

Introduction

Lead Authors: Ralph A. Kahn, NASA GSFC; Hongbin Yu, NASA GSFC/UMBC

Contributing Authors: Stephen E. Schwartz, DOE BNL; Mian Chin, NASA GSFC; Graham Feingold, NOAA ESRL; Lorraine A. Remer, NASA GSFC; David Rind, NASA GISS; Rangasayi Halthore, NASA HQ/NRL; Philip DeCola, NASA HQ

This report highlights key aspects of current knowledge about the global distribution of aerosols and their properties, as they relate to climate change. Leading measurement techniques and modeling approaches are briefly summarized, providing context for an assessment of the next steps needed to significantly reduce uncertainties in this component of the climate change picture. The present assessment builds upon the recent Inter-governmental Panel on Climate Change Fourth Assessment Report (IPCC AR4, 2007) and other sources.

1.1 Description of Atmospheric Aerosols

Although Earth's atmosphere consists primarily of gases, aerosols and clouds play significant roles in shaping conditions at the surface and in the lower atmosphere. Aerosols are liquid or solid particles suspended in the air, whose typical diameters range over four orders of magnitude, from a few nanometers to a few tens of micrometers. They exhibit a wide range of compositions and shapes, that depend on their origins and subsequent atmospheric processing. For many applications, aerosols from about 0.05 to 10 micrometers in diameter are of greatest interest, as particles in this size range dominate aerosol direct interaction with sunlight, and also make up the majority of the aerosol mass. Particles at the small end of this size range play a significant role in interactions with clouds, whereas particles at the large end, though much less numerous, can contribute significantly near dust and volcanic sources. Over the ocean, giant salt particles may also play a role in cloud development.

A large fraction of aerosols is natural in origin, including desert and soil dust, wildfire smoke, sea salt particles produced mainly by breaking bubbles in the spray of ocean whitecaps, and

volcanic ash. Volcanoes are also sources of sulfur dioxide, which, along with sulfur-containing gases produced by ocean biology and the decomposition of organic matter, as well as hydrocarbons such as terpenes and isoprene emitted by vegetation, are examples of gases that can be converted to so-called "secondary" aerosols by chemical processes in the atmosphere. Figure 1.1 gives a summary of aerosol processes most relevant to their influence on climate.

Table 1.1 reports estimated source strengths, lifetimes, and amounts for major aerosol types, based on an aggregate of emissions estimates and global model simulations; the ranges provided represent model diversity only, as the global measurements required to validate these quantities are currently lacking.

Aerosol optical depth (AOD) (also called aerosol optical thickness, AOT, in the literature) is a measure of the amount of incident light either scattered or absorbed by airborne particles. Formally, aerosol optical depth is a dimensionless quantity, the integral of the product of particle number concentration and particle extinction cross-section (which accounts for individual particle scattering + absorption),

along a path length through the atmosphere, usually measured vertically. In addition to AOD, particle size, composition, and structure, which are mediated both by source type and subsequent atmospheric processing, determine how particles interact with radiant energy and influence the heat balance of the planet. Size and composition also determine the ability of particles to serve as nuclei upon which cloud droplets form. This provides an indirect means for aerosol to interact with radiant energy by modifying cloud properties.

Among the main aerosol properties required to evaluate their effect on radiation is the *single-scattering albedo* (SSA), which describes the fraction of light interacting with the particle that is scattered, compared to the total that is scattered and absorbed. Values range from 0 for totally absorbing (dark) particles to 1 for purely scattering ones; in nature, SSA is rarely lower than about 0.75. Another quantity, the *asymmetry parameter* (g), reports the first moment of the cosine of the scattered radiation angular distribution. The parameter g ranges from -1 for entirely back-scattering particles, to 0 for isotropic (uniform) scattering, to +1 for entirely forward-scattering. One further quantity that

must be considered in the energy balance is the *surface albedo* (A), a measure of reflectivity at the ground, which, like SSA, ranges from 0 for purely absorbing to 1 for purely reflecting. In practice, A can be near 0 for dark surfaces, and can reach values above 0.9 for visible light over snow. AOD, SSA, g , and A are all dimensionless quantities, and are in general wavelength-dependent. In this report, AOD, SSA, and g are given at mid-visible wavelengths, near the peak of the solar spectrum around 550 nanometers, and A is given as an average over the solar spectrum, unless specified otherwise.

One of the greatest challenges in studying aerosol impacts on climate is the immense diversity, not only in particle size, composition, and origin, but also in spatial and temporal distribution.



About 10% of global atmospheric aerosol mass is generated by human activity, but it is concentrated in the immediate vicinity, and downwind of sources (e.g., Textor et al., 2006). These anthropogenic aerosols include primary (directly emitted) particles and secondary particles that are formed in the atmosphere. Anthropogenic aerosols originate from urban and industrial emissions, domestic fire and other combustion products, smoke from agricultural burning, and soil dust created by overgrazing, deforestation, draining of inland water bodies, some farming practices, and generally, land management activities that destabilize the surface regolith to wind erosion. The amount of aerosol in the atmosphere has greatly increased in some parts of the world during the industrial period, and the nature of this particulate matter has substantially changed as a consequence of the evolving nature of emissions from industrial, commercial, agricultural, and residential activities, mainly combustion-related.

One of the greatest challenges in studying aerosol impacts on climate is the immense diversity, not only in particle size, composition, and origin, but also in spatial and temporal distribution. For most aerosols, whose primary source is emissions near the surface, concentrations are greatest in the atmospheric boundary layer, decreasing with altitude in the free troposphere. However, smoke from wildfires and volcanic effluent can be injected above the boundary layer; after injection, any type of aerosol can be lofted to higher elevations; this can extend their atmospheric lifetimes, increasing their impact spatially and climatically.

