong global mitigation (M30) combined with moderate SRM that only shaves off residual warming, a controlled termination delivers moderate gains across the entire $c-\alpha_2$ range (figure 4(c); US\$77–122 trillion relative to M30 alone). These gains arise because decisive mitigation plus governed SRM largely eliminates abrupt warming catastrophes. However, the result hinges on ruling out abrupt termination (figure 5(d)). The M40 and M50 scenarios ($\times$ 10^{-5}f and S6) show that delaying mitigation

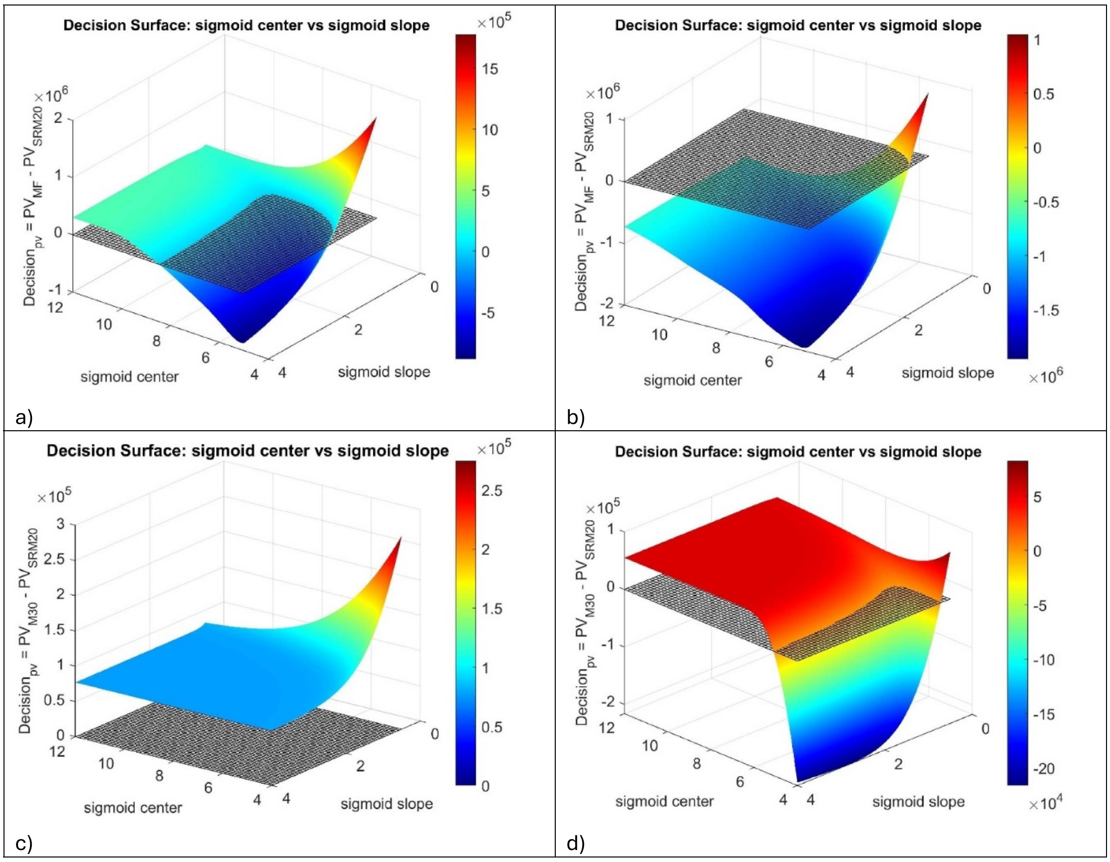

Figure 5. Decision surfaces comparing the SRM20 with the MF, and M30 scenarios a range of values of the sigmoid parameters: center (c) and slope (a). The upper panel presents the decision surfaces based on comparing the present value of MF and the expected present value of SRM20, while the lower panels show the decision surfaces for comparisons between SRM20 and M30. Figures in the left and right panels are calculated for a 15 year and 5 year window for recovering the masked warming, respectively. Positive values in the z-axis indicate regions of the decision surface in which SRM20 is preferred.

for longer periods of time after SRM deployment sharply raise SRM risk, especially when tipping can occur at low excess-rate values (low δ), even with a managed phase-out.

