Global indirect aerosol effects: a review

U. Lohmann and J. Feichter

1 ETH Institute of Atmospheric and Climate Science, Schafmattstr. 30, CH-8093 Zurich, Switzerland
2 Max Planck Institute for Meteorology, Bundesstr. 53, D-20146 Hamburg, Germany

Received: 7 October 2004 – Published in Atmos. Chem. Phys. Discuss.: 17 November 2004
Revised: 28 January 2005 – Accepted: 18 February 2005 – Published: 3 March 2005

Abstract. Aerosols affect the climate system by changing cloud characteristics in many ways. They act as cloud condensation and ice nuclei, they may inhibit freezing and they could have an influence on the hydrological cycle. While the cloud albedo enhancement (Twomey effect) of warm clouds received most attention so far and traditionally is the only indirect aerosol forcing considered in transient climate simulations, here we discuss the multitude of effects. Different approaches how the climatic implications of these aerosol effects can be estimated globally as well as improvements that are needed in global climate models in order to better represent indirect aerosol effects are discussed in this paper.

1 Introduction

Anthropogenic aerosol particles such as sulfate and carbonaceous aerosols have substantially increased the global mean burden of aerosol particles from preindustrial times to the present-day. Aerosol particles affect the climate system via the following physical mechanisms: First, they scatter and can absorb solar radiation. Second, they can scatter, absorb and emit thermal radiation. Third, aerosol particles act as cloud condensation nuclei (CCN) and ice nuclei (IN). The first two mechanisms are referred to as direct effects and are not subject of this paper but are discussed in detail in e.g., Haywood and Boucher (2000). The last one is referred to as indirect effect. It will be the subject of this review together with other atmospheric properties influenced by aerosols (e.g. semi-direct effect, suppression of convection). Even though the semi-direct effect is a consequence of the direct effect of absorbing aerosols, it changes cloud properties in response to these aerosols and therefore is part of this review on aerosol-cloud-interactions.

Clouds themselves are an important regulator of the Earth’s radiation budget. About 60% of the Earth’s surface is covered with clouds. Clouds cool the Earth-atmosphere system on a global average basis at the top-of-the-atmosphere. Losses of 48 W m$^{-2}$ in the solar spectrum are only partially compensated (30 W m$^{-2}$) by trapped infrared radiation. Measurements of the Earth Radiation Budget Experiment (ERBE) (Collins et al., 1994) indicate that small changes to macrophysical (coverage, structure, altitude) and microphysical properties (droplet size, phase) have significant effects on climate. For instance a 5% increase of the short-wave cloud forcing would compensate the increase in greenhouse gases between the years 1750–2000 (Ramaswamy et al., 2001). Consequently the growing interest in the impact of aerosols on climate stimulated the development of better physically based parameterizations in climate models. Nevertheless, the lack of understanding feedbacks of external forcings on clouds remains one of the largest uncertainties in climate modeling and climate change prediction (Cess et al., 1990; Houghton et al., 1996).

A summary of the different anthropogenic aerosol effects on clouds is given in Table 1 while the effects are discussed in detail in the subsequent chapters. Most transient climate model simulations allow for a cooling by aerosols in order to achieve good agreement with the observed temperature record. However, these studies usually ignore aerosol indirect effects beyond the Twomey effect (Roeckner et al., 1999; Boer et al., 2000). Here we illustrate that radiative forcings of other indirect aerosol effects exist and need to be considered in future transient simulations. A positive forcing is associated with a warming or energy gain of the Earth-atmosphere system while a negative forcing represents a cooling or energy loss. When available from the literature, we focus on the global aspect of these various anthropogenic indirect aerosol effects because a review of all regional studies on indirect aerosol effects would be beyond the scope of this study. We concentrate on studies that have been published since the 2001 International Panel on Climate Change (IPCC) report.

Correspondence to: U. Lohmann (ulrike.lohmann@env.ethz.ch)

© 2005 Author(s). This work is licensed under a Creative Commons License.
Table 1. Overview of the different aerosol indirect effects and range of the radiative budget perturbation at the top-of-the-atmosphere ($F_{T OA}$) [W m$^{-2}$], at the surface ($F_{S FC}$) and the likely sign of the change in global mean surface precipitation (P) as estimated from Fig. 2 and from the literature cited in the text.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Cloud type</th>
<th>Description</th>
<th>$F_{T OA}$</th>
<th>$F_{S FC}$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect aerosol effect for clouds with fixed water amounts (cloud albedo or Twomey effect)</td>
<td>All clouds</td>
<td>The more numerous smaller cloud particles reflect more solar radiation</td>
<td>$-0.5$</td>
<td>similar</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$-1.9$</td>
<td>to</td>
<td>$F_{T OA}$</td>
</tr>
<tr>
<td>Indirect aerosol effect with varying water amounts (cloud lifetime effect)</td>
<td>All clouds</td>
<td>Smaller cloud particles decrease the precipitation efficiency thereby prolonging cloud lifetime</td>
<td>$-0.3$</td>
<td>similar</td>
<td>decrease</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$-1.4$</td>
<td>to</td>
<td>$F_{T OA}$</td>
</tr>
<tr>
<td>Semi-direct effect</td>
<td>All clouds</td>
<td>Absorption of solar radiation by soot may cause evaporation of cloud particles</td>
<td>$+0.1$</td>
<td>larger</td>
<td>decrease</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$-0.5$</td>
<td>than</td>
<td>$F_{T OA}$</td>
</tr>
<tr>
<td>Thermodynamic effect</td>
<td>Mixed-phase clouds</td>
<td>Smaller cloud droplets delay the onset of freezing</td>
<td>?</td>
<td>?</td>
<td>increase or decrease</td>
</tr>
<tr>
<td>Glaciation indirect effect</td>
<td>Mixed-phase clouds</td>
<td>More ice nuclei increase the precipitation efficiency</td>
<td>?</td>
<td>?</td>
<td>increase</td>
</tr>
<tr>
<td>Rimming indirect effect</td>
<td>Mixed-phase clouds</td>
<td>Smaller cloud droplets decrease the riming efficiency</td>
<td>?</td>
<td>?</td>
<td>decrease</td>
</tr>
<tr>
<td>Surface energy budget effect</td>
<td>All clouds</td>
<td>Increased aerosol and cloud optical thickness decrease the net surface solar radiation</td>
<td>n/a</td>
<td>$-1.8$</td>
<td>decrease</td>
</tr>
</tbody>
</table>

2 Aerosol effects on water clouds

The IPCC Third Assessment Report concluded that the Twomey effect of anthropogenic aerosol particles amounts to 0 to $-2$ W m$^{-2}$ in the global mean (Ramaseswamy et al., 2001). The Twomey effect refers to the enhanced reflection of solar radiation due to the more but smaller cloud droplets in a cloud whose liquid water content remains constant (Twomey, 1959). Based on studies since the 2001 IPCC report as shown in Fig. 1, the upper negative bound is slightly reduced to $-1.9$ W m$^{-2}$. On the other hand, there is no climate model that suggests that the Twomey effect is close to zero, but the smallest cooling is $-0.5$ W m$^{-2}$ (Table 1).

In addition, the more but smaller cloud droplets reduce the precipitation efficiency and therefore enhance the cloud lifetime and hence the cloud reflectivity, which is referred to as the cloud lifetime or second indirect effect (Albrecht, 1989). This effect is estimated to be roughly as large as the Twomey effect as will be discussed below. Absorption of solar radiation by aerosols leads to a heating of the air, which can result in an evaporation of cloud droplets. It is referred to as semi-direct effect (Grasl, 1979; Hansen et al., 1997). This warming can partially offset the cooling due to the indirect aerosol effect. Conversely, as shown by Penner et al. (2003), Johnson et al. (2004) and indicated in Table 1 the semi-direct effect can result in a cooling depending on the location of the black carbon with respect to the cloud as discussed in chapter 6. Both the cloud lifetime effect and the semi-direct effect involve feedbacks because the cloud lifetime and cloud liquid water content change. Therefore they were not included in the radiative forcing bar chart of the IPCC (2001) assessment.

2.1 Evidence of aerosol effects on warm clouds from observational data

The indirect aerosol effect of changing cloud albedo and cloud lifetime due to anthropogenic emissions of aerosols and their precursors has been evaluated from observational studies, starting with observations of ship tracks perturbing marine stratus cloud decks off the coast of California, e.g. Ferek et al. (1998) and lately also over continental areas (Feingold et al., 2003; Penner et al., 2004). Investigations by Brenguier et al. (2000) and Schwartz et al. (2002) over the Atlantic Ocean showed that the cloud droplets were smaller in the polluted clouds than in the clean clouds. This contrast between polluted and clean clouds is partially offset because both papers found that the polluted clouds were thinner as they originated over the continents, which causes them to be drier than their counterpart marine clean clouds (Lohmann and Lesins, 2003). Since the cloud albedo depends on both the cloud droplet size and the cloud thickness.
These competing effects partially cancel each other making it more difficult to detect an indirect aerosol effect.

