Geoengineering: The world's largest control problem

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.1109/ACC.2014.6858658</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:23936193">http://nrs.harvard.edu/urn-3:HUL.InstRepos:23936193</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP</a></td>
</tr>
</tbody>
</table>
Geoengineering: the World’s largest control problem

Douglas G. MacMartin, Ben Kravitz, and David W. Keith

Abstract—Solar geoengineering (or Solar Radiation Management, SRM) refers to any intentional, large-scale manipulation of the Earth’s incoming solar radiation to offset some of the effects of anthropogenic greenhouse gases, reducing the associated risks from climate changes. Examples of such methods are injecting aerosols into the stratosphere or increasing marine cloud reflectivity, both of which would reflect some sunlight back to space. There are many serious concerns associated with any such approach, and also many challenges. One often overlooked aspect in geoengineering research is that this is a control problem, requiring (a) feedback of the observed climate state to manage significant uncertainty in both the radiative forcing and the climate’s dynamic response to this forcing, and (b) optimization of the distribution of radiative effect to minimize regional disparities as well as side-effects from the geoengineering implementation. We present recent progress on control for this challenging problem, building on [1, 2], and discuss open research gaps. This is the first time an explicit external feedback loop has been implemented in a fully coupled general circulation model of the Earth’s climate.

I. INTRODUCTION

Solar geoengineering might reduce risks from climate change [3–5], but research is required to understand the effectiveness and risks and to develop knowledge so that if it were ever deployed, it could be implemented as effectively as possible. The most frequently discussed option is injection of sulfate aerosols into the stratosphere to reflect a small fraction of incoming sunlight [4, 6]. This is motivated by the global cooling that follows large volcanic eruptions; e.g., after the Pinatubo eruption introduced more than 30 Tg of sulfate aerosols into the stratosphere in 1991, global mean temperatures temporarily decreased by roughly 0.5°C [7]. “Marine cloud brightening” [8] could similarly reflect sunlight; other options include space-based [9] (likely prohibitively expensive) and surface-based [10] methods. Approaches to remove CO₂ from the atmosphere are also generally described as “geoengineering”, but these yield only slow changes in climate and pose very different risks from solar geoengineering techniques; we do not discuss these.

Solar geoengineering could quickly reduce global mean temperatures at relatively low cost [11], but with a number of risks [12, 13]. Using geoengineering may lead to a delay in reducing CO₂ emissions; as a result, geoengineering could be required for centuries. Any technology will have some undesired side effects, e.g., stratospheric approaches will lead to changes in ozone and other chemistry [14]. Simulations with either a spatially-uniform aerosol layer or uniform solar reduction yield reasonable spatial compensation of temperature changes due to increased greenhouse gases [15, 16] but poorer compensation of precipitation; regional disparities in outcomes [17] could even lead to conflict. Further, the climate system is uncertain, requiring design for a system not fully understood, yet experimentation at a scale sufficient to measure climate response involves risks similar to deployment: if solar geoengineering is used, we must get it right the first time. These last two issues involving uncertainty and imprecise compensation can be at least partially addressed with control theory.

In addition to challenges that span physical, social, political, and ethical issues, geoengineering is a control problem that requires choosing the amount, spatial pattern, and temporal pattern of radiative forcing (or equivalently e.g. aerosol injection latitude and timing), to minimize the combined risks from both climate changes (temperature, precipitation, etc.) and from the implementation of geoengineering (less forcing results in lower risks, such as ozone depletion). These choices must be made despite significant model uncertainty, much of which is unresolved without large-scale, lengthy, and risky global experiments [18, 19].

Here we describe two key contributions. We first discuss global system dynamics and feedback algorithms to achieve a climate goal in the face of uncertainty (Section II; see also [2]). Next we introduce optimization of the spatial and seasonal pattern of forcing to minimize regional climate change or other metrics while constraining the amount of solar reduction used (Section III; see also [1]). Finally, we briefly discuss open research questions.

