An overview of geoengineering of climate using stratospheric sulphate aerosols

Philip J Rasch, Simone Tilmes, Richard P Turco, Alan Robock, Luke Oman, Chih-Chieh (Jack) Chen, Georgiy L Stenchikov and Rolando R Garcia

*Phil. Trans. R. Soc. A* 2008 **366**, 4007-4037

References

This article cites 60 articles, 6 of which can be accessed free
http://rsta.royalsocietypublishing.org/content/366/1882/4007.full.html#ref-list-1

Article cited in:
http://rsta.royalsocietypublishing.org/content/366/1882/4007.full.html#related-urls

Subject collections

Articles on similar topics can be found in the following collections

**climatology** (95 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](http://rsta.royalsocietypublishing.org/subscriptions)

To subscribe to *Phil. Trans. R. Soc. A* go to:
http://rsta.royalsocietypublishing.org/subscriptions
We provide an overview of geoengineering by stratospheric sulphate aerosols. The state of understanding about this topic as of early 2008 is reviewed, summarizing the past 30 years of work in the area, highlighting some very recent studies using climate models, and discussing methods used to deliver sulphur species to the stratosphere. The studies reviewed here suggest that sulphate aerosols can counteract the globally averaged temperature increase associated with increasing greenhouse gases, and reduce changes to some other components of the Earth system. There are likely to be remaining regional climate changes after geoengineering, with some regions experiencing significant changes in temperature or precipitation. The aerosols also serve as surfaces for heterogeneous chemistry resulting in increased ozone depletion. The delivery of sulphur species to the stratosphere in a way that will produce particles of the right size is shown to be a complex and potentially very difficult task. Two simple delivery scenarios are explored, but similar exercises will be needed for other suggested delivery mechanisms. While the introduction of the geoengineering source of sulphate aerosol will perturb the sulphur cycle of the stratosphere significantly, it is a small perturbation to the total (stratosphere and troposphere) sulphur cycle. The geoengineering source would thus be a small contributor to the total global source of ‘acid rain’ that could be compensated for through improved pollution control of anthropogenic tropospheric sources. Some areas of research remain unexplored. Although ozone may be depleted, with a consequent increase to solar ultraviolet-B (UVB) energy reaching the surface and a potential impact on health and biological populations, the aerosols will also scatter and attenuate this part of the energy spectrum, and this may compensate the UVB enhancement associated with ozone depletion. The aerosol will also change the ratio of diffuse to direct energy reaching

* Author for correspondence (pjr@ucar.edu).

One contribution of 12 to a Theme Issue ‘Geoscale engineering to avert dangerous climate change’.
the surface, and this may influence ecosystems. The impact of geoengineering on these components of the Earth system has not yet been studied. Representations for the formation, evolution and removal of aerosol and distribution of particle size are still very crude, and more work will be needed to gain confidence in our understanding of the deliberate production of this class of aerosols and their role in the climate system.

**Keywords:** climate change; geoengineering; sulphate aerosols; global warming

### 1. Introduction

The concept of ‘geoengineering’ (the deliberate change of the Earth’s climate by mankind; Keith 2000) has been considered at least as far back as the 1830s with J. P. Espy’s suggestion (Fleming 1990) of lighting huge fires that would stimulate convective updrafts and change rain intensity and frequency of occurrence. Geoengineering has been considered for many reasons since then, ranging from making polar latitudes habitable to changing precipitation patterns.

There is increasing concern by scientists and society in general that energy system transformation is proceeding too slowly to avoid the risk of dangerous climate change from humankind’s release of radiatively important atmospheric constituents (particularly CO₂). The assessment by the Intergovernmental Panel on Climate Change (IPCC 2007a) shows that unambiguous indicators of human-induced climate change are increasingly evident, and there has been little societal response to the scientific consensus that reductions must take place soon to avoid large and undesirable impacts.

To reduce carbon dioxide emissions soon enough to avoid large and undesirable impacts requires a near-term revolutionary transformation of energy and transportation systems throughout the world (Hoffert et al. 1998). The size of the transformation, the lack of effective societal response and the inertia to changing our energy infrastructure motivate the exploration of other strategies to mitigate some of the planetary warming. For this reason, geoengineering for the purpose of cooling the planet is receiving increasing attention. A broad overview to geoengineering can be found in the reviews of Keith (2000), WRMSR (2007), and the papers in this volume. The geoengineering paradigm is not without its own perils (Robock 2008). Some of the uncertainties and consequences of geoengineering by stratospheric aerosols are discussed in this paper.

This study describes an approach to cooling the planet, which goes back to the mid-1970s, when Budyko (1974) suggested that, if global warming ever became a serious threat, we could counter it with airplane flights in the stratosphere, burning sulphur to make aerosols that would reflect sunlight away. The aerosols would increase the planetary albedo and cool the planet, ameliorating some (but as discussed below, not all) of the effects of increasing CO₂ concentrations. The aerosols are chosen/designated to reside in the stratosphere because it is remote, and they will have a much longer residence time than tropospheric aerosols that are rapidly scavenged. The longer lifetime means that a few aerosols need be delivered per unit time to achieve a given aerosol burden, and that the aerosols will disperse and act to force the climate system over a larger area.

*Phil. Trans. R. Soc. A* (2008)
Sulphate aerosols are always found in the stratosphere. Low background concentrations arise due to transport from the troposphere of natural and anthropogenic sulphur-bearing compounds. Occasionally much higher concentrations arise from volcanic eruptions, resulting in a temporary cooling of the Earth system (Robock 2000), which disappears as the aerosol is flushed from the atmosphere. The volcanic injection of sulphate aerosol thus serves as a natural analogue to the geoengineering aerosol. The analogy is not perfect because the volcanic aerosol is flushed within a few years, and the climate system does not respond in the same way as it would if the particles were continually replenished, as they would be in a geoengineering effort. Perturbations to the system that might become evident with constant forcing disappear as the forcing disappears.

This study reviews the state of understanding about geoengineering by sulphate aerosols as of early 2008. We review the published literature, introduce some new material and summarize some very recent results that are presented in detail in the submitted articles at the time of the writing of this paper. In our summary we also try to identify areas where more research is needed.

Since the paper by Budyko (1974), the ideas generated there have received occasional attention in discussions about geoengineering (e.g. NAS92 1992; Turco 1995; Govindasamy & Caldeira 2000, 2003; Govindasamy et al. 2002; Crutzen 2006; Wigley 2006; Matthews & Caldeira 2007).

There are also legal, moral, ethical, financial and international political issues associated with a manipulation of our environment. Commentaries (Bengtsson 2006; Cicerone 2006; Kiehl 2006; Lawrence 2006; MacCracken 2006) to Crutzen

Figure 1. A schematic of the processes that influence the life cycle of stratospheric aerosols (adapted with permission from SPARC 2006).
(2006) address some of these issues and remind us that this approach does not treat all the consequences of higher CO₂ concentrations (such as ocean acidification; others are discussed in Robock 2008). Recently, climate modellers have begun efforts to provide more quantitative assessments of the complexities of geoengineering by sulphate aerosols and the consequences to the climate system (Rasch et al. 2008; Tilmes et al. 2008, submitted; Robock et al. 2008).

2. An overview of stratospheric aerosols in the Earth system

(a) General considerations

Sulphate aerosols are an important component of the Earth system in the troposphere and stratosphere. Because sulphate aerosols play a critical role in the chemistry of the lower stratosphere and occasionally, following a volcanic eruption, in the radiative budget of the Earth by reducing the incoming solar energy reaching the Earth surface, they have been studied for many years. A comprehensive discussion of the processes that govern the stratospheric sulphur cycle can be found in the recent assessment of stratosphere aerosols (SPARC 2006). Figure 1, taken from that report, indicates some of the processes that are important in that region.

Sulphate aerosols play additional roles in the troposphere (IPCC (2007a) and references therein). As in the stratosphere they act to reflect incoming solar energy (the ‘aerosol direct effect’), but also act as cloud condensation nuclei, influencing the size of cloud droplets and the persistence or lifetime of clouds (the ‘aerosol indirect effect’) and thus the reflectivity of clouds.

Although our focus is on stratospheric aerosols, one cannot ignore the troposphere, and so we include a brief discussion of some aspects of the tropospheric sulphur cycle also. A very rough budget describing the sources, sinks and transformation pathways¹ during volcanically quiescent times is displayed in figure 2. Sources, sinks and burdens for sulphur species are much larger in the troposphere than in the stratosphere. The sources of the aerosol precursors are natural and anthropogenic sulphur-bearing reduced gases (DMS, dimethyl sulphide; SO₂, sulphur dioxide; H₂S, hydrogen sulphide; OCS, carbonyl sulphide). These precursor gases are gradually oxidized (through both gaseous and aqueous reactions) to end products involving the sulphate anion (SO₄²⁻) in combination with various other cations. In the troposphere where there is sufficient ammonia, most of the aerosols exist in the form of mixtures of ammonium sulphate ((NH₄)₂SO₄) and bisulphate ((NH₄)HSO₄).