Aerosols are removed from the atmosphere primarily through cloud processing and wet

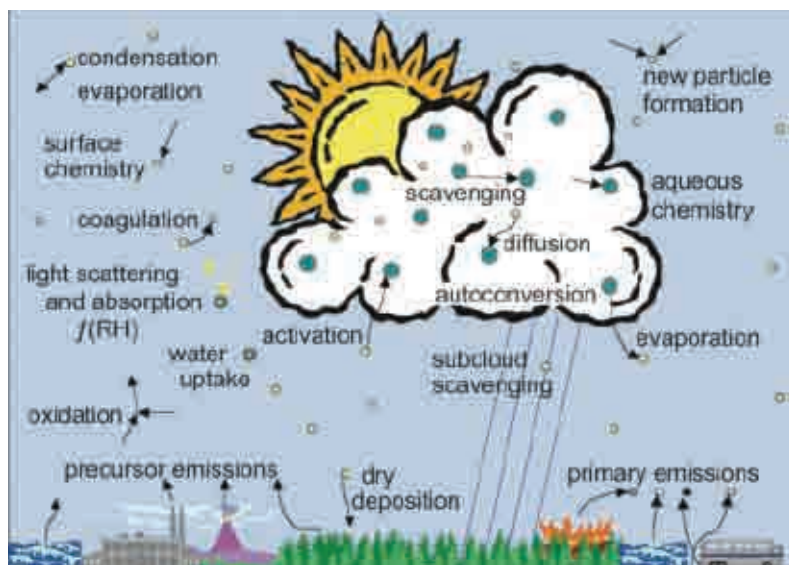


Figure 1.1. Major aerosol processes relevant to their impact on climate. Aerosols can be directly emitted as primary particles and can form secondarily by the oxidation of emitted gaseous precursors. Changes in relative humidity (RH) can cause particle growth or evaporation, and can alter particle properties. Physical processes within clouds can further alter particle properties, and conversely, aerosols can affect the properties of clouds, serving as condensation nuclei for new cloud droplet formation. Aqueous-phase chemical reactions in cloud drops or in clear air can also affect aerosol properties. Particles are ultimately removed from the atmosphere, scavenged by falling raindrops or settling by dry deposition. Modified from Ghan and Schwartz (2007).

Table 1.1. Estimated source strengths, lifetimes, mass loadings, and optical depths of major aerosol types. Statistics are based on results from 16 models examined by the Aerosol Comparisons between Observations and Models (AeroCom) project (Textor et al., 2006; Kinne et al., 2006). BC = black carbon; POM = particulate organic matter. See Chapter 3 for more details.

| Aerosol Type | Total source ¹ (Tg/yr ¹) | Lifetime (day) | Mass loading ¹ (Tg) | Optical depth @ 550 nm |
|----------------------|---|----------------|--------------------------------|--------------------------|
| | Median (Range) | Median (Range) | Median (Range) | Median (Range) |
| Sulfate ² | 190 (100-230) | 4.1 (2.6-5.4) | 2.0 (0.9-2.7) | 0.034 (0.015-0.051) |
| BC | 11 (8-20) | 6.5 (5.3-15) | 0.2 (0.05-0.5) | 0.004 (0.002-0.009) |
| POM ² | 100 (50-140) | 6.2 (4.3-11) | 1.8 (0.5-2.6) | 0.019 (0.006-0.030) |
| Dust | 1600 (700-4000) | 4.0 (1.3-7) | 20 (5-30) | 0.032 (0.012-0.054) |
| Sea salt | 6000 (2000-120000) | 0.4 (0.03-1.1) | 6 (3-13) | 0.030 (0.020-0.067) |
| Total | | | | 0.13 (0.065-0.15) |

¹ Tg (teragram) = 10¹² g, or million metric tons.

² The sulfate aerosol source is mainly SO₂ oxidation, plus a small fraction of direct emission. The organic matter source includes direct emission and hydrocarbon oxidation.

deposition in precipitation, a mechanism that establishes average tropospheric aerosol atmospheric lifetimes at a week or less (Table 1.1). The efficiency of removal therefore depends on the proximity of aerosols to clouds. For example, explosive volcanoes occasionally inject large amounts of aerosol precursors into the stratosphere, above most clouds; sulfuric acid aerosols formed by the 1991 Pinatubo eruption exerted a measurable effect on the atmospheric heat budget for several years thereafter (e.g., Minnis et al., 1993; McCormick et al., 1995; Robock, 2000, 2002). Aerosols are also removed by dry deposition processes: gravitational settling tends to eliminate larger particles, impaction typically favors intermediate-sized particles, and coagulation is one way smaller particles can aggregate with larger ones, leading to their eventual deposition by wet or dry processes. Particle injection height, subsequent air mass advection, and other factors also affect the rate at which dry deposition operates.

Despite relatively short average residence times, aerosols regularly travel long distances. For example, particles moving at mean velocity of 5 m s⁻¹ and remaining in the atmosphere for a week will travel 3000 km. Global aerosol observations from satellites provide ample evidence of this—Saharan dust reaches the Caribbean and Amazon basin, Asian desert dust and anthropogenic aerosol is found over the central Pacific and sometimes as far away as North America, and Siberian smoke can be deposited

in the Arctic. This transport, which varies both seasonally and inter-annually, demonstrates the global scope of aerosol influences.

As a result of the non-uniform distribution of aerosol sources and sinks, the short atmospheric lifetimes and intermittent removal processes compared to many atmospheric greenhouse trace gases, the spatial distribution of aerosol particles is quite non-uniform. The amount and nature of aerosols vary substantially with location and from year to year, and in many cases exhibit strong seasonal variations.

One consequence of this heterogeneity is that the impact of aerosols on climate must be understood and quantified on a *regional* rather than just a global-average basis. AOD trends observed in the satellite and surface-based data records suggest that since the mid-1990s, the amount of anthropogenic aerosol has decreased over North America and Europe, but has increased over parts of east and south Asia; on average, the atmospheric concentration of low-latitude smoke particles has increased (Mishchenko and Geogdzhayev, 2007). The observed AOD trends in the northern hemisphere are qualitatively consistent with changes in anthropogenic emissions (e.g. Streets et al., 2006a), and with observed trends in surface solar radiation flux (“solar brightening” or “dimming”), though other factors could be involved (e.g., Wild et al., 2005). Similarly, the increase in smoke parallels is associated with

The impact of aerosols on climate must be understood and quantified on a regional rather than just a global-average basis.

changing biomass burning patterns (e.g., Koren et al., 2007a).

1.2 The Climate Effects of Aerosols

Aerosols exert a variety of impacts on the environment. Aerosols (sometimes referred to particulate matter or “PM,” especially in air quality applications), when concentrated near the surface, have long been recognized as affecting pulmonary function and other aspects of human health. Sulfate and nitrate aerosols play a role in acidifying the surface downwind of gaseous sulfur and odd nitrogen sources. Particles deposited far downwind might fertilize iron-poor waters in remote oceans, and Saharan dust reaching the Amazon Basin is thought to contribute nutrients to the rainforest soil.