Figure 5 explores the joint sensitivity of the catastrophe-probability function ω to its threshold (c) and steepness (a) parameters, showing how improved knowledge of tipping-element characteristics can reshape the viability boundaries for SRM. While fixes the excess-rate at which $\omega = 50\%$, a controls how abruptly ω rises as a function of δ. The range of values chosen for this parameter is $a \in [0.5, 4]$, spanning very gradual ($a = 0.5$) to near step-like ($a = 4$) escalations. For instance, maintaining $c = 8$, would require a rate of warming 19 times the chosen safe rate to reach $\omega \approx 1$. For $a = 4$ the required rate would be only 10 times as large. Note that small values of a produce long tails, meaning non-zero catastrophe risk even at moderate rates.

Conditional on SRM governance eliminating the possibility of an abrupt termination, the SRM20 could be beneficial for most of the $c - a$ space when compared to the MF scenario, except when tipping can occur at both low rates (small c) and with a very sharp escalation (large a; figure 5(a)). In that region of values, even a gradual 15 year ramp-down can push post-termination warming across the tipping threshold, transforming SRM's benefits into severe losses. In contrast, if the increase in catastrophic damages is slow ($a < 1$), then SRM20 is preferrable as non-catastrophic damages would be larger under the FM scenario which sustains high rates of warming during the century. However, if the risk of abrupt termination exists, the vast majority of the $c - a$ space would lead to much larger losses than MF (figure 5(b)).

The combination of the M30 scenario and a moderate SRM deployment with slow phase out termination can eliminate the risks of catastrophic outcomes for the whole $c - a$ space considered in this analysis (figure 5(c)). Even in the case of possible abrupt termination (figure 5(d)), the deployment of SRM under the M30 scenario would be preferable to the GHG mitigation effort alone for wide regions of the parameter space. Nevertheless, this figure also reveals high risks when tipping points can occur at low warming levels, for which recent literature shows support (McKay *et al* 2022a, Terpstra *et al* 2025), and for almost any value of the *a* parameter. As discussed in the previous section, delaying global GHG mitigation entails higher risks from SRM deployment (figure S6).

3.5. Decision surfaces for fixed amounts of SRM deployment applied to different levels of warming

This subsection analyzes the effects of deploying various fixed levels of SRM cooling under each of the MF, M30, M40 and M50 emissions scenarios to determine in which cases SRM interventions are more likely to be risk-reducing than risk-amplifying. For the results presented below, reductions in global temperature of 0.3 °C, 1.0 °C and 2.0 °C were chosen to represent low, medium and high SRM interventions. This allows us to answer separately the following questions: (1) which baseline emissions scenario benefits most from SRM?, and (2) what would be the most risk-reducing level of SRM deployment in each emissions scenario?

Figure 6 shows a matrix of decision surfaces for the three fixed levels of SRM (columns) and the GHG emissions scenarios (rows), as a function of the probability of failure of SRM and the length of the window for recovering the masked warming. Percentages shown in each of the panels in this figure indicate the proportion of the chosen parameter space for which SRM scenario amplifies risk relative to the non-SRM scenario. These values are illustrative only and depend on the selected ranges for P_{failure} and SRM phase-out duration. Percentages are intended for comparing the relative change of the risk-amplifying region across emissions scenarios and SRM intervention levels. Comparing the values per column (figure 6) to address the first question, it can be seen that small amounts of SRM cooling may be risk-reducing in moderate emissions scenarios than in extreme ones. For example, for a 0.3 °C of SRM cooling and the M30 emissions scenario, there is only 9.7% of the combinations of parameter values (P_{failure} and SRM phase-out duration) that are risk-amplifying. However, in such instances the severity of losses can be very high. Note that catastrophic risk starts to dominate as baseline warming increases, even for small amounts of SRM deployment, and increasingly larger phase-out durations are required for SRM options to be economically viable. On the other hand, deploying the highest amount of SRM (2 °C cooling) produces decision surfaces with a majority of risk-amplifying results under any emissions scenario, ranging from 59.4% of the P_{failure} /phase-out combinations of values under the lowest GHG emissions scenario (M30) to 95.9% under the highest (MF).

The second question focuses on the tradeoff between damage reduction and catastrophic risk and helps identify the amount of SRM deployment that could lead to risk reduction for each emissions scenario. For all the emission scenarios considered, the lowest amount of SRM (0.3 °C) results in more parameter values combinations that are risk reducing than risk-amplifying. Nonetheless, the percentage of risk-increasing outcomes rapidly increases with baseline warming (up to 25.7% in the MF scenario) and the associated losses can be very large. In the case of SRM deployment for 1 °C of cooling, the M30 and M40 scenarios lead to considerably high percentages of risk-amplifying parameter combinations, but lower than 50%. For higher emissions scenarios and medium/high SRM interventions, most parameter combinations are risk amplifying. These results reinforce that the threshold of viable SRM interventions is lower when mitigation is weak.

