Aerosol Pollution Impact on Precipitation: A Scientific Review

Prof. Zev Levin, Chairman

Prof. William Cotton, Vice Chairman
## Lead Authors:

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andreae, Meinrat O.</td>
<td>Max Planck Institute for Chemistry, Mainz</td>
<td>Germany</td>
</tr>
<tr>
<td>Artaxo, Paulo</td>
<td>University of Sao Paulo</td>
<td>Brazil</td>
</tr>
<tr>
<td>Barrie, Leonard. A.</td>
<td>WMO</td>
<td>Switzerland</td>
</tr>
<tr>
<td>Brenguier, Jean-Louis</td>
<td>Meteo France</td>
<td>France</td>
</tr>
<tr>
<td>Cotton, William R.</td>
<td>Colorado State University</td>
<td>USA</td>
</tr>
<tr>
<td>Feingold, Graham</td>
<td>NOAA</td>
<td>USA</td>
</tr>
<tr>
<td>Gong, Sunling L.</td>
<td>Environment Canada</td>
<td>Canada</td>
</tr>
<tr>
<td>Hegg, Dean A.</td>
<td>University of Washington</td>
<td>USA</td>
</tr>
<tr>
<td>Kaufman, Yoram (Deceased)</td>
<td>NASA, Goddard</td>
<td>USA</td>
</tr>
<tr>
<td>Levin, Zev</td>
<td>Tel Aviv University</td>
<td>Israel</td>
</tr>
<tr>
<td>Lohmann, Ulrike</td>
<td>ETH, Zurich</td>
<td>Switzerland</td>
</tr>
<tr>
<td>Tanre, Didier</td>
<td>CNRS, University of Lille</td>
<td>France</td>
</tr>
<tr>
<td>Yuter, Sandra</td>
<td>North Carolina State University</td>
<td>USA</td>
</tr>
</tbody>
</table>

### Liaisons:

- L.A. Barrie (WMO)
- R. List (IUGG)
### Co-authors and contributors

<table>
<thead>
<tr>
<th>Authors</th>
<th>Chapters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arimoto, R.</td>
<td>4</td>
</tr>
<tr>
<td>Baltensperger, U.</td>
<td>5</td>
</tr>
<tr>
<td>Barrie, L. A.</td>
<td>3</td>
</tr>
<tr>
<td>Brenguier, J. L.</td>
<td>2</td>
</tr>
<tr>
<td>Bruintjes, R. T.</td>
<td>8</td>
</tr>
<tr>
<td>Cotton, W. R.</td>
<td>3, 6</td>
</tr>
<tr>
<td>Feichter, J.</td>
<td>3</td>
</tr>
<tr>
<td>Feingold, G.</td>
<td>2, 6</td>
</tr>
<tr>
<td>Guibert, S.</td>
<td>4</td>
</tr>
<tr>
<td>Hallett, J.</td>
<td>2, 6</td>
</tr>
<tr>
<td>Hobbs, P. V. (Deceased)</td>
<td>2, 8</td>
</tr>
<tr>
<td>Kinne, S.</td>
<td>4</td>
</tr>
<tr>
<td>Kloster, S.</td>
<td>3</td>
</tr>
<tr>
<td>Kreidenweis, S. M.</td>
<td>3, 4</td>
</tr>
<tr>
<td>Levin, Z.</td>
<td>2, 3, 5, 8</td>
</tr>
<tr>
<td>Li, S.M.</td>
<td>4</td>
</tr>
<tr>
<td>Lioussse, C.</td>
<td>3</td>
</tr>
<tr>
<td>Lohmann, U.</td>
<td>2</td>
</tr>
<tr>
<td>Radke, L. F.</td>
<td>3, 5</td>
</tr>
<tr>
<td>Schulz, M.</td>
<td>4</td>
</tr>
<tr>
<td>Shepherd, J. M.</td>
<td>5</td>
</tr>
<tr>
<td>Stier, P.</td>
<td>3</td>
</tr>
<tr>
<td>Strawbridge, K.</td>
<td>4</td>
</tr>
<tr>
<td>Sugimoto, N.</td>
<td>4</td>
</tr>
<tr>
<td>Terblanche, D.</td>
<td>8</td>
</tr>
<tr>
<td>Textor, C.</td>
<td>3, 4</td>
</tr>
<tr>
<td>Zhang, L.M.</td>
<td>4</td>
</tr>
</tbody>
</table>
# Scientific Reviewers

**Chairperson: Dr. George Isaac, Environment Canada**

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ayers, Greg</td>
<td>CSIRO Marine and Atmospheric Research</td>
<td>Australia</td>
</tr>
<tr>
<td>Barth, Mary</td>
<td>NCAR</td>
<td>USA</td>
</tr>
<tr>
<td>Bormann, Stephan</td>
<td>Johannes-Gutenberg-University</td>
<td>Germany</td>
</tr>
<tr>
<td>Choularton, Thomas</td>
<td>University of Manchester</td>
<td>UK</td>
</tr>
<tr>
<td>DeMott, Paul</td>
<td>Colorado State University</td>
<td>USA</td>
</tr>
<tr>
<td>Flossmann, Andrea</td>
<td>Laboratoire de Météorologie Physique/OPGC; Université Blaise Pascal/CNRS</td>
<td>France</td>
</tr>
<tr>
<td>Kahn, Ralph</td>
<td>Jet Propulsion Laboratory</td>
<td>USA</td>
</tr>
<tr>
<td>Khain, Alexander</td>
<td>The Hebrew University of Jerusalem</td>
<td>Israel</td>
</tr>
<tr>
<td>Leaitch, Richard</td>
<td>Environment Canada</td>
<td>Canada</td>
</tr>
<tr>
<td>Pandis, Spyros</td>
<td>University of Patras</td>
<td>Greece</td>
</tr>
<tr>
<td>Rosenfeld, Daniel</td>
<td>The Hebrew University of Jerusalem</td>
<td>Israel</td>
</tr>
<tr>
<td>Ryan, Brian</td>
<td>CSIRO Marine and Atmospheric Research</td>
<td>Australia</td>
</tr>
<tr>
<td>Twohy, Cynthia</td>
<td>Oregon State University</td>
<td>USA</td>
</tr>
<tr>
<td>Vali, Gabor</td>
<td>University of Wyoming</td>
<td>USA</td>
</tr>
<tr>
<td>Yau, Peter</td>
<td>McGill University</td>
<td>Canada</td>
</tr>
<tr>
<td>Zipser, Ed</td>
<td>University of Utah</td>
<td>USA</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS

PREFACE ........................................................................................................................................... 6
EXECUTIVE SUMMARY .................................................................................................................... 8
CHAPTER 1:  INTRODUCTION
Lead Authors: Leonard A. Barrie, Ulrike Lohmann, Sandra Yuter ................................................. 16
CHAPTER 2: PRINCIPLES OF CLOUD AND PRECIPITATION FORMATION
Lead Authors: William R. Cotton and Sandra Yuter ......................................................................... 31
CHAPTER 3: SOURCES AND NATURE OF ATMOSPHERIC AEROSOLS
Lead Authors: Meinrat O. Andreae, Dean A. Hegg, and Urs Baltensperger ................................. 69
CHAPTER 4: THE DISTRIBUTION OF ATMOSPHERIC AEROSOLS:
TRANSPORT, TRANSFORMATION AND REMOVAL
Lead Author: Sunling L. Gong & Leonard A. Barrie ........................................................................ 125
CHAPTER 5: IN SITU AND REMOTE SENSING TECHNIQUES FOR MEASURING
AEROSOLS, CLOUDS AND PRECIPITATION
Lead authors: Didier Tanré, Paulo Artaxo, Sandra Yuter, Yoram Kaufman ................................. 183
CHAPTER 6: EFFECTS OF POLLUTION AND BIOMASS AEROSOLS ON CLOUDS
AND PRECIPITATION; OBSERVATIONAL STUDIES
Lead Authors: Zev Levin, Jean-Louis Brenguier ............................................................................. 259
CHAPTER 7:  EFFECTS OF POLLUTION AEROSOL INCLUDING BIOMASS
BURNING ON CLOUDS AND PRECIPITATION: NUMERICAL MODELING
STUDIES
Lead authors: Graham Feingold, William R. Cotton, Ulrike Lohmann, Zev Levin .......... 302
CHAPTER 8: PARALLELS AND CONTRASTS BETWEEN DELIBERATE CLOUD
SEEDING AND AEROSOL POLLUTION EFFECTS
Lead Author: William R. Cotton ..................................................................................................... 342
CHAPTER 9:  SUMMARY .................................................................................................................. 364
CHAPTER 10: RECOMMENDATIONS ............................................................................................... 371
APPENDIX:  LIST OF ABBREVIATIONS USED IN THE REPORT ............................................. 377
REFERENCES ..................................................................................................................................... 381
PREFACE

Clouds are known to play a major role in climate through their direct interactions with solar radiation. In addition, precipitation from clouds is the overwhelmingly most important mechanism that replenishes ground water and completes the hydrological cycle. Changes in either the amounts and/or the spatial and temporal distribution of precipitation will have dramatic impacts on climate and on society. Increases or decreases in rainfall in one region could affect rainfall downwind. Similarly, changes in rainfall distribution will strongly affect semi-arid regions that are of dire need of water.

One of the factors that could contribute to cloud and precipitation modification is aerosol pollution from various sources such as urban air pollution and biomass burning. In 2003, the WMO and the IUGG recognized the potential danger from such effects and passed resolutions aimed at focusing attention to this issue. As a follow up to this resolution, the WMO and IUGG formed an international forum composed of a number of experts to review the state of the science and to identify areas that need further study.

Prof. Peter Hobbs from the University of Washington in Seattle was appointed as the chairman of this forum. Unfortunately, Prof. Hobbs passed away in mid 2005 and the responsibility for the report was transferred to Prof. Zev Levin (Chairman) and William Cotton (Co-chairman). The organization of this report reflects in a large part, the layout that was set up by Prof. Hobbs.

Dr. George Isaac from Environment Canada, the lead reviewer of this document, was responsible for the review process and the selection of the review panel. We would like to acknowledge the very important contributions to this document made by George and his team.

In 2006, just after the first version of the report was completed, Dr. Yoram Kaufman, the lead author of Chapter 5 died in a tragic traffic accident. We are thankful to Dr. Didier Tanre for agreeing to step into Yoram's big shoes and help us complete this report. During the preparation of the final draft of this report in 2006, Dr. Brian Ryan, one of the reviewers died.
suddenly from a stroke. We all appreciate Brian’s contributions to the report and his death is a great loss.

Since the report is aimed to both the general public and to experts in the field, we took upon ourselves to include chapters that discuss in detail the scientific background and the most recent findings in this field. Each chapter was written by a number of people under the leadership of the lead author. Since the contributors to each chapter were numerous, we listed them at the beginning of the report. It should be noted that some of the contributors provided input to more than one chapter. We would like to thank all of them for their contribution.

Prof. Zev Levin
Department of Geophysics and Planetary Science
Tel Aviv University
Israel

Prof. William R. Cotton
Department of Atmospheric Science
Colorado State University
Ft. Collins, CO
USA
EXECUTIVE SUMMARY

The WMO and IUGG resolutions handed a mandate to IAPSAG to review the effects of aerosol pollution, including biomass burning, on precipitation. The initial set up and framework of IAPSAG were illustrated by List [2004]. This document is a review of our knowledge of the relationship between aerosols and precipitation reaching the Earth's surface and it includes a list of recommendations that could help to advance our knowledge in this area.

**Background:** Life on Earth is critically dependent upon the continuous cycling of water between oceans, continents and the atmosphere. The atmospheric component of this cycle involves surface water moving into the atmosphere via wind-driven evaporation and biosphere-modulated evapotranspiration followed by transport and dispersion, multiple cycles of cloud formation/evaporation and ultimately removal as precipitation (rain or snow), fog water deposition or dew formation.

Precipitation (including rain, snow, and hail) is the primary mechanism for transporting water from the atmosphere back to the Earth's surface. It is also the key physical process that links aspects of climate, weather, and the global hydrological cycle. Although, precipitation is the parameter that has the most direct and significant influence on the quality of human lives in terms of the availability of water, the changes in temperature, past and future, have received the bulk of the public’s attention regarding global climate change. Precipitation is a more difficult component of the climate to assess because it is much more variable (locally, seasonally, and possibly long-term) than temperature. It is affected primarily by atmospheric dynamics; however, it is also influenced by cloud microphysical processes associated with aerosol properties, which are primarily responsible for cloud drop and ice crystal formation. Changes in precipitation regimes and the frequency of extreme weather events, such as floods, droughts, severe ice/snow storms, monsoon fluctuations and hurricanes are of great potential importance to life on the planet. Thus, by influencing the hydrological cycle, natural and anthropogenic changes in atmospheric aerosols might have important implications for precipitation, which in turn could feed back to climate changes.
**Aim:** The main goal of this document is to review what is known about the relationship between aerosols and precipitation reaching the Earth’s surface and to suggest ways to advance this knowledge.

**An overview of the report:** From an Earth Science perspective, a key question is how changes expected in climate will translate into changes in the hydrological cycle, and what trends may be expected in the future. With changes in climate, there will be changes in the atmospheric aerosol, whether from natural sources because of changes in winds or other atmospheric conditions, or whether from anthropogenic sources due to changes in activities as a response to climate change. The hydrological cycle is driven by a multiplicity of complex processes and interactions, many of which are inadequately understood and poorly represented in weather and climate models. We require a much better understanding and hence predictive capability of the moisture and energy storages and exchanges among the Earth’s atmosphere, oceans, continents and biological systems over a wide range of space scales from local to regional to hemispheric to global and on time scales from hours to decades. These elements are discussed in the Introduction chapter.

To set the stage for the discussion on aerosols and clouds, we briefly summarize the complex interactions between aerosols and clouds and the mechanisms that lead to precipitation formation. This is outlined in Chapter 2.

Chapters 3 and 4 discuss the nature and the global temporal and spatial distributions of aerosols, followed in Chapter 5 by a review of the measurement methods and instrumentation used to quantify precipitation.

In Chapters 6 and 7 we discuss the observations and model simulations of the effects of pollution aerosols on precipitation and identify some of the difficulties in estimating these effects. Finally Chapter 8 illuminates some of the parallels between the effects of pollution on precipitation and the attempts to artificially modify precipitation through cloud seeding.