These systematic differences in cloud thickness between clean and polluted clouds also affect the correlation between optical thickness and effective radius as investigated by Brenguier et al. (2003). This correlation is negative, as anticipated by Twomey (1977) if only cases of comparable values of geometrical thickness are considered. On the other hand, if the most polluted cases are also accounted for, the trend suggests a positive correlation, because the most polluted cloud systems sampled during ACE-2 were slightly drier, hence thinner, than the marine and intermediate cases. Likewise, Peng et al. (2002) showed that the slope between optical thickness and effective radius is positive for polluted clouds due to the increase in liquid water content and absence of drizzle size drops and vice versa for clean clouds.

Feingold et al. (2003) studied the indirect aerosol effect from ground-based remote sensing at the Atmospheric Radiation Measurement (ARM) site in Oklahoma using observations of subcloud Raman lidar aerosol extinction $\alpha$ at 355 nm and cloud droplet effective radius to define the aerosol indirect effect (IE) as the partial derivative of the logarithm of cloud droplet radius with respect to the logarithm of the aerosol extinction:

$$IE = -\frac{\partial \ln r_e}{\partial \ln \alpha}$$

Feingold et al. (2003) obtained IE values between 0.07 and 0.11 over the ARM site for liquid water paths between 100 and 130 g m$^{-2}$. They showed that for a homogeneous cloud with a constant liquid water content for which cloud optical depth is proportional to $N_d^{1/3}$, one can bracket IE to be between 0 and 0.33. This derivative of the indirect aerosol effect can be used for model validation. For example, model simulations by Lohmann and Lesins (2003) obtained larger slopes than observed, suggesting an overestimate of the indirect aerosol effect. There is, however, some uncertainty in this estimate related to different observing platforms. The estimate of the indirect effect by Feingold et al. (2003) at the ARM site is larger than estimated from POLDER satellite data by Bréon et al. (2002). Rosenfeld and Feingold (2003) pointed out that limitations of the POLDER satellite retrieval could explain this discrepancy. Penner et al. (2004) combined ARM data together with a Lagrangian parcel model at the ARM sites in Oklahoma as a surrogate for a polluted site and Alaska as a surrogate for a clean site to provide observational evidence of a change in radiative forcing due to the anthropogenic indirect aerosol effect.

Long-term observations from satellites over Europe and China show evidence for the semi-direct effect, i.e. a reduction in planetary albedo that can be attributed to absorbing aerosols in winter (Krüger and Graßl, 2002, 2004; Krüger et al., 2004). In summer, on the other hand, when more sulfate is produced, the Twomey effect is larger. So far, none of the techniques used to derive the indirect aerosol effect purely from observations permits estimates of the anthropogenic indirect aerosol effect globally.

2.2 Global estimates of indirect aerosol effects on warm clouds

Aerosol indirect effects are estimated from general circulation models (GCMs) by conducting a present-day simulation and a pre-industrial simulation in which the anthropogenic emissions are set to zero. The difference in the top-of-the-atmosphere radiation budget of these multi-year simulations is then taken to be the anthropogenic indirect aerosol effect. The aerosol mass or number is then either empirically related to the cloud droplet number concentration (Boucher and Lohmann, 1995; Menon et al., 2002a) or is obtained by using a physically-based parameterization (Abdul-Razzak and Ghan, 2002; Nenes and Seinfeld, 2003). Warm clouds form precipitation-size particles by the collision/coalescence process. In GCMs this is divided into the autoconversion (collisions and coalescence among cloud droplets) and the accretion of rain drops with cloud droplets. The former is either solely a function of the liquid water content (Sundqvist, 1978) and the cloud droplet size or concentration (Khairout-
2.3 Twomey effect

The estimates of the global mean Twomey effect and its division into the Northern and Southern Hemispheres are shown in Fig. 1. Note that the definition of the Twomey effect is not unique. While Chuang et al. (2002), Rotstayan and Liu (2003) and Quaas et al. (2004) define the Twomey effect as the net change in the shortwave flux at the top-of-the-atmosphere, Menon et al. (2002a) defined the Twomey effect in terms of the change in the net cloud radiative forcing at the top-of-atmosphere. The difference between these definitions is small because the contribution of the longwave radiation to the Twomey effect is below 0.1 W m\(^{-2}\) (Menon et al., 2002a; Rotstayan and Penner, 2001). Also the clear-sky radiation will remain the same in the absence of changes in temperature and differences in ice and snow cover between pre-industrial and present-day conditions. This latter constraint does not apply any longer when feedback processes are included. In general, the ratio of cooling of the Northern Hemisphere to the cooling of the Southern Hemisphere is larger when only sulfate aerosols are considered because biomass burning is only a minor source for sulfate but a large source for carbonaceous aerosols (Fig. 1). The lowest ratio is simulated by Quaas et al. (2004) who used the empirical Boucher and Lohmann (1995) relationship but using the maximum of the three hydrophilic species (sulfate, sea salt, organic carbon) instead of just sulfate aerosols as a surrogate for all species. Here sea salt has a more prominent role causing the anthropogenic emissions on the Southern Hemisphere to only play a minor role.

2.4 Twomey versus cloud lifetime effect

Climate model estimates of the cloud lifetime effect and the semi-direct aerosol effect are at least as uncertain as of the Twomey effect. As shown in Fig. 2, Kristjánsson (2002) and Williams et al. (2001) concluded that the Twomey effect at the top-of-the atmosphere is four times as important as the cloud lifetime effect whereas Lohmann et al. (2000), Ghan et al. (2001) and Quaas et al. (2004) simulated a cloud lifetime effect that is larger than the Twomey effect. This discrepancy is independent of the chemical nature of the anthropogenic aerosol species that are used in these different simulations. Likewise, the estimates of both indirect aerosol effects are smallest for the climate models that use the most anthropogenic species (Fig. 2). One reason for the large aerosol indirect effects obtained by Menon et al. (2002a) could be due to their empirical treatment between the aerosol mass and the cloud droplet number because sensitivity simulations by Lohmann et al. (2000) yielded a higher total indirect effect when an empirical relationship instead of a mechanistic relationship was used. However, other models that use an empirical relation such as Williams et al. (2001) obtain a smaller indirect aerosol effect. Another reason for the discrepancy between models could be the dependence of the indirect aerosol effect on the background aerosol concentration. Sensitivity studies by Lohmann et al. (2000) showed that reducing the minimum number of cloud droplets (which can be regarded as a surrogate for the background aerosol number concentration) from 40 cm\(^{-3}\) to 10 cm\(^{-3}\) increased the indirect aerosol effect from \(-1.1\) W m\(^{-2}\) to \(-1.9\) W m\(^{-2}\). Likewise differences in the cloud microphysics scheme, especially in the autoconversion rate, cause uncertainties in the indirect aerosol effect estimates (Lohmann and Feichter, 1997; Jones et al., 2001; Menon et al., 2002a, 2003).

Fig. 2. Global mean Twomey effect, lifetime effect, both effect and the ratio lifetime effect/Twomey effect of anthropogenic sulfate aerosols (red bars) from Williams et al. (2001), Rotstayan and Penner (2001), Ghan et al. (2001) and Jones et al. (2001), of anthropogenic sulfate and black carbon (green bars) from Kristjánsson (2002), of anthropogenic sulfate and organic carbon (blue bars) from Menon et al. (2002a); Quaas et al. (2004), of anthropogenic sulfate and black, and organic carbon (turquoise bars) from Lohmann et al. (2000); Takemura et al. (2005) and the mean plus standard deviation from all simulations (olive bars). The results from Menon et al. (2002a) and Ghan et al. (2001) are taken to be the averages of the simulations for only the Twomey effect and for both effects.
2.5 Land versus ocean

Most models suggest that the total indirect effect is at least as large over land as over the oceans (Fig. 3). The only exception are the simulations by Rostayn and Penner (2001) in which different empirical formulas were used for relating sulfate mass as a surrogate for all aerosols to cloud droplet number (relationship A from Boucher and Lohmann (1995)). This parameterization causes clouds over oceans to be more susceptible to increases in CCN. The flatter form of this parameterization over land than over ocean is broadly consistent with the idea that continental clouds are less susceptible to the effects of anthropogenic increases in CCN, because there are more natural CCN over land than over ocean. This conclusion is also consistent with the estimate of the total indirect aerosol effect as derived from combining POLDER satellite data and ECHAM4 GCM results (Lohmann and Lesins, 2002) that suggest a larger indirect aerosol effect over the oceans than over land. These satellite data, nevertheless, have to be viewed with caution, because the retrieval from the POLDER satellite is limited to clouds with a rather narrow cloud droplet size distribution that produce a glory (Rosenfeld and Feingold, 2003).

