Climate Forcing by Anthropogenic Aerosols

R. J. CHARLSON, S. E. SCHWARTZ, J. M. HALEs, R. D. CESS, J. A. COAKLEY, JR., J. E. HANSEN, D. J. HofMANN

Although long considered to be of marginal importance to global climate change, tropospheric aerosol contributes substantially to radiative forcing, and anthropogenic sulfate aerosol in particular has imposed a major perturbation to this forcing. Both the direct scattering of short-wavelength solar radiation and the modification of the shortwave reflective properties of clouds by sulfate aerosol particles increase planetary albedo, thereby exerting a cooling influence on the planet. Current climate forcing due to anthropogenic sulfate is estimated to be \(-1\) to \(-2\) watts per square meter, globally averaged. This perturbation is comparable in magnitude to current anthropogenic greenhouse gas forcing but opposite in sign. Thus, the aerosol forcing has likely offset global greenhouse warming to a substantial degree. However, differences in geographical and seasonal distributions of these forcings preclude any simple compensation. Aerosol effects must be taken into account in evaluating anthropogenic influences on past, current, and projected future climate and in formulating policy regarding controls on emission of greenhouse gases and sulfur dioxide. Resolution of such policy issues requires integrated research on the magnitude and geographical distribution of aerosol climate forcing and on the controlling chemical and physical processes.

The response of the Earth's climate to the perturbation in radiative forcing due to increased concentrations of infrared-active (greenhouse) gases is the subject of intense research because of the well-documented increases in concentrations of these gases over the industrial era and the recognition of the climatic importance of the radiative forcing associated with these increases. It is becoming apparent that anthropogenic aerosols exert a radiative influence on climate that is globally comparable to that of greenhouse gases but opposite in sign. However, this aerosol radiative influence has received much less attention than forcing by anthropogenic greenhouse gases. In view of the magnitude of aerosol influences on climate, it seems mandatory that these influences should be included in efforts to obtain accurate estimates of anthropogenic perturbations to the earth's radiation budget at present and over the industrial era. Such estimates are essential for (i) evaluating climate sensitivity from observed climate change, (ii) evaluating the performance of climate models, and (iii) reliably predicting potential future climate changes. In this article we describe the mechanisms by which anthropogenic aerosols perturb the global climate, provide estimates of the global-average magnitude of the aerosol perturbation in radiative forcing, outline the information required to describe the spatially nonuniform perturbation in climate models, assess where additional information is required, and suggest approaches to gaining this information.

Although it has long been recognized that tropospheric aerosols exert a cooling influence on climate because of their scattering of shortwave radiation and the resultant increase in planetary albedo (1), this influence has been widely assumed to be fairly uniform spatially and constant temporally [for example, (2–4)], and this perception has been reflected in most analyses of global climate change (5–8). However, industrial activities, especially emissions of SO\(_x\), which result in the formation of particulate sulfate (SO\(_2\)\(^{-}\)) compounds, contribute substantially to tropospheric aerosol, especially to submicrometer aerosol, which is effective in the scattering of shortwave radiation (9), and this aerosol is distributed quite nonuniformly over the earth and has substantially increased in concentration since around 1850 (10–12). Thus, there is strong reason to infer that anthropogenic sulfate aerosol substantially enhances local
Aerosol Mechanisms Affecting the Earth Radiation Budget

Radiative influences of aerosols on climate may be distinguished as direct, referring to scattering and absorption of radiation by the aerosol particles themselves, and indirect, referring to the influence of aerosols on cloud radiative properties. Aerosols may also affect cloud lifetimes.

Direct radiative influence. It has been recognized for over 60 years that aerosol particles influence the earth's radiative balance directly by backscattering and absorption of shortwave (solar) radiation (1), although regional quantification of this phenomenon has been achieved only slowly (18) and imperfectly. Aerosol particles also absorb longwave (infrared) radiation, but this effect is usually small because the opacity of aerosols decreases at longer wavelengths and because the aerosols are most concentrated in the lower troposphere, where the atmospheric temperature, which governs emission, is practically the same as the surface temperature (4, 7). In the unperturbed atmosphere, the principal aerosol constituents that contribute to light scattering are sulfate from biogenic gaseous sulfur compounds and organic carbon from partial atmospheric oxidation of gaseous biogenic organic compounds, such as terpenes (19). Sea salt and windblown soil dust contribute substantially at some locations, but their effects on the global climate are generally unimportant because the particles are large and usually short-lived and thus are transported only short distances. Submicrometer anthropogenic aerosols responsible for increased light scattering are produced in the atmosphere by chemical reactions of sulfur-, nitrogen-, and carbon-containing gases, predominantly sulfur. Light absorption is dominated by particles containing elemental carbon, C0, produced by incomplete combustion of carbonaceous fuel. The light-scattering effect is dominant at most latitudes, but absorption can dominate at high altitude, especially over highly reflective snow- or ice-covered surfaces (20-22). Other aerosol substances may also be locally important, especially those that are sporadic, such as those from volcanoes and wildfires as well as windblown dust from deserts.