**Summary:** From this report it is clear that aerosol-pollution affects the amount, and spatial and temporal distribution of clouds, in general agreement with our physical understanding of cloud processes. However, much uncertainty regarding these processes remains. Moreover, because of the complex interactions between environmental meteorological parameters,
aerosols, cloud microphysics and dynamics, it is difficult to establish clear causal relationships between aerosols and precipitation and to determine the sign of the precipitation change in a climatological sense. Based on many observations and model simulations the effects of aerosols on clouds are more clearly understood (particularly in ice-free clouds); the effects on precipitation are less clear.

Processes of aerosol effects on precipitation

Atmospheric aerosols are highly variable in space and time. A subset of atmospheric aerosols is responsible for the formation of drops (cloud condensation nuclei) and ice crystals (ice nuclei). The initial sizes and concentrations of these hydrometeors determine the efficiency with which precipitation will initially form. A subset of the cloud condensation nuclei population are called giant cloud condensation nuclei and/or ultragiant aerosol particles that have been observed to produce large drops that serve as embryos for initiating coalescence growth and accelerate precipitation development. Modeling studies suggest their importance increases with increasing pollution. There is a point of diminishing returns, however, since modeled precipitation amounts tend to decrease steadily with increasing pollution (all else being equal) and the absolute potential increase in rainfall is small.

Ice nuclei are a smaller sub-set of atmospheric aerosols than cloud condensation nuclei, however, their role in precipitation formation in certain clouds is critical. This is because of ice crystals grow fast to precipitation size embryos due to their lower saturation vapor pressure than water. The low concentrations of the ice nuclei and the number of different physical mechanisms (see Chapters 2 and 6) that they can take to form ice in clouds make it difficult to measure them. One would expect a good correlation between the concentrations of ice crystal and ice nuclei. However, measurements suggest that in many clouds the concentrations of ice crystals exceed that of the ice nuclei by several orders of magnitude. It is clear that the effects of aerosol pollution on ice formation in clouds is harder to quantify and with the recent improvement of measuring techniques, many more studies are required in laboratories and in different field environments.

Modifying the rate of growth of precipitating particles will likely affect the dynamics of the clouds, thus modifying the depth, size, lifetime and propagation of clouds or cloud systems. In fact modeling studies suggest that in deep convective clouds pollution not only affects
cloud and precipitation by ingesting aerosols into initial or primary cells, but also through the formation of secondary cells and cold-pool dynamics in a very nonlinear way and often many hours after initial convective cell formation.

Larger concentrations of CCN produce larger concentrations of smaller cloud drops that are slower to grow into raindrops. There are reports that show that reduction of cloud drop size also delays the formation of ice, which then forms at higher altitudes and lower temperatures. This may lead to suppression of precipitation in shallow and short-lived clouds, such as those that form during winter over topographical barriers. On the other hand, modeling and observational studies indicate that when deep convective clouds are prevalent, this delay in precipitation can lead to greater amounts of supercooled water being transported aloft and that once it freezes it can invigorate the storm dynamics. In some environments, the invigoration of the convective storm updrafts can lead to enhancement of surface precipitation, while in other environments reduced precipitation can result.

The evidence for aerosol effects on precipitation
Orographic clouds form when moist air is lifted over mountain barriers. This orographic lifting greatly amplifies precipitation over that which occurs for the prevailing weather systems over flat terrain. For many parts of the world orographic precipitation is the dominate water resource. Both observational and modeling studies suggest that orographic clouds are highly susceptible to modification of precipitation by pollution owing to the modest liquid water contents in them, the relatively short time the drops and ice crystals spend in the clouds and their large areal coverage. Some observational studies suggest that aerosol pollution may decrease orographic precipitation as much as 30% annually. In spite of many remaining uncertainties, this potential effect has major ramifications in watersheds, which depend upon orographic precipitation as the major water resource,

Global climate model (GCM) estimates of the change in global mean precipitation over the last 100 years due to the total aerosol effects (direct and indirect) vary between no change to -4.5%. A large contributor to this variability is the different methods of representing clouds and aerosol effects in the different models. The differences among models are amplified over land, ranging from -1.5% to -8.5%. Recent coupled ocean mixed layer GCM simulations show that the cooling due to the direct and indirect effects by aerosols reduce surface latent
and sensible heat transfer and, as a consequence, act to reduce surface evaporation and evapotranspiration. The concomitant reduction in atmospheric water vapor (a greenhouse gas) acts to counter greenhouse gas warming. On the other hand, a reduction in precipitation due to an increase in aerosol pollution leads to longer-lived clouds and more water vapor in the atmosphere. Thus, the potential influence of aerosols on climate could be far more significant than previously thought. Estimates of the consequences of coupled greenhouse gas warming and aerosol cooling in the future depend on uncertain estimates of future pollution emissions and greenhouse gas releases into the atmosphere, as well as the uncertainties associated with the parameterizations of cloud-aerosol interactions in GCMs.

The complexity of aerosol-precipitation interactions is highlighted in urban environments, where field measurements of the effects of urban pollution on precipitation from convective clouds downwind of cities such as St Louis, New York and Houston have failed to link the changes in precipitation to increases in aerosol pollution. Modeling studies suggest that urban land-use effects dominate over aerosol pollution on controlling the locations and amounts of precipitation. The importance of pollution aerosols to precipitation in urban regions depends on the unique properties of the meteorology of the urban environment, its interaction with rural physiography and urban land-use, and on the background aerosol concentrations and chemical properties.

It has also been shown that strongly absorbing pollution particles (e.g. black carbon) may reduce convection due to warming of the atmosphere, and reduction in surface fluxes. This may result in shallower clouds and in some cases, no clouds at all.

**Recommendations:**

- It is recommended that a series of international projects targeted toward unraveling the complex interactions among aerosols, clouds, and precipitation be implemented. A series of international workshops and field studies are needed to address the impacts on clouds and precipitation of aerosols from a range of sources including biomass burning, dust, and industrial pollution within different regional weather regimes in the tropics, middle latitudes, and the Polar Regions. All of these studies need to have an adequate scientific component, but should also deal with specific situations occurring in the developing world by addressing
the impact of pollution/precipitation on economic, sociological, environmental and health issues.

- It is recommended that WMO/IUGG take the lead in such projects together with other UN and International Organizations. Some of those projects could be sponsored and financially supported by the countries involved. For example, the effects of an evolving industrial economy such as China on precipitation should be studied. Similarly, a study of the effects of biomass burning and dust in some of the African regions would be highly valuable for our understanding of these complex issues. The WMO/IUGG can play a key coordination role in making sure the following recommendations are implemented.

- It is also proposed to consider a project in a country where the relationship between pollution and precipitation can be explored with the biggest arsenal of the newest scientific instrumentation.

- Since ice formation in clouds are not yet clearly understood, it is recommended that further laboratory studies and in-situ measurements be conducted to clarify the nucleation mechanisms.

- A Workshop should be held on developing improved instrumentation for measuring IN, small ice particles, and precipitation.

- There is a special need for more in-situ measurements to follow and link physical processes leading from aerosol effects on cloud growth and precipitation.

- Because both observational and modeling studies suggest that orographic clouds are highly susceptible to precipitation modifications due to pollution aerosols, and because the ramifications on water resources are large, it is recommended that focused, coordinated observational and modeling campaigns are implemented to study aerosol precipitation interactions for selected watersheds where pollution effects are likely to be large. Examples include the planned Convective and Orographically induced Precipitation Study (COPS) experiment in Germany in 2007, SUPERCIP in the Sierra Nevada of California and the Colorado River Basin in the western U.S. It is recommended that similar studies be conducted in regions such as China and India, where the effects of mega cities on precipitation could be evaluated. These studies should be basin wide and include ground-based, airborne, and satellite measurements of aerosol properties, in-situ microphysics measurements, surface and radar (particularly cloud radar) measurements of cloud structures and precipitation, basin-wide hydrological measurements of liquid precipitation, snowpack and runoff, etc.
A Workshop should be held to determine a strategy for a focused, coordinated observational and modeling campaign that could address the effects that aerosols have on orographic clouds. This might include additions to existing planned measurement campaigns or a new campaign(s).

Because cloud-resolving models suggest a significant dynamic response to pollution aerosols which then modifies simulated precipitation, particularly through secondary dynamic responses of clouds via cold pools and gravity waves, it is recommended that a coordinated observational and modeling campaign be organized to investigate the response of cloud systems to varying amounts and characteristics of pollution. These studies should include ground-based, airborne, and satellite measurements of aerosol properties, surface meteorological measurements, ground-based multiple Doppler radars, and airborne Doppler radars. These campaigns should be done in regions of significant biomass burning or urban pollution sources, and in other locations where the meteorology varies such that different cloud responses can be expected.

Because of the indications of strong influences of urban land-use and aerosol pollution on precipitation and lightning, it is recommended that a coordinated modeling and observational campaign be established in a number of large metropolitan regions where convection is prevalent. Crucial to site selection is the importance of finding locations where the local physiography (i.e. topography, land-water interfaces) is relatively simple and where generally widespread aerosol pollution is minimal. Measurements should include the list given above plus greater attention to surface measurements of fluxes and land-use properties.

It is recommended that aerosol-cloud-precipitation cloud-resolving-model intercomparison studies be implemented to serve as a stimulus for model refinement research and evaluations of model performance. These should be done for a variety of regimes including warm and cold-season orographic cloud models, shallow convection and deep convection, and stratiform cloud systems. These studies could be linked to some of the campaigns suggested above or to other research campaigns where implementation of an aerosol component could be accomplished easily and economically.

It is recommended that existing and new statistical methods be applied to current and future data sets to distinguish aerosol effects on precipitation from meteorological influences. Numerical models should play an important role in this process.
• It is urged that aerosol-cloud-precipitation specialists collaborate with global climate models, GCM, developers to refine the representation (parameterization) of aerosol-cloud-precipitation processes in GCMs.

• There is a strong need to assemble data sets or climatologies that can be used for the assessment of climate simulations with GCMs including aerosol-cloud-precipitation parameterizations. Some of the parameters that need evaluation include seasonal and annual precipitation amounts, regional precipitation climatologies, global and regional aerosol distributions, and TOA radiation budgets.
CHAPTER 1: INTRODUCTION

Lead Authors: Leonard A. Barrie, Ulrike Lohmann, Sandra Yuter

1.1 The Hydrological Cycle

Life on Earth is critically dependent upon the continuous cycling of water between the oceans, the continents and the atmosphere. The hydrological cycle depicted as a cartoon in Figure 1 is dynamic and complex. The atmospheric component of this cycle involves surface water moving into the atmosphere via wind-driven evaporation and biosphere-modulated evapo/transpiration followed by transport and dispersion, multiple cycles of cloud formation/evaporation and ultimately removal as precipitation (rain or snow) or by fog water deposition or dew formation.

Figure 1.1: Schematic of the hydrological cycling of water between atmosphere, land and oceans.
Precipitation (including rain, snow, and hail) is the primary mechanism for transporting water from the atmosphere back to the Earth's surface. It is also the key physical process that links aspects of climate, weather, and the global hydrological cycle. Moreover, precipitation is the parameter that has the most direct and significant influence on the quality of human lives in terms of the availability of fresh water.

Changes in precipitation regimes and the frequency of extreme hydrologic events, such as floods, droughts, severe ice/snow storms, monsoon fluctuations and hurricanes are of great potential importance. The hydrological cycle affects and interacts with other components of the climate system such as glaciers, seasonal snow cover, polar ice cover, vegetative cover and last but not least, atmospheric aerosols.

In the context of atmospheric science, aerosol is a collective name for suspended particulate matter that ranges in size from molecular clusters of 1 nm to giant particles of ~20 µm diameter. The variation of aerosol size distribution and composition in the atmosphere depends on a complex combination of processes involving primary (direct release to the atmosphere) and secondary (formed in the atmosphere from chemical gas to particle conversion) sources, transport, dispersion, cloud processing and removal by precipitation described in Chapters 2, 3 and 4. It is widely acknowledged that their presence in the atmosphere is essential to the initial formation of clouds and that we know much less about their influence on cloud growth and precipitation formation. The main goal of this document is to review what is known about the relationship between aerosols and precipitation reaching the Earth’s surface and to suggest ways to advance this knowledge.

From an Earth Science perspective, a key question is how changes expected in climate (e.g., warming or changes in atmospheric aerosols) will translate into changes in the hydrological cycle, and what trends may be expected in the future. The hydrological cycle is driven by a multiplicity of complex processes and interactions, many of which are inadequately understood and poorly represented in weather and climate models. We require a much better understanding and hence predictive capability of moisture and energy storages and exchanges among and within the Earth’s atmosphere, oceans, continents and biological systems over a wide range of spatial scales from local to regional to hemispheric to global and of time scales from hours to decades.
Figure 1.2 is a more quantitative depiction of the hydrological cycle shown in Figure 1.1. It shows the amounts of water in the various reservoirs and the rate of exchange between them. The hydrological cycle is closely linked to the atmospheric circulation and to temperature, which determines the maximum amount of water vapor in the atmosphere. However, it in turn affects other components of the weather and climate system through complex feedback mechanisms. For example, latent heat released by the condensation of water to form clouds is an important part of the atmospheric heat balance.

The atmosphere is by far the smallest reservoir of water. Although it accounts for only 0.3% of the total mass of the atmosphere, it is an important greenhouse gas forming a naturally varying background upon which the effects of anthropogenically influenced long-lived gases such as carbon dioxide and methane are superimposed.

The residence time of water in the atmosphere given by dividing the amount of water vapor in the atmosphere \((13 \times 10^{15} \text{ kg})\) by the annual mean precipitation rate \((506 \text{ kg/year})\) is 0.03 years (9 days). Nine days is by far the shortest residence time of water in any of its reservoirs. The residence time in lakes and rivers is on the order of a year and in the

![Figure 1.2: A schematic diagram of the hydrological cycle. The numbers in the boxes are estimates of the total amounts of water in the various reservoirs in units of \(10^{15} \text{ kg}\). The numbers alongside the arrows are estimates of average annual fluxes in units of \(10^{15} \text{ kg/year}\). Adapted from Oki [1999].](image-url)
Antarctic ice sheet is ~200,000 years. Over a period of 9 days, water vapor can be transported over thousands of kilometers.