2.6 Constraints on the indirect aerosol effect

The cooling from both indirect effects on water clouds of sulfate and carbonaceous aerosols has been estimated from climate models since the last IPCC report to be \(-1.2 \text{ to } -4.4 \text{ W m}^{-2}\) in the global mean (Ghan et al., 2001; Jones et al., 2001; Lohmann and Feichter, 2001; Williams et al., 2001; Menon et al., 2002a). This is larger than estimated from inverse calculations which start from historical climate record data of oceanic and atmospheric warming. They typically use ensembles of simulations with climate models of reduced complexity and estimate a smaller anthropogenic indirect aerosol effect within the range of \(0 \text{ to } -2 \text{ W m}^{-2}\) (Forrest et al., 2002; Knutti et al., 2002; Anderson et al., 2003). If internal variability is thought to be averaged out over the anthropocene, then the total aerosol effect can solely be deduced from the greenhouse gas forcing and the increase in land surface temperature and ocean heat content (Crum et al., and Ramanathan, 2003). They obtain a cooling effect of aerosols between \(-0.7 \text{ and } -1.7 \text{ W m}^{-2}\) similar to the values obtained by inverse models. The constraints from these inverse models or thermodynamic considerations are, however, not restricted to the indirect aerosol effect on water clouds only even though it is traditionally understood in this context. Instead, the range encompasses all indirect aerosol effects and other effects currently not included in climate models. We will revisit this issue in the conclusions and outlook section.

Sekiguchi et al. (2002) used different correlations between aerosol and cloud parameters derived from satellite remote sensing to estimate the radiative forcing of the aerosol indirect effect. Assuming that the column aerosol number concentration increased by 30%, the total global mean indirect effect on warm clouds is estimated to be between \(-0.6 \text{ and } -1.2 \text{ W m}^{-2}\). A smaller indirect aerosol effect is also obtained when constraining the total indirect aerosol effect by taking the difference in the slope of the cloud droplet effective radius-aerosol index relationship between the POLDER satellite data (Bréon et al., 2002) and the ECHAM GCM results into account (Lohmann and Lesins, 2002). This reduces the total global mean aerosol effect from \(-1.4 \text{ W m}^{-2}\) to \(-0.85 \text{ W m}^{-2}\). Indirect evidence for the existence of a cloud lifetime effect on a global scale was reached by Suzuki et al. (2004) when comparing simulations with and without a cloud lifetime effect with AVHRR satellite data of liquid water path as a function of column aerosol number (Nakajima et al., 2001).

2.7 Dispersion effect

Liu and Daum (2002) estimated that the magnitude of the Twomey effect can be reduced by 10–80% by including the influence that an increasing number of cloud droplets has
on the shape of the cloud droplet spectrum (dispersion effect). Taking this dispersion effect in global climate models into account, this reduction is rather moderate and amounts only to 15–35% (Peng and Lohmann, 2003; Rotstayn and Liu, 2003). Rotstayn and Liu (2005) obtained a similar reduction also for the cloud lifetime effect when including the dispersion effect.

2.8 Semi-direct aerosol effect

Lohmann and Feichter (2001); Kristjánsson (2002), and Penner et al. (2003) concluded that the semi-direct effect is only marginally important at the top of the atmosphere in the global mean whereas Jacobson (2002) pointed out that the climatic effect of black carbon is strongly positive. The influence of black carbon is dominated via its absorption of solar radiation within the atmosphere, which also leads to a large negative forcing at the surface. The net reduction in shortwave radiation at the surface from all aerosol direct and indirect effects is estimated to be between −1.8 and −4 W m⁻² (Ramanathan et al., 2001a; Lohmann and Feichter, 2001; Liepert et al., 2004).

Fig. 4. Schematic diagram of the warm indirect aerosol effect (solid arrows) and glaciation indirect aerosol effect (dotted arrows) (adapted from Lohmann, 2002a). CDNC denotes the cloud droplet number concentration and IP the number concentration of ice particles.

3 Aerosol effects on mixed-phase clouds

3.1 Aerosol effects on large-scale mixed-phase clouds

Since most precipitation originates via the ice phase (Lau and Wu, 2003), aerosol effects on ice clouds might have larger consequences for the hydrological cycle than aerosol effects on water clouds. Precipitation originating from supercooled liquid water clouds where the temperatures are too warm for homogeneous freezing of supercooled aerosols or cloud droplets to occur (Tₜ > −35°C) requires an aerosol surface to provide a substrate for ice initiation. This influence of aerosol particles on changing the properties of ice forming nuclei (IN) is poorly understood because of the variety of heterogeneous ice crystal nucleation modes. Aerosols can act as IN by coming into contact with supercooled cloud droplets (contact freezing), or by initiating freezing from within a cloud droplet by immersion or condensation freezing, or by acting as deposition nuclei. Ice nuclei that initiate freezing are also referred to as freezing nuclei. Contact nucleation is usually the most efficient process at slight supercoolings, while at lower temperatures immersion freezing can be more prevalent. Deposition nuclei are generally least efficient because the energy barrier that needs to be overcome for the phase change of water vapor to ice is larger than that required for the freezing nuclei modes. Unlike CCN, ice nuclei are generally insoluble particles, such as certain mineral dusts, soot, as well as some biological materials, e.g., Levin and Yankofsky (1983); Diehl et al. (2001); Gorbunov et al. (2001). Ice nuclei may lose their nucleability, if foreign gases such as sulfur dioxide (SO₂) or ammonia (NH₃) occupy their active sites (Pruppacher and Klett, 1997).

If some cloud droplets freeze in a supercooled water cloud, then ice crystals will grow at the expense of cloud droplets because of the lower saturation vapor pressure over ice than over water (the so-called Bergeron-Findeisen process). This leads to a rapid glaciation of the supercooled water cloud. Because the precipitation formation via the ice phase is more efficient than in warm clouds, these glaciated clouds have a shorter lifetime than supercooled water clouds (Rogers and Yau, 1989).

Lohmann (2002a) showed that if, in addition to mineral dust, a fraction of the hydrophobic soot aerosol particles is assumed to act as contact ice nuclei at temperatures between 0°C and −35°C, then increases in aerosol concentration from pre-industrial times to present-day pose a new indirect effect, a "glaciation indirect effect", on clouds as shown in Fig. 4. Here increases in contact ice nuclei in the present-day climate result in more frequent glaciation of supercooled clouds and increase the amount of precipitation via the ice phase. This reduces the cloud cover and the cloud optical depth of mid-level clouds in mid- and high latitudes of the Northern Hemisphere and results in more absorption of solar radiation within the Earth-atmosphere system. Therefore, this effect can at least partly offset the cloud lifetime effect. Laboratory
measurements by Gorbunov et al. (2001) yield evidence for hydrophilic soot as ice nuclei. In addition, evidence of effective ice nuclei was recently measured with the continuous flow diffusion chamber when sampling Asian dust particles (DeMott et al., 2003). In case of Saharan African dust, mildly supercooled clouds at temperatures between −5 to −9°C were already glaciated (Sassen et al., 2003).

Observations by Borys et al. (2003) in midlatitude orographic clouds show that for a given supercooled liquid water content, both the riming and the snowfall rates are smaller if the supercooled cloud has more cloud droplets as, for example, caused by anthropogenic pollution. Examination of this effect in global climate model simulations with pre-industrial and present-day aerosol concentrations showed that while the riming rate in stratiform clouds has indeed decreased due to the smaller cloud droplets in polluted clouds, the snowfall rate has actually increased (Fig. 5). This is caused by the pollution induced increase in aerosol and cloud optical thickness, which reduces the solar radiation at the surface and causes a cooling that favors precipitation formation via the ice phase (Lohmann, 2004).

3.2 Aerosol effects on deep convective clouds (thermodynamic effect)

Andronache et al. (1999) showed that an increase in sulfate loading during the TOGA-COARE experiment causes a significant decrease of the effective radius of cloud droplets (changes up to 2 µm on average) and an increase in the number concentration of cloud droplets of 5–20 cm⁻³ over a limited domain of 500 km. The change in the average net shortwave radiation flux above the clouds was estimated to be on average −1.5 W m⁻², with significant spatial and temporal variations. The changes in the average net longwave radiation flux above the clouds were negligible, but significant variations between −10 W m⁻² and 10 W m⁻² near the surface associated with changes in cloud water path of about 10–20% were simulated.

Rosenfeld (1999) and Rosenfeld and Woodley (2000) analyzed aircraft data together with satellite data suggesting that pollution aerosols suppress precipitation by decreasing cloud droplet size. This hypothesis was confirmed by a modeling study with a cloud resolving model by Khain et al. (2001) who showed that aircraft observations of highly supercooled water in deep convective clouds can only be reproduced if large concentrations of small droplets exist but not if the cloud is rather clean. Taking these results to the global scale, Nober et al. (2003) evaluated the sensitivity of the general circulation to the suppression of precipitation by anthropogenic aerosols by implementing a simple warm cloud microphysics scheme into convective clouds. They found large instantaneous local aerosol forcings reducing the warm phase precipitation, but the precipitation change at the surface was guided by feedbacks within the system. Hence, no estimate of the aerosol forcing on convective clouds can be given.