The stratospheric sulphur-bearing gases oxidize (primarily via reactions with the OH radical) to SO₂, which is then further oxidized to gaseous H₂SO₄. Stratospheric sulphate aerosols exist in the form of mixtures of condensed sulphuric acid (H₂SO₄), water and, under some circumstances, hydrates with nitric acid (HNO₃).

¹ Sulphur emissions and burdens are frequently expressed in differing units. They are sometimes specified with respect to their molecular weight. Elsewhere they are specified according to the equivalent weight of sulphur. They may be readily converted by multiplying by the ratio of molecular weights of the species of interest. We use only units of S in this paper, and have converted all references in other papers to these units. Also, in the stratosphere, we have assumed that the sulphate binds with water in a ratio of 75/25 H₂SO₄/water to form particles. Hence

\[ 3 \text{Tg SO}_4^{2-} = 2 \text{Tg SO}_2 = 1 \text{Tg S} \approx 4 \text{Tg aerosol particles}. \]
Although the OCS source is relatively small compared with other species, owing to its relative stability, it is the dominant sulphur-bearing species in the atmosphere. Oxidation of OCS is a relatively small contributor to the radiatively active sulphate aerosol in the troposphere, but it plays a larger role in the stratosphere where it contributes perhaps half the sulphur during volcanically quiescent conditions. Some sulphur also enters the stratosphere as SO$_2$ and as sulphate aerosol particles. The reduced sulphur species oxidize there and form sulphuric acid gas. The H$_2$SO$_4$ vapour partial pressure in the stratosphere—almost always determined by photochemical reactions—is generally supersaturated, and typically highly supersaturated, over its binary H$_2$O–H$_2$SO$_4$ solution droplets. The particles form and grow through vapour deposition, depending on the ambient temperature and concentrations of H$_2$O and H$_2$SO$_4$. These aerosol particles are then transported by winds (as are their precursors). Above the lower stratosphere, the particles can evaporate, and in the gaseous form the sulphuric acid can be photolysed to SO$_2$, where it can be transported as a gas, and may again oxidize and condense in some other part of the stratosphere. Vapour deposition is the main growth mechanism in the ambient stratosphere, and in volcanic clouds, over time.

Because sources and sinks of aerosols are so much stronger in the troposphere, the lifetime of sulphate aerosol particles in the troposphere is a few days, while that of stratospheric aerosol is a year or so. This explains the relatively smooth spatial distribution of sulphate aerosol and resultant aerosol forcing in the stratosphere, and much smaller spatial scales associated with tropospheric aerosol.

The net source$^1$ of sulphur to the stratosphere is believed to be of the order of 0.1 Tg S yr$^{-1}$ during volcanically quiescent conditions. A volcanic eruption completely alters the balance of terms in the stratosphere. For example, the eruption

---

Figure 2. A very rough budget (approx. 1 digit of accuracy) for most of the major atmospheric sulphur species during volcanically quiescent situations, following Rasch et al. (2000), SPARC (2006) and Montzka et al. (2007). Numbers inside boxes indicate species burden in units of Tg S, and approximate lifetime against the strongest source or sink. Numbers beside arrows indicate net source or sinks (transformation, transport, emissions, and deposition processes) in Tg S yr$^{-1}$.

Phil. Trans. R. Soc. A (2008)
of Mount Pinatubo is believed to have injected approximately 10 Tg S (in the form of SO₂) over a few days. This injection amount provides a source approximately 100 times that of all other sources over the year. The partial pressure of sulphuric acid gas consequently reaches much higher levels than those during background conditions. After an eruption, new particles are nucleated only in the densest parts of eruption clouds. These rapidly coagulate and disperse to concentration levels that do not aggregate significantly. Particle aggregation is controlled by Brownian coagulation (except perhaps under very high sulphur loadings). Coagulation mainly limits the number of particles, rather than the overall size of the particles, which depends more on the sulphur source strength (although considering the overall sulphur mass balance, the two processes both contribute). The particles’ growth is thus influenced by both vapour deposition and proximity to other particles.

The primary loss mechanism for sulphur species from the stratosphere is believed to be the sedimentation of the aerosol particles. Particle sedimentation is governed by Stokes’ equation for drag corrected to compensate for the fact that in the stratosphere at higher altitudes the mean free path between air molecules can far exceed the particle size, and particles fall more rapidly than they would otherwise. The aerosol particles settle out (larger particles settle faster), gradually entering the troposphere, where they are lost via wet and dry deposition processes.

Examples of the nonlinear relationships between SO₂ mass injection, particle size and visible optical depth as a function of time assuming idealized dispersion can be found in Pinto et al. (1998). These are detailed microphysical simulations, although in a one-dimensional model with specified dispersion. The rate of dilution of injected SO₂ is critical owing to the highly nonlinear response of particle growth and sedimentation rates within expanding plumes; particles have to be only 10 μm or less to fall rapidly, which greatly restricts the total suspended mass, optical depth and infrared effect. The mass limitation indicates that 10 times the mass injection (of say Pinatubo) might result in only a modestly larger visible optical depth after some months.

The life cycle of these particles is thus controlled by a complex interplay between meteorological fields (like wind, humidity and temperature), the local concentrations of the gaseous sulphur species, the concentration of the particles themselves and the size distribution of the particles.

In the volcanically quiescent conditions (often called background conditions), partial pressures of sulphur gases remain relatively low, and the particles are found to be quite small (Bauman et al. 2003), with a typical size distribution that can be described with a lognormal distribution with a dry mode radius, standard deviation and effective radius of 0.05/2.03/0.17 μm, respectively. After volcanic eruptions when sulphur species concentrations get much higher, the particles grow much larger (Stenchikov et al. 1998). Rasch et al. (2008) used numbers for a size distribution 6–12 months after an eruption for the large volcanic-like distribution of 0.376/1.25/0.43 μm following Stenchikov et al. (1998) and Collins et al. (2004). There is uncertainty in the estimates of these size distributions, and volcanic aerosol standard distribution σLN was estimated to range from 1.3 to greater than 2 in Steele & Turco (1997).

When the particles are small, they primarily scatter in the solar part of the energy spectrum, and play no role in influencing the infrared (long-wave) part of the energy spectrum. Larger particles seen after an eruption scatter and absorb in the solar wavelengths, but also absorb in the infrared (Stenchikov et al. 1998).
Thus small particles tend to scatter solar energy back to space. Large particles scatter less efficiently, and also trap some of the outgoing energy in the infrared. The size of the aerosol thus has a strong influence on the climate.

(b) Geoengineering considerations

To increase the mass and number of sulphate aerosols in the stratosphere, a new source must be introduced. Using Pinatubo as an analogue, Crutzen (2006) estimated a source of $5 \text{Tg S yr}^{-1}$ would be sufficient to balance the warming associated with a doubling of CO$_2$. Wigley (2006) used an energy balance model to conclude that approximately $5 \text{Tg S yr}^{-1}$ in combination with emission mitigation would suffice. These studies assumed that the long-term response of the climate system to a more gradual injection would be similar to the transient response to a Pinatubo-like transient injection. A more realistic exploration can be made in a climate system model (see §2d).

Rasch et al. (2008) used a coupled climate system model to show that the amount of aerosol required to balance the warming is sensitive to particle size, and that nonlinearities in the climate system mattered. Their model suggested that $1.5 \text{Tg S yr}^{-1}$ might suffice to balance the GHG warming, if the particles looked like those during background conditions (unlikely, as will be seen in §2c), and perhaps twice that would be required if the particles looked more like volcanic aerosols. Robock et al. (2008) used $1.5–5 \text{Tg S yr}^{-1}$ in a similar study, assuming larger particle sizes (which, as will be seen in §2c, is probably more realistic). They explored the consequences of injections in polar regions (where the aerosol would be more rapidly flushed from the stratosphere) and tropical injections.

All of these studies suggest that a source 15–30 times that of the current non-volcanic sources of sulphur to the stratosphere would be needed to balance warming associated with a doubling of CO$_2$. It is important to note that in spite of this very large perturbation to the stratospheric sulphur budget, it is a rather small perturbation to the total sulphur budget of the atmosphere. This suggests that the enhanced surface deposition (as for example ‘acid rain’) from a stratospheric geoengineering aerosol would be small compared with that arising from tropospheric sources globally, although it could be important if it occurred in a region that normally experienced little deposition from other sources.