Aerosols also interact strongly with solar and terrestrial radiation in several ways. Figure 1.2 offers a schematic overview. First, they scatter and absorb sunlight (McCormick and Ludwig, 1967; Charlson and Pilat, 1969; Atwater, 1970; Mitchell, Jr., 1971; Coakley et al., 1983); these are described as “direct effects” on shortwave (solar) radiation. Second, aerosols act as sites at which water vapor can accumulate during cloud droplet formation, serving as cloud condensation nuclei or CCN. Any change in number concentration or hygroscopic properties of such particles has the potential to modify the physical and radiative properties of clouds,

altering cloud brightness (Twomey, 1977) and the likelihood and intensity with which a cloud will precipitate (e.g., Gunn and Phillips, 1957; Liou and Ou 1989; Albrecht, 1989). Collectively changes in cloud processes due to anthropogenic aerosols are referred to as *aerosol indirect effects*. Finally, absorption of solar radiation by particles is thought to contribute to a reduction in cloudiness, a phenomenon referred to as the *semi-direct effect*. This occurs because absorbing aerosol warms the atmosphere, which changes the atmospheric stability, and reduces surface flux.

The primary direct effect of aerosols is a brightening of the planet when viewed from space, as much of Earth’s surface is dark ocean, and most aerosols scatter more than 90% of the visible light reaching them. The primary indirect effects of aerosols on clouds include an increase in cloud brightness, change in precipitation and possibly an increase in lifetime; thus the overall net impact of aerosols is an enhancement of Earth’s reflectance (shortwave albedo). This reduces the sunlight reaching Earth’s surface, producing a net climatic cooling, as well as a redistribution of the radiant and latent heat energy deposited in the atmosphere. These effects can alter atmospheric circulation and the water cycle, including precipitation patterns, on a variety of length and time scales (e.g., Ramanathan et al., 2001a; Zhang et al., 2006).

The primary direct effect of aerosols is a brightening of the planet when viewed from space. The primary indirect effects of aerosols on clouds include an increase in cloud brightness and possibly an increase in lifetime. The overall net impact of aerosols is an enhancement of Earth’s reflectance.

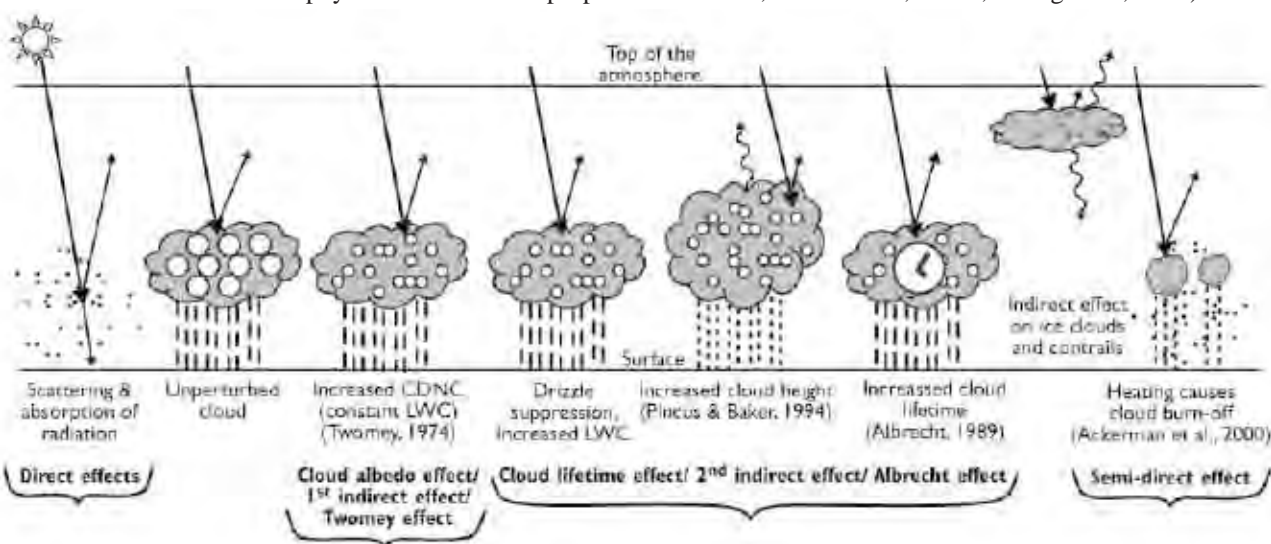


Figure 1.2. Aerosol radiative forcing. Airborne particles can affect the heat balance of the atmosphere, directly, by scattering and absorbing sunlight, and indirectly, by altering cloud brightness and possibly lifetime. Here small black dots represent aerosols, circles represent cloud droplets, straight lines represent short-wave radiation, and wavy lines, long-wave radiation. LWC is liquid water content, and CDNC is cloud droplet number concentration. Confidence in the magnitudes of these effects varies considerably (see Chapter 3). Although the overall effect of aerosols is a net cooling at the surface, the heterogeneity of particle spatial distribution, emission history, and properties, as well as differences in surface reflectance, mean that the magnitude and even the sign of aerosol effects vary immensely with location, season and sometimes inter-annually. The human-induced component of these effects is sometimes called “climate forcing.” (From IPCC, 2007, modified from Haywood and Boucher, 2000).)

Several variables are used to quantify the impact aerosols have on Earth's energy balance; these are helpful in describing current understanding, and in assessing possible future steps.

For the purposes of this report, *aerosol radiative forcing* (RF) is defined as the net energy flux (downwelling minus upwelling) difference between an initial and a perturbed aerosol loading state, at a specified level in the atmosphere. (Other quantities, such as solar radiation, are assumed to be the same for both states.) This difference is defined such that a negative aerosol forcing implies that the change in aerosols relative to the initial state exerts a cooling influence, whereas a positive forcing would mean the change in aerosols exerts a warming influence.

There are a number of subtleties associated with this definition:

(1) The initial state against which aerosol forcing is assessed must be specified. For direct aerosol radiative forcing, it is sometimes taken as the complete absence of aerosols. IPCC AR4 (2007) uses as the initial state their estimate of aerosol loading in 1750. That year is taken as the approximate beginning of the era when humans exerted accelerated influence on the environment.