4. Conclusions

Our results suggest that solar-radiation modification can lower expected climate damages only within a narrow intersection where three conditions coincide: (i) rapid and sustained global mitigation, (ii) a very low probability of abrupt SRM failure, and (iii) controlled, slow phase-out. Outside this joint governance space, SRM becomes a risk amplifier. Here we provide a coupled SPTP–GTP framework and a novel, rate-driven catastrophe damage function that together helped us to identify viability boundaries for SRM deployment and how its success and safe operation is intrinsically dependent on global mitigation.

We show that ignoring the potential for SPTP-GTP catastrophes component can systematically bias the assessment of the economic convenience for SRM deployment. Even with technical and operational reliability, non-zero failure probabilities arise from geopolitical shocks including nuclear wars and other global conflicts, terrorism, unforeseen side-effects of SRM, among others (Tang and Kemp 2021). These background risks put a lower bound to the probability of failure that cannot be engineered away or eliminated by SRM/GHG governance alone.

The idea and increasing relevance of SRM options in the policy and public discourse arise directly from the failure of governance of global mitigation of GHG. Paradoxically, the successful operational deployment of these options depends directly on fixing the governance problem of GHG mitigation. Both the SRM and GHG governance share similar challenges such as multilateral decision and political processes, strong institutional capacities, harmonizing diverging interests, ethical considerations, scientific uncertainty, as well as transparent monitoring and enforcing systems (UNEP 2023, Cherry *et al* 2025).