Khain et al. (2004)¹ postulate that smaller cloud droplets, such as originating from anthropogenic activity, would reduce the production of drizzle drops. When these droplets freeze, the associated latent heat release results in more vigorous convection. In a clean cloud, on the other hand, drizzle would have left the cloud so that less latent heat is released when the cloud glaciates resulting in less vigorous convection. Therefore, no squall line is formed with maritime aerosol concentrations, but the squall line arises under continental aerosol concentrations and results in more precipitation after 2 h of simulations with a detailed cloud microphysics model. More precipitation from polluted clouds was also simulated for different three-week periods over the Atmospheric Radiation Measurement Program (ARM) site in Oklahoma (Zhang et al., 2005) as well as for multicell cloud systems by (Seifert and Beheng, 2005).

On the contrary, cloud resolving model simulations of mixed-phase shallow cumuli revealed decreasing precipitation efficiency with increasing atmospheric concentrations of CCN because of the dominance of warm-rain processes in these simulations (Phillips et al., 2002). Likewise precipitation from single mixed-phase clouds is reduced under continental and maritime conditions when aerosol concentrations are increased (Khain et al., 2004¹).

Tropical biomass burning aerosols could have led to a reduction of ice crystal size in tropical deep convective clouds (Sherwood, 2002). These smaller and more numerous ice crystals would then lead to more scattering of solar radiation, i.e., exert a Twomey effect as discussed above. However, no global mean radiative forcing was deduced by Sherwood for this effect. He used this hypothesis to explain the increase in stratospheric water vapor, which by being a greenhouse gas, provides a positive radiative forcing that would partially offset the Twomey effect associated with the smaller ice crystal size in these deep convective clouds. This demonstrates the complex interactions between the different forcing agents that need to be understood and the difficulties to disentangle forcings and feedbacks. Both issues will be revisited in Sect. 7.

4 Aerosol effects on ice clouds

Condensation (con) trails left behind jet aircrafts form when hot humid air from jet exhaust mixes with environmental air of low vapor pressure and low temperature. The mixing is a result of turbulence generated by the engine exhaust. Contrails cannot be distinguished any longer from cirrus clouds once they lose their line-shape. While there are only a few general studies on aerosol effects on cirrus, many investigations analyzed the effect of aircraft emissions on climate. Therefore we will discuss these two effects separately below.

4.1 Aerosol effects on contrails

The IPCC aviation report (Penner et al., 1999) identified the effects of aircraft on upper tropospheric cirrus clouds as a potentially important climate forcing. One aspect may be described as the “direct” effect due to the formation of contrails as a result of supersaturated air from the aircraft. The “indirect” effect is due to the impact of an increase in FN (freezing nuclei) in the upper troposphere regions due to particulates from aircraft emissions. These FN act as nuclei for ice crystals which form cirrus clouds. First evidence of a climate effect by air traffic was provided by Boucher (1999) who used ship based measurements of cloud cover together with fossil fuel consumption data for aircraft to show that increases in air traffic fuel consumption in the 1980s are accompanied by an increase in cirrus cloudiness. Evaluation of a longer surface dataset over the United States from 1971 to 1995 by Minnis et al. (2004) confirmed an increase in cirrus over the northern oceans and the United States. Statistically significant increases in cirrus cloud cover of more than 2% per decade were found in the summertime over the North Atlantic and in the wintertime over North America by analyzing satellite data from the International Satellite Cloud Climatology Project (ISCCP) (Zerefos et al., 2003). A similar increase of 2% cirrus cloud cover per decade was identified from analyzing ISCCP data over Europe (Stordal et al., 2004).

Ponater et al. (2002) and Marquart et al. (2003) studied the climate effect of contrails using a global climate model, but so far related the contrail formation only to relative humidity but did not link it to aerosol properties. The study by Lohmann and Kärcher (2002) that parameterized homogeneous freezing of supercooled aerosols suggests that the impact of aircraft sulfur emissions on cirrus properties via homogeneous freezing of sulfate aerosols is small. Hence the question has been raised whether aircraft-generated black carbon particles serving as heterogeneous ice nuclei (Ström and Ohlsson, 1998) may have a significant impact on cirrus cloudiness and cirrus microphysical properties.

Hendricks et al. (2004) performed climate model simulations that revealed that the large-scale impact of aviation black carbon (BC) emissions on the upper troposphere/lower stratosphere (UTLS) BC mass concentration is small. Nevertheless, the simulations suggest a significant aviation impact on the number concentrations of UTLS BC particles and potential heterogeneous IN (BC and mineral dust particles). Large-scale increases of the potential heterogeneous IN number concentration of up to 50% were simulated. Provided that BC particles from aviation serve as efficient heterogeneous IN, maximum increases or decreases in ice crystal number concentrations of more than 40% were simulated assuming that the “background” (no aviation impact) cirrus cloud formation is dominated by heterogeneous or homogeneous nucleation, respectively (Hendricks et al., 2003).

4.2 Aerosol effects on cirrus clouds

An increase in the number of ice crystals in cirrus clouds would also exert a Twomey effect in the same way that the Twomey effect acts for water clouds. In addition, a change in the ice water content of cirrus clouds could exert a radiative effect in the infrared. The magnitude of these effects in the global mean is not known yet. Lohmann and Kärcher (2002) concluded that such an effect based solely on homogeneous freezing is small because the number of ice crystals is rarely limited by the number of supercooled aerosols. Exceptions are areas of large vertical updrafts, such as the upper tropical troposphere. A testbed for an aerosol effect on cirrus clouds is the Mt. Pinatubo eruption in 1991. Global climate model results suggest that effects from the Mt. Pinatubo eruption on clouds and climate considering only homogeneous freezing are small (Lohmann et al., 2003). These findings are consistent with the newer satellite analysis of three satellite-based cirrus datasets by Luo et al. (2002) who found that the Mt. Pinatubo volcanic aerosol did not have a significant systematic effect on cirrus cloud coverage and the brightness temperature difference, which is a surrogate for cloud optical thickness.

Kärcher and Lohmann (2003) developed a parameterization for heterogeneous immersion freezing of cirrus clouds.
They concluded that if only one type of ice nuclei with saturation ratios over ice larger than 1.3–1.4 triggers cirrus formation, then the influence of aerosols on cirrus clouds is still small. However, a much stronger indirect aerosol effect on cirrus clouds is possible if several ice nuclei types with different freezing thresholds compete during the freezing process. Moreover, ice nuclei can significantly enhance the frequency of occurrence of subvisible cirrus clouds, even when present at concentrations as low as 0.01 particles L\(^{-1}\) (Kärcher, 2004).

5 Aerosol induced changes of the surface energy budget and aerosol effects on precipitation

By increasing aerosol and cloud optical depth, human emissions of aerosols and their precursors cause a reduction of solar radiation at the surface (“solar dimming”). Such a reduction has been observed in the industrial regions of the Northern Hemisphere (Gillgen et al., 1998; Liepert, 2002; Stanhill and Cohen, 2001; Wild et al., 2004a). According to Liepert (2002) this decline of solar radiation from 1961 to 1990 amounts to 1.3% per decade over land surfaces of the Northern Hemisphere. In order for the surface energy balance to reach a new equilibrium state, the surface energy budget has to adjust:

\[
F_{sw} = F_{lw} + F_l + F_s + F_{cond}
\] (2)

Here \(F_{sw}\) is the net shortwave radiation available at the surface. This incoming energy has to be balanced by the net outgoing longwave radiation \(F_{lw}\), the latent heat flux \(F_l\), the sensible heat flux \(F_s\) and the conductive flux from below the surface \(F_{cond}\).

As shown in model simulations by Liepert et al. (2004) and Feichter et al. (2004) that use a global climate model coupled to a mixed-layer ocean model with increasing aerosol particles and greenhouse gases due to human activity, the decrease in solar radiation at the surface resulting from the increases in optical depth due to the direct and indirect anthropogenic aerosol effects is more important for controlling the surface energy budget than the greenhouse gas induced increase in surface temperature. The conductive flux from below the surface is negligible in the long-term mean. The three other components of the surface energy budget decrease in response to the reduced input of solar radiation. This mechanism could explain the observations of decreased pan evaporation over the last 50 years reported by Roderick and Farquhar (2002). As evaporation has to equal precipitation on the global scale, a reduction in the latent heat flux leads to a reduction in precipitation. Recent surface observations show that the decline in solar radiation at land surfaces disappears in the 1990s (Wild et al., 2004b)\(^2\). This is in agreement with recent emission trends in the “old” industrial regions in the northern hemisphere (Krüger and Graßl, 2002) as well as with long-term black carbon trends in the Canadian Arctic (Sharma et al., 2004). Thus, the increasing greenhouse effect may no longer be masked by an aerosol induced decline in solar radiation, resulting in the enhanced warming observed during the 1990s.