(2) A distinction must be made between aerosol RF and the *anthropogenic contribution* to aerosol RF. Much effort has been made to distinguishing these contributions by modeling and with the help of space-based, airborne, and surface-based remote sensing, as well as *in situ* measurements. These efforts are described in subsequent chapters.

(3) In general, aerosol RF and anthropogenic aerosol RF include energy associated with both the shortwave (solar) and the long-wave (primarily planetary thermal infrared) components of Earth's radiation budget. However, the solar component typically dominates, so in this document, these terms are used to refer to the solar component only, unless specified otherwise. The wavelength separation between the short- and long-wave components is usually set at around three or four micrometers.

(4) The IPCC AR4 (2007) defines radiative forcing as the net downward minus upward

irradiance at the tropopause due to an external driver of climate change. This definition excludes stratospheric contributions to the overall forcing. Under typical conditions, most aerosols are located within the troposphere, so aerosol forcing at TOA and at the tropopause are expected to be very similar. Major volcanic eruptions or conflagrations can alter this picture regionally, and even globally.

(5) Aerosol radiative forcing can be evaluated at the surface, within the atmosphere, or at top-of-atmosphere (TOA). In this document, unless specified otherwise, aerosol radiative forcing is assessed at TOA.

(6) As discussed subsequently, aerosol radiative forcing can be greater at the surface than at TOA if the aerosols absorb solar radiation. TOA forcing affects the radiation budget of the planet. Differences between TOA forcing and surface forcing represent heating within the atmosphere that can affect vertical stability, circulation on many scales, cloud formation, and precipitation, all of which are climate effects of aerosols. In this document, unless specified otherwise, these additional climate effects are not included in aerosol radiative forcing.

(7) Aerosol direct radiative forcing can be evaluated under cloud-free conditions or under natural conditions, sometimes termed "all-sky" conditions, which include clouds. Cloud-free direct aerosol forcing is more easily and more accurately calculated; it is generally greater than all-sky forcing because clouds can mask the aerosol contribution to the scattered light. Indirect forcing, of course, must be evaluated for cloudy or all-sky conditions. In this document, unless specified otherwise, aerosol radiative forcing is assessed for all-sky conditions.

(8) Aerosol radiative forcing can be evaluated instantaneously, daily (24-hour) averaged, or assessed over some other time period. Many measurements, such as those from polar-orbiting satellites, provide instantaneous values, whereas models usually consider aerosol RF as a daily average quantity. In this document, unless specified otherwise, daily averaged aerosol radiative forcing is reported.

(9) Another subtlety is the distinction between a "forcing" and a "feedback." As different parts of the climate system interact, it is often unclear

Aerosol radiative forcing is defined as the net energy flux (downwelling minus upwelling) difference between an initial and a perturbed aerosol loading state.



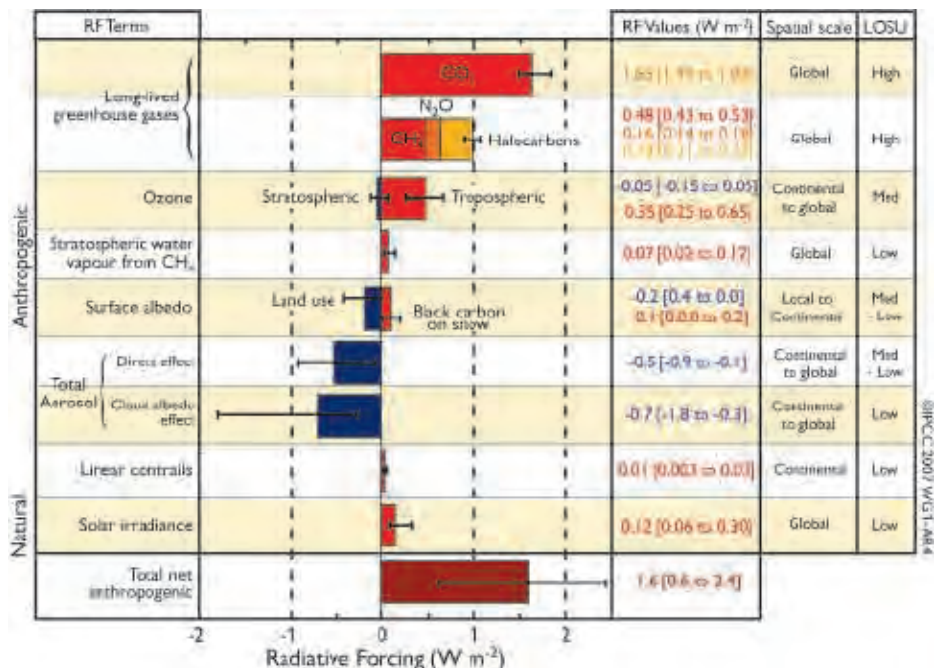


Figure 1.3a. (Above) Global average radiative forcing (RF) estimates and uncertainty ranges in 2005, relative to the pre-industrial climate. Anthropogenic carbon dioxide (CO₂), methane (CH₄), nitrous oxide (N₂O), ozone, and aerosols as well as the natural solar irradiance variations are included. Typical geographical extent of the forcing (spatial scale) and the assessed level of scientific understanding (LOSU) are also given. Forcing is expressed in units of watts per square meter ($W m^{-2}$). The total anthropogenic radiative forcing and its associated uncertainty are also given. Figure from IPCC (2007).

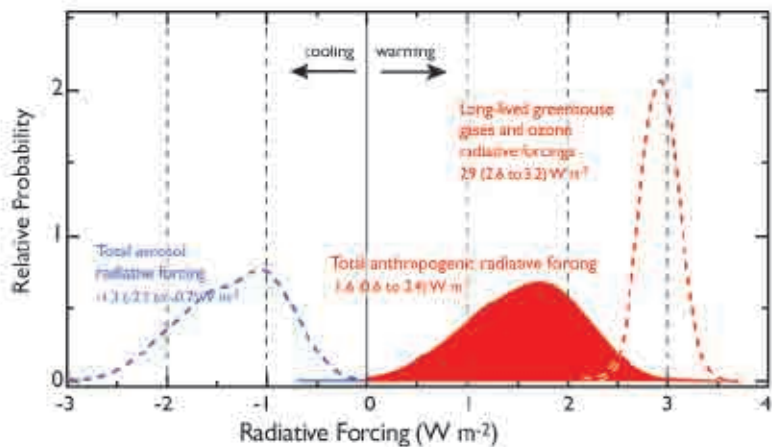


Figure 1.3b. (Left) Probability distribution functions (PDFs) for anthropogenic aerosol and GHG RFs. Dashed red curve: RF of long-lived greenhouse gases plus ozone; dashed blue curve: RF of aerosols (direct and cloud albedo RF); red filled curve: combined anthropogenic RF. The RF range is at the 90% confidence interval. Figure adapted from IPCC (2007).