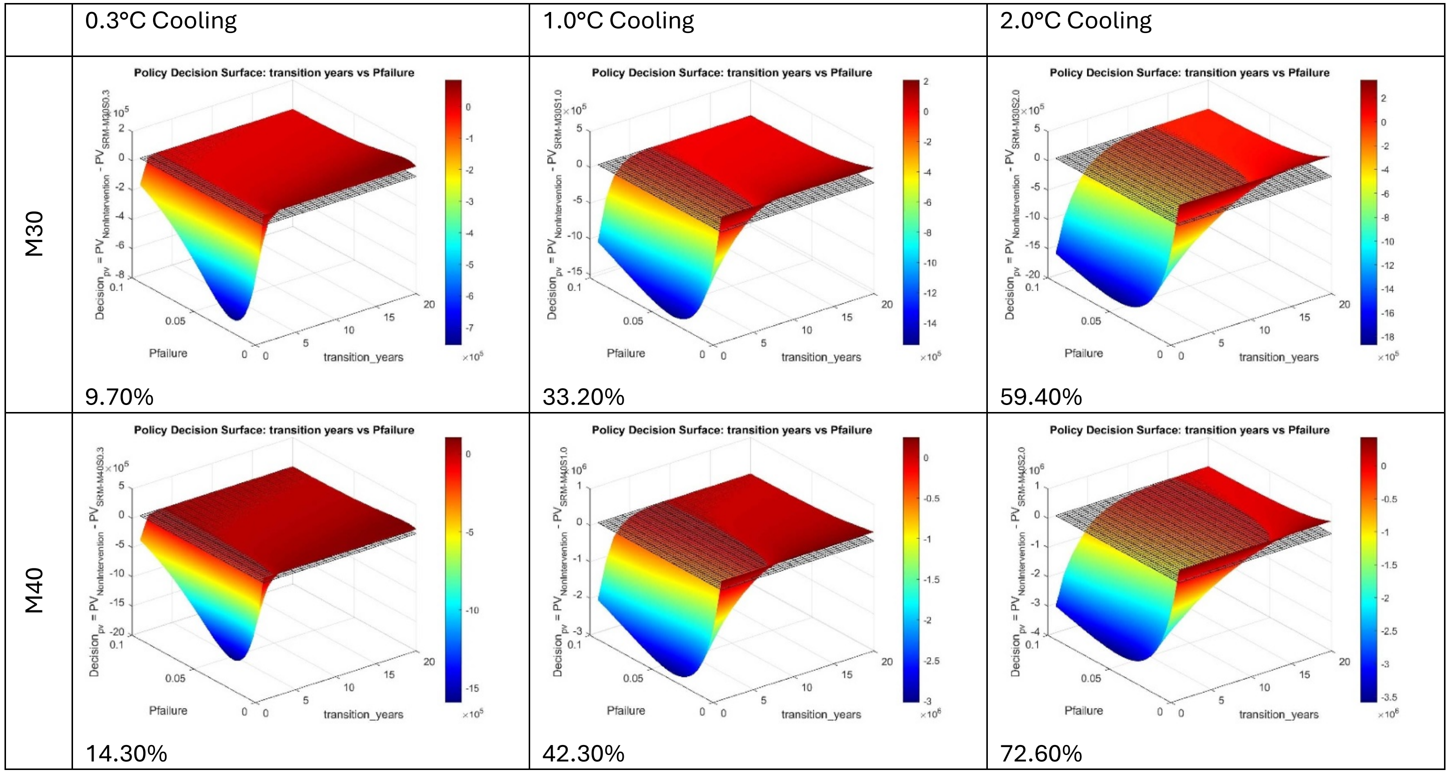

Figure 6. Decision surfaces for the probability of failure and transition years comparing fixed levels of SRM cooling for each of the four different scenarios of global GHG emissions. Columns show the decision surfaces for different levels of SRM cooling (0.3 °C, 1.0 °C, 2.0 °C) for each of the GHG emissions scenarios considered (M30, M40, M50 and MF). Rows show the decision surfaces for different GHG emissions scenarios (M30, M40, M50 and MF) and different levels of SRM cooling (0.3 °C, 1.0 °C, 2.0 °C). The percentage in each panel shows the proportion of the considered parameter space (i.e. 0%–10% for P_{failure} and 1–20 years for phase-out duration) for which SRM amplifies risk relative to the no-SRM scenarios. The units of the z axis are billion dollars.

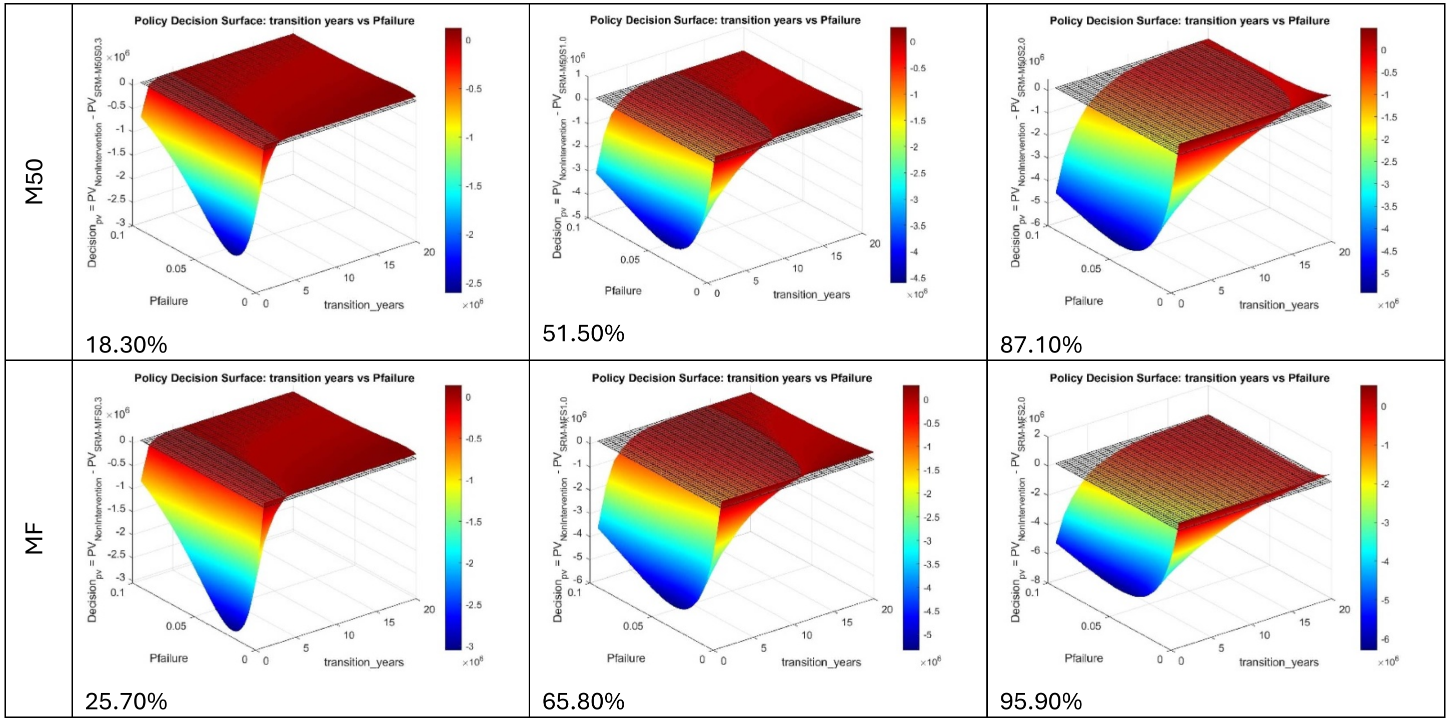

Figure 6. (Continued.)