There are competing issues in identifying the optimal way to produce a geoengineering aerosol. Since ambient aerosol can be a primary scavenger of new particles and vapours, their very presence limits new particle formation. When the stratosphere is relatively clean, the H$_2$SO$_4$ supersaturation can build up, and nucleation of new particles over time occurs more easily, with less scavenging of the new particles. Thus, the engineered layer itself becomes a limiting factor in the ongoing production of optically efficient aerosols.

Many of the earlier papers on geoengineering with stratospheric aerosols have listed delivery systems that release sulphur in very concentrated regions, using artillery shells, high flying jets, balloons, etc. These will release the sulphur in relatively small volumes of air. Partial pressures of sulphuric acid gas will get quite high, with consequences to particle growth and lifetime of the aerosols (see §2c for more detail).
An alternative would be to use a precursor gas that is quite long-lived in the troposphere but oxidizes in the stratosphere and then allow the Earth’s natural transport mechanisms to deliver that gas to the stratosphere, and diffuse it prior to oxidation. OCS might serve as a natural analogue to such a gas (Turco et al. 1980), although it is carcinogenic and a greenhouse gas.

Current sources of OCS are $\lesssim 1–2 \text{Tg S yr}^{-1}$ (Montzka et al. 2007). Perhaps 15 per cent of that is estimated to be of anthropogenic origin. Only approximately 0.03–0.05 \text{Tg S yr}^{-1} is estimated to reach the tropopause and enter the stratosphere (figure 2 and SPARC 2006). Residence times in the troposphere are estimated to be approximately 1–3 years, and much longer (3–10 years) in the stratosphere. Turco et al. (1980) speculated that if anthropogenic sources of OCS were to be increased by a factor of 10, then a substantial increase in sulphate aerosols would result. If we assume that lifetimes do not change (and this would require careful research in itself), then OCS concentrations would in fact need to be enhanced by a factor of 50 to produce a 1 \text{Tg S yr}^{-1} source.

It might also be possible to create a custom molecule that breaks down in the stratosphere that is not a carcinogen, but using less reactive species would produce a reservoir species that would require years to remove if society needs to stop production. Problems with this approach would be reminiscent of the climate impacts from the long-lived chlorofluorocarbons (CFCs), although lifetimes are shorter.

(c) Aerosol injection scenarios

An issue that has been largely neglected in geoengineering proposals to modify the stratospheric aerosol is the methodology for injecting aerosols or their precursors to create the desired reflective shield.

As exemplified in §2d, climate simulations to date have employed specified aerosol parameters, including size, composition and distribution, often with these parameters static in space and time. In this section, we consider transient effects associated with possible injection schemes that use aircraft platforms, and estimate the microphysical and dynamical processes that are likely to occur close to the injection point in the highly concentrated injection stream. There are many interesting physical limitations to such injection schemes for vapours and aerosols, including a very high sensitivity to the induced nucleation rates (e.g. homogeneous nucleation) that would be very difficult to quantify within injection plumes.

Two rather conservative injection scenarios are evaluated, both assume baseline emission equivalent to approximately 2.5 \text{Tg S yr}^{-1} (which ultimately forms approx. 10 \text{Tg of particles) as follows: (i) insertion of a primary aerosol, such as fine sulphate particles, using an injector mounted aboard an aircraft platform cruising in the lower stratosphere and (ii) sulphur-enhanced fuel additives employed to emit aerosol precursors in a jet engine exhaust stream. In each case injection is assumed to occur uniformly between 15 and 25 km, with the initial plumes distributed throughout this region to avoid hot spots. Attempts to concentrate the particles at lower altitudes, within thinner layers, or regionally—at high latitudes, for example—would tend to exacerbate problems in maintaining the engineered layer, by increasing the particle number density and thus increasing coagulation.
Our generic platform is a jet-fighter-sized aircraft carrying a payload of 10 metric tons of finely divided aerosol, or an equivalent precursor mass, to be distributed evenly over a 2500 km flight path during an 4-hour flight (while few aircraft are currently capable of sustained flight at stratospheric heights, platform design issues are neglected at this point). The initial plume cross section is taken to be 1 m², which is consistent with the dimensions of the platform. Note that, with these specifications, a total aerosol mass injection of 10 Tg of particles per year would call for 1 million flights, and would require several thousand aircraft operating continuously in the foreseeable future. To evaluate other scenarios or specifications, the results described below may be scaled to a proposed fleet or system.

(i) Particle properties

The most optically efficient aerosol for climate modification would have sizes, \( R_p \), of the order of 0.1 µm or somewhat less (here we use radius rather than diameter as the measure of particle size, and assume spherical, homogeneous particles at all times). Particles of this size have close to the maximum backscattering cross section per unit mass; they are small enough to remain suspended in the rarefied stratospheric air for at least a year and yet are large enough and thus could be injected at low enough abundances to maintain the desired concentration of dispersed aerosol against coagulation for perhaps months (although long-term coagulation and growth ultimately degrade the optical efficiency at the concentrations required—see below). As the size of the particles increases, the aerosol mass needed to maintain a fixed optical depth increases roughly as \( \sim R_p^3 \), the local mass sedimentation flux increases as \( \sim R_p^4 \), and the particle infrared absorptivity increases as \( \sim R_p^3 \) (e.g. Seinfeld & Pandis 1997). Accordingly, to achieve, and then stabilize, a specific net radiative forcing, similar to those discussed in §2d, larger particle sizes imply increasingly greater mass injections, which in turn accelerate particle growth, further complicating the maintenance of the engineered layer.

This discussion assumes a monodispersed aerosol. However, an evolving aerosol, or one maintained in a steady state, exhibits significant size dispersion. Upper-tropospheric and stratospheric aerosols typically have a lognormal-like size distribution with dispersion \( \sigma_{\text{LN}} \sim 1.6–2.0 \) (ln \( \sigma_{\text{LN}} \sim 0.47–0.69 \)). Such distributions require a greater total particle mass per target optical depth than a nearly monodispersed aerosol of the same mean particle size and number concentration. Accordingly, the mass injections estimated here should be increased by a factor of approximately 2, other things remaining equal (i.e. for \( \sigma_{\text{LN}} \sim 1.6–2.0 \), the mass multiplier is in the range of 1.6–2.6).

(ii) Aerosol microphysics

A bottleneck in producing an optically efficient uniformly dispersed aerosol—assuming perfect disaggregation in the injector nozzles—results from coagulation during early plume evolution. For a delivery system with the specifications given above, for example, the initial concentration of plume particles of radius \( R_{p0} = 0.08 \mu m \) would be approximately \( 1 \times 10^9 \text{ cm}^{-3} \), assuming sulphate-like particles with a density of \( 2 \text{ g cm}^{-3} \). This initial concentration scales inversely with the plume cross-sectional area, flight distance, particle specific density and
cube of the particle radius, and also scales directly with the mass payload. For example, if $R_{po}$ were 0.04 or 0.16 μm, the initial concentration would be approximately $1 \times 10^{10}$ or $1 \times 10^{8}$ cm$^{-3}$, respectively, other conditions remaining constant.

For an injected aerosol plume, the initial coagulation time constant is

$$t_{co} = \frac{2}{n_{po} K_{co}},$$

(2.1)

where $n_{po}$ is the initial particle concentration (#/cm$^3$) and $K_{co}$ is the self-coagulation kernel (cm$^3$ s$^{-1}$) corresponding to the initial aerosol size. For $R_{po} \sim 0.1 \mu m$, $K_{co} \sim 3 \times 10^{-9}$ cm$^3$ s$^{-1}$ (e.g. Turco et al. 1979; Yu & Turco 2001). Hence, in the baseline injection scenario, $t_{co} \sim 0.07-7$ s, for $R_{po} \sim 0.04-0.16$ μm, respectively. To assess the role of self-coagulation, these time scales must be compared with typical small-scale mixing rates in a stably stratified environment, as well as the forced mixing rates in a jet exhaust wake.

Turco & Yu (1997, 1998, 1999) derived analytical solutions of the aerosol continuity equation which describe the particle microphysics in an evolving plume. The solutions account for simultaneous particle coagulation and condensational growth under the influence of turbulent mixing, and address the scavenging of plume vapours and particles by the entrained background aerosol. A key factor—in addition to the previous specifications—is the growth, or dilution, rate of a plume volume element (or, equivalently, the plume cross-sectional area). The analytical approach incorporates arbitrary mixing rates through a unique dimensionless parameter that represents the maximum total number of particles that can be maintained in an expanding, coagulating volume element at any time. Turco & Yu (1998, 1999) show that these solutions can be generalized to yield time-dependent particle size distributions, and accurately reproduce numerical simulations from a comprehensive microphysical code. Although aerosol properties (concentration, size) normally vary across the plume cross section (e.g. Brown et al. 1996; Dürbeck & Gerz 1996), uniform mixing is assumed, and only the mean behaviour is considered.