Indirect radiative influence: Clouds. Cloud droplets form in the lower atmosphere by condensation of water on existing aerosol particles. Consequently, the concentration, size, and water solubility of the aerosol particles on which cloud droplets form (cloud condensation nuclei, CCN) have an immediate influence on the concentration and size of cloud droplets. These influences in turn affect the shortwave radiative properties of clouds that form on the aerosol, even in the absence of any influence on the macroscopic and thermodynamic properties of the clouds. Specifically, increased concentrations of CCN result in increased concentrations of cloud droplets, resulting in enhanced shortwave albedo of clouds (9, 13-15). In contrast, the perturbation in longwave absorption by tropospheric clouds arising from an increase in cloud-droplet concentration is much less, indeed negligible, because tropospheric clouds are already optically thick or nearly so at infrared wavelengths (23). Sulfate aerosols appear to dominate anthropogenic influences on CCN, although particles of smoke from biomass combustion, which are poorly characterized, may be important in some circumstances (24).

Cloud lifetime influence. A decrease in mean droplet size associated with an increase in cloud-droplet concentration is also expected to inhibit precipitation development and hence to increase cloud lifetimes (25-28). Such an enhancement of cloud lifetime and the resultant increase in fractional cloud cover would increase both the short- and longwave radiative influence of clouds. However, because this effect would predominantly influence low clouds, for which the shortwave influence dominates, the net effect would be one of a further cooling. Inhibited precipitation development might further alter the amount and vertical distribution of water and heat in the atmosphere and thereby modify the earth's hydrological cycle. Although these effects cannot yet be quantified, they have the potential of inducing major changes in global weather patterns as well as in the concentration of the dominant greenhouse gas itself, water vapor (27).

Anthropogenic Perturbation of Tropospheric Aerosol

Until recently, most investigations of the climate influence of tropospheric aerosols [for example, (2-4)] have focused on physical aspects, that is, the direct-scattering effect, and have viewed this aerosol as a "natural background" and therefore as a static feature of global radiative transfer, neglecting the chemical processes that produce the aerosol. A value of aerosol optical depth of about 0.12 is often used in radiative transfer models to represent such a natural background (3, 29). Shortwave scattering by such an aerosol is calculated to cause a 2 to 3 K surface cooling over the globe (4). However, compelling evidence requires abandoning the view that this aerosol represents a natural background. Anthropogenic SO2 emissions (primarily from fossil fuel combustion and metal smelting) have increased substantially over the industrial period, and such emissions now substantially exceed, on a global basis, natural emissions of sulfur-containing gases to the atmosphere [mainly dimethylsulfide (DMS) from marine phytoplankton] (12, 30, 31) (Fig. 1). Because residence times of SO2 and
SO$_4^{2-}$ in the troposphere are short (several days) compared to mixing times in the troposphere (several months within hemispheres; about a year between hemispheres), anthropogenic sulfate is concentrated mainly in industrial regions and areas downwind of them, largely in mid-latitudes. Indeed, concentrations of aerosol sulfate and of CCN within 1000 km or more downwind of regions of industrial emissions are an order of magnitude or more greater than those in remote regions (10, 12, 34, 35). The optically effective and cloud-nucleating fraction of the atmospheric aerosol in these regions is frequently dominated by sulfate particles (22, 35–37). Plumes of light-scattering aerosol extending from industrialized regions to the marine atmosphere have, like windblown dust from deserts, been discerned in satellite measurements (38).

Because more than 90% of industrial SO$_2$ is emitted in the Northern Hemisphere (NH) (30, 31, 39), anthropogenic sulfate is confined largely to the NH. Widespread distribution of anthropogenic aerosol sulfate in the NH is evidenced by elevated concentrations of non-sea-salt sulfate in aerosol and in precipitation at remote sites in the NH relative to those in the Southern Hemisphere (SH) and by increases in concentrations of sulfate in NH glacial ice over the past 100 years that closely match the record of anthropogenic emissions shown in Fig. 1 (11, 12, 32, 33). Mass concentrations of non-sea-salt aerosol sulfate appear to be enhanced by 30% or more over the natural background in much of the NH marine troposphere (12), although a complete database to support this contention is lacking (40). Furthermore, recent measurements (41) have shown that the aerosol optical depth in the unperturbed SH, δ = 0.05, is much lower than values that had been ascribed to natural background. Thus, much of the aerosol cooling formerly attributed to natural background must be due to anthropogenic aerosol.