Hence combinations like weak mitigation/strong SRM governance are internally inconsistent, whereas weak mitigation/ weak SRM governance are consistent with the past and current state of global climate policy, and define precisely the quadrant where our results show SRM intensifies risk. As such, unless the world can rapidly strengthen the very governance structures that are presently failing on emissions, SRM governance will probably replicate those failures, and could magnify global risk rather than reduce it. Strong evidence of global mitigation of GHG is a pre-requisite for SRM deployment and ensuring controlled phase-out is imperative.

Our findings have important implications for the SRM governance literature. The narrow viability boundaries we identify lend quantitative support to concerns raised by proponents of a non-use agreement, which contend that the systemic risks of SRM may be unacceptably high in a real-world context of geopolitical instability and failed cooperation (Biermann *et al* 2022). Conversely, for those who argue that governance is possible (Parson *et al* 2024), our model specifies the minimum necessary conditions for such governance: it must be robust enough to ensure near-zero failure probabilities and guarantee a slow phase-out, all while being coupled with successful global mitigation. However, this suggests that a world capable of the safe, long-term governance of SRM is likely a world that has already solved the mitigation problem, thereby obviating the need for SRM in the first place. Our results illustrate the need for SRM research and underscore that deployment should not be considered without a fundamental transformation in global climate governance. Since true climate catastrophes remain unobserved and SRM has never been field-tested at scale, studies about the potential consequences of these options are necessarily and considerably uncertain. The damage function should therefore be viewed as heuristic tools for helping thinking about SRM and tipping point dynamics, and numerical results as illustrative rather than precise projections. The qualitative insights on governance constraints and geophysical thresholds, however, are robust to wide parameter variation.

Future research should focus on investigating how the distribution of regional benefits and impacts could change with SRM, while accounting for the possibility of SPTP-GTP. Spatially explicit IAMs can take account the alteration of climate patterns caused by SRM and at the same time address intra-country differences in exposure, and even urban and rural areas (Estrada and Botzen 2021, Estrada *et al* 2025). Additional efforts should focus on the explicit modeling of SPTPs and inclusion on IAMs which could be helpful to identify conditions prone to generate them and to develop early warning signals.

Data availability statement

The data that supports the findings of this study are openly available in the supplementary files of this article.

Supplementary Figures available at https://doi.org/10.1088/2752-5295/ae33df/data1.

Acknowledgment

The authors thank Bárbara Cuesta-Castillo for her help with the literature review and preparing climate simulations.

Funding information

The authors acknowledge financial support from the Degrees Initiative under Grant SPF24MEX000.
The authors thank Bárbara Cuesta-Castillo for her help with the literature review and preparing climate simulations.

Author contributions

Francisco Estrada 0000-0002-1403-2180
Conceptualization (lead), Formal analysis (lead), Funding acquisition (lead), Investigation (lead), Methodology (lead), Project administration (lead), Resources (lead), Writing – original draft (lead), Writing – review & editing (lead)

Bernardo A Bastien-Olvera 0000-0002-3734-4267
Conceptualization (equal), Formal analysis (equal), Investigation (equal), Methodology (supporting), Writing – original draft (equal), Writing – review & editing (equal)

Oscar Calderon-Bustamante 0000-0002-0715-6871
Conceptualization (equal), Formal analysis (equal), Investigation (equal), Methodology (supporting), Project administration (supporting), Software (lead), Writing – review & editing (supporting)

Miguel A Altamirano
Conceptualization (equal), Formal analysis (equal), Investigation (equal), Methodology (supporting), Writing – review & editing (equal)

Rodrigo Muñoz-Sánchez 0000-0002-9200-0838
Conceptualization (equal), Formal analysis (equal), Investigation (equal), Methodology (supporting), Writing – review & editing (equal)

Juan Moreno-Cruz 0000-0003-3707-142X
Conceptualization (equal), Formal analysis (supporting), Investigation (equal), Methodology (supporting), Writing – review & editing (equal)