(iii) **Quiescent injection plumes**

An otherwise passive (non-exhaust) injection system generally has limited turbulent energy, and mixing is controlled more decisively by local environmental conditions. If the quiescent plume is embedded within an aircraft wake, however, the turbulence created by the exhaust, and wing vortices created at the wingtips, can have a major impact on near-field mixing rates (e.g. Schumann et al. 1998). For a quiescent plume, we adopt a linear cross-sectional growth model that represents a small-scale turbulent mixing perpendicular to the plume axis (e.g. Justus & Mani 1979). Observations and theory lead to the following empirical representation for the plume volume:

$$V(t)/V_0 = (1 + t/\tau_{mix}),$$

(2.2)

where $V$ is the plume volume element of interest (equivalent to the cross-sectional area in the near-field), $V_0$ is its initial volume and $\tau_{mix}$ is the mixing time scale. For the situations of interest, we estimate $0.1 \leq \tau_{mix} \leq 10$ s.
Following Turco & Yu (1999, eqn (73)), we find for a self-coagulating primary plume aerosol

\[ N_p(t) = N_{po} Z_{1}^1 C_f^m \ln \left( \frac{1}{C_f^c} \right), \]  

where \( N_p \) is the total number of particles in the evolving plume volume element at time \( t \), and \( N_{po} \) is the initial number. We also define the scaled time, \( f_c = t / t_{co} \), and scaled mixing rate, \( f_m = \tau_{mix} / \tau_{co} \). The local particle concentration is

\[ n_p(t) = \frac{N_p(t)}{V(t)}. \]

In figure 3, predicted changes in particle number and size are illustrated as a function of the scaled time for a range of scaled mixing rates. The ranges of parameters introduced earlier result in an approximate range of \( 0.014 \leq f_m \leq 140 \). At the lower end, prompt coagulation causes only a small reduction in the number of particles injected, while at the upper end reductions can exceed 90 per cent in the first few minutes. Particle self-coagulation in the plume extending over longer time scales further decreases the initial population—by a factor of a 1000 after one month in the most stable situation assumed here, but by only some 10s of per cent for highly energetic and turbulent initial plumes.

The dashed line in figure 3 shows the effect of coagulation at the ‘unit mixing time’, at which the plume volume has effectively doubled. Clearly, prompt coagulation significantly limits the number of particles that can be injected into the ambient stratosphere when stable stratification constrains early mixing. Initial particle concentrations in the range of approximately \( 10^{10} - 10^{11} \text{ cm}^{-3} \) would be rapidly depleted, as seen by moving down the unit mixing time line in figure 3 (further, \( 10^{11} \text{ cm}^{-3} \) of 0.08 \( \mu \text{m} \) sulphate particles exceed the density of...

Figure 3. Evolution of the total concentration of particles \( N_p \) and the mass-mean particle radius \( R_p \) in an expanding injection plume. Both variables are scaled against their initial values in the starting plume. The time axis \( (f_c = t / \tau_{co}) \) is scaled in units of the coagulation time constant \( \tau_{co} \). Each solid line, corresponding to a fixed value of \( f_m \), gives the changes in \( N_p \) and \( R_p \) for a specific mixing time scale \( \tau_{mix} \) measured relative to the coagulation time scale \( \tau_{co} \) or \( f_m = \tau_{mix} / \tau_{co} \). The heavy dashed line shows the changes at the unit mixing time, for which \( f_c = f_m \) when the plume cross-sectional area has roughly doubled; the longer the mixing time scale, the greater the reduction in particle abundance and particle radius.

\[ \text{Phil. Trans. R. Soc. A} \ (2008) \]
stratospheric air). A consequence of prompt coagulation is that it is increasingly
difficult to compensate for plume coagulation (at a fixed mass injection rate) by
reducing the starting particle size. Initial particle concentrations could
simultaneously be reduced to offset coagulation, but the necessary additional
flight activity would affect payload and/or infrastructure. It is also apparent that
rapid mass injections in the forms of liquids or powders for the purpose of
reducing flight times would lead to mass concentrations greatly exceeding those
assumed above (generally \(<1 \times 10^{-4} \, \text{g cm}^{-3}\) ), causing large particle or droplet
formation and rapid fallout.

(iv) Aerosol injection in aircraft jet exhaust

The effects of high-altitude aircraft on the upper troposphere and lower
stratosphere have been extensively studied, beginning with the supersonic
transport programmes of the 1970s and extending to recent subsonic aircraft
impact assessments (under various names) in the USA and Europe (e.g.
NASA-AEAP 1997). These projects have characterized aircraft emissions and jet
plume dynamics, and developed corresponding models to treat the various
chemical, microphysical and dynamical processes.

Enhancing aircraft fuel with added sulphur compounds (H₂S, Sₙ) could
increase the particle mass in a jet wake. It is well established that ultrafine
sulphate particles are generated copiously in jet exhaust streams during flight
(e.g. Fahey et al. 1995). The particles appear to be nucleated by sulphuric acid on
ions (hereafter chemiions, e.g. Yu & Turco (1997, 1998b) formed in the
combustion process of jet engines by radical reactions. Sulphuric acid is a
by-product of sulphur residues in the fuel (typically less than 1% sulphur by
weight); most of this fuel sulphur is emitted as SO₂. The fraction emitted as
H₂SO₄ decreases as the fuel sulphur content increases, and accounts for roughly 2
per cent of the total sulphur as the fuel sulphur content approaches
approximately 1 per cent.

The concentrations of chemiions in jet emissions are strongly limited by ion–
ion recombination along the engine train to approximately \(1 \times 10^9 \, \text{cm}^{-3}\) at the
exit plane (e.g. Arnold et al. 2000). Considering a variety of direct measurements
of particles in jet wakes, Kärcher et al. (2000) showed that chemion nucleation is
consistent with the observed relative constancy of the ultrafine volatile (non-
soot) particle emission factor, \(E_p \sim 1–2 \times 10^{17} \, \text{kg}^{-1} \, \text{fuel}\) (where it should be noted
that the concentrations of soot particles are typically less than 1 per cent of the
total number of particles emitted). \(E_p\) is quite insensitive to the fuel sulphur
content, a fact that is also consistent with a chemion nucleation source. While
vapour trails formed in jet wakes can significantly modify the injected particle
properties (e.g. Yu & Turco 1998a), condensation trails are extremely rare under
normally dry stratospheric conditions.

If we imagine enhanced jet fuel sulphur contents of 5 per cent by weight
(10–100 times current amounts) for geoengineering purposes, then the annual
consumption of approximately 50 Tg of such fuel during stratospheric flight
(approx. half the amount used by current commercial aviation) could emit up to
2.5 Tg of sulphur that would eventually generate roughly 10 Tg of sulphate
aerosol. The total number of particles emitted—for \(E_p \sim 1 \times 10^{17} \, \text{kg}^{-1} \, \text{fuel}\)—
would amount to approximately \(5 \times 10^{27}\). This number, uniformly dispersed over

Phil. Trans. R. Soc. A (2008)
a 10-km thick layer from 15 to 25 km, yields an average concentration of approximately $1 \times 10^3 \text{ cm}^{-3}$ with a particle radius of roughly 0.06 $\mu$m; in other words, an ideal geoengineered solar shield. These estimates (i) assume no unexpected chemistry or microphysics in the early wake that would alter the emission factor significantly, (ii) allow for an ideal distribution of sulphate mass among the particles, and (iii) ignore coagulation following emission.

The mixing rates in a jet wake are very rapid. Schumann et al. (1998) fit a wide range of exhaust plume observations in the upper troposphere and lower stratosphere with a ‘universal’ mixing curve. We use their result in the form,

$$V/V_0 = 100 t^{0.8}; \quad t \geq 0.0032 \text{ s}. \quad (2.4)$$

Equation (2.4) describes, roughly, plume dilution starting at the exhaust exit prior to mixing with turbine bypass air, through the jet zone, vortex region and into the ambient mixing regime. Schumann et al. (1998) state that the fit is best between 1 and 50 s. For the approximately $1 \times 10^9 \text{ cm}^{-3}$ incipient particles in the initial exhaust stream, the extent of self-coagulation can be projected using the more general analytical approach discussed earlier (Turco & Yu 1999). Thus, even at $10^5$ s, approximately three-quarters of the initial particles remain (compared with an estimated 0.0007% if mixing were completely suppressed). Clearly, prompt coagulation is not an issue in a jet exhaust plume.