which elements are “causes” of climate change (forcings among them), which are responses to these causes, and which might be some of each. So, for example, the concept of aerosol effects on clouds is complicated by the impact clouds have on aerosols; the aggregate is often called aerosol-cloud interactions. This distinction sometimes matters, as it is more natural to attribute responsibility for causes than for responses. However, practical environmental considerations usually depend on the net result of all influences. In this report, “feedbacks” are taken as the consequences of changes in surface or atmospheric temperature, with the understanding that for some applications, the accounting may be done differently.

In summary, aerosol radiative forcing, the fundamental quantity about which this report is written, must be qualified by specifying the initial and perturbed aerosol states for which the radiative flux difference is calculated, the altitude at which the quantity is assessed, the wavelength regime considered, the temporal averaging, the cloud conditions, and whether total or only human-induced contributions are considered. The definition given here, qualified as needed, is used throughout the report.

Although the possibility that aerosols affect climate was recognized more than 40 years ago, the measurements needed to establish the magnitude of such effects, or even whether

specific aerosol types warm or cool the surface, were lacking. Satellite instruments capable of at least crudely monitoring aerosol amount globally were first deployed in the late 1970s. But scientific focus on this subject grew substantially in the 1990s (e.g. Charlson et al., 1990; 1991; 1992; Penner et al., 1992), in part because it was recognized that reproducing the observed temperature trends over the industrial period with climate models requires including net global cooling by aerosols in the calculation (IPCC, 1995; 1996), along with the warming influence of enhanced atmospheric greenhouse gas (GHG) concentrations – mainly carbon dioxide, methane, nitrous oxide, chlorofluorocarbons, and ozone.

Improved satellite instruments, ground- and ship-based surface monitoring, more sophisticated chemical transport and climate models, and field campaigns that brought all these elements together with aircraft remote sensing and *in situ* sampling for focused, coordinated study, began to fill in some of the knowledge gaps. By the Fourth IPCC Assessment Report, the scientific community consensus held that in global average, the sum of direct and indirect top-of-atmosphere (TOA) forcing by anthropogenic aerosols is negative (cooling) of about -1.3 W m^{-2} (-2.2 to -0.5 W m^{-2}). This is significant compared to the positive forcing by anthropogenic GHGs (including ozone), about $2.9 \pm 0.3 \text{ W m}^{-2}$ (IPCC, 2007). However, the spatial distribution of the gases and aerosols are very different, and they do not simply exert compensating influences on climate.

The IPCC aerosol forcing assessments are based largely on model calculations, constrained as much as possible by observations. At present, aerosol influences are not yet quantified adequately, according to Figure 1.3a, as scientific understanding is designated as “Medium - Low” and “Low” for the direct and indirect climate forcing, respectively. The IPCC AR4 (2007) concluded that uncertainties associated with changes in Earth’s radiation budget due to anthropogenic aerosols make the largest contribution to the overall uncertainty in radiative forcing of climate change among the factors assessed over the industrial period (Figure 3b).

Although AOD, aerosol properties, aerosol vertical distribution, and surface reflectivity all contribute to aerosol radiative forcing, AOD

usually varies on regional scales more than the other aerosol quantities involved. *Forcing efficiency* (E_r), defined as a ratio of direct aerosol radiative forcing to AOD at 550 nm, reports the sensitivity of aerosol radiative forcing to AOD, and is useful for isolating the influences of particle properties and other factors from that of AOD. E_r is expected to exhibit a range of values globally, because it is governed mainly by aerosol size distribution and chemical composition (which determine aerosol single-scattering albedo and phase function), surface reflectivity, and solar irradiance, each of which exhibits pronounced spatial and temporal variations. To assess aerosol RF, E_r is multiplied by the ambient AOD.

Figure 1.4 shows a range of E_r , derived from AERONET surface sun photometer network measurements of aerosol loading and particle properties, representing different aerosol and surface types, and geographic locations. It demonstrates how aerosol direct solar radiative forcing (with initial state taken as the absence of aerosol) is determined by a combination of aerosol and surface properties. For example, E_r due to southern African biomass burning smoke is greater at the surface and smaller at TOA than South American smoke because the southern African smoke absorbs sunlight more strongly, and the magnitude of E_r for mineral dust for several locations varies depending on the underlying surface reflectance. Figure 1.4 illustrates one further point, that the radiative forcing by aerosols on surface energy balance can be much greater than that at TOA. This is especially true

Reproducing the observed temperature trends over the industrial period with climate models requires including net global cooling by aerosols in the calculation.

The radiative forcing by aerosols on surface energy balance can be much greater than that at the top of the atmosphere.

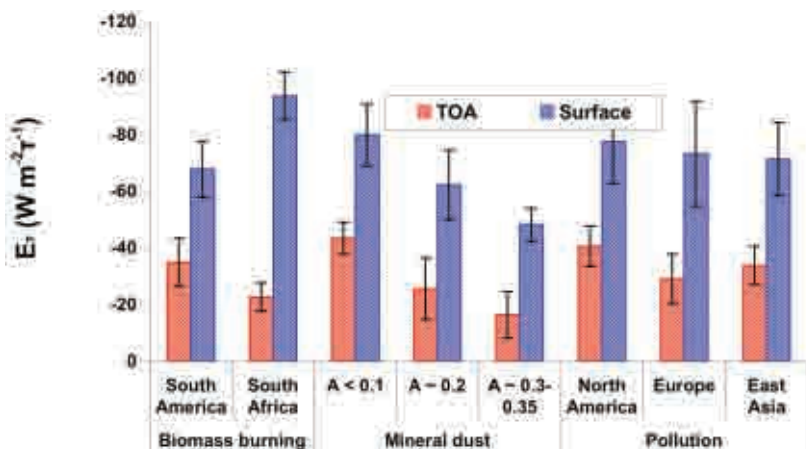


Figure 1.4. The clear-sky forcing efficiency E_r , defined as the diurnally averaged aerosol direct radiative effect (W m^{-2}) per unit AOD at 550 nm, calculated at both TOA and the surface, for typical aerosol types over different geographical regions. The vertical black lines represent \pm one standard deviation of E_r for individual aerosol regimes and A is surface broadband albedo. (adapted from Zhou et al., 2005).