Wouter Botzen 0000-0002-8563-4963
Conceptualization (equal), Formal analysis (supporting), Investigation (equal), Methodology (supporting), Writing – review & editing (equal)

References

Aaheim A, Romstad B, Wei T, Kristjánsson J E, Muri H, Niemeier U and Schmidt H 2015 An economic evaluation of solar radiation management *Sci. Total Environ.* **532** 61–69
Aldy J E *et al* 2021 Social science research to inform solar geoengineering *Science* **374** 815–8
Anthoff D, Estrada F and Tol R S J 2016 Shutting down the thermohaline circulation *Am. Econ. Rev.* **106** 602–6
Asayama S and Hulme M 2019 Engineering climate debt: temperature overshoot and peak-shaving as risky subprime mortgage lending *Clim. Policy* **19** 937–46
Avin S, Wintle B C, Weitzdörfer J, Héigeartaigh S S Ó, Sutherland W J and Rees M J 2018 Classifying global catastrophic risks *Futures* **102** 20–26
Bathiany S, Hidding J and Scheffer M 2020 Edge detection reveals abrupt and extreme climate events *J. Clim.* **33** 6399–421
Baum C M, Fritz L, Low S and Sovacool B K 2024 Public perceptions and support of climate intervention technologies across the Global North and Global South *Nat. Commun.* **15** 1–15
Baum S, de Neufville R and Barrett A 2018 A model for the probability of nuclear war *SSRN Electron. J* (available at: https://papers.ssrn.com/abstract=3137081)
Baur S, Nauels A, Nicholls Z, Sanderson B M and Schleussner C F 2023 The deployment length of solar radiation modification: an interplay of mitigation, net-negative emissions and climate uncertainty *Earth Syst. Dyn.* **14** 367–81
Biermann F *et al* 2022 Solar geoengineering: the case for an international non-use agreement *Wiley Interdiscip. Rev. Clim. Change* **13** e754
Brent K, Simon M and McDonald J 2024 From informal to formal governance of solar radiation management *Clim. Policy* **25** 947–64
Burke M, Hsiang S M and Miguel E 2015 Global non-linear effect of temperature on economic production *Nature* **527** 235–9
Cai Y, Lenton T M and Lontzek T S 2016 Risk of multiple interacting tipping points should encourage rapid CO_2 emission reduction *Nat. Clim. Change* **6** 520–5
Carlson C J, Colwell R, Hossain M S, Rahman M M, Robock A, Ryan S J, Alam M S and Trisos C H 2022 Solar geoengineering could redistribute malaria risk in developing countries *Nat. Commun.* **13** 1–9
Cherry T L, Kallbekken S, McEvoy D M and Siu W Y 2025 The strategic and governance implications of solar radiation modification: perspectives from delegates of international climate negotiations *Environ. Res. Lett.* **20** 014053
Clark B 2023 How to argue about solar geoengineering *J. Appl. Phil.* **40** 505–20
Climate Overshoot Commission 2023 Reducing the risks of climate overshoot 126 (available at: www.overshootcommission.org/_files/ugd/0c3b70_bab3b3c1cd394745b387a594c9a68e2b.pdf)
Corry O 2017 The international politics of geoengineering: the feasibility of Plan B for tackling climate change *Secur. Dialogue* **48** 297–315
Diaz D and Moore F 2017 Quantifying the economic risks of climate change *Nat. Clim. Change* **7** 774–82
Drijfhout S, Bathiany S, Beaulieu C, Brovkin V, Claussen M, Huntingford C, Scheffer M, Sgubin G and Swingedouw D 2015 Catalogue of abrupt shifts in intergovernmental panel on climate change climate models *Proc. Natl Acad. Sci.* **112** E5777–86
Estrada F and Botzen W J W 2021 Economic impacts and risks of climate change under failure and success of the paris agreement *Ann. New York Acad. Sci.* **1504** 95–115
Estrada F, Lupi V, Botzen W J W and Tol R S J 2025 Urban and non-urban contributions to the social cost of carbon *Nat. Commun.* **16** 4193
Estrada F, Tol R S J and Gay-García C 2015 The persistence of shocks in GDP and the estimation of the potential economic costs of climate change *Environ. Model. Softw.* **69** 155–65
Fan Y, Tjiputra J, Muri H, Lombardozzi D, Park C E, Wu S and Keith D 2021 Solar geoengineering can alleviate climate change pressures on crop yields *Nat. Food* **2** 373–81
Felgenhauer T, Bala G, Borsuk M E, Camilloni I, Wiener J B and Xu J 2025 Practical paths to risk-risk analysis of solar radiation modification *Oxford Open Clim. Change* **5** kgaf012
Forster P M *et al* 2025 Indicators of global climate change 2024: annual update of key indicators of the state of the climate system and human influence *Earth Syst. Sci. Data* **17** 2641–80
Gay B A, Mandrake L, Miner K R and Miller C E 2025 Assessing the impacts of mitigation and geoengineering intervention scenarios on Earth system dynamics and climatological variability with multimodal simulations *Sci. Rep.* **15** 1–25