(v) **Longer term plume processing**

The extended microphysical processing of an injection plume can be critical owing to the long induction time before the plume becomes widely dispersed as a part of the background aerosol. Yu & Turco (1999) studied the far-wake regime of jet exhaust for upper tropospheric conditions to estimate the yield of cloud condensation nuclei from volatile aircraft particulate emissions. In their simulations, the background aerosol surface area density (SAD) ranged from 12.7 to 18.5 $\mu$m$^2$ cm$^{-3}$ for summer conditions. The resulting scavenging of fresh plume particles amounted to approximately 95 per cent after 10 days (that is, the effective emission index was decreased by a factor of 20). Moreover, only approximately 1 in 10 000 of the original particles had grown to 0.08 $\mu$m at that time, corresponding to a fuel sulphur content of 0.27 per cent by weight, with 2 per cent emitted as H$_2$SO$_4$. For a geoengineering scheme with 5 per cent fuel sulphur, although the primary exhaust sulphuric acid fraction would probably be less than 1 per cent, the initial growth rate of the chemiions would probably be accelerated.

At typical mixing rates, background aerosol concentrations would be present in an injection plume within a minute or less. The natural stratosphere has an ambient aerosol concentration of $1-10 \text{ cm}^{-3}$, with an effective surface area of less than $1 \mu$m$^2$ cm$^{-3}$. However, in a geoengineered stratosphere, at the desired baseline optical depth, a SAD greater than $10 \mu$m$^2$ cm$^{-3}$ would prevail. Further, any attempt to concentrate the engineered layer regionally or vertically, or both, would greatly exacerbate both self-coagulation and local scavenging.

The coagulation kernel for collisions of the background engineered particles (assuming a minimum radius of approx. 0.1–0.2 $\mu$m following ageing) with jet exhaust nanoparticles of approximately 10–80 nm is approximately $1 \times 10^{-7}$ to $4 \times 10^{-9} \text{ cm}^3 \text{s}^{-1}$, respectively (Turco et al. 1979). Using a mean scavenging
kernel for growing jet particles of approximately $2 \times 10^{-8} \text{ cm}^3 \text{s}^{-1}$, and a background concentration of $120 \text{ cm}^{-3}$ (determined for a doubling of the mass injection rate to maintain the optical depth, see below), the estimated scavenging factor is $\exp(-2.5 \times 10^{-6} t)$. After 1 day, the reduction in number is a factor of approximately 0.80, and over 10 days, approximately 0.1, consistent with the result of Yu & Turco (1999). Keeping in mind that the optical requirements of the engineered layer are roughly based on total cross section (ignoring infrared effects), while the scavenging collision kernel is also approximately proportional to the total background particle surface area (for the particle sizes relevant to this analysis), larger particles imply a lower concentration (and greater injection mass loading) but about the same overall scavenging efficiency.

The background aerosol will also affect the partitioning of any injected vapours between new and pre-existing particles. Considering the injection of \( \text{SO}_2 \) in jet exhaust as an example, it should be noted that \( \text{SO}_2 \) oxidation in the stabilized plume should occur over roughly a day, unless oxidants are purposely added to the plume. By this time the \( \text{SO}_2 \) would be so dilute and relative humidity so low that additional nucleation would be unlikely.

At approximately 1 day, the residual plume exhaust particles may have achieved sizes approaching 0.05 \( \mu \text{m} \) (Yu & Turco 1999). Then, considering the considerably larger surface area of the background aerosol, only a fraction of the available precursor vapours would migrate to new particles, with the rest absorbed on pre-existing aerosol. Using an approach similar to that in Turco & Yu (1999), we infer that the jet-fuel sulphur injection scenario partitions roughly 20 per cent of the injected sulphur into new particles, with the rest adding to the background mass. Considering the higher fuel sulphur content, and reduced number of condensation sites, the residual injected plume particles could grow on average to approximately 0.08 \( \mu \text{m} \). While this is a desirable size, the effective emission index is an order of magnitude below that needed to maintain the desired layer under the conditions studied. Either the fuel sulphur content or fuel consumption could be doubled to regain the overall target reflectivity. Nevertheless, as the expanding injection plumes merge and intermix following the early phase of coagulation scavenging, the aerosol system undergoes continuing self-coagulation as the layer approaches, and then maintains, a steady state. The consequences of this latter phase are not included in these estimates.

(vi) **Summary**

A primary conclusion of the present analysis is that the properties of aerosols injected directly into the stratosphere from a moving (or stationary) platform, or in the exhaust stream of a jet aircraft, can be severely affected by prompt and extended microphysical processing as the injection plume disperses, especially owing to self-coagulation and coagulation scavenging by the background aerosol. Early coagulation can increase mass requirements by a factor of 2 or more primarily because increased particle size leads to reduced optical efficiency. In addition, the resulting dispersion in particle sizes implies even greater mass injections by up to a factor of approximately 2. Thus, consideration of particle aggregation and size dispersion increases, at least by several fold, the estimated engineering and infrastructure development effort needed to produce a required net solar forcing. We wish to emphasize that these calculations are merely one
exploration of an idealized set of delivery scenarios. Many others are possible, and would require similar sets of calculations, and, if deemed promising, far more elaborate studies.

(d) Global modelling

Most of the studies mentioned in the previous sections calibrated their estimates of the climate response to geoengineering aerosol (Crutzen 2006; Wigley 2006) based upon historical observations of the aerosol produced by volcanic eruptions. Crutzen and Wigley focused primarily upon the surface temperature cooling, resulting from the aerosol’s shielding effect. Trenberth & Dai (2007) analysed historical data to estimate the role of the shielding on the
hydrological cycle, and concluded that there would be a substantial reduction in precipitation over land, with a consequent decrease in runoff and river discharge to the ocean.

The analogy between a volcanic eruption and geoengineering via a sulphate aerosol strategy is imperfect. The aerosol forcing from an eruption lasts a few years at most, and eruptions occur only occasionally. There are many timescales within the Earth system, and their transient response to the eruption is not likely to be the same as the response to the continuous forcing required to counter the warming associated with greenhouse gases. Furthermore, we have no precise information on the role the eruptions might have on a world warmer than today. For example, the response of the biosphere to a volcanic eruption might be somewhat different in a warmer world than it is today. It is thus of interest to explore the consequences of geoengineering using a climate model—a tool (albeit imperfect) that can simulate some of the complexities of the Earth system—and ask how the Earth’s climate might change if one could successfully introduce particles into the stratosphere.

Govindasamy & Caldeira (2000, 2003), Govindasamy et al. (2002) and Matthews & Caldeira (2007) introduced this line of exploration, mimicking the impact of stratospheric aerosols by reducing the solar constant to diminish the energy entering the atmosphere (by 1.8%). These studies are discussed in more detail elsewhere in this volume, so we will not review them further here.

Rasch et al. (2008) used a relatively simple representation of the stratospheric sulphur cycle to study this problem. The aerosol and precursor distributions’ evolution is controlled by production, transport and loss processes as the model atmosphere evolves. The aerosols are sensitive to changes in model climate and this allows some feedbacks to be explored (for example changes in temperature of the tropical tropopause, and lower stratosphere, and changes to cross tropopause transport). Their model used a ‘bulk’ aerosol formulation carrying only the aerosol mass (the particle size distribution was prescribed). They used an atmosphere ocean general circulation model, a coupled variant of the NCAR community atmosphere model (CAM3; Collins et al. 2006), coupled to a slab ocean model (SOM). The model was designed to produce a reasonable climate for the troposphere and middle atmosphere. The use of a SOM with a thermodynamic sea ice model precluded a dynamic response from the ocean and sea ice, which requires a more complex model such as that of Robock et al. (2008) discussed below.

The model was used to explore the evolution of the sulphate aerosol and the climate response to different amounts of precursor injection, and the size of the aerosol. SO$_2$ was injected uniformly and continuously in a 2 km thick region at 25 km between 10$^\circ$ N and 10$^\circ$ S. Owing to the difficulties of modelling the particle size evolution discussed in §2c, the study assumed the distribution to be either ‘small’ such as that seen during volcanically quiescent situations or ‘large’ such as particles seen following an eruption. Figure 4 shows the aerosol distribution and radiative forcing for an example simulation (assuming a 2 Tg S yr$^{-1}$ source and particle size similar to a volcanic aerosol). We have chosen to focus on the June, July, August season to highlight some features that disappear when displaying annual averages. The aerosol is not distributed uniformly in space and time. The mass of aerosol is concentrated in equatorial regions near the precursor injection source region, and in polar regions in areas where air densities are
higher, and mixing into the troposphere is less than the mid-latitudes and sub-tropics, where relatively rapid exchange with the troposphere takes place. Aerosol burdens are the highest in the winter hemisphere, but because solar insolation is lower there, radiative forcing is also lower than that in the summer hemisphere. Maximum radiative forcing occurs in the high latitudes of the summer hemisphere, acting to effectively shield the high latitudes resulting in a substantial recovery of sea ice compared with the $2\times CO_2$ scenario (Rasch et al. 2008).