In regions having high concentrations of anthropogenic aerosol, aerosol forcing is much stronger than the global average, and can exceed the magnitude of greenhouse gas warming.



Radiative heating of the atmosphere by absorbing particles can change the atmospheric temperature structure, affecting vertical mixing, cloud formation and evolution, and possibly large-scale dynamical systems.

when the particles have SSA substantially less than 1, which can create differences between surface and TOA forcing as large as a factor of five (e.g., Zhou et al., 2005).

Table 1.2 presents estimates of cloud-free, instantaneous, aerosol direct RF dependence on AOD, and on aerosol and surface properties, calculated for three sites maintained by the US Department of Energy's Atmospheric Radiation Measurement (ARM) program, where surface and atmospheric conditions span a significant range of natural environments (McComiskey et al., 2008a). Here aerosol RF is evaluated relative to an initial state that is the complete absence of aerosols. Note that aerosol direct RF dependence on individual parameters varies considerably, depending on the values of the other parameters, and in particular, that aerosol RF dependence on AOD actually changes sign, from net cooling to net warming, when aerosols reside over an exceedingly bright surface. Sensitivity values are given for snapshots at fixed solar zenith angles, relevant to measurements made, for example, by polar-orbiting satellites.

The lower portion of Table 1.2 presents upper bounds on instantaneous measurement uncertainty, assessed individually for each of AOD, SSA, g , and A , to produce a 1 W m^{-2} top-of-atmosphere, cloud-free aerosol RF accuracy. The values are derived from the upper portion of the table, and reflect the diversity of conditions captured by the three ARM sites. Aerosol RF sensitivity of 1 W m^{-2} is used as an example; uncertainty upper bounds are obtained from the partial derivative for each parameter by neglecting the uncertainties for all other parameters. These estimates produce an instantaneous AOD measurement uncertainty upper bound between about 0.01 and 0.02, and SSA constrained to about 0.02 over surfaces as bright or brighter than the ARM Southern Great Plains site, typical of mid-latitude, vegetated land. Other researchers, using independent data sets, have derived ranges of E_{τ} and aerosol RF sensitivity similar to those presented here, for a variety of conditions (e.g., Christopher and Jones, 2008; Yu et al., 2006; Zhou et al., 2005).

These uncertainty bounds provide a baseline against which current and expected near-future instantaneous measurement capabilities are assessed in Chapter 2. Model sensitivity is usually evaluated for larger-scale (even global)

and longer-term averages. When instantaneous measured values from a randomly sampled population are averaged, the uncertainty component associated with random error diminishes as something like the inverse square root of the number of samples. As a result, the accuracy limits used for assessing more broadly averaged model results corresponding to those used for assessing instantaneous measurements, would have to be tighter, as discussed in Chapter 4.

In summary, much of the challenge in quantifying aerosol influences arises from large spatial and temporal heterogeneity, caused by the wide variety of aerosol sources, sizes and compositions, the spatial non-uniformity and intermittency of these sources, the short atmospheric lifetime of most aerosols, and the spatially and temporally non-uniform chemical and microphysical processing that occurs in the atmosphere. In regions having high concentrations of anthropogenic aerosol, for example, aerosol forcing is much stronger than the global average, and can exceed the magnitude of GHG warming, locally reversing the sign of the net forcing. It is also important to recognize that the global-scale aerosol TOA forcing alone is not an adequate metric for climate change (NRC, 2005). Due to aerosol absorption, mainly by soot, smoke, and some desert dust particles, the aerosol direct radiative forcing at the surface can be much greater than the TOA forcing, and in addition, the radiative heating of the atmosphere by absorbing particles can change the atmospheric temperature structure, affecting vertical mixing, cloud formation and evolution, and possibly large-scale dynamical systems such as the monsoons (Kim et al., 2006; Lau et al., 2008). By realizing aerosol's climate significance and the challenge of charactering highly variable aerosol amount and properties, the US Climate Change Research Initiative (CCRI) identified research on atmospheric concentrations and effects of aerosols specifically as a top priority (NRC, 2001).

1.3. Reducing Uncertainties in Aerosol-Climate Forcing Estimates

Regional as well as global aerosol radiative effects on climate are estimated primarily through the use of climate models (e.g., Penner et al., 1994; Schulz et al., 2006). These numerical models are evaluated based on their ability to simulate the aerosol- and cloud-related processes that affect climate for current and past

Table 1.2. Top-of-atmosphere, cloud-free, instantaneous direct aerosol radiative forcing dependence on aerosol and surface properties. Here TWP, SGP, and NSA are the Tropical West Pacific island, Southern Great Plains, and North Slope Alaska observation stations maintained by the DOE ARM program, respectively. Instantaneous values are given at specific solar zenith angle. Upper and middle parts are from McComiskey et al. (2008a). Representative, parameter-specific measurement uncertainty upper bounds for producing 1 W m^{-2} instantaneous TOA forcing accuracy are given in the lower part, based on sensitivities at three sites from the middle part of the table.

| Parameters | TWP | SGP | NSA |
|--|-------|-------|-------|
| Aerosol properties (AOD, SSA, g), solar zenith angle (SZA), surface albedo (A), and aerosol direct RF at TOA (F): | | | |
| AOD | 0.05 | 0.1 | 0.05 |
| SSA | 0.97 | 0.95 | 0.95 |
| g | 0.8 | 0.6 | 0.7 |
| A | 0.05 | 0.1 | 0.9 |
| SZA | 30 | 45 | 70 |
| $F (\text{W m}^{-2})$ | -2.2 | -6.3 | 2.6 |
| Sensitivity of cloud-free, instantaneous, TOA direct aerosol radiative forcing to aerosol and surface properties, W m^{-2} per unit change in property: | | | |
| $\partial F / \partial (\text{AOD})$ | -45 | -64 | 51 |
| $\partial F / \partial (\text{SSA})$ | -11 | -50 | -60 |
| $\partial F / \partial g$ | 13 | 23 | 2 |
| $\partial F / \partial A$ | 8 | 24 | 6 |
| Representative measurement uncertainty upper bounds for producing 1 W m^{-2} accuracy of aerosol RF: | | | |
| AOD | 0.022 | 0.016 | 0.020 |
| SSA | 0.091 | 0.020 | 0.017 |
| g | 0.077 | 0.043 | |
| A | 0.125 | 0.042 | 0.167 |

conditions. The derived accuracy serves as a measure of the accuracy with which the models might be expected to predict the dependence of future climate conditions on prospective human activities. To generate such predictions, the models must simulate the physical, chemical, and dynamical mechanisms that govern aerosol formation and evolution in the atmosphere (Figure 1.1), as well as the radiative processes that govern their direct and indirect climate impact (Figure 1.2), on all the relevant space and time scales.