Some studies (for example, (5)) have used the fraction of the total aerosol source strength as a basis for estimating the magnitude of the anthropogenic fraction of radiative forcing by tropospheric aerosols. This approach is likely to underestimate the importance of aerosols derived from industrial SO$_2$ for three reasons. (i) Sulfate particles from SO$_2$ oxidation are much smaller than mechanically produced materials such as soil dust and sea salt. The lifetime of sulfate in the troposphere (several days) is therefore long compared to that of the main mass of soil dust and sea salt (a day or less). Thus, mass concentration in the troposphere is more sensitive to source strength for the long-lived SO$_4^{2-}$ than for the short-lived substances. (ii) The optical-scattering efficiency per unit mass concentration of submicrometer sulfates is an order of magnitude greater than that of supermicrometer dust (9, 22, 42). (iii) Unlike soil dust, the key sulfate compounds [H$_2$SO$_4$ to (NH$_4$)$_2$SO$_4$] are very hygroscopic (43); hence, at relative humidities typical of the lower troposphere (75 to 80%), these particles have roughly twice the optical scattering per unit mass concentration as dry particles (22, 44, 45). Thus, their effect on aerosol optical depth is disproportionately large compared to their dry mass concentration.

Although not explicitly coupled to chemical sources, the indirect radiative effect of anthropogenic sulfate aerosol and its global implications are suggested by satellite observations of cloud brightening in polluted air (15) and by increased cloud-top albedo and increased concentrations of cloud droplets associated with exhaust plumes of ships at sea (ship trails) (14). High CCN concentrations downwind of major SO$_2$ emission regions also are circumstantial evidence that atmospheric chemical reactions of SO$_2$ pollution produce CCN (34). The physical and climatic effects have only recently been connected to the atmospheric chemical reactions that produce the cloud-active particles (46).

Although the climatic significance of anthropogenic SO$_2$ emissions is only now coming to be recognized, many features of the overall picture relating these emissions to climate forcing have been recognized and relatively well understood for some time. It has long been recognized that pollution aerosol could cause significant radiative forcing (47), but the physical effect was not connected to the sources and chemical processes that maintain the aerosol concentration field. The atmospheric chemistry community recognized the role of chemical reactions of sulfur gases as a major source of aerosol particles (48) but did not appreciate the climatic impact of such an aerosol. Although satellite images (38, 49) and networks of chemical and optical measurements (50) showed that the horizontal extent of the influence of anthropogenic SO$_2$ on aerosol optical depth was 1000 to 2000 km or more, the climatic influence of such a widespread perturbation in optical depth was not appreciated. Researchers who recognized chemical processes as the source of climate forcing aerosols (51) failed to extend their results to the hemispheric or global scale. Researchers linking climate influences to the optical properties of anthropogenic sulfate aerosol (2) did not consider chemical processes. A large body of research addressing the transformation, transport, and deposition of anthropogenic SO$_2$ and SO$_4^{2-}$ (for reviews, see, for example, (52, 53)) focused on the chemical and biological consequences of these emissions, largely on local and regional scales, rather than on any climate influence. Microphysical evidence, such as the morphology of sampled aerosol particles and the volatility of CCN (36), indicates that the number population of marine CCN is dominated by sulfate particles. Nevertheless, researchers demonstrating the dominance of natural marine aerosol production by oxidation of DMS and the global climatic influence of this aerosol either failed to account for the role of anthropogenic emissions or intentionally excluded consideration of these emissions (46, 54).

Although all of the components of the linkage of anthropogenic emissions to climate forcing rest on a substantial body of research, the picture as a whole is only now emerging. This picture, together with the substantial excess of anthropogenic gases in comparison to natural sulfur gases entering the NH atmosphere in recent decades (Fig. 1), presents a compelling argument that anthropogenic sulfate aerosol has exerted an influence on climate over this period, certainly in and downwind of industrial regions where an increase in the SO$_2$ emission has occurred, and perhaps throughout the NH.
Global-Mean Radiative Forcing Perturbation