While the largest forcing in the annually averaged sense occurs in equatorial regions, the seasonal forcing is the largest in the summer hemisphere. The most sensitivity in the response occurs at the poles, consistent with the general behaviour of climate models to uniform radiative forcing from greenhouse gases (IPCC 2007a), and also to the response to volcanic eruptions (Robock 2000), and to simpler explorations of geoengineering (Govindasamy & Caldeira 2000). Stratosphere–troposphere exchange (STE) processes respond to greenhouse gas forcing and interact with geoengineering. Nonlinear feedbacks modulate STE processes and influence the amount of aerosol precursor required to counteract CO$_2$ warming. Rasch et al. (2008) found that approximately 50 per cent more aerosol precursor must be injected than would be estimated if STE processes did not change in response to greenhouse gases or aerosols. Aerosol particle size was also found to play a role. Roughly double the aerosol mass is required to counteract greenhouse warming if the aerosol particles were as large as those seen during volcanic eruptions because larger particles are less effective at scattering incoming energy and trapping some of the outgoing energy. An estimate of 2 Tg S yr$^{-1}$ was considered to be more than enough to balance the warming in global-mean terms from a doubling of CO$_2$ if particles were small (probably unlikely), but insufficient if the particles were large. Small particles were optimal for geoengineering through radiative effects, though they also provided more surface area for chemical reactions to occur. The reduced single scattering albedo of the larger particles and increased absorption in the infrared regime lessen the impact of the geoengineering, making large particle sulphate less effective in cooling the planet. That study also indicated the potential for ozone depletion. Ozone depletion issues are discussed in more detail in §2d(i).

A typical surface temperature change from present day to a doubling of current carbon dioxide levels (denoted $2\times CO_2$) scenario is shown in figure 5, along with the result of geoengineering at 2 Tg S yr$^{-1}$ (assuming a volcanic-sized particle). The familiar CO$_2$ warming signal, particularly at high latitude, is evident, with a substantial reduction resulting from geoengineering. The simulation uses an emission rate that is not sufficient to completely counter-balance the warming. Geoengineering at this amplitude leaves the planet 0.25–0.5 K warmer than present over most of the globe, with the largest warming remaining at the winter pole. It is also straightforward to produce an emission that is sufficient to over-cool the model (e.g. Rasch et al. 2008). The polar regions and continents show the most sensitivity to the amplitude of the geoengineering.

Robock et al. (2008; hereafter referred to as the ‘Rutgers’ study) moved to the next level of sophistication in modelling geoengineering on the climate system. They used the GISS atmospheric model (Schmidt et al. 2006) and included a similar formulation for sulphate aerosols (Oman et al. 2005, 2006a,b) with a

*Phil. Trans. R. Soc. A* (2008)
substantially lower horizontal ($4 \times 5^\circ$) and vertical (23 layers to 80 km) spatial resolution than Rasch et al. (2008). Instead of using a slab ocean and sea ice model, they included a full ocean and sea ice representation. While Rasch et al. (2008) examined the steady-state response of the system for present and doubled CO$_2$ concentrations, Robock et al. (2008) explored solutions with transient CO$_2$ forcings using an IPCC A1B scenario with transient greenhouse gas forcing. They examined the consequences of injections of aerosol precursors at various altitudes and latitudes to a 20-year burst of geoengineering, between 2010 and 2030. We focus on two of their injection scenarios: (i) an injection of 2.5 Tg S yr$^{-1}$ in the tropics at altitudes between 16 and 23 km; and (ii) an injection of 1.5 Tg S yr$^{-1}$ at latitude 68$^\circ$N between 10 and 15 km. They chose a dry mode radius of 0.25 m, intermediate to the ranges explored in the Rasch et al. (2008) study. The mid-latitude injection produces a shorter lifetime for the aerosol, and concentrates its impact on the Arctic, although, as they show (and as seen below), it has global consequences. This type of geoengineering scenario shares some commonalities with scenarios described by Caldeira elsewhere in this volume. Robock et al. (2008) also showed that geoengineering is able to return sea ice, surface temperature and precipitation patterns to values closer to the present day values in a climate system model.

As an example, we show changes in precipitation for a few scenarios from Rasch et al. (2008) and Robock et al. (2008) in figure 6, again for a JJA season. Because the signals are somewhat weaker than evident in the surface temperature changes shown above, we have hatched areas where changes exceed 2 s.d. of an ensemble of control simulations to indicate differences that are likely to be statistically important. Figure 6a,b shows results from the NCAR model from Rasch et al. (2008), and figure 6c,d (labelled Rutgers) shows results from the GISS model as described in Robock et al. (2008).

As noted in IPCC (2007b), projections of changes from forcing agents to the hydrologic cycle through climate models are difficult. Uncertainties are larger than in projections of temperature, and important deficiencies remain in the simulation of clouds, and tropical precipitation in all climate models, both...
regionally and globally, so results from models must be interpreted carefully and viewed cautiously. Nevertheless, climate models do provide information about the fundamental driving forces of the hydrologic cycle and its response to changes in radiative forcing (e.g. Annamalai et al. 2007).

The NCAR results (figure 6a), consistent with IPCC (2007c) and the 20+ models summarized there, suggest a general intensification in the hydrologic cycle in a doubled CO₂ world with substantial increases in regional maxima (such as monsoon areas) and over the tropical Pacific, and decreases in the subtropics. Geoengineering (figure 6b, in this case not designed to completely compensate for the CO₂ warming) reduces the impact of the warming substantially. There are many fewer hatched areas, and the white regions indicating differences of less than 0.25 mm d⁻¹ are much more extensive.

The Rutgers simulations show a somewhat different spatial pattern, but, again, the perturbations are much smaller than those evident in an ‘ungeoengineered world’ with CO₂ warming. Figure 6c shows the precipitation

*Phil. Trans. R. Soc. A* (2008)
distributions for the polar injection; figure 6d shows the distributions for the equatorial injection. Both models show changes in the Indian and SE Asian monsoon regions, and common signals in the equatorial Atlantic. There are few common signals between the NCAR and Rutgers estimates. Robock et al. (2008) have emphasized that the perturbations that remain in the monsoon regions after geoengineering are considerable and expressed concern that these perturbations would influence the lives of billions of people. This would certainly be true. However, it is important to keep in mind that: (i) the perturbations after geoengineering are smaller than those without geoengineering; (ii) the remaining perturbations are less than or equal to 0.5 mm d$^{-1}$ in an area where seasonal precipitation rates reach 6–15 mm d$^{-1}$; (iii) the signals differ between the NCAR and Rutgers simulations in these regions; and (iv) monsoons are a notoriously difficult phenomenon to model (Annamalai et al. 2007). These caveats only serve to remind the reader about the importance of a careful assessment of the consequences of geoengineering, and the general uncertainties of modelling precipitation distributions in the context of climate change.

(i) Impact on chemistry and the middle atmosphere

Historically, most attention has focused on the surface chemistry responsible for chlorine activation and ozone depletion taking place on Polar Stratospheric Clouds, but ozone loss also occurs on sulphate aerosols, and this is evident following volcanic eruptions (Solomon 1999; Stenchikov et al. 2002). Ozone depletion depends upon a complex interaction between meteorological effects (for example temperature of the polar vortex, frequency and occurrence of sudden warmings), stratospheric photochemistry and, critically, halogen concentrations connected with the release of CFCs in the last few decades. Reductions in the ozone column following Pinatubo of 2 per cent in the tropics and 5 per cent in higher latitudes were observed when particle SAD exceeded 10 ($\mu$m)$^2$ cm$^{-3}$ (e.g. Solomon 1999). Rasch et al. (2008) noted that regions with high aerosol SAD associated with geoengineering sulphate aerosol were coincident with cold temperatures (figure 4) and indicated concern that ozone depletion might be possible, at least until most active chlorine has been flushed from the stratosphere (thought to occur after approx. 2050). Recently, Tilmes and colleagues have begun to explore some aspects of ozone depletion associated with geoengineering, and we summarize some of that work here.