II. FEEDBACK

Models of the climate response to greenhouse gases (or solar reductions) do not all agree; see, e.g., Figure 1, which compares 24 fully coupled atmosphere-ocean general circulation models (AOGCMs) from the Coupled Model Intercomparison Project, phase 5 (CMIP5). Model discrepancies are even larger for regional precipitation responses [16]. Furthermore, in addition to uncertainty in the climate response, there is uncertainty in the radiative forcing from any particular geoengineering implementation (e.g., how much sunlight is reflected by stratospheric aerosols depends on particle size, which depends on coagulation; there may also be some influence on cirrus clouds that could change the net radiative effect [20]). Open-loop strategies will therefore result in significant error in meeting specific climate objectives, and feedback of the observed climate state must be an essential element of any solar geoengineering strategy [30].

D. G. MacMartin is in the Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125
Email: macmardg@cds.caltech.edu
B. Kravitz is in the Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, WA 99352
D. W. Keith is in the School of Engineering and Applied Sciences and Kennedy School of Government, Harvard University, Cambridge, MA 02139
A. Dynamic Model

A box diffusion model represents the dynamic behavior of global mean temperature well enough to design the control algorithm. This model consists of heat diffusion into an effectively semi-infinite reservoir (deep ocean), with a surface layer of fixed heat capacity (land, atmosphere, and surface ocean). A semi-infinite diffusion model matches the global mean temperature response in the HadCM3L GCM over a wide range of frequencies [21]; and also fits the transient response of most CMIP5 models [22]. Including a surface layer better predicts the short time-scale response. For radiative forcing $F(t)$, the surface temperature $T(t)$ and deep ocean temperature $T_d(z,t)$ satisfy

$$C \frac{dT}{dt} = F - \lambda T + \beta \frac{\partial T_d}{\partial z} \bigg|_{z=0}$$

$$\frac{\partial T_d}{\partial t} = \kappa \frac{\partial^2 T_d}{\partial z^2},$$

with boundary condition $T_d(0,t) = T(0,t)$ (taking the top of the deep ocean as $z = 0$). The parameter $\lambda$ describes the natural climate feedback (the change in radiation due to a change in surface temperature), $C = \rho H$ is the surface layer heat capacity per unit area, $\kappa$ the thermal diffusivity, and $\beta = c\rho c$ for density $\rho$ and specific heat capacity $c$.

Taking Laplace transforms [2,23,24], the transfer function from radiative forcing perturbation $F(s)$ to temperature anomaly $T(s)$ is

$$G(s) = \frac{T(s)}{F(s)} = \frac{1}{\lambda + \beta(s/\kappa)^{1/2} + Cs}$$

Equation (3) is compared with the calculated HadCM3L frequency response in Fig. 2; the latter was computed by introducing 1% sinusoidal variations in solar forcing [21].

In this model, a 2.3% solar reduction would offset the 3.7 W m$^{-2}$ radiative forcing from a doubling of atmospheric CO$_2$ [1]. Using this to convert solar reduction into radiative forcing, the best fit to the calculated frequency response gives $\lambda = 1.2$ W m$^{-2}$ K$^{-1}$, $\tau = \beta^2/(\lambda^2 \kappa) = 13$ years, and $C = 3.2 \times 10^9$ J m$^{-2}$ K$^{-1}$.

The step response in Fig. 1 illustrates the robustness challenge. The steady-state response varies by more than a factor of two, and the dynamics also change. Some models are better fit with a much larger heat capacity [25], which would increase the high frequency phase lag in Fig. 2.

B. Feedback Overview

The block diagram in Fig. 3 illustrates the coupled human-climate feedback system. The dynamic system characterized by $G(s)$ describes the global mean temperature response to imposed radiative forcing $F$, including forcing associated with anthropogenic climate change $F_d$, solar geoengineering $F_r$, and perturbations $w$ responsible for natural variability (which has a $1/f$ power spectrum, so $w$ is approximately white). The solar geoengineering forcing includes the best estimate $F$ of the radiative forcing required to maintain $T = T_{ref}$ in the presence of $F_d$, as well as the component $F_r$ that corrects for errors in this estimate based on feedback of the observed climate state.