On a regional scale, smoke from sugarcane and forest fires was shown to reduce cloud droplet sizes and therefore tends to inhibit precipitation (Warner and Twomey, 1967; Warner, 1968; Eagan et al., 1974). Heavy smoke from forest fires in the Amazon Basin has been observed to increase cloud droplet number concentrations and to reduce cloud droplet sizes (Reid et al., 1999; Andreae et al., 2004). Andreae et al. (2004) suggested that this delayed the onset of precipitation from 1.5 km above cloud base in pristine clouds to more than 5 km in polluted clouds, and to more than 7 km in pyro-clouds. They suggested also that elevating the onset of precipitation released latent heat higher in the atmosphere and allowed invigoration of the updrafts, causing intense thunderstorms and large hail. Together, these processes might affect the water cycle, the pollution burden of the atmosphere, and the dynamics of atmospheric circulation. Also, satellite data revealed plumes of reduced cloud particle size and suppressed precipitation originating from some major urban areas and from industrial facilities such as power plants (Rosenfeld, 2000). However, precipitation from similar polluted clouds over oceans appears to be much less affected, possibly because giant sea salt nuclei override the precipitation suppression effect of the large number of small pollution nuclei (Feingold et al., 1999; Rosenfeld et al., 2002). Here, large droplets initiated by large sea salt aerosols may grow to precipitation size by collecting small cloud droplets, thereby cleansing the air. If these giant CCN are however covered by film-forming compounds, then their impact would be less than previously estimated (Medina and Nenes, 2004).

Observed precipitation trends over land for the period 1900–1998 show a complex pattern in the tropics but, indicating, for instance, a drying of the Sahel in North Africa (Hulme et al., 1998). Dry conditions in the Sahel are associated with a near-global, quasi-hemispheric pattern of contrasting sea surface temperature anomalies (cooler in the northern hemisphere and warmer in the southern hemisphere). Using a global climate model/mixed-layer ocean model Williams et al. (2001) and Rotstyn and Lohmann (2002) showed that the dynamical and hydrological changes in this region in response to the indirect effect of anthropogenic sulfate aerosols are similar to the observed changes that have been associated with the Sahelian drought (Pollard et al., 1986; Giannini et al., 2003). This is, in the model the anthropogenic aerosol cooling dominates on the Northern Hemisphere, which causes a southward shift of the intertropical convergence zone. If, on the contrary, the Northern Hemisphere surface temperature is increased more than the Southern Hemisphere surface temperature due to the increase in
fossil fuel combustion of black carbon, then the intertropical convergence zone shifts northward, strengthens the Indian summer monsoon and increases the rainfall in the Sahel (Roberts and Jones, 2004). Because the shortwave radiation at the surface decreases over the Northern Hemisphere both due to anthropogenic sulfate as well as due to fossil fuel black carbon, the surface shortwave radiation is not a good indicator of the expected changes in circulation and precipitation.

6 Aerosol effects on the vertical stability of the atmosphere

Changes in the atmospheric lapse rate modify the longwave emission, affect the water vapor feedback (Hu, 1996) and the formation of clouds. Observations and model studies show that an increase in the lapse rate produces an amplification of the water vapor feedback (Sinha, 1995). Model simulations by Feichter et al. (2004) show that aerosol cooling extends up to the tropopause with a maximum in the boundary layer of the northern mid and high latitudes. In the tropics aerosol cooling is at maximum in the upper troposphere (Fig. 6). The overall effect of the aerosol forcing is a cooling near the surface in the polluted regions of the Northern Hemisphere that stabilizes the lower atmosphere whereas the near surface changes in temperature are smaller in the tropics and the mid-latitudes of the Southern Hemisphere. The implications of these aerosol induced lapse rate changes on other climate feedbacks such as the water vapor feedback are, however, not quantified yet.

Likewise, a destabilization of the atmosphere above the boundary as a result of black carbon heating within the boundary layer was obtained in a climate model study by Menon et al. (2002b). Their GCM was driven by the observed aerosol optical depths over India and China. If the aerosols were assumed to be absorbing, the atmospheric stability above the boundary layer was reduced, resulting in enhanced vertical motion. This affected the large-scale circulation and produced precipitation pattern in China that resembled those associated with the floods and droughts that China has experienced in recent years.

Over the Indian Ocean region during the dry winter monsoon season it has been estimated that anthropogenic aerosols especially the highly absorbing aerosols can decrease the average solar radiation absorbed by the surface in the range of 15 to 35 W m⁻² (Ramanathan et al., 2001b). This results in an increase in the atmospheric heating between the surface and 3 km altitude by up to 60 to 100%. Similar perturbations in the atmosphere have been observed over other regions namely East Asia, South America, sub-Saharan Africa, which are subjected to large loading of absorbing aerosols. Such a perturbation imposed over the Indian Ocean in the 15° S–40° N and 50–120° E region can lead to a large regional cooling at the surface in the range of 0.5 to 1 K accompanied by a warming of the lower troposphere by about 1 K as has been deduced from a GCM study with fixed sea surface temperatures (Chung et al., 2002). This vertical heating gradient alters the latitudinal and inter-hemispheric gradients in solar heating and these gradients play a prominent role in driving the tropical circulation (Ramanathan et al., 2001b) and determining the amount of precipitation (Chung and Zhang, 2004).

Additionally, temperature changes due to absorbing aerosols can cause the evaporation of cloud droplets (semi-direct effect), as was shown in a large eddy model simulation study that used black carbon concentrations measured during the Indian Ocean Experiment (Ackerman et al., 2000). Recent LES simulations by Johnson et al. (2004) obtained a positive semi-direct effect only when the absorbing aerosol layer was situated within the boundary layer but obtained a negative semi-direct effect when the absorbing aerosol layer is situated above the cloud layer. The negative semi-direct effect is caused by the stronger inversion, which reduces the cloud-top entrainment rate resulting in a shallower, moister boundary layer with a higher liquid water path. Koren et al. (2004) deduced a strong anti-correlation of cloud cover and aerosol optical depth from MODIS satellite data over the Amazon. The data could suggest that smoke aerosols are responsible for the evaporation of cloud droplets.

7 Indirect aerosol effect – forcing or response?

Even though anthropogenic aerosol effects on the hydrological cycle through the aerosol lifetime effect or the surface energy budget involve fast feedbacks within the climate system and are therefore not considered a forcing in the “classical” sense, they pose a “forcing” on the hydrological cycle. Various ideas were brought forward to extend the classical
forcing concepts in order to include processes that immediately imply a feedback, such as the semi-direct effect or the indirect aerosol effect.

7.1 Classical definition of forcing

IPCC has summarized the global and annual mean radiative forcing from 1750 to the present-day due to recognized human-related and natural processes. It includes the contribution of the well-mixed greenhouse gases (GHG), ozone, aerosols, aviation induced climate perturbations, and the solar contribution. Here the term radiative forcing of the climate system is defined as: “The radiative forcing of the surface-troposphere system due to the perturbation in or the introduction of an agent is the change in net irradiance at the tropopause after allowing for stratospheric temperatures to readjust to radiative equilibrium, but with the surface and tropospheric temperatures and state held fixed at the unperturbed values.” The definition of forcing is thus restricted to changes in the radiation balance of the Earth-troposphere system imposed by external factors, with no changes in stratospheric dynamics, without any surface and tropospheric feedbacks in operation, and with no dynamically-induced changes in the amount and distribution of atmospheric water (as termed adjusted forcing $F_a$ by Hansen et al. (2002), see Fig. 7).

The instantaneous forcing $F_i$ (Fig. 7) is the radiative forcing at the top of the atmosphere $F$, that needs to be used to estimate the change in surface temperature $\Delta T_{sfc}$ in transient models. These two quantities are linked by the climate sensitivity parameter $\lambda$ defined as: $\lambda = \Delta T_{sfc} / F_i$. Provided that forcings due to different agents can be added linearly and that the climate sensitivity is constant, the response of the surface temperature due to a wide range of forcings can easily be estimated.

The concept of climate sensitivity holds for agents such as long-lived greenhouse gases and the direct effect of scattering aerosols. However, it breaks down once an absorbing aerosol such as black carbon is considered because its top-of-the-atmosphere forcing may be small, but because it absorbs large amounts of solar radiation, its surface forcing is disproportionately larger than one would extrapolate from the top-of-the-atmosphere forcing (Lohmann and Feichter, 2001; Ramanathan et al., 2001a). This definition of forcing also precludes all aerosol effects that comprise microphysically-induced changes in the water substance. Thus approaches have been developed which differ from a pure or instantaneous forcing in that fields other than the initially perturbed quantity have been allowed to vary.

7.2 New approaches

One new avenue for calculating the top-of-the-atmosphere forcing is the concept of fixed sea surface temperature forcing (see Fig. 7c), as introduced by Cess et al. (1990) and first applied to studies of the Twomey effect by Boucher and Lohmann (1995). Here small changes in land surface temperature could impose a climate response (Hansen et al., 20053). This concept was extended by Shine et al. (2003) to fixing land surface temperatures in addition to sea surface temperatures (fixed ground temperature forcing, see also Hansen et al. (2005)3). They argue that changes in temperatures over land and ocean are related and thus it is more consistent to fix surface temperatures globally. Doing so enables to separate between forcings that change atmospheric parameters and those that invoke surface temperature changes. For example, the forcing of black carbon on reducing cloudiness

---

(semi-direct effect) can now be isolated from the changes in surface temperature caused by black carbon.