Tilmes et al. (2007) estimated Arctic ozone depletion for the 1991–1992 winter following the eruption of Mt Pinatubo based on satellite observations and aircraft and balloon data, and found enhanced ozone loss in connection with enhanced SAD. They used an empirical relationship connecting meteorological conditions and ozone depletion to estimate 20–70 Dobson units (DU, a unit of mass of ozone in a column) of extra ozone depletion from the volcanic aerosols in the Arctic for the two winters following the eruption.

Tilmes et al. (2008) estimated the impact of geoengineered aerosols for future halogen conditions using a similar empirical relationship, but this time including aerosol loading and changing halogen content in the stratosphere. They based their estimates of ozone depletion on an extrapolation of present meteorological condition into the future, and assumptions about the amount and location of the geoengineering aerosol. They predicted a substantial increase of chemical ozone
depletion in the Arctic polar regions, especially for very cold winters, and a delay of 30–70 years in the recovery of the Antarctic ozone hole. This estimate of ozone depletion might be considered high, because they chose to impose a forcing (through the geoengineering aerosol) that was sufficient to counter the warming from a doubling of CO₂ even during the early twenty-first century, at a time when the halogen content is high and the CO₂ concentrations are relatively low, so a strong geoengineering is not required. However, even after 2050, accounting for the projected decline in halogens and the increase in CO₂ concentrations, they found a substantial depletion of ozone in polar regions, especially for very cold winters.

Tilmes et al. (submitted) extended their earlier calculation by using one of the aerosol distributions calculated in Rasch et al. (2008) to explore the impact of geoengineered sulphate aerosols. Rather than estimating ozone depletion using the empirical relationships, the study used the interactive chemistry climate model Whole Atmosphere Chemistry Climate Model (WACCM). The configuration included an explicit representation of the photochemistry relevant to the middle atmosphere (Kinnison et al. 2007), and a SOM. This model allows a first-order response of the troposphere to greenhouse warming, to changes to the middle atmosphere chemical composition and circulation structures, and exposes the interaction between the chemistry and dynamics. As in Tilmes et al. (2008), sulphur loading appropriate to substantially counteract the warming associated with a doubling of CO₂ was assumed, resulting in a possible overestimation of the impact of geoengineering before 2050.

Two simulations of the time period 2010–2050 were performed as follows: (i) a baseline run without geoengineering aerosols; and (ii) a simulation containing geoengineering aerosols. For the baseline run, monthly mean background values of aerosols were assumed to match background SAGEII estimates (SPARC 2006). For the geoengineering run, a repeating annual cycle of aerosols derived from the run scenario labelled as ‘volc2’ from Rasch et al. (2008) was employed. That scenario assumed aerosols with a particle size distribution similar to that following a volcanic eruption, and an aerosol burden produced from a 2 Tg S yr⁻¹ injection of SO₂. Both model simulations used the IPCC A1B greenhouse gas scenario and changing halogen conditions for the stratosphere. In the model simulations, the halogen content in the stratosphere was assumed to decrease to 1980 values by approximately 2060 (Newman et al. 2006). The study thus explored the impact of geoengineering during a period with significant amount of halogens in the stratosphere so that ozone depletion through surface chemistry is important.

In addition to the desired cooling of the surface and tropospheric temperatures, the enhanced sulphate aerosols in the stratosphere directly influence middle atmosphere temperatures, chemistry and wind fields. The temperature-dependent heterogeneous reaction rates in the stratosphere affect the amount of ozone. Ozone plays an important role in the energy budget of the stratosphere, absorbing incoming solar energy and outgoing energy in the infrared. It therefore influences temperatures (and indirectly the wind field), especially in polar regions. Additional aerosol heating also results in warmer temperatures in the tropical lower stratosphere (between 18 and 30 km). This results in an increase of the temperature gradient between tropics and polar regions (as mentioned in Robock 2000). As a consequence, the polar vortex becomes stronger and colder, and the Arctic polar vortex exists longer with geoengineering than without, which influences polar ozone depletion.
In the tropics and mid-latitudes, enhanced heterogeneous reactions cause a slight increase of ozone owing to the shift of the NO\textsubscript{x}/NO\textsubscript{y} equilibrium towards NO\textsubscript{y} in the region of high-aerosol loading with an increase of ozone loss rates above and below this layer owing to the higher temperatures in the geoengineering run. On average, the column ozone increases by 2–3 per cent maximum at 20–30\textdegree north and south. In polar regions, an increase in heterogeneous reaction rates has a more severe impact on the ozone layer.

Figure 7. Partial chemical ozone depletion between 350 and 550 K in (a) the Arctic vortex core in April and (b) the Antarctic vortex core in October. The depletion is estimated using the baseline model run (black diamonds), the geoengineering run (red diamonds) and observations (Tilmes et al. 2006, 2007), as displayed in figure 7. Estimates for present day depletion are indicated in black triangles. Estimates for the control simulations and geoengineered atmosphere are shown in black and red diamonds, respectively.

In the tropics and mid-latitudes, enhanced heterogeneous reactions cause a slight increase of ozone owing to the shift of the NO\textsubscript{x}/NO\textsubscript{y} equilibrium towards NO\textsubscript{y} in the region of high-aerosol loading with an increase of ozone loss rates above and below this layer owing to the higher temperatures in the geoengineering run. On average, the column ozone increases by 2–3 per cent maximum at 20–30\textdegree north and south. In polar regions, an increase in heterogeneous reaction rates has a more severe impact on the ozone layer.

Chemical ozone loss in the polar vortex between early winter and spring can be derived for both model simulations. These results can be compared with estimates derived from observations between 1991 and 1992 and 2004 and 2005 for both hemispheres (Tilmes et al. 2006, 2007), as displayed in figure 7. Estimates for present day depletion are indicated in black triangles. Estimates for the control simulations and geoengineered atmosphere are shown in black and red diamonds, respectively.

The WACCM model does a relatively good job of reproducing the ozone depletion for the Antarctic vortex (figure 7b). Ozone loss decreases linearly with time (black diamonds), and year to year variability in the model is similar to that of the observations. The WACCM model suggests a 40–50 DU increase in ozone depletion in the Antarctic vortex owing to geoengineering.
The model reproduces the depletion and variability much less realistically in the Arctic (figure 7a). Averaged temperatures in the simulated vortex are similar to observations, but the model does not reproduce the observed chemical response in ozone. The simulated polar vortex is 2–5° (in latitude) too small and the vortex boundary is not as sharp as that seen in the observations. The ozone depletion starts later in the winter owing to warmer temperatures in the beginning of the winter and there is less illumination at the edge of the smaller vortex (necessary to produce the depletion). Chemical ozone depletion for the WACCM3 baseline run in the Arctic is less than half of that derived from observations. Underestimates of bromine concentrations may also contribute to the underestimation of chemical ozone loss.

Examples of spatial changes in ozone depletion are shown in figure 8, which displays the difference between the baseline and geoengineering runs for Antarctica (figure 8a,b) for two winters with similar temperature structure and
for the coldest winter of each simulation in the Arctic (figure 8c,d). Antarctic winters show approximately 30 DU smaller column ozone values for the geoengineering simulation. Larger ozone losses occur over a wider area of the vortex for the geoengineering model run.

The geoengineering simulation suggests that ozone depletion will be somewhat larger in the Arctic. The amplitude of the variability in ozone depletion is increased in the geoengineering simulation, and colder vortex temperatures occur during winter and spring. The coldest three winters of the geoengineering run are 1–2.5°C colder than the coldest winter in the baseline run (between 20 and 25 km in altitude) between mid-December and March. The warm Arctic winter in the baseline case shows little ozone depletion (figure 8c). The colder temperatures and larger vortex in the geoengineering run result in increased depletion compared with the control. The Arctic ozone column falls below 250 DU in the vortex core and reaches latitudes of 70°N. Note that the inability of the model to reproduce the chemical signal of observed ozone depletion in the unperturbed calculation means one must be cautious in interpreting the model estimates for the Arctic.

These studies (Tilmes et al. 2008, submitted) thus indicate that geoengineering may have a significant impact on the ozone layer, with a possible decrease in ozone by an additional 40–50 DU when geoengineering is employed, and a possible delay of a few decades before the ozone recovery would begin again. More precise quantification of these effects will require a better specification of the aerosol evolution, and more realistic representations for the model dynamics and chemistry.