Estimates of the global-average perturbations in radiative forcing induced by anthropogenic aerosols because of direct light scattering and enhanced cloud albedo permit a comparison of the climate influence of these aerosols with that due to increases in concentrations of CO₂ and other anthropogenic greenhouse gases. A doubling of global concentrations of CO₂ is thought (8) to contribute an increase in longwave radiative forcing of the troposphere of ~4.4 W m⁻², and the increase from the preindustrial concentration of 280 ppm to the 1990 value of 353 ppm represents an increase of ~1.25 W m⁻². The total increase in forcing by all anthropogenic greenhouse gases over this period is estimated to have been 2 to 2.5 W m⁻² (5, 8). We use these comparisons as a yardstick by which to assess the importance of the aerosol forcing, but because of their different natures, the radiative forcing by anthropogenic sulfate aerosol should not be viewed as a simple compensation of greenhouse gas forcing.

Direct forcing. The area mean shortwave forcing ΔF_R, resulting from an increase in aerosol concentration can be expressed as

$$\Delta F_R = -\frac{1}{4} \frac{F_T}{\Delta R_0} (1 - A_0) \Delta R_a$$  

(1)

where (1/4 F_T) is the global mean top-of-the-atmosphere (TOA) radiative flux, A_0 is the fractional cloud cover, and ΔR_a is the perturbation in planetary mean albedo due to enhanced aerosol concentration. The factor (1 - A_0) is introduced because the albedo is enhanced mainly for non-cloud-covered portions of the planet. The negative sign denotes that the forcing is an effective cooling tendency. For an optically thin, light-scattering aerosol (optical depth δ_a << 1), ΔR_a is linear in the optical depth due to that substance and can be represented (45) as

$$\Delta R_a = 27 \Delta \delta_a$$  

(2)

where T is the fraction of incident light that is transmitted by the atmospheric layer above the aerosol layer (that is, not absorbed by water vapor), R_a is the mean albedo of the underlying surface, and δ_a is the mean aerosol optical depth of the aerosol. Combination of Eqs. 1 and 2 yields for sulfate aerosol

$$\Delta F_R = -\frac{1}{4} \frac{F_T}{\Delta \delta_a} (1 - A_0) \Delta \delta_a$$  

(3)

Loading of sulfate aerosol enters this equation by the dependence of the area mean extinction on the area mean column burden of anthropogenic sulfate aerosol, B_{SO₄²⁻} (mol m⁻²),

$$\Delta \delta_a = \alpha_{SO₄²⁻} f(R) B_{SO₄²⁻}$$  

(4)

where α_{SO₄²⁻} is the molar scattering cross section of sulfate at a reference low relative humidity (30%), which in industrial regions exhibits a near-universal value of ~500 m² mol⁻¹ (~5 m² per g of SO₄²⁻) (22, 37, 42, 55, 56), and f(R) accounts for the relative increase in scattering due to the increase in particle size associated with deliquescence or hygroscopic accretion of water with increasing relative humidity (44, 45). The sulfate burden is related to parameters involving sources and removal of atmospheric sulfate:

$$B_{SO₄²⁻} = \frac{Q_{SO₄} Y_{SO₄²⁻} \tau_{SO₄²⁻}}{A}$$  

(5)

where \(Q_{SO₄}\) (in moles of S per year) is the source strength of anthropogenic SO₂, \(Y_{SO₄²⁻}\) is the fractional yield of emitted SO₂ that reacts to produce SO₄²⁻ aerosol, \(\tau_{SO₄²⁻}\) (years) is the sulfate lifetime in the atmosphere or in the geographical region under consideration, and A is the area of the geographical region to which the material is presumed to be confined, for example, the entire earth, the NH, or some smaller region.

Reasonable values of the parameters in Eqs. 1 to 5 (Table 1) (22, 37, 44, 45, 57–61), which take into account the geographic distribution of cloud cover and surface albedo, place the global-mean direct radiative forcing due to sulfate aerosol at ~1 W m⁻², uncertain to a factor of 2. This value is comparable in magnitude to the forcing ascribed to the ~25% increase in CO₂ concentration above the preindustrial value. Despite the considerable uncertainty in the input parameters and consequently in the estimated forcing, the calculation serves to establish the significance of the direct aerosol forcing. Central to this approach is separation of the forcing by anthropogenic SO₄²⁻ aerosol from that either by the total anthropogenic aerosol or by the totality of tropospheric aerosol. This separation assumes that the SO₄²⁻ forcing is proportional to concentrations of anthropogenic SO₄²⁻. The resulting forcing can thus be viewed as the partial derivative of radiative forcing with respect to anthropogenic sulfate. Because more than 90% of anthropogenic SO₂ is emitted in the NH (31, 39), the resultant forcing is exerted predominantly in the NH, with a magnitude roughly twice the global-average value.