Another link between aerosols and the global hydrological cycle is by aerosols acting as cloud condensation nuclei (CCN) and ice nuclei (IN). The concentrations and effectiveness of these particles influence the formation of cloud drops and ice crystals and the resulting precipitation. Since in many parts of the world precipitation originates via the ice phase, IN from both natural and anthropogenic sources might affect precipitation. As discussed in Chapter 7, results of simulations by several cloud-resolving models and general circulation models suggest that changes in ice nuclei concentrations can have a major impact on both precipitation and on the surface energy budget, sometimes leading to surface warming and other times leading to surface cooling [Carrio et al., 2005; Lohmann, 2002].

A potentially important link between aerosols and the global hydrological cycle is through the surface energy budget. As aerosol and cloud optical depths increase, less solar radiation reaches the surface. This is balanced by decreased outgoing energy from the surface, either in the form of terrestrial longwave radiation or by reduced latent and sensible heat fluxes. As shown from coupled global climate model/mixed-layer ocean model simulations, the decrease in solar radiation seems to be more important in controlling the amount of evaporation in the global mean than the increase in surface temperature resulting from the increase in greenhouse gases [Liepert et al., 2004]. Aerosols could impact global atmospheric circulations through several mechanisms including changes in cloud cover, ocean and land-surface heating, rate of latent heat release, and precipitation efficiency.

### 1.2 Global Distribution of Clouds and Precipitation

Clouds cover about 60% of the Earth's surface [Rossow et al., 1993] and through scattering of solar radiation, play an important role in the radiative balance of the Earth. At any given location, daily cloud and precipitation variability can be large since clear sky conditions can transition to overcast skies and rain within a few hours. When temporal averages of one month or more are calculated, climatological patterns of cloud and precipitation start to emerge. Large-scale atmospheric circulation yields regional differences in cloud and precipitation distributions. The atmospheric circulation is driven by the combined effects of
the latent heat released as water vapor condenses into cloud droplets as well as the direct radiative forcing of water vapor, greenhouse gases and aerosols, the indirect forcing of aerosols through their effects on clouds [IPCC 2001] and the varying surface reflectivity to solar radiation.

For a given region, the largest source of variability of clouds and precipitation is usually the seasonal cycle. Monsoons are seasonal circulations driven by temperature contrasts between ocean and land surfaces. Monsoon circulations reverse direction with the seasons. Interannual variations in large-scale atmospheric oscillations such as El-Nino-Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO) are associated with interannual variability of global precipitation distribution.

*Hartmann et al.* [1992] used the International Satellite Cloud Climatology Project (ISSCP) cloud cover, cloud-top height and cloud-type datasets [*Rossow and Schiffer*, 1991] to yield the global patterns of clouds subdivided by cloud type (Figures 1.3-1.5).
Figure 1.3. Geographic distribution of cloud fractional coverage for June, July, August 1983-1990 for (a) high thin cloud (cloud top height > 440 hPa and optical depth < 9.38), (b) high thick cloud (cloud top height > 440 hPa and optical depth > 9.38), (c) Low cloud (cloud top height < 440 hPa) and (d) total cloud. Plots courtesy of D. L. Hartmann, University of Washington.

High-cloud tops occur in regions with large-scale ascent in the tropics and in midlatitude storm tracks associated with baroclinic waves. Optically thick high clouds occur where convection is most active and cover a slightly smaller area of the globe than high, optically thin clouds. Low clouds are predominately oceanic and occur most frequently over the subtropical eastern ocean margins and over middle and high latitude oceans. The low clouds over the subtropical eastern oceans are associated with large-scale subsidence (descending branch of the Hadley cell) and lower than average sea surface temperature which yield
stratocumulus clouds trapped below an inversion. Low clouds over middle latitude oceans are usually stratus [Norris 1998].

Figure 1.4: As in Figure 1.3 except for December, January, February 1983-1990. Plots courtesy of D. L. Hartmann, University of Washington.

Total cloud cover is greatest over middle latitude oceans in both the northern and southern hemispheres (Figure 1.5). Minima in cloud cover occur over subtropical desert regions but smaller cloud amounts also occur over the Caribbean Sea and the southern subtropical zones of the Pacific, Atlantic and Indian oceans [Hartmann, 1994]. Cloud cover has seasonal maxima in regions of intense tropical convection such as the Bay of Bengal in June-August (Figure 1.3) and over the Amazon and Indonesia in December-February (Figure 1.4).
Precipitation forms within clouds, so it is not surprising that the global distribution of precipitation (Figure 1.6) has many similarities to the global distribution of clouds with high tops (Figure 1.5a). However, there are important differences between the cloud and precipitation maps related to the weak instantaneous correlations of cloud top properties and surface rainfall (see Chapters 5, 6 and 7).
Global average precipitation rate is estimated to be 2.6 mm day\(^{-1}\), with higher average values over the ocean (2.8 mm day\(^{-1}\)) as compared to land (2.1 mm day\(^{-1}\)). On average, three quarters of the global precipitation falls over the world’s oceans [Adler et al., 2003].

The zonally averaged precipitation (Figure 1.7) is closely related to the global circulation pattern and cloud pattern with peak values at the latitude of the ITCZ for the ocean, peak values over land associated with Amazonia and Indonesia, and secondary maxima at middle latitudes associated with the storm tracks of baroclinic disturbances.
Figure 1.7. Zonally averaged annual mean GPCP version 2 climatology (mm day$^{-1}$) calculated for 2.5° x 2.5° grid boxes. Total (solid line), ocean-only grid boxes (dotted line), Land-only grid boxes (dashed line) Gaps in high latitudes occur where there are no land or ocean grid boxes. From Adler et al. [2003].

The relative contributions of liquid-phase versus ice-phase precipitation mechanisms to the total precipitation at given locations need to be known in order to determine the potential impacts of aerosols on precipitation since the aerosols have different impacts on the two precipitation mechanisms. These contributions are simple to estimate for shallow clouds that contain only liquid-phase precipitation but are very difficult to determine for mixed-phase clouds, which yield the vast majority of global precipitation.

In addition to the limitations arising from incomplete understanding of precipitation processes, direct determination of inadvertent effects of aerosols is also problematic due to large uncertainties in precipitation measurements (Chapter 5). The current accuracy and precision of precipitation measurements (rain and snowfall) in many parts of the globe makes the detection of effects of both deliberate and inadvertent aerosol seeding of clouds on surface precipitation impossible.
Global estimates of tropical precipitation between 30° N and 30° S have the lowest uncertainties as compared to middle latitude and polar precipitation estimates (Section 5.5), so we will focus our discussion of global precipitation variability to tropical areas. Figure 1.8 shows monthly global anomalies from a 23-yr tropical climatological mean value based on GPCP superimposed on 12-month running mean. Also overlaid are the Niño 3.4 SST index and months with significant volcanic eruptions. Monthly anomalies are typically smaller over ocean, usually < 0.4 mm day⁻¹, compared to larger variations occurring over land [Adler et al., 2003, Haddad et al., 2004].

![Figure 1.8. Tropical (30° N to 30° S) averages of monthly precipitation anomalies (mm day⁻¹) for (top) total, (middle) ocean, and (bottom) land. Vertical dashed lines indicate months of significant volcanic eruptions. Black curves in all three panels indicate the Niño 3.4 SST index in °C. From Adler et al. [2003].](image)

Diverse studies have shown that ENSO is a major factor in the interannual variability of sea surface temperatures, large-scale atmospheric circulation, and the global precipitation distribution [U.S. CLIVAR Pan American Implementation Panel, 2002]. Dry periods over the Indonesia and Malaysia region, South America and Africa are associated with the warm phase of ENSO (Figure 1.9). These are some of the same regions where biomass burning aerosols are prevalent. The impact of major volcanic events on precipitation is difficult to
isolate from that of ENSO [Adler et al., 2003]. Removal of the ENSO signature on precipitation is a necessary step in assessing the long-term impact of aerosols on precipitation.

In contrast to many climate predictions, the 23-year record examined by Adler et al. [2003] in Figure 1.8 shows no noticeable trend in global or tropical precipitation. However, the size of the predicted precipitation increase associated with global warming is smaller in magnitude than the interannual variations associated with ENSO and is likely not detectable in such a short record [Adler et al, 2003].

![Figure 1.8](image)

**Figure 1-9.** Nature of global El Nino impacts during 1997-1998. Brown indicates dry regions while green indicates wet. From U.S. CLIVAR Pan American Implementation Panel [2002].

### 1.3 Global Aerosol Distributions

The spatial distribution of aerosols as represented by average monthly aerosol optical depth in Figure 1.10 illustrates the complexity of the aerosol and the fact that it is transported long distances before being deposited. Details of these aerosols are given in Chapters 3 and 4. The main sources are divisible into natural soil dust, sea salt, marine biogenic sulfur, terrestrial biogenic and boreal biomass burning while anthropogenic are those from industry, fossil fuel combustion and human-activity related biomass burning. Each type has its own characteristic sources, size distribution and effectiveness as CCN and IN.
Figure 1.10  Long term mean AOD in the northern hemisphere months of (a) May (b) July and c) Sept. compiled by combining data from satellites with surface-based aerosol sunphotometer data. Satellite observations were validated using ground-based AOD observations [Kinne, 2007].
1.4 Aerosol-Precipitation Interactions: An Inherent Part of Climate Change

Aerosols, cloud properties and precipitation are recognized by the United Nations Framework Convention on Climate Change and the Global Climate Observing System as essential climate variables [GCOS-107, 2006]. IPCC [2001] acknowledges the importance of anthropogenic aerosol impacts on clouds through effects on cloud albedo (indirect effect #1). The best estimate of this indirect cloud albedo effect due to anthropogenic aerosols from pre-industrial times to the present-day from different climate models varies between –0.5 to –1.9 W m\(^{-2}\) [Lohmann and Feichter, 2005]. Secondly, a reduction of the drizzle production due to the more and smaller cloud droplets may prolong of cloud lifetimes and hence cloud coverage (cloud lifetime or indirect effect #2). The cloud lifetime effect, the semi-direct effect (heating of the air due to absorption of solar radiation by absorbing aerosol) and aerosol-ice cloud interactions in response to anthropogenic aerosols are considered to be part of the climate response rather than radiative forcing [IPCC, 2001]. Nevertheless, the total effect resulting from these aerosol-cloud interactions can be estimated from climate model simulations. These simulations are conducted such that one multi-year simulation uses pre-industrial aerosol and their precursor emissions (i.e. the anthropogenic sources are switched off) and another multi-year simulation that uses present-day aerosol emissions. Aerosols also tend to reduce the net radiation reaching the surface and thereby lower surface temperatures indirectly affecting the hydrological cycle and atmospheric circulation. The impact of aerosols on the radiation balance and on the global mean precipitation is discussed in chapter 7.

1.5 Unraveling the Aerosol-Precipitation Factor in Long Term Observations: Ongoing Experiments of Opportunity

Changes taking place in certain regions of the globe, historically and in the future, may provide an opportunity to separate the aerosol effect on precipitation from other factors given a good experimental design and adequate systematic observations for analysis. These studies may also be able address the distributions of air pollution and fresh water which have economic, sociological, environmental and health impacts. It is recommended that WMO/IUGG participate in designing such field studies, together with other UN, international and national organizations.
• Indian subcontinent: documented changes in precipitation in an evolving industrial economy dominated by the Indian Monsoon
• China: documented changes in precipitation in an evolving industrial economy.
• Amazonia: a history of biomass burning in Amazonia with simultaneous research studies
• Influence of Saharan Dust and biomass burning on the African monsoon and Atlantic Tropical Storms development: a possibility of testing the effects on storm development and accompanying effects on precipitation
• Volcanic eruptions: massive perturbations of climate and general circulation through aerosols.

1.6 The Structure of This Review

This review addresses current knowledge of how natural and anthropogenic aerosols might affect clouds, precipitation and hence the hydrological cycle and climate. Diagnosis of the link between aerosol pollution and precipitation has not been possible due to the lack of studies. Thus, the report focuses on available published evidence indicating possible links and on identifying effective ways to proceed in filling gaps in our understanding. In subsequent chapters, we discuss in some detail, the formation mechanisms of clouds and precipitation (Chapter 2), the sources and cloud relevant physical/chemical characteristics of atmospheric aerosols (Chapter 3), their global distribution and presence in various regions (Chapter 4), ground-based, airborne and satellite techniques for observing the effects of aerosols on clouds and precipitation (Chapter 5), the observed and model simulations of the effects of aerosols on clouds and precipitation (Chapter 6 and 7) and lessons learned about inadvertent aerosol effects on clouds and precipitation from many decades of attempts to modify clouds and precipitation by cloud seeding (Chapter 8). Finally, in Chapter 9 and 10 respectively, we summarize the main points of the report and add a list of recommendations to increase knowledge of the effects of aerosols on clouds and precipitation.
In this chapter we provide an overview of the basic physical processes responsible for the formation of clouds and precipitation. A number of important concepts are discussed and terms defined which will be used in later chapters. For more detail on these topics the reader is referred to textbooks by Pruppacher and Klett [1997], Rogers and Yau [1989], relevant chapters in Wallace and Hobbs [2006], Cotton and Anthes [1989], and Houze [1993] and review articles by Stewart [1985] and Cantrell and Heymsfield [2005].

2.1 Formation and Structure of Clouds

Different cloud types are defined according to the phases of water present and the temperature of cloud top [AMS Glossary]. If all portions of a cloud have temperatures greater than 0°C it is referred to as a warm cloud, or a liquid phase cloud. In clouds extending above the 0°C level, precipitation may form either by ice phase or droplet coalescence processes. Ice-crystal clouds consist entirely of ice crystals. Analogously, a water cloud is composed entirely of liquid water drops. A mixed-phase cloud contains both water drops (supercooled at temperatures below 0°C) and ice crystals, without regard to their actual spatial distributions (coexisting or not) within the cloud. Most convective clouds extending into air colder than about −10°C are mixed clouds, though the proportion of ice crystals to water drops may be small until the cloud builds to levels of still lower temperature. Provided the temperature is not below about –40°C, supercooled droplets may coexist with ice particles. However, the liquid phase within unactivated aerosols (haze particles) may coexist with ice particles to very low temperatures.