Gupta A and Möller I 2019 De facto governance: how authoritative assessments construct climate engineering as an object of governance *Environ. Polit.* **28** 480–501
Halstead J 2018 Stratospheric aerosol injection research and existential risk *Futures* **102** 63–77
Harding A R, Ricke K, Heyen D, MacMartin D G and Moreno-Cruz J 2020 Climate econometric models indicate solar geoengineering would reduce inter-country income inequality *Nat. Commun.* **11** 1–9
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
IPCC 2018 Impacts of 1.5 °C of global warming on natural and human systems *Glob. Warm. 1.5 °C. An IPCC Spec. Rep. impacts Glob. Warm. 1.5 °C above pre-industrial levels Relat. Glob. Greenh. gas Emiss. pathways, Context Strength. Glob. response to Threat Clim. Chang*
Irvine P J *et al* 2017 Towards a comprehensive climate impacts assessment of solar geoengineering *Earth's Future* **5** 93–106
Irvine P J, Sriver R L and Keller K 2012 Tension between reducing sea-level rise and global warming through solar-radiation management *Nat. Clim. Change* **2** 97–100
Irvine P, Emanuel K, He J, Horowitz L W, Vecchi G and Keith D 2019 Halving warming with idealized solar geoengineering moderates key climate hazards *Nat. Clim. Change* **9** 295–9
Jebari J *et al* 2021 From moral hazard to risk-response feedback *Clim. Risk Manage.* **33** 100324
Jinnah S and Dove Z 2025 Solar radiation management: a history of the governance and political milestones *Environ. Sci. Atmos.* **5** 656–73
Jones A *et al* 2013 The impact of abrupt suspension of solar radiation management (termination effect) in experiment G2 of the Geoengineering Model Intercomparison Project (GeoMIP) *J. Geophys. Res. Atmos.* **118** 9743–52
Kalkuhl M and Wenz L 2020 The impact of climate conditions on economic production. Evidence from a global panel of regions *J. Environ. Econ. Manag.* **103** 102360
Keys P W, Barnes E A, Diffenbaugh N S, Hurrell J W and Bell C M 2022 Potential for perceived failure of stratospheric aerosol injection deployment *Proc. Natl Acad. Sci. USA* **119** e2210036119
Kotz M, Levermann A and Wenz L 2024 The economic commitment of climate change *Nature* **628** 551–7
Lawrence M G, Schäfer S, Muri H, Scott V, Oschlies A, Vaughan N E, Boucher O, Schmidt H, Haywood J and Scheffran J 2018 Evaluating climate geoengineering proposals in the context of the Paris Agreement temperature goals *Nat. Commun.* **9** 3734
Lenton T M, Held H, Kriegler E, Hall J W, Lucht W, Rahmstorf S and Schellnhuber H J 2008 Tipping elements in the Earth's climate system *Proc. Natl Acad. Sci. USA* **105** 1786–93
Lenton T M, Rockström J, Gaffney O, Rahmstorf S, Richardson K, Steffen W and Schellnhuber H J 2019a Climate tipping points—too risky to bet against *Nature* **575** 592–5
Link P M and Tol R S J 2011 Estimation of the economic impact of temperature changes induced by a shutdown of the thermohaline circulation: an application of FUND *Clim. Change* **104** 287–304
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
Markusson N, McLaren D and Tyfield D 2018 Towards a cultural political economy of mitigation deterrence by negative emissions technologies (NETs) *Glob. Sustain.* **1** e10
Mastrandrea M D and Schneider S H 2001 Integrated assessment of abrupt climatic changes *Clim. Policy* **1** 433–49
Matthews H D and Caldeira K 2007 Transient climate-carbon simulations of planetary geoengineering *Proc. Natl Acad. Sci. USA* **104** 9949–54
McKay D I A, Staal A, Abrams J F, Winkelmann R, Sakschewski B, Loriani S, Fetzer I, Cornell S E, Rockström J and Lenton T M 2022a Exceeding 1.5 °C global warming could trigger multiple climate tipping points *Science* **377** eabn7950
Meinshausen M, Nicholls Z, Lewis J, Gidden M, Vogel E, Freund M and Wang R 2020 The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500 *Geosci. Model Dev.* **13** 3571–605