More elaborate calculations, both for an assumed distribution of extinction (4, 6) that depends on latitude but is longitudinally invariant and for spatially varying concentrations of sulfate (45), yield forcings similar to those of the above global-average model. The regional energy budget calculations also show that although the forcing due to anthropogenic sulfate is distributed broadly across the NH, the radiative forcing varies considerably with location. In the cloud-top albedo R_{CT} due to an increase in droplet number concentration N under the assumption that the liquid-water volume

**Table 1.** Evaluation of global-mean direct radiative forcing due to anthropogenic sulfate (Eqs. 1 through 5). Double underline indicates evaluation from preceding quantities.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Value</th>
<th>Units</th>
<th>Relative uncertainty (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q_{SO₂}</td>
<td>90 × 10^{12}</td>
<td>g of sulfur per year</td>
<td>15</td>
<td>*</td>
</tr>
<tr>
<td>Y_{SO₂}</td>
<td>2.8 × 10^{12}</td>
<td>mol of sulfur per year</td>
<td>50</td>
<td>†</td>
</tr>
<tr>
<td>\tau_{SO₂}</td>
<td>0.02</td>
<td>year</td>
<td>50</td>
<td>‡</td>
</tr>
<tr>
<td>A</td>
<td>5 × 10^{14}</td>
<td>m²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B_{SO₂}</td>
<td>4.6 × 10^{3}</td>
<td>(g SO₂ per m²) m⁻²</td>
<td>40</td>
<td>(22, 37, 45, 58)</td>
</tr>
<tr>
<td>α_{SO₂}</td>
<td>5</td>
<td>m² (g SO₂⁻)⁻¹</td>
<td>40</td>
<td>(22, 37, 45, 58)</td>
</tr>
<tr>
<td>f(R)</td>
<td>1.7</td>
<td></td>
<td>20</td>
<td>(44, 45)</td>
</tr>
<tr>
<td>δ_{SO₂}</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F_T</td>
<td>1.370</td>
<td>W m⁻²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 − A</td>
<td>0.4</td>
<td></td>
<td>10</td>
<td>(59)</td>
</tr>
<tr>
<td>1 − R_a</td>
<td>0.85</td>
<td></td>
<td>10</td>
<td>(60)</td>
</tr>
<tr>
<td>\delta_a</td>
<td>0.29</td>
<td></td>
<td>25</td>
<td>(61)</td>
</tr>
</tbody>
</table>

*From Fig. 1. †Mean dry-deposition velocity of SO₂, 0.5 cm s⁻¹; height of mixed layer, 2 km; δ_{SO₂} = 1 × 10^{14} cm s⁻¹; mean OH concentration, 1 × 10^6 cm⁻³; in-cloud conversion assumed equal to that of gas phase (45). These parameter values are consistent with results of more detailed calculations [for example, (57)]. ‡SO₂, 1 week, representing the frequency of encountering precipitation removal.
fraction $L$ is unchanged by the perturbation in $N$; any increase in $L$
due to inhibition of precipitation would enhance the perturbation in
radiative forcing. An approximate analytical expression for the
reflectance of a nonabsorbing, horizontally homogeneous cloud is
(62, 63)

$$R_{CT} = \frac{\beta(\mu_0) R_C/\mu_0}{1 + \beta(\mu_0) \delta C/\mu_0}$$  \hspace{1cm} (6)

Here $\mu_0$ is the cosine of the solar zenith angle, $\beta(\mu_0)$ is the fraction
of sunlight incident to the cloud at angle $\mu_0$ that is scattered in the
upward direction for single-particle scattering, and $\delta C$ is the optical
depth of the cloud,

$$\delta C = \frac{\pi^2}{3} Q_{ext} N z_{CT}$$  \hspace{1cm} (7)

where $z_{CT}$ is the thickness of the cloud, $r_e$ is the effective drop radius,
and $Q_{ext}$ is the average extinction efficiency. For cloud droplets of
radius much greater than the wavelength of visible light, $Q_{ext}$ may
be approximated as a constant, $Q_{ext} \approx 2$. The upward fraction $\beta$ is
a slowly varying function of $\mu_0$ and $\delta C$ (64).