If the dynamics of $G(s)$ and radiative forcing $F_d$ were known, then $F$ could be chosen so that $T$ only differs from $T_{ref}$ by natural variability; in reality, using the best estimate of $F$ will yield some error that could be corrected with feedback. We define $F_r = F_d + F$ as the residual radiative forcing. The temperature error relative to the desired temperature, $T_e = T - T_{ref}$, can then be expressed as

$$T_e = \frac{G(s)}{1 + G(s)K(s)}(F_r + w) = G_{fb}(s)(F_r + w)$$

where $G_{fb}(s) = G(s)S(s)$. The static equivalent is well known in climate literature [26, 27]. However, recognizing that this predicts the transient as well as steady-state response is novel in climate analysis, as is the possibility for the feedback $K(s)$ to be chosen as opposed to being a property
of the climate system. Furthermore, the waterbed effect
\[ \int_0^\infty \log |S(i\omega)|d\omega = 0 \]  
(5)
is important in understanding the dynamic behavior resulting from solar geoengineering. The feedback acts on temperature errors due to both greenhouse gases and natural variability. Frequencies where \(|S(i\omega)| > 1\) correspond to amplification of natural climate variability, which can be particularly significant if time delay is introduced.

From Fig. 2, PI control \(K(s) = k_p + k_i/s\) is sufficient for this application (see the Bode plot in Figure 4). This yields
\[ G_{fb}(s) = \frac{s}{k_i + s(\lambda + k_p + \beta(s/k)^{1/2} + Cs)} \]  
(6)
and zero steady-state error; proportional gain alone is equivalent to a change in intrinsic climate feedback parameter \(\lambda\).

C. Feedback in a GCM

The results in Figure 5, based on simulations in [2], represent the first ever explicit feedback loop implemented in a fully-coupled AOGCM of the Earth’s climate. We use HadCM3L, which has resolution of 3.75° in longitude by 2.5° in latitude in both the atmosphere and ocean, with 19 vertical levels in the atmosphere and 20 in the ocean [29].

There are three take-away messages from Fig. 5. Most important, though unsurprising, is that properly designed feedback of the “observed” temperature can maintain the global mean temperature at a desired target value, despite uncertainty in both the model and the forcing. Second (also unsurprising to anyone experienced in designing controllers) is the importance of proper representation of system dynamics. A poorly designed controller can result in near-oscillatory behavior, illustrated here by increasing the time delay, resulting in poor phase margin (see Fig. 4). The third important result is that the dynamic behavior resulting from implementing feedback in a complex climate model is well predicted from the simple box-diffusion model. This is illustrated by comparing the ratio of the closed-loop and open-loop amplitude spectra with the predicted sensitivity function.

While an important first step, there are several details yet to be filled in. First, the control gains were tuned using the frequency response of the same model that was used to evaluate performance. This is clearly not an option if geoengineering is ever implemented in the real world. A second climate model could serve as a proxy for the unknown real-world dynamics, to verify that a controller designed on one model is robust to the inter-model uncertainty [30]. Second, a PI controller illustrates the potential and also some dynamic characteristics associated with implementing feedback, but more complicated controllers may be appropriate. Adaptation could be used to better learn the dynamics as time progresses [31], or model-predictive control to enforce constraints on the amount and rate of climate change and on the rate of change of applied forcing. Finally, while many climate impacts can be related to global mean temperature, a more realistic goal would consider other variables such as precipitation or sea ice, as well as regional effects. Including multiple objectives will require managing more than a single degree of freedom. The next section introduces this as an optimization problem; future work will integrate feedback into the multiple degree of freedom optimization problem to manage uncertainty.

III. OPTIMIZATION

The second application of control theory to solar geoengineering we discuss is optimization to (a) improve the compensation between the climate effects of greenhouse gases and those due to solar geoengineering, (b) maximize the benefit with minimum solar reduction, and (c) consider trade-offs between different metrics. Some of this appears in [1].