2.1.1 Dynamical Aspects of Cloud Formation

Clouds form when the saturation vapor pressure becomes smaller than the actual partial pressure of the water vapor in the air. The difference is condensed in the form of liquid water or ice, depending on the temperature. The saturation vapor pressure decreases when the temperature decreases (Clausius-Clapeyron). The most common way clouds form is
therefore when a buoyant parcel of air is lifted (convective ascent) and cooled by adiabatic expansion. For an ascending parcel inside a cloud, the temperature decreases following a moist adiabatic lapse rate which is slightly less (0.65 °C per 100 m) than in clear air adiabatic ascent (1 °C per 100 m), because of the latent heat released by condensation. The corresponding rate of condensation depends on the temperature and pressure of the cloud cell. For example at 900 m and 20 °C, it is on the order of 2 g/kg per km of ascent. When the cloud base temperature and pressure are determined, the mixing ratio of condensed water at any level above cloud base can be derived as the difference between the water vapor mixing ratio at cloud base and the saturation water vapor mixing ratio at that level. This is referred to as the adiabatic water-mixing ratio. The actual condensed water-mixing ratio is generally lower than the adiabatic value. At cloud base, the condensation of the available water vapor is not instantaneous and the actual water vapor partial pressure can be momentarily higher than the saturation vapor pressure leading to supersaturation. Supersaturation plays a critical role near cloud base for the activation of CCN and IN that initiate cloud droplets (Sec. 2.1.2.1) or ice crystals (Sec. 2.1.2.2). Further, up in the cloud, when cloud particles are numerous and big enough, the rate of condensation keeps pace with the production of supersaturation, so that supersaturations remain steady or decline. The actual condensed water content in a convective cloud is generally lower than the adiabatic one, because the ascending air from the cloud base is continuously mixed with drier air entrained from outside the cloud. The condensed water content, either liquid or ice, is a key parameter for precipitation formation. Precipitation is most likely to form in the regions of largest condensed mixing ratio, i.e. in the least diluted cloud cells.

A second critical parameter is the Lagrangian time scale of the cloud particles, which is the time it takes a parcel of air containing an ensemble of cloud particles to form when entering a cloud, and evaporate or sublimate when the parcel reaches cloud top. Liquid phase precipitation is produced when the numerous but very small droplets, which have grown by condensation, collide and coalesce into a few much larger raindrops with a significant fall speed. Small ice phase particles can grow to precipitation size by any combination of vapor deposition, collection of small water droplets (riming) or collection of other ice particles (aggregation). These processes require significant time, sometimes comparable to or exceeding the lifetime of a cloud. If the Lagrangian time scale of the cloud particles is
shorter than the time needed to produce precipitation, all the condensed water vapor may evaporate in the atmosphere without any precipitation reaching the surface.

The principal types of ascent, each of which produces distinctive cloud forms, are:

*Local ascent of warm, buoyant air parcels in a conditionally unstable environment*, which produces convective clouds. Because of the release of latent heat, a cloud convective parcel becomes warmer, hence positively buoyant, with respect to its environment. Air parcels within convective cells are thus able to rise into the atmosphere, until they reach a stable boundary such as a temperature inversion or lose their buoyancy by mixing. Convective clouds have diameters from about 0.1 to 10 km and air ascends in them with velocities up to a few meters per second, although updraft speeds of several tens of meters per second can occur within small volumes of large convective clouds. Within stronger updrafts, ascents of a few kilometers typically produce condensed water mixing ratios of a few grams per kilogram. Mixing ratios of more than 10 g/kg are possible in very strong updrafts within deep cumulonimbus clouds.

The lifetimes of convective clouds range from minutes to several hours. The particles Lagrangian time scale though may be shorter. For example, in a shallow cumulus cloud with a depth of ~1.5 km and characteristic updraft speeds of 3 m s\(^{-1}\), the Lagrangian time scale is \(t_p = 1500 \text{ m}/3 \text{ m s}^{-1} = 500 \text{ s} \approx 8 \text{ min}\). This represents the time available for initiation of precipitable particles. Once initiated, precipitation may continue over the remaining lifetime of the cloud. In a towering cumulus cloud with depth of ~10 km and updraft speeds of ~15 m s\(^{-1}\), the Lagrangian time scale is \(t_p = 10,000 \text{ m}/15 \text{ m s}^{-1} \approx 660 \text{ sec} \approx 11 \text{ mins}\), only slightly longer than shallow cumulus clouds. The main advantage that a towering cumulus cloud experiences over that of a shallow cloud in forming precipitation particles is associated with the greater amounts of condensate that is produced in deeper clouds. Because precipitation growth by collection is a non-linear function of the amount of condensate in a cloud [Kessler, 1969; Manton and Cotton, 1977], precipitation growth proceeds quite rapidly in cumulonimbus clouds relative to low liquid water content cumulus clouds. Supercell storms (~12 km depth) have lifetimes of several hours but their strong updrafts (~40 m s\(^{-1}\)), yield a Lagrangian time scale of only \(t_p = 12,000 \text{ m}/40 \text{ m s}^{-1} = 300 \text{ s} = 5 \text{ mins}\), which is shorter than that for shallow cumulus clouds. A characteristic
feature of supercell storms is the bounded weak echo region where updrafts are so strong that there is not sufficient time to produce radar-detectable precipitation elements at mid levels of the storm [Browning and Ludlam, 1962; Marwitz, 1972]

_Forced lifting of stable air_ to produce layer or stratiform clouds. Such clouds can occur at altitudes from ground level to the tropopause and extend over hundreds of thousands of square kilometers. Lifting rates range from a few centimeters per second to ~10 cm s\(^{-1}\). Layer clouds generally exist over periods of tens of hours and their Lagrangian time scales are \(\sim 10^4\) s (\(\sim 3\) hrs). Due to their long Lagrangian time scales, precipitation is likely in these clouds in spite of their small water contents (\(\sim 0.5 \text{ g kg}^{-1}\)).

_Forced lifting of air as it passes over hills or mountains_ to produce orographic clouds. Updraft velocities depend upon the speed and direction of the wind and the height of the barrier, and they can be several meters per second. Water contents are typically a few tenths of a gram per cubic meter of air, depending upon the altitude at which the air enters the cloud upwind and its maximum altitude above the mountain top. Orographic clouds may be quite transitory, but if winds are steady they may last for many hours. However, the relevant time scale that determines the time available for precipitation formation is not the time that it takes a parcel to ascend from cloud base to cloud top. Instead, it is the time that it takes a parcel to transect from the upwind lateral boundary to its downwind boundary, as shown in Figure 2-1. For example, for a 15 m s\(^{-1}\) wind speed and a cloud of 18,000 m lateral extent, \(t_p = 18,000 \text{ m}/15 \text{ m s}^{-1} = 1200 \text{ s} = 20 \text{ min}\). This time scale is longer than that for cumulus clouds but considerably shorter than that for layer clouds. Since the liquid water contents of stable, wintertime orographic clouds are low, production of precipitation requires efficient conversion of cloud droplets to precipitation. These conditions can occur in shallow maritime clouds passing over mountains near shore. Lagrangian time scales are longer for deeper mixed and ice-phase orographic clouds as compared to shallower liquid phase clouds. Note that the adiabatic liquid water content for orographic clouds is not determined by the ascent of a cloud parcel from cloud base to cloud top. Instead, it is determined by the vertical displacement of a parcel of air as it transects the cloud over the mountain barrier.
Vertical ascent is not the only way of lowering the temperature and the water vapor saturation pressure. *Cloud top radiative cooling* can lead to destabilization of cloudy layers. This is the main driving force for marine stratocumulus clouds. It is also important in fogs, stratus clouds and cirrus clouds. Longwave radiative flux divergence at cloud top creates cooling of the air, which, in turn, produces higher density air parcels that descend through the cloud layer causing vertical mixing. In the case of a marine stratocumulus layer in which the sea surface temperatures are slightly warmer than the overlying air, the descending cool air parcels can descend through most of the unstable sub-cloud layer, which enhances vertical mixing through the depth of the cloudy boundary layer. Typical lifetimes of stratus and stratocumulus clouds are ~6 to 12 h. The Lagrangian time scale for 1000 m deep clouds having vertical velocities of ~0.1 m s\(^{-1}\) is \(t_p = 1000 \text{ m} / 0.1 \text{ m s}^{-1} = 10^4 \text{ s} \approx 3 \text{ h}\). Thus, in spite of the fact that they have liquid water contents of only ~0.1 g kg\(^{-1}\), the long Lagrangian time scales permit the formation of precipitation, in the form of drizzle, in stratus and stratocumulus clouds. Because of cloud top radiative cooling, the temperature at the top of a convective cloud may decrease more than expected from the adiabatic lapse rate, hence leading to a superadiabatic water mixing ratios.

As noted earlier, the above concepts of a cloud are based on simple “back-of-the-envelope” calculations. In general, clouds are very turbulent and thus the time-scales for precipitation
formation can be much longer than simple Lagrangian parcel estimates. Likewise, the liquid water contents of clouds can be quite variable. In cumulus clouds, it is not uncommon to find regions of high liquid water content, say 0.5 to 1.0 gram per cubic meter next to regions with hardly any condensate only tens of meters away. Often cumulus clouds exhibit considerable asymmetry in structure with upshear parts of the cloud experiencing little mixing and turbulence and downshear portions experiencing very large mixing and turbulence. In such a cloud, hydrometeors undergo growth on the upshear side and experience evaporation and turbulence on the downshear.

Thus far, we have only talked about individual cumulus clouds and the lifetimes of individual cells. But in fact, precipitating clouds can affect neighboring clouds by exciting gravity waves, by latent heating induced buoyancy bores and by producing air chilled by evaporation of precipitation particles in the subcloud layer. As illustrated in Figure 2-2 even small cumuli that produce precipitation that evaporates in the sub-cloud layer can form a cold pool. The chilled air is denser than surrounding air and spreads out beneath

![Figure 2-2.](image)

(a) Illustration of droplets settling from the upper levels of a cloud, thus reducing the amount of liquid water content or water-loading burden on the cloud. (b) Illustration of the formation of an evaporatively-chilled layer near the surface, which can lift surrounding moist air sometimes to the LCL and LFC. (From Cotton [1990]).
the cloud where it can lift the surrounding air sometimes enough to generate new cloud cells. In some cases neighboring precipitating clouds can produce cool outflows that run into each other and cause lifting of the air leading to the merger of the neighboring clouds (see Figure 2-3. The merged cloud cell is often wider and deeper than the parent clouds and is more likely to produce rain. As illustrated in Figure 2-4 the cold pool and leading edge gust front are important to the maintenance of an ordinary thunderstorm as an efficient engine. The fact that precipitating clouds alter their local environment is important for understanding how pollutants can influence clouds (Chapter 7) and how seeding clouds can perhaps alter precipitation on time scales greatly exceeding the lifetime of individual clouds (Chapter 8).

Figure 2-3. Schematic illustration relating downdraft interaction to bridging and merger in case of light wind and weak shear. (From Simpson et al, [1980]).
2.1.2 Liquid Phase Clouds

2.1.2.1 Cloud Droplet Formation

The supersaturations required to nucleate drops by the chance collisions of water vapor molecules (i.e., homogeneous nucleation) greatly exceed the observed supersaturations in the atmosphere. Consequently, droplets do not form in natural clouds by homogeneous nucleation. Instead, they form by heterogeneous nucleation onto atmospheric aerosol.

Köhler [1926] first determined the equilibrium vapor pressure above small solution droplets (Figure 2-5). It can be seen from the curves shown in Figure 2-5 that below a certain droplet size, the relative humidity of the air adjacent to a solution droplet is less
than that which is in equilibrium with a plane surface of pure water at the same temperature (i.e., 100%). As a droplet increases in size, the lowering of the equilibrium vapor pressure above its surface due to the dissolved material becomes increasingly less and the equilibrium vapor pressure over a small curved droplet (Kelvin curvature effect) becomes the dominant influence.

Suppose a particle of NaCl with mass $10^{-19}$ kg is placed in air with a water supersaturation of 0.4% (indicated by the dashed line in Figure 2-5). As can be seen from Figure 2-5, the solution droplet will experience a supersaturation, and the droplet would grow by condensation. As it does so, the supersaturation adjacent to the surface of this solution droplet will initially increase, but even at the peak in its Köhler curve the supersaturation adjacent to the droplet is less than the ambient supersaturation. Consequently, the droplet will grow over the peak in its Köhler curve and down the right-hand side of this curve to form a fog or cloud droplet. A droplet that has passed over the peak in its Köhler curve and continues to grow is said to be activated.

![Figure 2-5. Variations of the relative humidity and supersaturation adjacent to droplets of (1) pure water (blue), and adjacent to solution droplets containing the following fixed masses of salt: (2) $10^{-19}$ kg of NaCl (red), (3) $10^{-18}$ kg of NaCl (orange), (4) $10^{-17}$ kg of NaCl (brown), (5) $10^{-19}$ kg of $(NH_4)_2SO_4$ (green), and (6) $10^{-18}$ kg of $(NH_4)_2SO_4$ (violet). Note the discontinuity in the ordinate at 100% relative humidity. Adapted from Rasool [1973].](image)
Now consider a particle of \((\text{NH}_4)_2\text{SO}_4\) with mass \(10^{-19}\) kg that is placed in the same ambient supersaturation of 0.4%. In this case, condensation will occur on the particle and it will grow as a solution droplet. At point A the supersaturation adjacent to the droplet is equal to the ambient supersaturation. If the droplet at A should grow slightly, the supersaturation adjacent to it would increase above the ambient supersaturation, and therefore the droplet would evaporate back to point A. If the droplet at A should evaporate slightly, the supersaturation adjacent to it would decrease below the ambient supersaturation, and the droplet would grow by condensation back to A in Figure 2-5. Hence, in this case, the solution droplet at A is in stable equilibrium with the ambient supersaturation. If the ambient supersaturation were to change a little, the location of A in Figure 2-5 would shift and the equilibrium size of the droplet would change accordingly. Droplets in this state are said to be unactivated or haze droplets. Haze droplets in the atmosphere can considerably reduce visibility by scattering light.

A sub-set of the atmospheric aerosol discussed in Chapter 3 serves as particles upon which water vapor condenses to form droplets that are activated and grow by condensation to form cloud droplets at the supersaturations achieved in clouds (~0.1–1%). These particles are called cloud condensation nuclei (CCN). It follows from the above discussion that the larger the size of a particle with a given chemical composition, the more readily it is wetted by water, and the greater its solubility, the lower will be the supersaturation at which the particle can serve as a CCN. For example, to serve as a CCN at 1% supersaturation, completely wettable but water insoluble particles need to be at least \(~0.1\ \mu\text{m}\) in radius, whereas soluble particles can serve as CCN at 1% supersaturation even if they are as small as \(~0.01\ \mu\text{m}\) in radius. Most CCN consist of a mixture of soluble and insoluble inorganic and organic components (called internally-mixed nuclei). The solubility of a particle has an important effect on its effectiveness as a CCN. For example, the initial minimum dry radius of a particle that is activated by a supersaturation of 0.1% is 0.075 \(\mu\text{m}\) if the particle is completely soluble. However, if the ratio of the soluble mass to the total mass of the particle is only 0.2, the particle would need to have a dry radius of 0.13 \(\mu\text{m}\) to be activated by a supersaturation of 0.1%.