The component of the indirect aerosol effect related to changes in precipitation efficiency (the cloud lifetime effect) is presently evaluated in climate models as the difference in net radiation at the top-of-the-atmosphere between a present-day and a preindustrial simulation using fixed sea surface temperatures. Rotstayn and Penner (2001) have shown that for the global-mean direct and Twomey indirect effects, the fixed sea surface temperature forcing differed by less than 10% from the corresponding pure forcing. Therefore the authors concluded that evaluation of the globally averaged cloud lifetime effect as a fixed sea surface temperature forcing is satisfactory.

Joshi et al. (2003) and Hansen and Nazarenko (2004) introduced the concept of efficacies of different forcing agents. “Efficacy” is defined as the ratio of the climate sensitivity parameter \( \lambda \) for a given forcing agent to that for a given change in CO\(_2\) (\( E = \lambda_i / \lambda_{CO_2} \)) as shown in Table 2. The efficacy is then used to define an effective forcing \( F_e = F E \). Table 2 evaluates this concept by summarizing the forcings, responses, efficacies and effective forcings of different forcing agents from equilibrium climate models simulations coupled to a mixed-layer ocean. Scattering sulfate aerosols are less efficient than well-mixed greenhouse gases in changing the surface temperature for a given forcing whereas the direct effect of all aerosols has an efficacy that is slightly larger than one (Table 2). If the direct and indirect effects on water clouds from all aerosol species are considered, the efficacy may be larger or smaller as compared to CO\(_2\) (Table 2).

### 7.3 Additivity of the different forcing agents

One implicit assumption is that forcings are additive, which may not be true. For instance, Rotstayn and Penner (2001) showed that the combined Twomey and cloud lifetime effect is smaller than the sum of the individual effects (Table 2). Feichter et al. (2004) showed that the global warming estimated from a global climate model/mixed layer ocean model due to a combined aerosol and greenhouse gas forcing is significantly smaller (0.57 K) than that obtained by adding the individual changes (0.85 K). Even more drastically, the positive global hydrological sensitivity per 1 K surface temperature change (\( \Delta P / \Delta T \)) in the scenarios with only aerosol forcing or greenhouse gas forcing changes into a negative hydrological sensitivity when both the aerosol and greenhouse gas forcings are applied simultaneously (see Table 3). Likewise if all aerosol effects and the greenhouse gas forcings are combined, then the efficacy is larger than expected from averaging the efficacy for greenhouse gases and for all aerosol effects individually. Therefore the concept of efficacies adds no benefit when trying to understand the non-linearity between forcing and response.

This is in contradiction to the results by Gillett et al. (2004) and Matthews et al. (2004) who did not find any non-linearities when adding the natural forcings volcanic

<table>
<thead>
<tr>
<th>Experiment</th>
<th>( F_i )</th>
<th>( T_{sfc} )</th>
<th>( \lambda )</th>
<th>( E )</th>
<th>( F_e )</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>WM-GHGs: 1860–1990</td>
<td>2.12</td>
<td>1.82</td>
<td>0.86</td>
<td>1</td>
<td>2.12</td>
<td>based on Roeckner et al. (1999)</td>
</tr>
<tr>
<td>2xCO(_2)</td>
<td>3.48</td>
<td>3.52</td>
<td>1.01</td>
<td>1</td>
<td>3.48</td>
<td>Rotstayn and Penner (2001)</td>
</tr>
<tr>
<td>1880–2000</td>
<td>2.52</td>
<td>1.21</td>
<td>0.48</td>
<td>1</td>
<td>2.52</td>
<td>Hansen et al. (2005)(^3)</td>
</tr>
<tr>
<td>Direct effect: SO(_4)</td>
<td>−0.34</td>
<td>−0.24</td>
<td>0.71</td>
<td>0.83</td>
<td>−0.28</td>
<td>based on Roeckner et al. (1999)</td>
</tr>
<tr>
<td>SO(_4)</td>
<td>−0.80</td>
<td>−0.55</td>
<td>0.69</td>
<td>0.68</td>
<td>−0.55</td>
<td>Rotstayn and Penner (2001)</td>
</tr>
<tr>
<td>all AP</td>
<td>−0.54*</td>
<td>−0.28</td>
<td>0.52</td>
<td>1.08</td>
<td>−0.58</td>
<td>Hansen et al. (2005)(^3)</td>
</tr>
<tr>
<td>Twomey effect: SO(_4)</td>
<td>−0.89</td>
<td>−0.78</td>
<td>0.87</td>
<td>1.01</td>
<td>−0.90</td>
<td>based on Roeckner et al. (1999)</td>
</tr>
<tr>
<td>SO(_4)</td>
<td>−1.35</td>
<td>−1.14</td>
<td>0.84</td>
<td>0.83</td>
<td>−1.12</td>
<td>Rotstayn and Penner (2001)</td>
</tr>
<tr>
<td>all AP</td>
<td>−0.77*</td>
<td>−0.38</td>
<td>0.49</td>
<td>1.03</td>
<td>−0.79</td>
<td>Hansen et al. (2005)(^3)</td>
</tr>
<tr>
<td>Lifetime effect: SO(_4)</td>
<td>−1.32*</td>
<td>−1.04</td>
<td>0.79</td>
<td>0.78</td>
<td>−1.03</td>
<td>Rotstayn and Penner (2001)</td>
</tr>
<tr>
<td>all AP</td>
<td>−1.01*</td>
<td>−0.45</td>
<td>0.45</td>
<td>0.93</td>
<td>−0.94</td>
<td>Hansen et al. (2005)(^3)</td>
</tr>
<tr>
<td>Total indirect: SO(_4)</td>
<td>−2.57*</td>
<td>−2.24</td>
<td>0.87</td>
<td>0.86</td>
<td>−2.21</td>
<td>Rotstayn and Penner (2001)</td>
</tr>
<tr>
<td>All aerosol effects (direct and indirect on water clouds)</td>
<td>−1.40*</td>
<td>−0.87</td>
<td>0.62</td>
<td>0.72</td>
<td>−1.01</td>
<td>Feichter et al. (2004), ( F_i ) from Lohmann and Feichter (2001)</td>
</tr>
<tr>
<td></td>
<td>−2.34*</td>
<td>−1.15</td>
<td>0.49</td>
<td>1.02</td>
<td>−2.40</td>
<td>Hansen et al. (2005)(^3)</td>
</tr>
</tbody>
</table>
aerosols, solar insolation variability and orbital changes and the anthropogenic forcings greenhouse gases and sulfate aerosols. Therefore, it appears that the non-linearity is caused by the either absorbing aerosols or aerosol effects on the hydrological cycle, because an enhanced cloud lifetime and reduced precipitation efficiency reduces the wet removal of aerosols, thus, prolonging the lifetime of aerosols. Along the same lines Hansen et al. (2005) conclude that the cooling estimated from the sum of the indirect effects of the individual aerosol components is smaller than the cooling from all four aerosol types (sulfates, organic carbon, nitrates and black carbon). This is caused by the saturation of the indirect effect, i.e. its sublinear increase with increasing aerosol number concentration (Boucher and Pham, 2002). If, on the other side, it is argued that the cloud lifetime effect is not a forcing but encompasses a feedback, the forcing part may be additive. This, however, cannot be disentangled because the forcing of the cloud lifetime effect cannot be separated from its feedback. Along the same lines, Harshvardhan et al. (2002) concluded that in order to extract the indirect effect from observations, particularly those based on regional and global data sets, the response of cloud systems to their thermodynamic environments cannot be discarded. Thus, indirect aerosol forcing and cloud feedback are intimately coupled.

8 Feedbacks of clouds on aerosols

Baker and Charlson (1990) discussed feedbacks of clouds on aerosols in terms of two stable CCN concentration regimes in the cloud-topped boundary layer. The stable low CCN concentration regime prevails over the oceans. It consists of a balance between drizzle as the major aerosol sink and moderate aerosol production from marine sources. The stable high CCN concentration regime prevents drizzle formation, allowing aerosol concentrations to be enriched. This scenario is more typical for continental aerosols.

Lohmann and Feichter (1997) showed that the sulfate burden increased by 50% when feedbacks with clouds are taken into account. In this positive feedback loop, more sulfate aerosols decrease the precipitation formation rate, which in turn increases the lifetime of sulfate and results in more long-range transport of sulfate to remote regions where wet removal is less efficient. If, in addition, aerosol effects on mixed-phase clouds are taken into account, then a negative feedback loop can be established. If a fraction of the anthropogenic black carbon acts as contact ice nuclei, the precipitation formation via the ice phase is enhanced, removing aerosols from the atmosphere. Depending on the fraction of black carbon as contact ice nuclei, the anthropogenic aerosol burden can be reduced between 38% and 58% as compared to the simulation where black carbon does not act as a contact nuclei (Lohmann, 2002a).