The radiative forcing due to greenhouse gases is relatively
uniform in space and time (latitude and season), while solar geoengineering is most effective in summer in the respective hemisphere [15]. Furthermore, the different mechanisms of radiative forcing have different relative impacts on temperature vs precipitation [32]. As a result, spatially and seasonally uniform solar reductions do not fully compensate climate changes due to greenhouse gases in any model [16]. The compensation can be improved by optimizing the distribution of solar reduction. This might be achieved, for example, by varying the latitude and time of year of stratospheric aerosol injection. Here we explore the potential benefit for the solar reduction patterns in Fig. 6. In practice, neither these patterns nor the resulting climate response will be known, and feedback would be needed to manage this uncertainty.

We again use HadCM3L, and compute the climate response for each spatial and temporal pattern in Fig. 6. Assuming linear superposition to estimate the response to any combination of these patterns (eq. (7)) is a good approximation [33, 34]. A pre-industrial CO$_2$ (278 ppm) simulation defines the “baseline” or desired climate, and a 2×CO$_2$ case used to define the perturbed climate that geoengineering is intended to compensate. For each applied radiative forcing pattern, we compute the monthly means over the second half of a 200-year simulation to minimize errors in estimating the long time-scale behavior.

The monthly-mean temperature and precipitation differences between the 2×CO$_2$ and baseline cases for all $m$ grid cells can be concatenated into a vector $b \in \mathbb{R}^{24m}$. Temperatures and precipitations are normalized by the standard deviation of their interannual natural variability [17, 34]; this measures change compared to “normal” variability. Similarly, the response to the $n$ patterns of solar reduction can be described by $A \in \mathbb{R}^{24m \times n}$. Define $u \in \mathbb{R}^n$ as the forcing amplitude corresponding to each pattern. Assuming linearity, the residual response is then

$$z = b - Au$$

Minimizing a quadratic function of climate changes relative to a pre-industrial baseline while constraining the average insolation reduction yields a constrained least-squares problem that can be solved using quadratic programming:

$$J^* = \min_u \left\{ u^T (A^T A) u + 2 b^T A u \right\} \quad \text{with } Cu \leq 0 \quad (8)$$

In addition to the constraint on average solar reduction, the constraint that we can only reduce sunlight, not increase it, can also be written as $Cu \leq 0$, with each row of $C$ enforcing this at a particular latitude and time of year. (It is sufficient to enforce this at only a few times and latitudes.) It is also straightforward to include in the same framework variables one might wish to maximize, such as Arctic sea ice extent.

We also consider the forcing pattern that minimizes the worst-case climate change in any grid-cell:

$$J^* = \min_u \max_i |(b - Au)_i| \quad \text{with } Cu \leq 0 \quad (9)$$

where $(\cdot)_i$ is the $i$th element of the vector. Figure 8 is obtained by optimizing as in (8) with constraints $Au \leq \alpha - b$ and $-Au \leq \alpha + b$, where $\alpha$ is the allowable worst-case.

Allowing the distribution of solar reduction to vary in space and time improves the compensation of climate change...
due to greenhouse gases in this model, and also allows different trade-offs to be made depending on the weighting of different objectives in the optimization. Results in [1] illustrate several trade-offs, including (a) the ability to reduce the worst-case climate change over any region while still reducing the global rms climate change: this is relevant for minimizing regional disparities, (b) the trade-off between minimizing global rms temperature changes and global rms precipitation changes, and (c) the trade-off between minimizing global rms temperature and precipitation changes and maximizing Arctic sea ice. The last of these is reproduced in Fig. 7 for illustration. The trade-off is shown as a function of the average solar reduction, a proxy for the negative impacts associated with implementing geoengineering (e.g., decreased ozone from stratospheric aerosols). For example, restoring pre-industrial sea ice in this model requires ~2.5% reduction in sunlight if the reduction is spatially and seasonally uniform, but the same sea ice extent can be obtained with only 0.5% solar reduction if solar changes are concentrated optimally in space and time.