Another aspect of CCN activity that is often overlooked is the wettability of the aerosol particle (i.e., the ability of water to spread out over the surface of the particle) as measured
by the contact angle of water on the particle. It is generally assumed that particles are completely wettable. However, this is by no means always the case [e.g., Knight, 1971]. For example, in the atmosphere there are a number of organic materials that are not wettable. To the extent that a substance has a non-zero contact angle, its ability to serve as a CCN will be hindered.

The surface tension of the solution formed by condensation onto a soluble particle will also affect the subsequent growth of the solution droplet. This is because for the Kelvin effect the energy barrier that has to be overcome for a droplet to be activated varies as the third power of the surface tension [Wallace and Hobbs, 2006]. The chemical components, concentrations, solubilities, and surface tensions of solution droplets that form in various environments are not well documented, and the effects of these parameters on droplet activation, individually and in combination are not well understood, but it is generally believed that surface tension effects are small.

Another factor of importance during droplet nucleation and vapor deposition growth is the diffusivity of vapor molecules near the surface of small droplets. Ordinarily diffusivities of water vapor that can be found in chemical handbooks do not take into account the kinetic effects of vapor diffusion when vapor molecules are within the mean free path distance of the droplet surface. To account for kinetic effects, the vapor diffusivities are modified with a so-called accommodation coefficient. There is strong disagreement in the literature regarding the value of the accommodation coefficient for atmospherically relevant conditions [Shaw and Lamb, 1999; Laaksonen et al., 2005]. For example, the existence of film forming compounds in the aerosol may significantly reduce the mass accommodation. This parameter has a very strong effect on the number of droplets activated, with the latter tending to decrease with increasing values of this coefficient. Its effect is much stronger than many effects associated with changes in aerosol composition mentioned above [e.g., Kreidenweis et al., 2003].

2.1.2.2 Cloud Condensation Nuclei

Because of the uncertainties in predicting the CCN nucleating ability of atmospheric particles, an empirical approach is generally taken by measuring the concentration of particles that serve as CCN at various prescribed supersaturations; this is called the CCN
supersaturation spectrum. The concentrations of CCN active at various supersaturations can be measured with a thermal gradient diffusion chamber, or other devices based on similar principles. A diffusion chamber may take a variety of geometric configurations, with the essential feature being that particles are statically or dynamically (with continuous flow) exposed to a steady supersaturation field created either by wetted plates held at different temperatures or by a streamwise gradient of temperature. By varying the temperature difference between the plates, it is possible to produce maximum supersaturations in the chamber that range from a few tenths of 1% to a few percent, which are similar to the inferred supersaturations that activate droplets in clouds [Wieland, 1956]. Worldwide measurements of CCN concentrations have not revealed any systematic latitudinal or seasonal variations. At a given location, CCN vary by several orders of magnitude with time, depending on the proximity of sources, wind direction, air mass type, precipitation and cloudiness [e.g., Twomey, 1960; Jiusto, 1966; Radke and Hobbs, 1969]. Near the earth's surface continental air masses generally contain larger concentrations of CCN than clean marine air masses (Figure 2-6). For example, the concentration of CCN in the continental air mass over the Azores, depicted in Figure 2-6, is about ~300 cm$^{-3}$ at 1% supersaturation, while in the marine air mass over Florida it is ~100 cm$^{-3}$, and in clean Arctic air it is only ~30 cm$^{-3}$. The ratio of CCN (at 1% supersaturation) to the total number of particles in the air (CN) is ~0.2–0.6 in marine air; in continental air this ratio is generally less than ~0.01 but can rise to ~0.1. The very low ratio of CCN to CN in continental air is attributable to the large number of very small particles, which are not activated at low supersaturations. Concentrations of CCN over land decline by about a factor of five between the planetary boundary layer and the free troposphere [Squires and Twomey, 1966; Hoppel et al., 1973; Hobbs et al., 1985]. Over the same height interval, concentrations of CCN over the ocean remain fairly constant, or may even increase with height reaching a maximum concentration just above the mean cloud height [Hoppel et al., 1973; Hudson, 1983; Hegg et al., 1990]. Ground-based measurements indicate that there is a diurnal variation in CCN concentrations at some locations, with a minimum at about 6 a.m. and a maximum at about 6 p.m. [Twomey and Davidson, 1970, 1972].
The observations described above provide clues as to the origins of natural CCN. First of all it appears that the land acts as a major source of CCN, because the concentrations of CCN are generally high over land and decrease with altitude. Some of the soil particles and dusts that enter the atmosphere probably serve as CCN, but they do not appear to be a dominant source. The rate of production of CCN (active at a supersaturation of 0.5%) from burning vegetable matter is on the order of $10^{12}-10^{15}$ per kg of material consumed [Eagan et al., 1974b]. Thus, forest fires are a prolific source of CCN [Twomey, 1960; Twomey and Warner, 1967; Warner and Twomey, 1967; Warner, 1968, Hobbs and Radke, 1969; Woodcock and Jones, 1970; Stith et al., 1981]. Although sea-salt particles enter the air over the oceans, they do not appear to be a dominant source of CCN, even over the oceans [Twomey, 1968, 1971; Radke and Hobbs, 1969; Dinger et al., 1970; Hobbs, 1971], although, because of their solubilities and large sizes, they may enhance precipitation by serving as giant CCN.

There appears to be a widespread and probably a fairly uniform source of CCN over both oceans and land, the nature of which has not been definitely established. A likely candidate is gas-to-particle conversion, which can produce particles up to a few tenths of a micrometer in diameter that can act as CCN if they are soluble and wettable. Gas-to-particle conversion mechanisms that require solar radiation might be responsible for the
observed peak in CCN concentrations at ~6 p.m. Most CCN consist predominately of sulfates, although some organic material is usually present. Over the oceans, organic sulfur from the ocean (in the form of the gases dimethylsulfide (DMS) and methane sulfonic acid (MSA)) provides a source of CCN, with the DMS and MSA being converted to sulfate in the atmosphere \cite{Charlson et al., 1987; Hoppel, 1987; Luria et al., 1989; Gras, 1990; Hegg et al., 1991a,b; Ayers and Gras, 1991; Ayers et al., 1991; Pandis et al., 1994}. Evaporating clouds also release sulfate particles that are somewhat larger than non-cloud processed aerosol because of the additional material deposited onto them in clouds \cite{Radke and Hobbs, 1969; Easter and Hobbs, 1974; Hegg et al., 1980; Hegg and Hobbs, 1981; Hoppel et al., 1986; Birnili et al., 1999; Feingold and Kriedenweis, 2002}. It has been estimated that ~80% of sulfate mass globally is formed in the aqueous phase, and the remainder from the gas phase. As expected from the Köhler curves, larger sulfate particles serve more efficiently as CCN \cite{Hegg, 1985; Leaitch et al., 1986; Lelieveld et al., 1997}. The production of sulfates (and maybe other soluble materials), followed by the release of these particles when the droplets evaporate, is an important mechanism for increasing the efficiency of CCN \cite{Twomey and Wocjiewski, 1969; Hobbs, 1971; Easter and Hobbs, 1974}. There are also a number of anthropogenic sources of CCN, which are discussed in Chapters 3, 4 and 6.

2.1.2.3 Effects of CCN on the Microphysical Structures of Clouds

In a cloud, we are concerned with the growth of a large number of droplets in a rising parcel of air. As the parcel rises it expands, cools, and eventually reaches saturation with respect to liquid water. Further uplift and adiabatic cooling produces supersaturations that initially increase in proportion to the updraft velocity. CCN are activated as the supersaturation rises starting with the most efficient CCN. When the rate at which water vapor in excess of saturation, is equal to the rate at which water vapor condenses onto the CCN and droplets, the supersaturation in the cloud reaches a maximum value. The concentration of cloud droplets is determined at this stage (which generally occurs less than 100 m above cloud base) and is equal to the concentration of CCN activated by the peak supersaturation that has been attained. Subsequently, the growing droplets consume water vapor faster than it is made available by the cooling of the air, so the supersaturation begins to decrease. The haze droplets then begin to evaporate while the activated droplets continue to grow by condensation. The rate of growth of a droplet by condensation is inversely
proportional to its radius, therefore the radius of the smaller activated droplets grow faster than that of the larger droplets. Consequently, in this simplified model, the size distributions of the droplets in the cloud become increasingly narrower with time (that is, the droplets approach a monodispersed distribution). This sequence of events is illustrated by the results of theoretical calculations shown in Figure 2-7.

**Figure 2-7.** Theoretical computations of the growth of cloud condensation nuclei by condensation in a parcel of air rising with a speed of 60 cm s⁻¹. A total of 500 CCN cm⁻³ was assumed with \(\frac{im}{M_s}\) values as indicated; \(m\) the mass of material dissolved in the droplet, \(M_s\) the molecular weight of the material, and \(i\) its van’t Hoff factor (i.e., the number of ions produced by each molecule of the material when it dissolves). Thus, \(\frac{im}{M_s}\) is the effective number of kilomoles of the material in the dissolved droplet. One can think of the terms \(im\) as representing the particular chemical properties of the salt. Note how the droplets that have been activated (brown, blue and purple curves) approach a monodispersed size distribution after just 100 s. (i.e. ~60 m above cloud base). The variation with time of the supersaturation of the air parcel is also shown (dashed red line). From Wallace and Hobbs [2006], based on data from Howell [1949].

Comparisons of cloud droplet size distributions measured a few hundred meters above the bases of non-precipitating warm cumulus clouds with droplet size distributions computed assuming growth by condensation for about 5 min show good agreement. The droplets produced by condensation during this time period extend only up to about 10 µm in radius. Moreover, as mentioned above, the rate of increase in the radius of a droplet growing by condensation is inversely proportional to the drop radius, therefore, the rate of growth decreases with time. It is clear, therefore, as first noted by Reynolds [1877], that growth by condensation alone in warm clouds is much too slow to produce raindrops with radii of
several millimeters. Yet rain does form in clouds that contain only water drops. The enormous increases in size required to transform cloud droplets into raindrops is illustrated by the scaled diagram shown in Figure 2-8. For a cloud droplet 10 µm in radius to grow to a raindrop 1 mm in radius requires an increase in volume of one millionfold! However, only about one droplet in a million (about 1 liter\(^{-1}\)) in a cloud has to grow by this amount for the cloud to rain. The mechanism responsible for the selective growth of a few droplets into raindrops in warm clouds is discussed in Section 2.2.1.

**Figure 2-8.** Relative sizes of cloud droplets and raindrops; \(r\) is the radius in micrometers, \(n\) the number per liter of air, and \(v\) the terminal fall speed in centimeters per second. The circumference of the circles are drawn approximately to scale, but the black dot representing a typical CCN is twenty-five times larger than it should be relative to the other circles. From Wallace and Hobbs [2006].

Since the concentration of CCN generally increases in passing from oceanic air sheds to continental air sheds to urban environments, the concentrations of cloud droplets likewise increase, and the droplet growth rate with altitude above cloud base decreases, at least for non-precipitating clouds [e.g., Gerber, 1996]. For example, droplet concentrations in non-polluted, non-precipitating marine cumulus clouds are generally less than 100 cm\(^{-3}\) while they can reach values over 1000 cm\(^{-3}\) in a polluted environment [Squires, 1958], see Figure 2-9. As we will see in Section 2.2.1, clouds with large concentrations of droplets are more colloidal stable and less likely to precipitate than clouds with small concentrations of droplets.

The effects of CCN from anthropogenic sources on cloud structures and precipitation are discussed in Chapter 5 and 6.
Figure 2-9. Cloud droplet number distributions measured in stratocumulus clouds in the vicinity of the Azores by the FSSP-100 (circles) and PMS 1D (diamonds) cloud probes, averaged over 15 km of flight path for case 1—clean marine air (red symbols and curve), and averaged over 4 km of flight path for case 2—continentally influenced air (blue symbols and curve). The vertical bars are the geometric standard deviations of the droplet concentrations. From Garrett and Hobbs [1995].

2.1.3 Ice Phase Processes

2.1.3.1 Nucleation of Ice

Ice particles can form either homogeneously or heterogeneously on some form of ice nuclei (IN). Homogeneous nucleation can take place either directly from the vapor or by freezing of cloud droplets. However, homogeneous nucleation of ice crystals from the vapor, or the chance formation of an embryo of ice-like structure of critical size, requires very high supersaturations with respect to ice and such low temperatures that it does not take place in the troposphere. On the other hand, homogeneous freezing of supercooled droplets by the chance formation of a cluster of ice-like embryos can occur in the atmosphere.

For homogeneous freezing to occur, enough ice-like water molecules must come together within the droplet to form an embryo of ice large enough to survive and grow. If an ice embryo within a droplet exceeds a certain critical size, its growth will produce a decrease in the energy of the system. On the other hand, any increase in the size of an ice embryo smaller than the critical size causes an increase in total energy. In the latter case, from an energetic point of view, the embryo is likely to break up.
Since the numbers and sizes of the ice embryos that form by chance aggregations increase with decreasing temperature, below a certain temperature (which depends on the volume of water considered) freezing by homogeneous nucleation becomes a virtual certainty. Homogeneous nucleation occurs in about one second at about -41°C for droplets about 1 µm in diameter, and at about -35°C for drops 100 µm in diameter. An analogous freezing process occurs for unactivated droplets or haze particles at temperatures below -40°C, a process for ice formation in cirrus clouds (DeMott, 2002). Hence, in the atmosphere, homogeneous nucleation by freezing generally occurs only in high clouds or high latitudes.

If a droplet contains a rather special type of particle, called a freezing nucleus, it may freeze by a process known as heterogeneous nucleation in which water molecules in the droplet collect onto the surface of the particle to form an ice-like structure that may increase in size and cause the droplet to freeze. Since the formation of the ice structure is aided by the freezing nucleus, and the ice embryo also starts off with the dimensions of the freezing nucleus, heterogeneous nucleation can occur at much higher temperatures than homogeneous nucleation.