To examine the perturbation in $R_{CT}$ due to a perturbation in $N$,
we note that for a cloud of fixed $L$, $r_e$, and $N$ are not independent but
are related approximately by $L = (4/3)\pi r_e^3 N$, hence

$$\delta C = \frac{\pi Q_{ext}}{4} \left(\frac{3L}{4\pi}\right)^{2/3} N^{1/3}$$  \hspace{1cm} (8)

Equations 6 and 8 permit calculation of the cloud albedo for a given
physical thickness of cloud, liquid-water volume fraction, and
cloud-droplet number density (Fig. 2). Cloud-top droplets are most
sensitive to a change in $N$ when $0.3 \leq R_{CT} \leq 0.7$. In this range a 1% increase
in $N$ leads to an increase in $R_{CT}$ of $\approx 10^{-4}$.

The perturbation in global-mean radiative forcing that would result from a
change in cloud-top albedo of marine stratus and stratocumulus clouds was assessed by
Charlson et al. (46). The change in TOA albedo for a given change in cloud-top albedo $\Delta R_{CT}$ is

$$\Delta R_{TOA} = 0.8 \Delta R_{CT}$$  \hspace{1cm} (9)

If the fraction of the globe that is covered by marine stratiform
clouds, $A_{str}$, remains unchanged by the increase in $N$, then the change in the global-mean albedo is

$$\Delta R_{GM} = A_{str} \Delta R_{TOA}$$  \hspace{1cm} (10)

The resulting change in solar radiative forcing is given by the change in
global-mean albedo times the global-mean TOA radiative flux
(1/4 $F_T$):

$$\Delta F_C = \frac{1}{4} F_T \Delta R_{GM}$$  \hspace{1cm} (11)

The perturbation in forcing responds linearly to the perturbation in
cloud-top albedo and can thus be displayed as a linear dependence on
log $N$ (Fig. 3) (65); a uniform global increase of droplet
concentration of 30% in only marine stratus and stratocumulus clouds would decrease global-mean shortwave radiative forcing by
$\approx 2$ W m$^{-2}$ (46, 66). On the basis that mass concentrations of
non-sea-salt sulfate aerosol at remote locations in the NH exceed
those in the SH by $\approx 30$%, and assuming that the relative change in
number concentration is the same as the relative change in mass
concentration, Schwartz (12) suggested that the present-day mean
NH perturbation in cloud radiative forcing due to anthropogenic sulfa
tes is approximately $\approx 2$ W m$^{-2}$. For an assumed global-mean
enhancement in $N$ of 15% (Table 2), the calculated global-mean
perturbation in shortwave radiative forcing by clouds is approximately
$\approx 1$ W m$^{-2}$. By far the greatest uncertainty in this estimate is the
degree that cloud-droplet number concentration is enhanced,

which is uncertain by at least a factor of 2 (67). Substantially greater
perturbations, $\approx 10^4$ W m$^{-2}$ or more, would be expected in regions
directly influenced by industrial emissions, where CCN number
concentrations are commonly enhanced by a factor of 5 or more
above the natural background (34, 35).

Describing Anthropogenic Aerosol Forcing in Climate Models

Although a strong case is made that anthropogenic sulfate aerosols exert an important influence on the earth's radiative balance, there remain large gaps in the knowledge necessary to quantify this influence in models of current, past, or future climate. To incorporate the influence of aerosols, researchers will need to develop a basis for coupling the chemical processes that produce the aerosols and the meterological processes that distribute and remove them to the physical properties of the aerosols that influence radiative transfer and cloud microphysical properties. The compo
tents constituting the overall phenomenon are indicated in Fig. 4.

Because the optical and cloud-nucleating properties of an aerosol depend on the size and composition of the particles constituting the aerosol, the influence of the aerosol cannot be uniquely and unequivocally ascribed to specific chemical substances. Rates of new particle formation and of the time evolution of the size distribution of aerosols, which affect their optical and cloud-nucleating proper
ties, are highly dependent on the rate of production of condensible material and on the concentration and size distribution of aerosol particles already present. A description of these rates is uncertain to orders of magnitude and is a major unsolved problem.
Table 2. Evaluation of perturbation in global-mean cloud radiative forcing due to anthropogenic sulfate (Eqs. 6 through 11). Double underbar indicates evaluation from preceding quantities.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Value</th>
<th>Units</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N/N_0)</td>
<td>1.15</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>(dR_{COA}/dR_{CT})</td>
<td>0.083</td>
<td></td>
<td>Fig. 3</td>
</tr>
<tr>
<td>(\Delta R_{CT})</td>
<td>0.012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta R_{COA}/\Delta R_{CT})</td>
<td>0.8</td>
<td></td>
<td>(46)</td>
</tr>
<tr>
<td>(\Delta R_{GIM})</td>
<td>0.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(A_{max})</td>
<td>0.3</td>
<td></td>
<td>†</td>
</tr>
<tr>
<td>(\Delta F_{IM})</td>
<td>0.003</td>
<td>W m(^{-2})</td>
<td></td>
</tr>
<tr>
<td>(F_T)</td>
<td>1370</td>
<td>W m(^{-2})</td>
<td></td>
</tr>
<tr>
<td>(\Delta F_{IC})</td>
<td>-1</td>
<td>W m(^{-2})</td>
<td></td>
</tr>
</tbody>
</table>