Table 3. Hydrological sensitivity estimated from different pairs of coupled equilibrium ECHAM4 climate model/mixed-layer ocean simulations with an interactive aerosol module. Changes in global mean 2-m temperature $\Delta T$ [K], precipitation $\Delta P$ [%] and the hydrological sensitivity (change in precipitation per unit degree change in temperature) $\Delta P/\Delta T$ [%/K] between a simulation for present-day conditions (representative for 1985) and of a simulation for pre-industrial conditions (representative for 1860) are given (Feichter et al., 2004).

<table>
<thead>
<tr>
<th>Experiment</th>
<th>$\Delta T$</th>
<th>$\Delta P$</th>
<th>$\Delta P/\Delta T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>present-day aerosol conc., varying GHG conc.</td>
<td>1.3</td>
<td>2.3</td>
<td>1.7</td>
</tr>
<tr>
<td>Varying aerosol conc., fixed GHG conc.</td>
<td>$-0.9$</td>
<td>$-3.5$</td>
<td>$3.9$</td>
</tr>
<tr>
<td>Varying aerosol and GHG conc.</td>
<td>0.6</td>
<td>$-1.1$</td>
<td>$-1.9$</td>
</tr>
</tbody>
</table>

9 Uncertainties and needs for improvements in the representation of aerosol effects on clouds in global climate models

In order to narrow down the uncertainties associated with the indirect aerosol effects on climate, general circulation models need to be improved in many aspects:

9.1 Representation of aerosols

Since the pioneering study by Langner and Rodhe (1991) who used a coarse horizontal resolution chemical transport model based on climatological meteorology, the complexity of the aerosol precursor chemistry, of the treatment of transport processes, of the parameterization of particle dry deposition and wet removal has been increased. Recently attempts have been undertaken to calculate not just the aerosol mass but also the particle number concentration by parameterizing aerosol formation and dynamical processes. Two kinds of aerosol dynamics models were developed: modal schemes and bin schemes. Modal schemes have been applied for mineral dust (Schulz et al., 1998) and for sulfuric acid, soot and seasalt (Wilson et al., 2001; Ghan et al., 2001; Vignati et al., 2004). Bin schemes have been applied for sulfate (Adams and Seinfeld, 2002), for mineral dust studies (Tegen and Fung, 1994), for sea salt aerosols (Gong et al., 1997) and for sulfate and seasalt (Gong and Barrie, 2003; Spracklen et al., 2005).

Most of the earlier studies concerned with the effect of aerosol particles on the climate system have just taken sulfate particles into account or have considered sulfate as a surrogate for all anthropogenic aerosols (Jones et al., 1994; Boucher and Lohmann, 1995). Lately most major GCMs include also carbonaceous aerosols, dust and sea salt (for state of model development see: AEROCOM model intercomparison: http://nansen.ipsl.jussieu.fr/AEROCOM/ and Kinne et al., 2003). Simulating nitrate aerosols is more difficult because of their semi-volatile nature (Adams et al., 2001; Met-
zger et al., 2002). Aside from physical and microphysical processes the lack of time-resolved and accurate emission inventories introduces large uncertainties (Bond et al., 2004). In particular, biogenic sources and emissions from biomass burning are highly uncertain. Both biogenic and biomass burning emissions depend on environmental conditions (e.g., weather) and exhibit high interannual variability, which has not been taken into account by climate studies. Probably the largest uncertainty is associated with organic aerosols because current measurement techniques cannot identify all organic species (Kanakidou et al., 2004; O’Dowd et al., 2004). Thus, sources are not well identified and the chemical pathways in the atmosphere are complex and simulations are CPU-time consuming. Organic aerosols can either result from primary production or can originate from gas-to-particle conversion (secondary production). The estimates of the global emission strength of these volatile organic carbon aerosols are a major wildcard in simulations of future scenarios. Advances in measurement techniques for volatile aerosols will have to precede any improvement in modeling capabilities.

There is increasing evidence that aerosol particles are predominantly a conglomerate of different internally mixed chemical substances (Murphy and Thomson, 1997; Cziczo et al., 2004; Kojima et al., 2004). In contrast most GCMs still treat aerosols as external mixtures in terms of their optical properties (e.g., Feichter et al., 1997) because internal mixtures have more degrees of freedoms, are more complex and require an added computational burden. The mixing state of aerosols (externally versus internally mixed) is not only crucial for their optical properties (Haywood and Shine, 1995; Jacobson, 2001; Lesins et al., 2002) but also for their ability to act as CCN. That is, a slight coating of an only moderate soluble organic species can drastically increase its ability to act as a CCN, (e.g., Broekhuizen et al., 2004; Lohmann et al., 2004). Therefore, treating the degree of mixing properly is essential for aerosol processing in GCMs, including aerosol-cloud interactions. It is the route that needs to be taken in order to improve the treatment of aerosols in GCMs. Advanced aerosol modules in some GCMs have been expanded to include aerosol mixtures (Ghan et al., 2001; Stier et al., 2004; Easter et al., 2004). In modal representations of the aerosol size distribution it is important to predict aerosol number as well as mass, so that processes that influence aerosol mass only do not affect aerosol number, and processes that influence aerosol number only do not affect aerosol mass for each mode. The importance of this was demonstrated by Ghan et al. (2001); Zhang et al. (2002); Stier et al. (2004).

9.2 Cloud droplet formation

Linking aerosol particles to cloud droplets is probably the weakest point in estimates of the indirect aerosol effects. In order to treat cloud droplet formation accurately, the aerosol number concentration, its chemical composition and the vertical velocity on the cloud scale need to be known. Abdul-Razzak and Ghan (2000) developed a parameterization based on Köhler theory that can describe cloud droplet formation for a multi-modal aerosol. This approach has been extended by Nenes and Seinfeld (2003) to include kinetic effects, such that the largest aerosols do not have time to grow to their equilibrium size. Also, the competition between natural and anthropogenic aerosols, such as between sulfate and sea salt, as CCN needs to be considered (Ghan et al., 1998; O’Dowd et al., 1999).

Organic carbon is an important cloud condensation nuclei, especially if it is surface active (Shulman et al., 1996; Nenes et al., 2002; Russell et al., 2002). Facchini et al. (1999) indicated that a lowering of the surface tension of some surface-active organic aerosols as obtained from fog water samples would enhance the cloud droplet number concentration, cloud albedo and, hence, could lead to a negative forcing of up to \(-1\) W m\(^{-2}\). On the other hand, Feingold and Chuang (2002) suggested that amphiphilic film-forming compounds retard cloud droplet formation. The delayed activation enables the growth of a mode of larger drops that formed earlier on and therefore leads to an increase in dispersion, and in drizzle formation. Chemical effects of the same order as the indirect effect were pointed out by Nenes et al. (2002) and as large as unresolved cloud dynamics by Lance et al. (2004). While the effect of surface active organics has recently been included in the parameterization of cloud droplet formation by Abdul-Razzak and Ghan (2004), other effects of organics, such as their film-forming ability are not considered yet.

In order to apply one of these parameterizations, the updraft velocity relevant for cloud formation needs to be known. Some GCMs apply a Gaussian distribution (Chuang et al., 1997) or use the turbulent kinetic energy as a surrogate for it (Lohmann et al., 1999). Other GCMs avoid this issue completely and use empirical relationships between aerosol mass and cloud droplet number concentration instead (Menon et al., 2002a). This method is limited because of scarce observational data base. At present, relationships can only be derived between sulfate aerosols, sea salt, organic carbon and cloud droplet number, but no concurrent data for dust or black carbon and cloud droplet number are available yet. This does not imply that diagnostic schemes are inferior for present-day climate simulations. Menon et al. (2003), for example, showed that mechanistic schemes presented no advantage over diagnostic scheme when compared to observations taken during the Second Aerosol Characteristic Experiment (ACE-2). However, because of their greater universality, physically based approaches should be used in future studies of aerosol-cloud-interactions and future climate simulations.
9.3 Treatment of large-scale clouds

Since the first IPCC assessment, large improvements in the description of cloud microphysics for large-scale clouds have been made. Whereas early studies diagnosed cloud amount based on relative humidity, most GCMs now predict cloud condensate in large-scale clouds. The degree of sophistication varies from predicting the sum of cloud water and ice (Rasch and Kristjánsson, 1998) to predicting cloud water, cloud ice, snow and rain as separate species (Fowler et al., 1996). Because the aerosol indirect effect is based on the change in cloud droplet number concentration, some GCMs predict cloud droplet number concentrations in addition to the cloud water mass mixing ratio using one of the above described physically based aerosol activation schemes as a source term for cloud droplets (Ghan et al., 1997a; Lohmann et al., 1999). Likewise the number of ice crystals needs to be predicted in addition to the ice water mass mixing ratio in order to estimate the effect of aerosols on mixed-phase and ice clouds (Ghan et al., 1997b; Lohmann, 2002b). It has been shown by Gierens and Spichtinger (pers. comm. 2005) that only two-moment schemes are able to account for the size-dependent sedimentation rate, which leads to important differences in the cloud vertical structure, cloud lifetime and cloud optical properties. While two-moment schemes are superior to one-moment schemes, they are inferior to size-resolved cloud microphysics. The latter are, however, not computationally affordable at present.