Here we further evaluate whether similar performance improvement can be obtained with only spatially uniform (but seasonally varying) or seasonally uniform (but spatially varying) solar reduction. Figure 8 demonstrates that while some improvement in worst-case climate changes are possible with only seasonal variation, the biggest benefit comes from allowing the solar reduction to vary as a function of latitude; introducing both provides only minor further improvements. For different objectives, different degrees of freedom are important to include. The sea-ice improvement of Fig. 7 requires both spatial and seasonal variations in solar reduction, although some benefit is again obtained with only spatial (concentrated in the Arctic).

Fig. 7. Trade-off between maximizing Arctic sea-ice extent and minimizing global-rms temperature and precipitation changes (from [1]) if the solar reduction is optimized in space and time. The trade-off is shown for increasing uniform solar reduction (solid black line) and non-uniform reduction weighting only sea ice (upper dashed line), only temperature and precipitation (lower dashed line), or some combination (connecting lines, at constant average solar reduction). Non-uniform solar reduction improves outcomes (as measured here), and can also result in climates that are not achievable with only uniform reductions.

Fig. 8. Optimization of different objectives, evaluating the benefit of spatially- or seasonally-varying solar reductions. The top panel considers the trade-off between minimizing the global rms departure from pre-industrial temperature and precipitation and the regional worst-case, addressing regional disparities. The lower panel illustrates the potential increase in sea ice as a function of average solar reduction; for the case with all degrees of freedom, this corresponds to the upper dashed lines in Fig. 7.

IV. CHALLENGES AND OPEN QUESTIONS

Several open issues have been noted, including using feedback (Sec. II) to manage uncertainty with multiple degrees of freedom (Sec. III). As a concrete example, an initial target for geoengineering might be the Arctic [35], to reduce risks associated with permafrost thaw, sea ice loss, or Greenland ice-sheet melt. Implementing such a scenario without adverse tropical precipitation effects [36] would require a counterbalancing southern hemisphere solar reduction. This would almost certainly require feedback to manage uncertainty. Also noted is the need for more complex control algorithms incorporating learning, and constraints both on climate variables and to ensure implementation is feasible. Controls engineers may also be useful in designing tests to maximize signal-to-noise ratio for system identification [19].

Climate changes occur over decadal and longer timescales. While stratospheric aerosol concentrations could not be rapidly modulated, marine cloud brightening could be modulated at time scales of order one week, opening up the possibility of responding at the time-scale of hurricane evolution, for example [37]. (Hurricane control has been suggested earlier, but without a realistic mechanism for excitation [38]). Spatial and temporal control of solar reduction could also be used to alter the dynamics of El Niño [39]. Averaging over space and time tends to reduce model uncertainty, and thus attempts to control local dynamics at these shorter timescales will be challenging.
More broadly, introducing a feedback loop around a climate model as in Sec. II is relevant beyond geoengineering for understanding human-climate interactions (e.g., incorporating gradual learning to change predicted optimal pathways for reducing emissions), understanding natural climate feedbacks (e.g., permafrost thaw leading to CO₂ and methane release leading to further warming), or better understanding of model behavior (e.g., automated model tuning).

V. SUMMARY COMMENTS

For a given level of climate change, there is clearly less climate risk if this is reached through reduced greenhouse gas emissions rather than geoengineering. However, because of technical and societal inertia, atmospheric CO₂ concentrations are likely to continue to increase, and for a given atmospheric concentration, there may be less risk if some amount of geoengineering is used rather than none. While there are many concerns associated with geoengineering, it is essential to research the problem to know what can and cannot be achieved, and to be prepared to implement it intelligently rather than in a crisis. Control and dynamics are essential elements of geoengineering research, indeed, this is the largest control problem ever considered!

ACKNOWLEDGMENTS

BK is supported by the Fund for Innovative Climate and Energy Research (FICER). The Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute under contract DE-AC05-75RL01830. DGM thanks Carnegie Institution of Science and Ken Caldeira in particular for hospitality and computing resources used in this work.

REFERENCES