*Based on the comparison of mass concentrations of non-sea-salt sulfate aerosol at remote locations of the NH and SH and the assumption that CCN concentration scales linearly with non-sea-salt sulfate mass (12). † Fraction of hemisphere occupied by nonoverlapped marine stratus and stratoalbta (46).

The indirect (cloud) forcing is more difficult to assess than the direct forcing. Additional key phenomena that must be considered are (i) the relations between mass and number concentrations and composition of precloud aerosol concentrations and the number, concentration, and size distribution of cloud droplets formed in air containing the aerosol; and (ii) the dependence of aerosol perturbation to cloud radiative forcing on the nature of the cloud fields. The nature of the relation between CCN concentrations and sulfate concentrations, which is likely to be nonlinear (28, 68), is not established (69).

The knowledge required to evaluate the influence of anthropogenic aerosols on cloud lifetimes, atmospheric dynamics, and the hydrological cycle includes the dependence of precipitation formation processes on cloud microphysical properties, such as number concentration and size distributions of cloud droplets. These secondary phenomena are considerably more difficult to quantify than the radiative properties, and additional research is required to accurately describe them in climate models (69).

Because anthropogenic aerosols are short-lived in the atmosphere and their sources are geographically localized, their concentration is not uniformly distributed. It is therefore necessary to specify pertinent aerosol properties as functions of location and time rather than merely as spatial- or temporal-average quantities. Vertical-distribution information is required because light scattering and absorption are altitude-dependent, as are cloud properties. Horizontal-distribution information is essential because of the variability of cloud properties, aerosol removal processes, and surface albedo, and because of nonlinearities in the dependence of cloud radiative properties on CCN concentrations. Temporal-distribution information is needed because of the covariance of aerosol concentrations and removal processes with diurnaly, synoptically, and seasonally varying features of the earth-atmosphere system. Once the spatial and temporal dependences of the pertinent aerosol properties are known, these dependences can be incorporated in dynamical models of the global atmosphere to evaluate the magnitude of the perturbation in forcing and in associated climatic changes due to anthropogenic aerosols.

In view of these details of the aerosol forcing problem and the major unsolved problems currently precluding a purely theoretical, model-based description of this forcing, a semiempirical approach seems most likely to succeed in the near-term. Such an approach would utilize empirically determined aerosol size distributions, dependence of size distribution on relative humidity, and relations between size distributions of aerosols and their optical and cloud nucleation properties, in conjunction with models that explicitly calculate aerosol mass loading and composition, taking into account sources of precursor gases, transport, gas-to-particle conversion, and deposition. This approach could thus build on the foundation that has been developed in studies of the atmospheric chemistry of sulfur compounds conducted in support of issues such as acid deposition and visibility reduction (for example, (53)).

Better quantification of aerosol radiative forcing and of its relation to anthropogenic emissions requires global representative measurements of aerosol and cloud properties. Long-term monitoring of the aerosol optical depth and of effective particle size can be achieved by satellites equipped with spectropolarimeters to measure...
the radiance and polarization of reflected sunlight in the spectral range from the near-ultraviolet to the near-infrared (70). Simultaneous monitoring of global cloud properties, specifically, cloud height, cloud optical depth, and particle phase and effective size, may be obtained from the same satellites. Such empirical aerosol and cloud data, including year-to-year changes and long-term trends, could provide a wide variety of tests for models of aerosol and cloud physics. However, evaluation and improvement of inferences drawn from satellite measurements require that the measurements are tied to concurrent ground-based measurements of optical properties and to in situ surface and aircraft measurements of chemical and microphysical properties (71). The combination of such measurements would permit rigorous testing of model calculations of forcing necessary to give confidence in their use as climate models.