3. Summary, discussion and conclusions

Geoengineering by stratospheric aerosols as a possible means of mitigating the climate change associated with increased greenhouse gases has been reviewed. Sulphate aerosols in the stratosphere will increase the reflectivity of the planet and counteract some of the effects of CO₂ warming. Part of the attraction of using stratospheric aerosols arises because volcanic eruptions form a natural but imperfect analogue to geoengineering. Observations following major volcanic eruptions have demonstrated that sulphate aerosol, in sufficient amounts, will cool the planet, and that the Earth system can survive this kind of perturbation. Although the topic has been discussed over the last 30 years, only very recently have attempts been made to understand the interactions between various components of the climate system using modern tools for understanding climate consequences. These tools provide opportunities to quantify the interactions and consequences, and to explore those consequences on time scales that are much longer than the influence of a single volcanic eruption.

We have shown that state-of-the-art climate models used to simulate the Earth system produce the intended physical response to geoengineering, i.e. the Earth does cool, and many components of the system return to a state more like an unperturbed Earth. However,

— Our studies have shown that the delivery of aerosols or their precursors, at least using our hypothetical aircraft, is a formidable task. For the conservative scenarios we have explored, it would take of the order of a million flights of 4-hour duration (2500 km) per year to deliver the nominal...
amount of aerosol ($10 \text{Tg particles yr}^{-1} = 2.5 \text{Tg S yr}^{-1}$) needed to balance the warming associated with increasing greenhouse gas emissions. These numbers are still quite rough, and it is possible that up to four times as much sulphur might be required. We have not investigated the entire spectrum of delivery systems. The issues and methodology we have suggested may be relevant to other proposed delivery systems (artillery shells, balloons, hoses, other aircraft), although details will certainly be different. It may be possible to design more efficient methods for delivery, but all will require careful attention to detail and the difficulty of designing a system that produces particles of the right size over broad regions of the stratosphere should not be underestimated.

— Although it is possible to cool the Earth to approximately the same globally averaged surface temperature, it is not likely that all aspects of the physical system will return to a state such as that prior to human-induced CO$_2$ increases. It is important to emphasize the uncertainties in our characterization of these issues. We have made initial exploratory forays into understanding the consequences of geoengineering, but much work remains to be done. The high sensitivity of polar regions to processes regulating energy in and out of the system would make it difficult to reproduce precisely the seasonal cycle of the polar climate for a pre-industrial (or even present day) world with geoengineering.

A recent study by Stenchikov et al. (2006) showed that models have difficulty in capturing the regional response of the climate system to volcanic eruptions. They argued that volcanoes' influence on the Arctic Annular Oscillation is associated with the extra heating in the equatorial lower stratosphere, changing the temperature gradient in the lower stratosphere vortex and producing stronger westerlies and a winter warming over northern Eurasia and North America. Models identified in that paper (which were reviewed in the IPCC report) tended to underestimate, and misplace the Northern Hemisphere winter surface temperature warming seen over Siberia in the observations following an eruption. This suggests that while the zeroth-order response of a surface cooling is likely to be robust, the first-order response of other components of the climate systems is a difficult problem and that model regional responses to stratospheric forcing changes must be viewed with caution.

As discussed in §2d, there are also hints that rainfall patterns would be different from an undisturbed Earth, although it is likely that they would be much closer to that distribution than in a world with $2\times$CO$_2$ and no geoengineering.

— An increase in aerosol burden is likely to increase ozone depletion. We have shown that current chemistry climate models have difficulty in reproducing quantitatively the dynamics and chemistry of the arctic middle atmosphere. Better coupled chemistry climate models would allow an improved estimate of ozone, sulphate aerosol and dynamical interactions. The first step is to improve the models' capability in reproducing present day ozone representation, particularly for the Northern Hemisphere.

— Reductions in ozone will lead to increases in solar ultraviolet-B radiation reaching the Earth's surface with a potential impact on human health (Ambach & Blumthaler 1993; Madronich & de Grujil 1993) and biological populations (Blaustein et al. 1994). The increase in UV associated with ozone depletion could be compensated for by increased light extinction and attenuation by the aerosol
cloud itself; Vogelmann et al. (1992) and Wetzel et al. (2003) explored the compensation between these effects. Vogelmann et al. (1992) studied the effect for volcanic eruptions and concluded that, for stratospheric aerosol with an optical depth of 0.1–0.2 (approx. the value required for geoengineering), ozone and aerosol effects approximately compensated. At higher aerosol amounts, the aerosol attenuation did not balance the enhancement from ozone, and UV was enhanced at the surface. This kind of calculation should be repeated with a focus upon geoengineering and global warming, since ozone distributions and aerosol spatial and particle size distributions might differ significantly for geoengineering scenarios compared with their volcanic eruption counterpart.

Gu et al. (2003) showed that volcanic aerosols from the Pinatubo eruption substantially increase diffuse radiation worldwide, with a resulting enhancement to photosynthesis and uptake of CO$_2$. The same effect is to be anticipated with the geoengineering shield. Govindasamy et al. (2002) explored some aspects of interactions between the physical Earth system and the biosphere. They showed that stabilizing the temperature but not CO$_2$ induced a change in Net Primary Productivity. Their study had a number of limitations as follows: (i) they used a prescribed CO$_2$ concentration, eliminating important feedbacks, (ii) they did not use a biospheric model that included nutrient limitation, (iii) they did not include an ocean biosphere, and (iv) their model was not sensitive to changes in the ratio of direct to diffuse radiation.

While ecosystems can survive occasional volcanic eruptions, it is not clear whether the consequences to ecosystems would be from long-term changes in direct/diffuse energy, or increases in UV radiation. These issues argue for more attention on the consequences of stratospheric aerosols to ecosystems.

The change in ratio of direct to diffuse radiation will also have an effect on solar energy production with technologies that make use of solar concentrators. Advances in solar energy production which operate efficiently in the presence of diffuse radiation are also possible, but a different technology is needed. Characterizing the consequences of geoengineering for these technologies is worthwhile.

As mentioned in §§1 and 2b,d, larger aerosol particles exhibit significant absorption in the infrared part of the energy spectrum. The cooling resulting from the scattering of incoming solar energy is thus partly compensated for by the absorption in the infrared. The proclivity of this geoengineering method to form large particles makes it a less efficient solution than it would be if small particles were easily generated and maintained.

There are also occasional concerns voiced about increases to acid rain from this type of geoengineering. We have shown that, although the perturbations to the stratospheric sulphur cycle are quite large (increasing the background sources there by a factor of 15–30), they are perhaps 2 per cent of the total (troposphere + stratosphere) sulphur sources. Therefore it is unlikely that geoengineering will have a significant impact on acid deposition and the global increment could easily be balanced by a small reduction in tropospheric emissions. On the other hand, it is possible that the deposition of the geoengineering aerosol could influence a region that normally sees little sulphate deposition from tropospheric sources if it occurs there. This should be looked into.
— It is obvious that current models of the sulphur cycle could be substantially improved. It would be desirable to move beyond the bulk aerosol formulations used here to models that included the evolution of the particle size distribution, accounting explicitly for aerosol growth and coagulation. This would include a mechanism to move from the source as determined by the delivery system, to evolution within the plume and finally to scales resolved by a global model.
— It is clear that this geoengineering method will not alleviate the problems engendered by absorption of CO$_2$ in the oceans, with a resulting decrease in ocean pH.

Substantial reductions in greenhouse gas emissions must take place soon to avoid large and undesirable climate impacts. This study has reviewed one technique that might be used in a planetary emergency to mitigate some of the effects of a projected global warming. We emphasize that, while the studies highlighted here are a step along the way, we believe no proposal (including the ideas explored here) has yet completed the series of steps required for a comprehensive and thoroughly studied geoengineering mitigation strategy occurring in the peer reviewed literature (Cicerone 2006). Our review of studies of geoengineering by sulphate aerosols suggests it will ameliorate some consequences of global warming. The study highlights some positive aspects of the strategy. However, many uncertainties remain in understanding the influence of geoengineering on the climate system (particularly on aspects related to likely impacts on the biosphere). More work is required to understand the costs, benefits and risks involved, and to reconcile the legal, political and ethical issues of geoengineering.

We thank Alvia Gaskill, Ken Caldeira and three anonymous referees for their constructive suggestions. Dennis Shea kindly provided computer scripts facilitating some of the model analysis. Work by R.P.T. was supported by NSF grant ATM-07-30761. A.R., L.O. and G.L.S. were supported by NSF grant ATM-0730452. P.J.R. and S.T. were supported by NCAR and the NSF. C.-C.C. was supported by a grant from the U. Calgary. NCAR simulations were performed at NCAR. Rutgers simulations were carried out at NASA GISS. Model development and computer time at GISS were supported by NASA climate modelling grants. The National Center for Atmospheric Research is sponsored by the National Science Foundation. P.J.R. thanks Diane Pendlebury for her help in redrafting figure 1.

References


*Phil. Trans. R. Soc. A* (2008)