If the particle that initiates freezing is contained within the droplet, it is called an immersion freezing nucleus. However, cloud droplets may also be frozen if a suitable particle in the air comes into contact with the droplet, in which case freezing is said to occur by contact freezing, and the particle is referred to as a contact nucleus. Laboratory experiments suggest that some particles can cause a drop to freeze by contact freezing at temperatures several degrees higher than if they were embedded in the drop [Fletcher, 1962; Levkov, 1971; Gokhale and Spengler, 1972; Pitter and Pruppacher, 1973].

Recent laboratory experiments by Durant and Shaw [2005] showed that as droplets evaporate, embedded aerosol particles become increasingly likely to penetrate the air-water interface layer and promote freezing. They interpreted this as evidence that contact freezing may be just as effective from the inside-out as from the outside-in.

Certain particles in the air also serve as centers upon which ice can form directly from the vapor phase. These particles are referred to as deposition nuclei. Ice can form by deposition provided that the air is supersaturated with respect to ice and the temperature is low
enough. If the air is supersaturated with respect to water, a suitable particle may serve either as a \textit{condensation-freezing} nucleus (in which case liquid water first condenses onto the particle and subsequently freezes) or as a \textit{deposition nucleus} (in which case there is no intermediate liquid phase, at least on the macroscopic scale). It is generally thought that condensation-freezing is preferred at smaller supercoolings and large supersaturations, while deposition is preferred at large supercoolings and small supersaturation. In practice, it is not easy to distinguish between deposition and condensation-freezing modes.

2.1.3.2 Properties of Ice Nuclei

If we refer to an ice-nucleating particle in general, without specifying its mode of action, we call it an \textit{ice nucleus} (IN). However, it should be kept in mind that the temperature at which a particle can cause ice to form depends, in general, upon the mechanism by which the particle nucleates the ice as well as upon the previous history of the particle. The basic distinction that has to be made is whether nucleation is from the vapor or from the liquid phase [Vali, 1985].

Particles with molecular spacings and crystallographic arrangements similar to those of ice (which has a hexagonal structure) tend to be effective as ice nuclei, although this is neither a necessary nor a sufficient condition for a good ice nucleus. Most effective ice nuclei are virtually insoluble in water. Some inorganic soil particles (mainly clays) can nucleate ice at fairly high temperatures (i.e., above -15°C), and they probably play an important role in nucleating ice in clouds. Desert dust is known to be very good ice nuclei [DeMott et al., 2003a]. For example, in one study, 87% of the snow crystals collected on the ground had clay mineral particles at their centers and more than half of these were kaolinite [Kumai, 1951]. Of course, it is only circumstantial evidence that they served as ice nuclei. Many organic materials are effective ice nucleators [Schnell and Vali, 1976a]. Decayed plant leaves contain copious ice nuclei, some active as high as -4°C. Ice nuclei active at -4°C have also been found in seawater rich in plankton. In addition, some plant pathogenic bacteria have also been found to be effective ice nuclei at temperatures as high as -2°C [Vali et al., 1976; Yankofsky et al., 1981; Levin and Yankofsky, 1983; Levin et al., 1987]. Recently Von Blohn et al. [2005] have identified pollen as good ice nuclei at warm temperatures.
In some cases, after a particle has served as an ice nucleus and all of the visible ice is then evaporated from it, but the particle is not warmed above -5°C or exposed to a relative humidity with respect to ice of less than 35%, it may subsequently serve as an ice nucleus at a temperature a few degrees higher than it did initially [Roberts and Hallett, 1968]. This is referred to as preactivation. Thus, for example, preactivated nuclei may be transferred between sequential wave clouds and act at a higher temperature for the second and later clouds. However, preactivation is lost if the initially activated ice nucleus is either heated above at least –5°C and/or is dehydrated to less than ~35% relative humidity with respect to ice. Because preactivation can be lost if the particles are warmed or dehydrated prior to testing in an IN counter, this behavior creates a significant dilemma for IN measurements in almost any sampling scenario.

2.1.3.3 Variations of ice nuclei and ice particle concentrations

Ice nucleus concentrations can sometimes vary by several orders of magnitude over several hours (see also Chapters 5 and 6). Historically, ice nuclei concentrations have been assumed as predictors of ice particle concentrations. Laboratory studies in the 1940’s and 1950’s using systems for low temperature processing of ice nuclei collected on filters indicated that the variability in ice nuclei concentration was strongly a function of temperature. Fletcher [1962] derived an empirical relationship relating the increase of concentration of ice nuclei with decreasing temperature:

\[
\ln N = a(T_1 - T)
\]  

(2.1)

where N is the concentration of active ice nuclei per liter of air, T is the air temperature, and \(T_1\) is the temperature at which one ice nucleus per liter is active (typically about -20°C) and \(a\) varies from about 0.3 to 0.8. For \(a = 0.6\), eqn. (2.1) predicts that the concentration of ice nuclei increases by about a factor of 10 for every 4°C decrease in temperature. In urban air, the total concentration of aerosol is on the order of \(10^8\) liter\(^{-1}\) and only about one particle in \(10^8\) acts as an ice nucleus at -20°C.

Later laboratory studies [e.g. Huffman and Vali, 1973; Huffman, 1973a,b] could not reproduce Fletcher’s \(N-T\) empirical relationship. Measurements in a variety of field campaigns using a continuous flow diffusion chamber exhibited considerable variability in IN concentration [DeMott et al., 2003a; Rogers et al., 1998; Rogers et al., 2001; Prenni et al., 2007]. The variability in ice nuclei concentration is likely not a simple function of
temperature but also depends on thermodynamical, dynamical and aerosol characteristics [DeMott et al., 1994].

The activity of a particle as a condensation-freezing or a deposition nucleus depends not only on the temperature but also on the supersaturation of the ambient air. The effect of supersaturation on measurements of ice nucleus concentrations is shown in Figure 2-10, where it can be seen that at a constant temperature the greater the supersaturation with respect to ice the more particles serve as ice nuclei. The empirical equation to the best-fit line to these measurements (the red line in Figure 2-10) is

\[
N = \exp \left\{ a + b \left[ 100 (S_i - 1) \right] \right\}
\]  

(2.2)

where \(N\) is the concentration of ice nuclei per liter, and \(S_i\) is the supersaturation with respect to ice, \(a = -0.639\) and \(b = 0.1296\) [Meyers et al., 1992]. These measurements were obtained using a continuous flow diffusion chamber (CFDC), which limited data suggests exhibits roughly a factor of ten higher concentrations of IN at warmer temperatures than older devices such as the filter-processing systems. Recognizing the need to allow vertical and horizontal variations in IN concentrations in mesoscale model simulations, Cotton et al. [2003] modified equation (2.2) to include the prognostic variable \(N_{IN}\):

\[
N_i = N_{IN} \exp \left[ 12.96 (S_i - 1) \right],
\]  

(2.3)

where \(T < -5^\circ C\). The variable \(N_{IN}\) can be deduced from continuous flow diffusion chamber data and used as a forecast variable in regional simulations [i.e. van den Heever et al., 2006; van den Heever and Cotton, 2007]. Equations (2.1), (2.2) and (2.3) only represent initial ice particle formation on IN and do not necessarily represent actual ice particle concentrations because other processes such as ice multiplication (see Section 2.1.3.4), sedimentation, breakup, and advection can greatly influence the concentrations of ice particles. Note that equation (2.3) allows for both horizontal and vertical variations in IN. Because the aerosol contributing to IN are large and large aerosol generally decrease with height [Georgii and Kleinjung, 1968; DeMott et al., 2003b], we expect that IN concentrations generally decrease with height as well.
Figure 2-10. Ice nucleus concentration measurements versus ice supersaturation; temperatures are noted alongside each line. All data were obtained at ground level. The red line is eqn. (4.2). Blue squares are data from Rogers [1993]; green squares are from Al-Naimi and Saunders [1985]. From Wallace and Hobbs [2006]’s adaptation of the Meyers et al.[1992] figure.

One should not therefore be surprised that many observations of ice particle concentrations do not show a good correlation with temperature [Gultepe et al., 2001; Field et al., 2005]. Gultepe et al. [2001], for example, compiled data from the glaciated regions of stratus clouds for a number of field campaigns and found that ice crystals smaller than 1000 µm diameter do not show a good correlation with temperature and that the concentrations of these smaller ice particles varied up to three orders of magnitude, for a given temperature (see Figure 2-11). On the other hand, measurements of ice particle concentrations in wave clouds where only “initial” ice particles are likely [Cooper and Vali, 1981], showed ice particle concentrations increasing with decreasing temperature.

In summary, we see that there remain many unanswered questions regarding the concentrations of ice nuclei, their composition, their activation relative to different environmental factors, and their relationship to ice crystal concentrations. Work has been hampered by severe difficulties in precise measurement of IN (see Chapters 5 and 6). Moreover, current field deployable devices for measuring IN do not take into account activation of IN by contact freezing. Fletcher’s or Meyers empirical curves are still used and often misused in numerical models, ranging from cloud-resolving models to GCMs. These curves at best only represent the concentrations of ice particles initially formed in clouds and probably rarely represent ice particle concentrations in most cloud systems.
2.1.3.4 Concentrations of Ice Particles in Clouds: Ice Multiplication

As we have seen, it has become increasingly evident that concentrations of ice crystals in "real" clouds are not always represented by the concentrations of IN measured or expected to be activated in such environments. In particular, it has been found that at temperatures warmer than \(-10^\circ\text{C}\), the concentration of ice crystals can exceed the concentration of IN activated at cloud top temperature by as much as three or four orders of magnitude [Braham, 1964; Koenig, 1963; Mossop, 1970; Mossop and Ono, 1969; Mossop et al., 1967,

Some explanations or hypotheses that have been proposed to account for the high ice particle concentrations observed in some clouds are as follows:

- Ice multiplication by fracturing of fragile ice crystals, which may breakup during collision with each other. [Vardiman, 1978].
- Fragmentation of large drops during freezing. [Mason and Maybank, 1960].
- Secondary ice particle formation during ice particle riming. (Hallett and Mossop, 1974; Mossop and Hallett, 1974].
- Enhanced ice nucleation in the presence of spuriously high supersaturations. [Hobbs and Rangno, 1985].
- Secondary ice particle generation during evaporation of ice particles (Oraltay and Hallett, 1989; Dong et al., 1994).

Of these processes, the one that has been given the most attention and quantified in models is secondary ice particle formation by the rime-splinter process. Laboratory studies by Hallett and Mossop (1974) and Mossop and Hallett (1974), confirmed by Goldsmith et al. (1976) have indicated that copious quantities of splinters are produced during ice particle riming under very selective conditions. These conditions are:

- Temperature in the range of -3 to -8C.
- A substantial concentration of large cloud droplets (D > 24 µm).
- Large droplets coexisting with small cloud droplets (D < 12.3 µm).

An optimum average splinter production rate of 1 secondary ice particle for 250 large droplet collisions occurred at a temperature of -5C.

This process is consistent with observations of the greatest departure from IN estimates of ice crystals when clouds contain graupel particles and frozen raindrops and is consistent with field observations (Hobbs and Cooper (1987)).
There is much indirect or inferential evidence that evaporation enhances ice crystal concentrations. This evidence is perhaps more intriguing than it is compelling. Some field studies have related unusually high ice nuclei numbers or unusual increases in ice crystal numbers to circumstances in which clouds were evaporating. Cooper [1995], for example, found a 100-fold increase in ice crystal concentrations in the evaporation region of orographic layer clouds. The largest ice enhancements in the Cooper study were observed in clouds with temperatures approaching the onset temperature for homogeneous freezing. Smaller enhancements were found in warmer clouds and no enhancements were found warmer than about -20°C. Further evidence of the possible role of evaporation nucleation has been presented by Field et al. [2001] and Cotton and Field [2002]. They show observational evidence and supporting parcel modeling calculations from wave cloud studies that suggest ice had to form close to the downstream edge of wave cloud. Ice production coincident with the start of the liquid cloud, or earlier, would have suppressed the observed liquid cloud. Some of the observations of rapid ice crystal concentration enhancement versus expected IN concentrations in cumulus cloud studies of Hobbs and Rangno [1985; 1990] and Rangno and Hobbs [1994] were also observed to originate in close proximity to regions of cloud evaporation. Stith et al. [1994] followed the development of ice in a cumulus turret near its top at -18°C. During the updraft stages, low ice concentrations were observed in the turret (similar to what would be expected from primary ice nucleation), but during the downdraft stages, the ice concentrations increased by an order of magnitude. This observation cannot be explained by rime splintering.

In summary, it is unlikely that all primary and secondary ice-forming processes have been quantitatively identified. Other mechanisms may sometimes operate, but their exact nature remains a mystery. Moreover, our ability to measure small ice crystals, in particular, has significant errors and need improvement. Consequently, there are large uncertainties associated with our ability to simulate the affect of aerosols on the initiation of ice, and subsequent impacts on precipitation. This remains as one of the critical problems in cloud physics.
2.2 Formation of Precipitation

Precipitation-sized particles can form by two mechanisms, known as the collision-coalescence and the ice particle mechanisms. The collision-coalescence mechanism can operate in clouds containing droplets, whether they are situated above or below the 0°C level. The ice particle mechanism can operate only in clouds with temperatures below 0°C and containing some ice particles.

2.2.1 The Collision-Coalescence Mechanism

This mechanism for rainfall production involves the initial formation of a few cloud drops with radii of about 20 µm by condensation onto CCN, followed by their rapid and substantial growth as they fall through the air and collide with smaller cloud droplets. If a drop increases in radius from 20 µm to a radius typical of raindrops, say 2 mm, its volume increases by a factor of one million. In other words, about one million cloud droplets have to combine to form one raindrop!

Due to the flow of air around a falling drop, a drop will not collect all of the droplets that lie in its path. Calculated values of the collision efficiency, defined as the ratio of the cross-sectional area over which droplets are collected to the geometric cross-sectional area of the collector drop, are shown in Figure 2-12. This figure shows that the collision efficiency is negligible (<0.1) until a collector drop attains a radius of about 20 µm. Hence, the need for a few droplets to grow to 20 µm in radius by condensation if a cloud is to form raindrops by the collision-coalescence mechanism. As the droplet grows by condensation alone, the rate of increase in the radius slows (Figure 2-13). In contrast, after a droplet has reached a radius of 20 µm its collision efficiency increases rapidly with increasing size (Figure 2-13), so the droplet grows increasingly fast by collision. Moreover, as droplets get larger their cross sectional areas and fall velocities increase thereby increasing the collection kernel between large and small drops. Were it not for the fact that growth by collisions takes over as growth by condensation becomes negligible, many clouds would not rain.
Figure 2-12. Calculated values of the collision efficiency, $E$, for collector drops of radius $r_1$ with droplets of radius $r_2$. From Pruppacher and Klett [1997].