Summary

Taken as an integrated body of information, the arguments and data presented above lead to the conclusion that anthropogenic sulfate produces a mean radiative forcing of the NH climate that is comparable in magnitude but opposite in sign to the anthropogenic perturbation in forcing by greenhouse gases. The direct aerosol and indirect cloud components each contribute a globally averaged forcing of roughly −1 W m⁻² that is confined predominantly to the NH. However, aerosol forcing differs greatly in nature from greenhouse gas forcing. Such differences make it mandatory that aerosol influences are explicitly represented in climate models rather than simply represented as a scalable anti-greenhouse influence (72). The large uncertainties in the magnitude and geographical distribution of aerosol forcing, much larger than those associated with greenhouse gas forcing, indicate the need for substantial research to improve the description of aerosol forcing. Among the components of such a research program would be a new set of satellites dedicated to questions of radiative forcing and able to provide the needed global coverage and high frequency of sampling. Connection of this global data set to local and column integral optical, chemical, and microphysical properties requires a set of ground-based observing sites carefully located to delineate the key characteristics of air masses that are either influenced or unaffected by anthropogenic sulfate. Understanding the chemical and physical processes that produce the aerosol and control its relevant properties requires both laboratory and field studies. These research activities must be integrated. Ground-based, airborne, satellite, and laboratory data acquisition activities should be coordinated, and modeling is needed at a variety of scales, from the mesoscale to the global scale, to aid in the design of measurement strategies, to provide a framework for data analysis, and to examine the climatic impact of anthropogenic aerosols.

Key questions remain unanswered regarding the climate influence of anthropogenic aerosol and its relation to the climate influence of increased concentrations of greenhouse gases: To what degree do such physically, altitudinally, geographically, and temporally different forcings compensate each other? Do such disparate forcings produce identical but opposite meteorological and climatological responses? Have the radiative effects of sulfate aerosol decreased the magnitude of warming from anthropogenic greenhouse gases or delayed its onset or both? Or are there other climatic influences of anthropogenic aerosol that might have offset the radiative influences described here? To what extent will reductions in SO₂ emissions for control of acid deposition exacerbate the greenhouse effect? Perhaps the sobering reality of unanswered questions as these will remind the research and policy communities that relating climate response to anthropogenic perturbations is still a long way from being an exercise in engineering design.
Addictive Drugs: The Cigarette Experience

THOMAS C. SCHELLING

Cigarettes are among the most addictive substances of abuse and by far the most deadly. In this country smokers know it and try to stop. Their success has been dramatic but partial and excruciatingly slow, and until recently quite uncoerced by government. Cigarettes and nicotine have characteristics distinct among addictive drugs, and some of these help explain why efforts to quit smoking are so often frustrated. Nicotine itself is the most interesting chemical in the treatment of addiction and, in some forms, can pose a dilemma: compromise by settling for pure nicotine indefinitely, or stay with cigarettes and keep trying to quit. Nicotine is not alone among addictive drugs in becoming increasingly identified with the poorer classes.

HALF THE MEN WHO EVER SMOKE IN THIS COUNTRY have quit, and nearly half the women. At the end of World War II, three-quarters of young men smoked; the fraction is now less than a third and going down. Fifty million people have quit smoking, and another 50 million who would have become smokers since 1945 did not.

This dramatic abandonment of a life-threatening behavior was entirely voluntary. Until recently there was virtually no regulation of smoking by any level of government. The situation changed sharply in the late 1980s after dramatic changes in smoking behavior were well under way.

Surveys documented that the public was aware of the risks (1). Ninety percent or more answered yes to whether smoking caused cancer and heart disease. The facts were impressive. In 1982 the Surgeon General estimated 130,000 premature cancer deaths, in 1983 170,000 deaths from heart disease, and in 1984 50,000 deaths from lung disease (2). The total was later increased to more than 400,000.

Where do people learn about these dangers? Newspapers reported the annual reports of the surgeon general, but smoking was rarely news and inherently a dull subject. Only recently have city ordinances, airline restrictions, liability suits, advertising bans, and excise taxes made cigarettes occasional front-page news. Magazines rarely mention smoking; some of the most popular magazines report more than 25% of their advertising revenues from cigarettes.

The only emphatic repetitive communications about the hazards of smoking are the advertisements on billboards, and in magazines and newspapers. For two decades the central theme has been tar and nicotine. The message sent is that lighter cigarettes are safer but the message received must also be that smoking is dangerous. It is anybody's guess whether the cumulative impact is to entice people into smoking and to keep them smoking, or to drive home the deadly message about tar and nicotine.

No surgeon general has ever publicized the benefits of lower tar and nicotine, but the tar and nicotine have fallen by half. Smokers can infer that the government would not require labeling unless tar and nicotine made a difference.

Thirty years ago smoking was not much associated with social

The author is in the Department of Economics and the School of Public Affairs, University of Maryland, College Park, MD 20742, and was formerly director of the Institute for the Study of Smoking Behavior and Policy.