Some models simulate aerosol emissions, transports, chemical processing, and sinks, using atmospheric and possibly also ocean dynamics generated off-line by separate numerical systems. These are often called Chemistry and Transport Models (CTMs). In contrast, General Circulation Models or Global Climate Models (GCMs) can couple aerosol behavior and dynamics as part of the same calculation, and are capable of representing interactions between aerosols and dynamical aspects of the climate system, although currently many

of them still use prescribed aerosols to study climate sensitivity.

The IPCC AR4 total anthropogenic radiative forcing estimate, shown in Figure 1.3, is 1.6 W m^{-2} from preindustrial times to the present, with a likely range of 0.6 to 2.4 W m^{-2} . This estimate includes long-lived GHGs, ozone, and aerosols. The increase in global mean surface temperature of 0.7°C , from the transient climate simulations in response to this forcing, yields a transient climate sensitivity (defined as the surface temperature change per unit RF) over the industrial period of 0.3 to $1.1^\circ\text{C}/(\text{W m}^{-2})$.

Under most emission scenarios, CO_2 is expected to double by the latter part of the 21st century. A climate sensitivity range of 0.3 to $1.1^\circ\text{C}/(\text{W m}^{-2})$ translates into a future surface temperature increase attributable to CO_2 forcing at the time of doubled CO_2 of 1.2 to 4.7°C . Such a range is too wide to meaningfully predict the climate response to increased greenhouse gases (e.g., Caldeira et al., 2003). As Figure 1.3 shows, the largest contribution to overall uncertainty in estimating the climate response is from aerosol RF.

The key to reducing uncertainty in the role of aerosols in climate is to much better represent the processes that contribute to the aerosol climate effects in models. This report highlights three specific areas for continued, focused effort: (1) improving measurement quality and coverage, (2) achieving more effective use of measurements to constrain model simulations and to test model parameterizations, and (3) producing more accurate representation of aerosols and clouds in models. This section provides a brief introduction to the current state of aerosol measurements and model representations of aerosol processes, as they relate to assessing aerosol impacts on climate. More complete discussion of these topics and assessment of possible next steps are given in Chapters 2, 3, and 4.

Improving measurement quality and coverage. Aerosol mass concentration, size and composition distributions, and absorption properties, as functions of location and time, are the main aerosol-specific elements of CTMs. They depend on primary particle and precursor gas emissions, on gas-to-particle conversion processes, on transport, humidification and cloud

processing, and removal mechanisms. Satellite instruments, surface-based networks (*in situ* and remote sensing), and research aircraft all contribute quantitative measurements of aerosol properties and/or distributions that can be used to help constrain models, as well as to test and refine the model representations of processes that govern aerosol life cycles. As described in Chapter 2, the current situation reflects the significant progress that has been made over the past decade in satellite, airborne, ground-based and laboratory instrumentation, actual measurements available from each of these sources, remote sensing retrieval methods, and data validation techniques.

However, each type of measurement is limited in terms of the accuracy, and spatial and temporal sampling of measured quantities. At present, satellite passive imagers monitor AOD globally up to once per day, with accuracies under cloud-free, good but not necessarily ideal viewing conditions of about 0.05 or $(0.1$ to $0.2) \times \text{AOD}$, whichever is larger, for vegetated land, somewhat better over dark water, and less well over bright desert (e.g., Kahn et al., 2005a; Remer et al., 2005). Reliable AOD retrieval over snow and ice from passive remote sensing imagers has not yet been achieved. From space, aerosol vertical distribution is provided mainly by lidars that offer sensitivity to multiple layers, even in the presence of thin cloud, but they require several weeks to observe just a fraction of a percent of the planet.

From the expansive vantage point of space, there is enough information to identify column-average ratios of coarse to fine AOD, or even aerosol air mass types in some circumstances, but not sufficient to deduce chemical composition and vertical distribution of type, nor to constrain light absorption approaching the ~ 0.02 SSA sensitivity suggested in Section 1.2.

As a result, it is difficult to separate anthropogenic from natural aerosols using currently available satellite data alone, though attempts at this have been made based on retrieved particle size and shape information (see Chapter 2). At present, better quantification of anthropogenic aerosol depends upon integrating satellite measurements with other observations and models. Aircraft and ground-based *in situ* sampling can help fill in missing physical and chemical detail, although coverage is very limited in

The key to reducing uncertainty in the role of aerosols in climate is to much better represent the processes that contribute to the aerosol climate effects in models.



both space and time. Models can contribute by connecting observed aerosol distributions with likely sources and associated aerosol types. Surface remote-sensing monitoring networks offer temporal resolution of minutes to hours, and greater column AOD accuracy than satellite observations, but height-resolved particle property information has been demonstrated by only a few cutting-edge technologies such as high-spectral-resolution lidar (HSRL), and again, spatial coverage is extremely limited.

Even for satellite observations, sampling is an issue. From the passive imagers that provide the greatest coverage, AOD retrievals can only be done under cloud-free conditions, leading to a “clear-sky bias,” and there are questions about retrieval accuracy in the vicinity of clouds. And retrievals of aerosol type from these instruments as well as from surface-based passive remote sensing require at least a certain minimum column AOD to be effective; the thresholds depend in part on aerosol type itself and on surface reflectivity, leading to an “AOD bias” in these data sets.

Other measurement-related issues include obtaining sufficiently extensive aerosol vertical distributions outside the narrow sampling beam of space-based, airborne, or ground-based lidars, retrieving layer-resolved aerosol properties, which is especially important in the many regions where multiple layers of different types are common, obtaining representative *in situ* samples of large particles, since they tend to be under-sampled when collected by most aircraft inlets, and acquiring better surface measurement coverage over oceans.