Another factor effecting droplet growth by collection is the efficiency with which colliding drops coalesce. There is a good database of the efficiencies of drop collision efficiencies [Beard and Ochs, 1993] but somewhat poorer knowledge of the coalescence efficiencies [Whelpdale and List, 1971; Levin et al., 1973; Beard and Ochs, 1993; Ochs et al., 1995], between drops over the range 1-300 microns. It is generally assumed that coalescence efficiencies are close to unity for small droplet collisions.

The basic problem is how do drops grow to a radius of 20 µm or greater fast enough to allow precipitation growth during the lifetime of clouds?

For an idealized cloud, the mean volume droplet radius at any height above cloud base, can be derived from the adiabatic water content at that height, $q_c(h)$, and the number concentration of activated nuclei $N$ as

$$R_v = \left(\frac{q_c(h)}{4/3\pi \rho_w N}\right)^{1/3} \quad (2.4)$$

For example, at a droplet concentration of 100 cm$^{-3}$, $q_c$ must reach a value of 3.2 g m$^{-3}$ (slightly less than 2 km of vertical ascent), for the mean volume radius to reach 20 µm. Since after CCN activation, the droplet spectrum approaches a monodispersed distribution, there are no droplets significantly bigger than $R_v$. The adiabatic assumption provides an estimate of the maximum droplet size, so that clouds growing in a polluted environment
(N>500 cm⁻³) should not be able to generate 20 µm radius droplets during their lifetime, based on the simplistic adiabatic model alone.

Four mechanisms have been hypothesized to explain how larger droplets or precipitation embryos can form:

- Role of giant cloud condensation nuclei.
- Turbulence influences on condensation growth.
- Turbulence influence on droplet collision and coalescence.
- Radiative cooling of drops to form precipitation embryos.

![Figure 2-13](image)

**Figure 2-13.** Schematic curves of the growth of a drop (a) by condensation (blue curve) and (b) by collection of droplets (red curve).

### 2.2.1.1 Role of Giant Cloud Condensation Nuclei (GCCN)

Observations reported by Woodcock [1953], Nelson and Gokhale [1968], Hindman [1975], Johnson [1976; 1982], and Hobbs et al. [1980] have shown the presence of potentially significant concentrations of aerosol particles of sizes as large as 100 µm. Their concentrations are ~10⁻³ cm⁻³ [Woodcock, 1953], that is, about one in 10⁵ or 10⁶ CCN are giant particles. Nevertheless, these particles can have a significant effect on the development of precipitation by serving as coalescence embryos [Johnson, 1982; Feingold et al., 1999; Yin et al., 2000a]. The droplet that forms is large enough for coalescence to start immediately even before the droplet reaches its critical size based on the Köhler
equation. This can occur if the nuclei are completely soluble (e.g., sea-salt particles) or are mixed particles with a soluble coating [e.g., mineral dust with a coating of sulfate, Levin et al., 1996] or are very large and wettable. The presence of GCCN on precipitation formation has been investigated in a number of cloud resolving models, which show that their contribution to rain formation may be appreciable in polluted clouds but has little influence in clouds forming in clean airmasses [i.e. Feingold et al., 1999; Khain et al., 2000].

2.2.1.2 Turbulence influences on Condensation

While the adiabatic model assumes a uniform updraft and a constant droplet concentration, convective clouds are in fact made of a series of updrafts of diverse intensities. The least vigorous updraft at cloud base also produces the lowest droplet concentration during activation of the available CCN. Turbulence further contributes to entrainment of dry environmental air in clouds, hence reducing LWC below the adiabatic value and diluting the droplet concentration below its initial value after CCN activation. Convective cells with a significantly reduced concentration might generate bigger droplets than adiabatic cores, if they are further experiencing a convective ascent [Baker et al., 1980; Telford and Chai., 1980, and Telford et al., 1984]. Turbulence however contributes also to the continuous mixing of the convective cells, hence smoothing out their differences in term of concentration and droplet growth.

Airborne measurements performed in cumulus clouds with the high resolution droplet spectrometer Fast-FSSP (see Chapter 6) reveal the occurrence of very narrow droplet spectra in cloud cores [Brenguier and Chaumat, 2001]. However, droplet spectra are often much broader than the narrow adiabatic reference, with droplet sizes extending from zero to the maximum predicted by the adiabatic model. This reflects the impact of mixing between convective cells that have experienced various levels of dilution with the entrained air. There are also airborne observations in stratocumulus clouds showing negative correlations between droplet concentration and droplet sizes [see for example Fig. 7 in Pawlowska and Brenguier, 2000]. This corroborates the hypothesis that fluctuations of the updraft intensity at cloud base or ascent of diluted cloud cells might contribute to the formation of droplets bigger than the adiabatic prediction based on the mean droplet concentration. In deeper clouds, airborne cloud traverses at a given level all look rather
similar in term of droplet concentration and sizes, suggesting that concentration/size correlations progressively dissipate due to continuous mixing.

Numerical simulations [Vaillancourt et al., 2002; Shaw et al., 1998] suggest that turbulence may also generate concentration fluctuations at the microscale by inertia in the regions of high vorticity, hence leading to superadiabatic droplet growth in the microcells with the lowest concentrations. This hypothesis was however contradicted by in situ measurements of the droplet spatial distribution at the microscale [Chaumat and Brenguier, 2001].

In summary, there is no observational evidence that turbulence contributes to the formation of precipitation embryos by enhanced condensation growth. In fact, stochastic processes induced by turbulence are not likely to enhance droplet growth because condensation is a cumulative process. To experience superadiabatic growth, droplets need to remain isolated in regions of higher supersaturation for a significant part of their ascent. The odds of this are low as mixing continuously redistributes droplets in the cloud. Turbulence is more likely to affect a discontinuous process like collision since once droplets coalesce, they cannot be separated.

2.2.1.3 Turbulence Influence on Droplet Collision and Coalescence

Turbulence can influence the collision and coalescence process in three ways:

- By enhancing collision efficiencies.
- By enhancing the collection kernels.
- By producing inhomogeneities in droplet concentration.

The collision efficiencies we discussed earlier were calculated in laminar or stagnant flow. In turbulent flow droplets will be accelerating and thereby be able to cross streamlines more efficiently than in laminar flow resulting in enhanced collision efficiencies. Large droplets, having more inertia will be affected more by turbulence than smaller drops. Calculations by Koziol and Leighton [1996] suggest that this effect is small for droplets smaller than 20µm diameter.
However, turbulence can also cause fluctuations on vertical fall speeds and horizontal motions, such that the collection kernel is enhanced [Pinsky and Khain, 1997; Khain and Pinsky, 1997], relative to that defined in laminar flow.

Because the collection rate is proportional to the square of droplet concentrations, inhomogeneities in droplet concentrations due to turbulence can produce enhanced regions of collection where the droplet concentrations are locally enhanced in, say regions of low vorticity [Pinsky and Khain, 1997]. This hypothesis has been examined in a cloud-resolving model by Khain and Pokrovsky [2005].

2.2.1.4 Radiative cooling of drops to form precipitation embryos

Consider a population of droplets that resides near cloud top for a sufficiently long time. Those droplets will emit radiation to space quite effectively if the atmosphere above is relatively dry and cloud free. As a result, the droplets will be cooler than they would without considering radiative effects. This means that the saturation vapor pressure at the surface of the droplet will be lowered and the droplets will grow faster.

But radiation cooling is proportional to the cross sectional area of a droplet so that its effect is much greater on larger droplets than small ones [Roach, 1976; Barkstrom, 1978; Guzzi and Rizzi, 1980; Austin et al., 1995]. In fact, Harrington et al. [2000] have shown that in a marine stratocumulus environment, when droplets are competing for a limited supply of water vapor, the larger droplets grow so rapidly by radiative enhancement that droplets smaller than 10µm in radius evaporate producing a bimodal size spectrum. This process is only effective in clouds where droplets reside near cloud top for time scales of 12 min or longer such as fogs, stratus, and stratocumulus. Cumulus clouds with vigorous overturning “expose” droplets to space for too short a time.

2.2.1.5 Mature phase of liquid phase precipitation formation

Once precipitation embryos form, they collect smaller droplets rapidly and they then transform into drizzle and raindrops in a matter of minutes if the liquid water content in clouds is great enough. The final size spectrum of the raindrops is determined by the liquid water content in the clouds as well as their trajectories through updrafts and downdrafts in the cloud. Moreover, raindrops may break up either spontaneously or by colliding with
other raindrops and breaking up. Thus, the final size-spectrum of raindrops is influenced by
the breakup process as well.

In summary, the ultimate control on the initiation, evolution, and intensity of rainfall from
warm clouds is the temporal-spatial structure of a cloud's updraft/downdraft and liquid
water content (LWC). These properties along with the initial activated droplet population,
provide the boundary conditions (i.e. time scales) for particle growth, that determine the
intensity and duration of precipitation. That is, each droplet must first undergo nucleation,
then condensation growth, followed by some mechanism of forming collection embryos,
then stochastic collection, followed by continued raindrop collection of smaller drops and
then breakup. All these steps take time. Thus critical to the formation of mature rain is
whether or not the cloud lifetime and parcel lifetimes are long enough to allow all these
steps to operate.

In simplest terms, clouds most likely to produce the most precipitation by exclusively
liquid phase processes are maritime and warm-based. By maritime, we mean clouds
forming in a clean atmosphere with low CCN concentrations. If a cloud is warm-based then
its saturation mixing ratio at cloud base is high. For example, a cloud with a base
temperature of 24°C has a saturation mixing ratio of roughly 20g/kg, whereas one with a
base temperature of 5°C has a saturation mixing ratio of less than 6g/kg. Thus if a cloud
with a warm base has a depth of, say, 3km it has a much greater potential of condensing a
large amount of condensate than a cloud of similar depth having a higher, cold cloud base.
We sometimes refer to warm-based, maritime clouds as being colloidally unstable [see
Cotton and Anthes, 1989].

By contrast, a cloud, which is cold-based, and forms in a continental or an aerosol-polluted
environment has a much smaller chance of producing precipitation and is said to be
colloidally stable. The potential for liquid phase precipitation formation in most clouds
will be between these two extremes.

2.2.2 Ice Particle Mechanisms

In regions of clouds lower than 0°C, supercooled water droplets can be present with or
without coexisting ice particles. In Section 2.1.3, we discussed the nucleation of ice in
clouds and ice multiplication mechanisms. Here we are concerned with the growth of ice particles to precipitation sizes in mixed clouds.

As early as 1789 Benjamin Franklin [1789] wrote “It is possible that, in summer, much of what is rain, when it arrives at the surface of the Earth, might have been snow, when it began its descent, but being thawed, in passing through the warm air near the surface, it is changed from snow to rain.” This idea was not developed until the early twentieth century, when Wegener [1911] noted that ice particles would grow preferentially by vapor deposition in a mixed cloud. Subsequently, Bergeron [1933] and Findeisen [1938] developed this idea in a more quantitative manner.

In a mixed cloud dominated by supercooled droplets, the air is close to saturation with respect to liquid water and is therefore supersaturated with respect to ice. For example, air saturated with respect to liquid water at –10ºC is supersaturated with respect to ice by 10%, and at –20 ºC, the air is supersaturated with respect to ice by 21%. These supersaturations with respect to ice are much greater than the supersaturations of cloudy air with respect to liquid water, which rarely exceed 1%. Consequently, in mixed clouds dominated by supercooled water droplets, ice particles will grow from the vapor phase much more rapidly than droplets. This can lead to a variety of ice crystal types, some of which may be large enough to precipitate. This process is generally referred to the “vapor growth stage” of ice particle growth.

In a mixed-phase cloud, ice particles can also grow by colliding with supercooled droplets, which then freeze onto them. This growth process, which is referred to as rimeing, can produce heavily rimed ice particles, graupel particles and, if the cloud is deep enough, contains sufficient supercooled water, and updrafts are strong, hailstones can form. Finally, ice particles in a cloud may collide and aggregate with each other, leading to larger particles, which are called aggregates. Unfortunately, the collision and coalescence efficiencies of ice particles rimeing cloud droplets are not well documented. Even more poorly quantified are the collision and coalescence efficiencies, and collection kernels among ice particles undergoing aggregation.

The growth of ice particles, first by deposition of water vapor followed by rimeing and/or aggregation, can produce precipitation-sized particles. If temperatures at ground level are
below 0ºC, these particles will reach the ground as snow; if the surface temperature is above 0ºC, the ice particles will partially melt, or melt completely, and reach the ground as wet snow or rain. An increase in the concentration of cloud droplets accompanied by decreases in the average size of the droplets (which can be caused by an increase in the concentration of CCN) may decrease growth by riming due to the lower collection efficiency of smaller droplets [Borys et al., 2000, 2003].

In deep convective clouds, the rapidity of glaciation is dependent upon the presence of drizzle drops and large supercooled raindrops [Cotton, 1972; Koenig and Murray, 1976; Scott and Hobbs, 1977]. When CCN concentrations are high, liquid phase precipitation collection processes are suppressed and supercooled raindrops are few in number, small ice crystals must first grow to several hundred micrometers in diameter before they begin collecting cloud droplets. Then the riming process proceeds slowly until the ice particles have grown to millimeter-size. Furthermore, since riming is suppressed in clouds forming in air masses with high CCN concentrations, secondary ice particle production by the rime-splintering process [Hallett and Mossop, 1974; Mossop and Hallett, 1974] is suppressed as well.