9.4 Treatment of convective clouds

There is currently a big discrepancy between the degree of sophistication in cloud microphysics in large-scale clouds (see above) and a very rudimentary treatment of cloud microphysics in convective clouds. Recently there is evidence emerging that biomass burning affects convective clouds (Rosenfeld, 1999; Sherwood, 2002; Roberts et al., 2003), which requires improvements in the treatment of convective clouds. A first study to this effect was conducted by Nober et al. (2003) as discussed above. They basically decreased the precipitation efficiency for warm cloud formation in convective clouds depending on the cloud droplet number concentration. Zhang et al. (2005) took this approach a step further and introduced the same microphysical processes (autoconversion, freezing, aggregation, etc.) that are considered in large-scale clouds into convective clouds as well. An alternative avenue is to represent small-scale and mesoscale processes provided by a cloud-resolving model embedded in each column of a large-scale model, also known as the superparameterization (Grabowski, 2004).

9.5 Subgrid-scale variability and radiative transfer

A new approach to account for unresolved spatial variability and microphysical process rates is to consider probability distribution functions of the respective quantities (Pincus and Klein, 2000; Tompkins, 2002). This approach has been extended to account for subgrid-scale variability in cloud cover and cloud condensate in radiative transfer through inhomogeneous cloud fields by Pincus et al. (2003). Such a treatment is necessary because errors originating from treating clouds as plane parallel homogeneous clouds can overpredict the Twomey effect by up to 50% (Barker, 2000).

10 Conclusions

In summary, aerosol effects on clouds can be divided into the radiative effects and the effects on the hydrological cycle:

10.1 Aerosol radiative effects

The aerosol radiative effects can be further divided into those that exert a positive perturbation on the radiation budget and those that exert a negative perturbation:

- Both the Twomey and the cloud lifetime effect act to cool the Earth-atmosphere system by increasing cloud optical depth and cloud cover, respectively. This reduces the net solar radiation at the top-of-the atmosphere as well as at the surface.

- Carbonaceous aerosols and dust exert a positive forcing at the top-of-the atmosphere, at least in regions with high surface albedo, and can thus directly warm the atmosphere. This effect can be amplified if absorption of solar radiation of these aerosol particles occurs within cloud droplets (Chylek et al., 1996). The resulting increase in temperature reduces the relative humidity and may result in the evaporation of cloud droplets. The reduced cloud cover and cloud optical depth will in turn further amplify warming of the Earth-atmosphere system.

- Another way in which aerosols could contribute to a warming is by decreasing cloud amount due to increasing precipitation. As more aerosols generally lead to more and smaller cloud droplets, this effect is not very likely to happen but may occur if a few anthropogenic aerosol were to act as giant nuclei or as ice nuclei.

Traditionally estimates of the direct and indirect aerosol forcing are based on model studies in which the radiative forcing is the difference to an aerosol (component) free or less loaded (e.g. pre-industrial) reference state. Thus, a direct validation with measurements of the aerosol radiative forcing is basically impossible. Only a combination of satellite data with model simulations can advance the pure model based estimates of global indirect aerosol effects. While some pioneering studies using this approach exist (Lohmann and Lesins, 2002; Quaas et al., 2004; Suzuki et al., 2004), much more research needs to be done here.
Some attempts have been made to estimate the total radiative forcing since pre-industrial times including all quantitative radiative forcing estimates and uncertainties included in the IPCC 2001 bar chart (Boucher and Haywood, 2001; Schwartz, 2004). The resulting total radiative forcing (well-mixed greenhouse gases, solar activity, ozone, direct aerosol effects and Twomey effect) has a 75–97% probability of being positive. These estimates neglect the cloud lifetime effect and aerosol effects on mixed-phase and ice clouds because our knowledge about these latter effects is not sufficient to predict their magnitudes yet. Probability ranges, however, can be estimated from inverse simulations or thermodynamic considerations (Forest et al., 2002; Knutti et al., 2002; Anderson et al., 2003; Crutzen and Ramanathan, 2003). They limit the sum of all indirect aerosol effect to between 0 to $-2 \text{ W m}^{-2}$. Given that the GCM estimates for the Twomey effect alone amount to $-0.5$ to $-1.9 \text{ W m}^{-2}$ (Table 1), either all other aerosol indirect effects cancel each other or the Twomey effect is smaller than current climate models predict. Thus uncertainties in aerosol forcing must be reduced at least three-fold for uncertainties in climate sensitivities to be meaningful reduced and bounded (Schwartz, 2004).

A problem for assessing the aerosol indirect effect from data is the shift in aerosol and aerosol precursor emissions. The decrease in emissions in eastern Europe in the 1990s was used by Krüger and Gräßl (2002) to investigate the indirect versus the semi-direct effect. In future the main emission centers will shift from the traditional industrial centers in mid-latitudes of the Northern Hemisphere to the subtropics and tropics. Kristjánsson (2002) predicted that the global mean aerosol radiative forcing remains the same in 2100, but this is only one climate model study so far.

Finally, aerosol radiative forcing is regionally highly variable and differs also in sign from region to region (Ramanathan et al., 2001a; Chameidis et al., 2002). Thus, it is questionable whether the global mean change in surface temperature is sufficient to characterize the radiative impact of aerosols. Changes in the hydrological cycle caused by aerosols are probably more important than the mere temperature change because they have consequences for fresh water supply and food production among others.

10.2 Aerosol effects on the hydrological cycle

Here the following effects can be distinguished:

- Aerosols may change the occurrence and frequency of convection and thus could be responsible for droughts and flood simultaneously.
- Aerosols may cause reductions in the net solar radiation reaching the surface. In particular the direct sunlight is reduced. Thus, even in a greenhouse gas induced warmer climate the evaporation could decrease and the hydrological cycle could be expected to slow down.
- Aerosol induced cooling can have consequences in other parts on the world. It is believed that the cooling of the Northern Hemisphere causes a southward shift of the intertropical convergence zone, which could have been partly responsible for the Sahelian drought.
- Anthropogenic aerosols could influence mixed-phase clouds by retarding the onset of freezing due to their smaller size (thermodynamic process), by acting as ice nuclei in the different freezing modes and hence speeding up the Bergeron-Findeisen process (glaciation effect) and by the reducing the riming process (see also Table 1). These aerosol influences are not studied well enough to predict their sign yet. However, these aerosol effects may suggest a mechanism for a decreasing cloud water content with increasing aerosol load.

We are entering a new area of aerosol research by investigating the interactions between aerosols and the hydrological cycle. Research in this area started with cloud seeding research, as summarized in the overview article by Bruintjes (1999). Investigations in cloud seeding research could benefit from satellite-based microphysical retrievals that can be combined with in situ cloud sampling to monitor the effects of natural and anthropogenic aerosol or hygroscopic seeding material on cloud droplet size evolution, and the effects of ice-forming nuclei on ice-particle concentrations, both of which determine the efficiency of precipitation formation. The cloud seeding community, however, has traditionally not been interested in the climate impact of anthropogenic aerosols or their effect on the global hydrological cycle, but has focused on the influence of aerosols on precipitation on a local to regional scale. Thus, a knowledge exchange between the two research communities would be beneficial.

Our knowledge about aerosol effects on clouds and the hydrological cycle is still very rudimentary. The observations for the hydrological cycle are less complete than for global mean temperature and the physical constraints are weaker so that it will be substantially harder to quantify the range of possible changes in the hydrological cycle (Allen and Ingram, 2002). Therefore, clearly more research in terms of field experiments, laboratory studies and modeling efforts is needed in order to understand and quantify the effect of anthropogenic aerosols on clouds and the hydrological cycle. This is especially important because cloud feedbacks in climate models still represent one of the largest uncertainties.
As shown by Stocker et al. (2001) there is still no consensus on whether clouds provide a negative or positive climate feedback in response to a doubling of carbon dioxide. It is largely because of these uncertainties in cloud feedback that the uncertainty range of the increase in the global mean surface temperature in response to a doubling of carbon dioxide varies between 1.5 and 4.5 K. The cloud feedback problem thus has to be solved in order to assess the aerosol indirect forcing more reliably.

Acknowledgements. The authors thank S. Kinne, L. Rotstayn, S. Ghan and one anonymous reviewer for useful comments and suggestions.

Edited by: K. Carslaw

References


www.atmos-chem-phys.org/acp/5/715/
Atmos. Chem. Phys., 5, 715–737, 2005


SRef-ID: 1680-7375/acp/2004-4-5855


Krüger, O., Marks, R., and Graßl, H.: Influence of pollution on cloud reflectance, J. Geophys. Res., 109,