Achieving more effective use of measurements to constrain models. Due to the limitations associated with each type of observational data record, reducing aerosol-forcing uncertainties requires coordinated efforts at integrating data from multiple platforms and techniques (Seinfeld et al., 1996; Kaufman et al., 2002a; Diner et al., 2004; Anderson et al., 2005a). Initial steps have been taken to acquire complementary observations from multiple platforms, especially through intensive field campaigns, and to merge data sets, exploiting the strengths of each to provide better constraints on models (e.g., Bates et al., 2006; Yu et al., 2006; Kinne et al., 2006; see Chapter 2, Section 2.2.6). Advanced instrument concepts, coordinated measurement strategies, and retrieval

techniques, if implemented, promise to further improve the contributions observations make to reducing aerosol forcing uncertainties.

Producing more accurate representation of aerosols in models. As discussed in Chapter 3, models, in turn, have developed increasingly sophisticated representations of aerosol types and processes, have improved the spatial resolution at which simulations are performed, and through controlled experiments and inter-comparisons of results from many models, have characterized model diversity and areas of greatest uncertainty (e.g., Textor et al., 2006; Kinne et al., 2006).

A brief chronology of aerosol modeling used for the IPCC reports illustrates these developments. In the IPCC First Assessment Report (1990), the few transient climate change simulations that were discussed used only increases in greenhouse gases. By IPCC Second Assessment Report (1995), although most GCMs still considered only greenhouse gases, several simulations included the direct effect of sulfate aerosols. The primary purpose was to establish whether the pattern of warming was altered by including aerosol-induced cooling in regions of high emissions such as the Eastern U.S. and eastern Asia. In these models, the sulfate aerosol distribution was derived from a sulfur cycle model constrained by estimated past aerosol emissions and an assumed future sulfur emission scenario. The aerosol forcing contribution was mimicked by increasing the surface albedo, which improved model agreement with the observed global mean temperature record for the final few decades of the twentieth century, but not for the correct reasons (see Chapter 3).

The IPCC Third Assessment Report (TAR, 2001) report cited numerous groups that included aerosols in both 20th and 21st century simulations. The direct effect of sulfate aerosols was required to reproduce the observed global temperature change, given the models’ climate sensitivity and ocean heat uptake. Although most models still represented aerosol forcing by increasing the surface albedo, several groups explicitly represented sulfate aerosols in their atmospheric scattering calculations, with geographical distributions determined by off-line CTM calculations. The first model calculations that included any indirect effects of aerosols on clouds were also presented.

Due to the limitations associated with each type of observational data record, reducing aerosol-forcing uncertainties requires coordinated efforts at integrating data from multiple platforms and techniques.



Continued progress with measurement, modeling, and at the interface between the two, promises to improve estimates of aerosol contribution to climate change.



The most recent IPCC assessment report (AR4; 2007) summarized the climate change experiments from more than 20 modeling groups that this time incorporated representations of multiple aerosol species, including black and organic carbon, mineral dust, sea salt and in some cases nitrates (see Chapter 3). In addition, many attempts were made to simulate indirect effects, in part because the better understood direct effect appeared to be insufficient to properly simulate observed temperature changes, given model sensitivity. As in previous assessments, the AR4 aerosol distributions responsible for both the direct and indirect effect were produced off-line, as opposed to being run in a coupled mode that would allow simulated climate changes to feed back on the aerosol distributions.

The fact that models now use multiple aerosol types and often calculate both direct and indirect aerosol effects does not imply that the requisite aerosol amounts and optical characteristics, or the mechanisms of aerosol-cloud interactions, are well represented. For example, models tend to have lower AOD relative to measurements, and are poorly constrained with regard to speciation (see Table 3.2 and Figure 3.1 in Chapter 3). To bridge the gap between measurements and models in this area, robust relationships need to be established for different aerosol types, connecting the AOD and types retrieved from spacecraft, aircraft, and surface remote sensing observations, with the aerosol mass concentrations that are the fundamental aerosol quantities tracked in CTMs and GCMs.

As detailed below, continued progress with measurement, modeling, and at the interface between the two, promises to improve estimates of aerosol contributions to climate change, and to reduce the uncertainties in these quantities reflected in Figure 1.3.

1.4 Contents of This Report

This report assesses current understanding of aerosol radiative effects on climate, focusing on developments of aerosol measurement and modeling subsequent to IPCC TAR (2001). It reviews the present state of understanding of aerosol influences on Earth's climate system, and in particular, the consequences for climate change of their direct and indirect effects. This report does not deal with several natural

forcings that involve aerosols. Stratospheric aerosols produced by large volcanic eruptions exert large, short-term effects which are particularly important for characterizing climate system response to forcing, and the effects of recent eruptions (e.g. Pinatubo) are well documented (e.g., Minnis et al., 1993; McCormick et al., 1995; Robock et al., 2002). However these effects are intermittent and have only short-term environmental impacts (ca. 1 year). Galactic cosmic rays, modulated by the 11-year solar cycle, have been reported to correlate with the total cloud cover (e.g., Svensmark and Friis-Christensen, 1997), possibly by aiding the nucleation of new particles that grow into cloud condensation nuclei (e.g., Turco et al., 1998). However, the present mainstream consensus is that these phenomena exert little to no effect on cloud cover or other cloud properties (e.g., Lockwood and Fröhlich, 2008; Kristjánsson et al., 2008).

The Executive Summary reviews the key concepts involved in the study of aerosol effects on climate, and provides a chapter-by-chapter summary of conclusions from this assessment. Chapter 1 provides basic definitions, radiative forcing accuracy requirements, and background material on critical issues needed to motivate the more detailed discussion and assessment given in subsequent chapters.

Chapter 2 assesses the aerosol contributions to radiative forcing based on remote sensing and *in situ* measurements of aerosol amounts and properties. Current measurement capabilities and limitations are discussed, as well as synergy with models, in the context of the needed aerosol radiative forcing accuracy.

Model simulation of aerosols and their direct and indirect effects are examined in Chapter 3. Representations of aerosols used for IPCC AR4 (2007) climate simulations are discussed, providing an overview of near-term modeling option strengths and limitations for assessing aerosol forcing of climate.

Finally, Chapter 4 provides an assessment of how current capabilities, and those within reach for the near future, can be brought together to reduce the aerosol forcing uncertainties reported in IPCC AR4 (2007).