Meinshausen M, Raper S C B and Wigley T M L 2011 Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6—Part 1: model description and calibration *Atmos. Chem. Phys.* **11** 1417–56
Moreno-Cruz J, McEvoy D M, McGinty M and Cherry T L 2024 The Economics and Governance of Solar Geoengineering *Rev. Environ. Econ. Policy* **19** 1–24
Nicholls Z *et al* 2021 Reduced complexity model intercomparison project phase 2: synthesizing earth system knowledge for probabilistic climate projections *Earth's Future* **9** e2020EF001900
Nijsse F, Sharpe S, Sahastrabuddhe R, Lynch C, Akther A, Lam A and Lenton T 2025 How a positive tipping point cascade in power, transport and heating can accelerate the low-carbon transition (available at: www.researchsquare.com)
Nordhaus W 2019 Economics of the disintegration of the Greenland ice sheet *Proc. Natl Acad. Sci.* **116** 201814990
Office of the European Union 2024 *Solar Radiation Modification* (Publications Office of the European Union) 10.2777/391614
Parker A and Irvine P J 2018 The risk of termination shock from solar geoengineering *Earth's Futur.* **6** 456–67
Parson E A, Buck H J, Jinnah S, Moreno-Cruz J and Nicholson S 2024 Toward an evidence-informed, responsible, and inclusive debate on solar geoengineering: a response to the proposed non-use agreement *Wiley Interdiscip. Rev. Clim. Change* **15** e903
Pindyck R S 2013 Climate change policy: what do the models tell us? *J. Econ. Lit.* **51** 860–72
Rabitz F 2019 Governing the termination problem in solar radiation management *Environ. Polit.* **28** 502–22
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
Ross A and Damon Matthews H 2009 Climate engineering and the risk of rapid climate change *Environ. Res. Lett.* **4** 045103
Schimel D and Miller C 2023 Do two climate wrongs make a right? *AGU Adv.* **4** e2023AV001020
Spaiser V *et al* 2024 Negative social tipping dynamics resulting from and reinforcing Earth system destabilization *Earth Syst. Dyn.* **15** 1179–206
Stephens J C, Kashwan P, McLaren D and Surprise K 2021 The risks of solar geoengineering research *Science* **372** 1161
Stern N 2013 The structure of economic modeling of the potential impacts of climate change: grafting gross underestimation of risk onto already narrow science models *J. Econ. Lit.* **51** 838–59
Tang A and Kemp L 2021 A fate worse than warming? stratospheric aerosol injection and global catastrophic risk *Front. Clim.* **3** 720312
Terpstra S, Falkena S K J, Bastiaansen R, Bathiany S, Dijkstra H A and von der Heydt A S 2025 Assessment of abrupt shifts in CMIP6 models using edge detection *AGU Adv.* **6** e2025AV001698
Tol R S J 2024 A meta-analysis of the total economic impact of climate change *Energy Policy* **185** 113922

Trisos C H, Amatulli G, Gurevitch J, Robock A, Xia L and Zambri B 2018 Potentially dangerous consequences for biodiversity of solar geoengineering implementation and termination *Nat. Ecol. Evol.* **2** 475–82

UNEP 2023 One atmosphere: an independent expert review on solar radiation modification research and deployment (available at: https://wedocs.unep.org/20.500.11822/41903)

van der Wijst K I *et al* 2023 New damage curves and multimodel analysis suggest lower optimal temperature *Nat. Clim. Chang.* **13** 434–41 (available at: https://www.nature.com/articles/s41558-023-01636-1)

Vellinga M and Wood R A 2002 Global climatic impacts of a collapse of the atlantic thermohaline circulation *Clim. Change* **54** 251–67

Weitzman M L 2009 On modeling and interpreting the economics of catastrophic climate change *Rev. Econ. Stat.* **91** 1–19

Weitzman M L 2011 Fat-tailed uncertainty in the economics of catastrophic climate change *Review of Environmental Economics and Policy* vol 5 (The University of Chicago Press) pp 275–92

Weitzman M L 2012 GHG targets as insurance against catastrophic climate damages *J. Public Econ. Theory* **14** 221–44