If the updrafts and liquid water contents in deep cumulonimbi are large enough, graupel particles, frozen raindrops and large aggregates can serve as embryos for hailstone formation [Heymsfield et al., 1980]. At first, the density of graupel particles is low as the collected frozen droplets are loosely compacted on the surface of the graupel particle. As the ice particle becomes larger, it falls faster, sweeps out a larger cross-sectional area, and its growth by collection of supercooled droplets increases proportionally. As the growth rate increases, the collected droplets may not freeze instantaneously upon impact, and thus flow over the surface of the hailstone filling in the gaps between collected droplets. The density of the ice particle, therefore, increases close to that of pure ice as the dense hailstone falls still faster, growing by collecting supercooled droplets as long as the cloud liquid water content is large. The ultimate size of the hailstone is determined by the amount of supercooled liquid water in the cloud and the time that the growing hailstone can remain in the high rainwater region. The time that a hailstone can remain in the high liquid water content region, in turn, is dependent on the updraft speed and the fall speed of the ice particle. If the updraft is strong, say 35-40 m/s, and the particle fall-speed through the air is only on the order of 1-2 m/s, then the ice particle will be rapidly transported into the anvil...
of the cloud before it can take full advantage of the high liquid water content region. The ideal circumstance for hailstone growth is that the ice particle reaches a large enough size as it enters the high liquid water content region of the storm so that the ice particle fall speed nearly matches the updraft speed [Foote, 1984]. In such a case, the hailstone will only slowly ascend or descend while it collects cloud droplets at a very high rate. Eventually, the hailstone fall speed will exceed the updraft speed or it will move into a region of weak updraft or downdraft. The size of the hailstone reaching the surface will be greatest if the large airborne hailstone settles into a vigorous downdraft, as the time spent below the 0C level will be lessened and the hailstone will not melt very much. Thus, a particular combination of airflow and particle growth history is needed to produce large hailstones.

In summary, hail growth is most likely if the convective storm updrafts are strong, if the liquid water contents are large, the storm is deep and long-lived, and the melting level is low relative to the ground [Foote, 1984]. There is some limited modeling work that suggests the concentrations of CCN and IN are factors in hail formation [Danielsen, 1977].

### 2.2.3 Summary of Cloud Microphysical Processes

In preceding sections, we have seen that the evolution of precipitation in clouds can take on a variety of forms and involve numerous physical processes. The evolution of ice-phase precipitation processes is greatly dependent upon the prior or concurrent evolution of the liquid-phase. These processes, in turn, are dependent upon the characteristics of the air mass (i.e. aerosols), the liquid-water production of the cloud, the vertical motion of air within the cloud, the turbulent structure, and the time scales of the cloud. Illustrated in Figure 2-14 are the different precipitation paths that may occur depending upon whether the cloud is a cold-based continental cloud versus a warm-based maritime cloud. Remember from earlier discussions that we use the term maritime cloud to represent a very clean airmass and continental to represent one with much higher CCN concentrations. A polluted cloud would have still higher CCN concentrations. The figure does not note the speed by which these regimes can produce precipitation. We have pointed out earlier that a cloud with a vigorous warm rain process or what we refer to in the figure as a warm-based maritime cloud will produce precipitation much faster than a cold-based continental cloud.
The rapidity of glaciation of a warm-based maritime cloud is much faster than the cold-based maritime cloud since the presence of drizzle drops and supercooled raindrops once frozen can rapidly transform a cloud from an all water cloud to an ice-dominated cloud. This should not be interpreted to mean that the largest precipitation elements such as hail would occur in a warm-based maritime cloud. In fact, just the opposite can take place as a vigorous precipitation process lower in the cloud can deplete supercooled water amounts higher up and lower the trajectory of precipitation elements leading to smaller sized hailstones (see Chapter 8).

This diagram is used to illustrate two distinctly different regimes. In fact, there is a continuum of cloud types between these two extreme states. The heaviest precipitating clouds, and generally most efficient, are those that are warm-based and form in maritime airmasses. By contrast, clouds developing in a polluted airmass should be less efficient in producing precipitation than similar clouds with the same cloud base temperatures. This simple reasoning is valid only for single clouds and storms. As we will see in Chapter 7, in some cases a suppression or retardation of precipitation in a primary convective cell could lead to a transformation of a cloud into a long-lived squall line through the interaction of cold-pools with wind shear leading to self-propagating storm.
2.3 Precipitation Efficiency

The term *precipitation efficiency* has many definitions in the literature and generally refers to the fraction of either the water vapor input at cloud base or condensed water in cloud that falls out as precipitation. Instantaneous values of PE vary from near zero early in the lifetime of a cloud before precipitation has commenced to values exceeding 100% during the dissipation stages of a storm, when cloud-base moisture fluxes are near zero [Market et al., 2003]. Doswell et al. [1996] suggest that PE is most meaningful when averaged over the storm lifetime. Market et al. [2003] propose that it is best to define a volume around a moving system and employ storm-relative winds in evaluating PE. In this way, storm-averaged PE can be obtained for a moving system. PE estimates from different studies are

---

**Figure 2.14.** Flow diagram describing microphysical processes, including different paths for precipitation formation. (Adapted from Braham [1968]).
problematic to compare because of differing definitions of PE, different data sources such as aircraft in situ versus radar measurements, and different time and spatial averaging [Hobbs et al., 1980].

Fankhauser [1988] presented one of the most detailed studies of thunderstorm precipitation efficiency. Data were taken from seven storms during the Cooperative Convective Precipitation Experiment (CCOPE) using aircraft, rawinsondes, a surface mesonetwork and radar histories and calculated values ranging from 19% to 47%. Various environmental quantities, such as kinetic energy, cloud base area and height, and cloud base mixing ratio were found to be factors that had a strong positive correlation to PE levels. He found that variables such as the bulk Richardson number, the ratio of buoyant energy to the amount of wind shear, and convective available potential energy (CAPE), a measure of potential updraft strength, did not correlate well with PE.

Another strategy for estimating precipitation efficiencies is to use drying ratio (DR) as proposed by Smith et al. [2003].

\[
DR = \frac{\text{total precipitation}}{\text{vapor flux}}
\]  

(2.5)

DR encapsulates the moisture budget of the large-scale air mass transformation. Because PE requires an estimate of vertical air velocity at cloud base, it is subject to rather large errors in those measurements. Since DR uses the ratio of precipitation to water vapor flux, it is easier to quantify than PE. Moreover, DR can be estimated using hydrogen and oxygen isotope analysis, which permits evaluation using streamflow or sapwood collected near a stream. Isotope analysis indicated a DR value of 43% for the combination of the coastal and Cascade mountain ranges in western Oregon and 48% for the southern Andes [Smith et al. 2005; Smith and Evans, 2006].

To date no one has made estimates of storm-averaged PE or DR for storms developing in environments containing different cloud-nucleating aerosol concentrations. Such work should be encouraged as a means of determining the impact of aerosols on precipitation while accounting for different water vapor and vertical air motion conditions.
3.1 Introduction

In view of the nonlinear and spatiotemporally complex interactions between aerosols, clouds and precipitation, it is quite inadequate to represent aerosols by the conventional approach of atmospheric pollutant cycles, i.e., by specifying sources, burdens and sinks in terms of masses of particular aerosol components or “species”. The basic information required for modeling the role of aerosols in cloud and precipitation processes is the 4-D distribution of the number concentration of particles that can be activated at the supersaturations occurring in clouds (Cloud Condensation Nuclei, CCN) or that can initiate the formation of ice particles, especially at temperatures above the level of homogeneous freezing of cloud droplets (Ice Nuclei, IN). This information is difficult or impossible to derive from conventional studies of aerosol cycles, such as those that formed the basis of the section on aerosol sources and burdens in the IPCC-TAR report [Penner et al., 2001]. Several problems present themselves:

First, the very concept of "source strength" is difficult to define for some aerosol types, since the secondary aerosol species (e.g., sulfates, secondary organics) are not directly emitted, but are formed in the atmosphere from gaseous precursors. For these species, the “source” is a set of atmospheric reactions, not a surface emission process, and therefore the production rate depends sensitively on the efficiency of these reactions. These “sources” cannot be derived from measurements, but have to be inferred from measurements of the source strength of precursors, the (always incomplete) knowledge of atmospheric transformations, and a formal or informal inversion of measurements of atmospheric burdens. The source estimates for these aerosols therefore depend strongly on how the source processes are represented in the models.

Second, the conventional mass-based treatment of aerosol fluxes does not take into account the fact that the properties of aerosol particles can vary strongly with particle size, and that the cloud-physical effects of aerosols depend on number concentrations, not on mass concentrations. For example, the largest number concentration of cloud-active seasalt particles resides in the submicron size fraction, which represents only a very minor mass fraction of the seasalt aerosol. On the other hand, this fraction of the seasalt aerosol has quite
different atmospheric sink mechanisms and lifetime than the coarse seasalt mode, which dominates the mass source strength and burden of the aerosol component. Similar considerations apply to mineral dust and primary biogenic aerosols (see Chapter 4).

A third problem arises from the fact that aerosol species often combine to form mixed particles, with properties and atmospheric lifetimes different from those of their components. Even though the direct radiative effects of these “internal mixtures” may be substantially different from those of “external” mixtures [Jacobson, 2001; Chung and Seinfeld, 2002], this distinction has been ignored in many analyses of the global aerosol system. However, for the effects of aerosols on clouds, even relatively small amounts of soluble material in or on otherwise insoluble particles can cause large differences, and the mixing state of the aerosol cannot be ignored.

It follows from these considerations, that a meaningful analysis of the sources and burdens of cloud-active particles cannot be achieved through emission inventories, as it has been done typically for atmospheric trace gases. Rather, it must take the form of an atmospheric model that contains emissions of precursor gases and primary aerosol species, and accounts for transformations and sink terms of precursor species as well as the various aerosol components. The model must carry either explicitly or in parameterized form information about the number/size distribution of soluble and wettable materials, ice-nucleating ability, and the interactions between gas phase and various condensed phases. Fundamentally, therefore, the issue of aerosol sources cannot be separated from an analysis of transport and atmospheric processes and must be treated interactively within the framework of a numerical model that represents these processes. Unfortunately, this type of approach is still in its infancy. One of the most complete analyses (suitable for long-term climate simulations) to date is represented in the aerosol-climate model ECHAM5-HAM [Stier et al., 2005]. In this model, the major aerosol components sulfate, particulate organic matter (POM), black carbon (BC), seasalt, and mineral dust are determined through an aerosol module that is interactively connected to a climate model. In a somewhat complementary approach, the distributions of ammonium sulfate and nitrate aerosols have been modeled using prescribed meteorology and a detailed description of gas/aerosol partitioning [Metzger et al., 2002a].
In the following sections, we will present an updated inventory of emission fluxes, production rates, and burdens for the major aerosol species, providing information about particle numbers when possible. For selected aerosol types, we will provide spatial information in the form of 2D distribution maps. Ultimately, however, it must be acknowledged that this information just represents a convenient visualization of the spatiotemporally resolved distributions, which can be fully represented only in the form of aerosol modules in coupled aerosol-climate models. In addition, these data are results from one particular global aerosol model (ECHAM5-HAM) and do thus include the uncertainties connected to such simulations [Textor et al., 2005].

3.2 Natural and Anthropogenic Sources of Atmospheric Aerosols

Aerosol sources are classified into primary and secondary types. Primary aerosols are those that are emitted into the atmosphere directly as condensed solids or liquids, whereas secondary aerosols are formed within the atmosphere from gaseous precursors. Sea salt, mineral dust and soot particles are clearly primary particles, whereas organic particles from the oxidation of volatile organic compounds (VOC) and sulfates from the oxidation of SO\(_2\) or other sulfur-containing gases are clearly secondary. There is a “grey zone” between primary and secondary in the case of some low-volatility organic compounds that condense onto aerosol near, but not directly within, emitting sources, such as some hydrocarbons in vehicular exhaust or condensable from biomass burning. We will discuss these materials together with the strictly primary emissions, taking into account that the definition of secondary aerosol formation involves chemical transformation from more volatile precursors [Fuzzi et al., 2006].

3.2.1 Soil Dust

Soil or mineral dust is the most conspicuous aerosol component in satellite imagery of global aerosol distributions (see Figure 3-1). Many estimates of source strength and burden have been published; much of the older work has been summarized by Duce [1995], while the more recent literature has been reviewed by Tegen [2003] and Zender and Tegen [2004]. The estimates of global annual source fluxes are in the range of 1000-2150 Tg a\(^{-1}\), and have not changed much since the reviews of Andreae [1995] and IPCC-TAR [2001]. There has been
tremendous improvement, however, in the sophistication of the dust deflation and transport models used and in the techniques used for validation of the models, especially the use of remote sensing [Israelevich et al., 2002; Tegen et al., 2002; Werner et al., 2002; Luo et al., 2003; Zender et al., 2003; Miller et al., 2004; Tegen et al., 2004]. The range of burdens predicted by these models, 8-36 Tg, remains still quite large, mostly due to differences in the way deposition processes are treated in the models. These differences result to a substantial extent from different assumptions regarding the size distribution of the dust aerosol.

There are major discrepancies in the size distributions of the mineral dust aerosol predicted by the models and observed by various techniques. Near the sources, the mass distribution peaks at large sizes, 10 µm or greater, in observation and models. At the source, about 10% of the dust is in the submicron fraction [Tegen et al., 2002; Luo et al., 2003]. More relevant at the regional and global scales is the size distribution in the atmospheric burden after some atmospheric transport and processing has taken place. In the models, mass median diameters are typically 2.5 µm or less. In a study using the NASA GISS atmospheric general circulation model (AGCM) with interactive treatment of dust flux and climate, Miller et al. [2004] obtain a substantially lower dust MMD (~1.6 µm) than in an off-line tracer model (~ 3 µm). Observations show at least twofold higher sizes than the model predictions, with some internal disagreement between different measurement approaches. Measurements of dust MMD by aerodynamic methods give a mean of 4.5±1.3 µm, in good agreement with the results from optical inversion methods (5±1.5 µm), but considerably below values obtained by optical counter methods [Reid et al., 2003]. Since the current dust models are being constrained by data obtained by various combinations of these observational approaches, and in turn used to predict radiative forcing, substantial problems related to non-physical parameterizations may result.

Detailed analysis of dust source areas in recent years has shown that there are areas with particularly strong dust emission potential (“preferential source areas”), in particular dry lake beds in arid areas [Israelevich et al., 2002; Prospero et al., 2002; Tegen et al., 2002; Zender et al., 2003]. These regions are also of special relevance to the issue of aerosol/cloud effects, as they contain a higher fraction of clay-size (1 µm or less) particles and therefore have the potential of releasing an over-proportionally higher number fraction of particles compared to the coarser desert dust particles. They also include soluble residues from the evaporation of
the lake water, which renders them highly suitable as CCN [Formenti et al., 2003]. Clay minerals are also known to be effective ice nuclei, although, presumably, these minerals from dry lakes may behave quite differently than ones from higher deserts due to the additional soluble content.

Dust emission shows substantial variability at all time scales from the last glacial maximum to the anthropogenically influenced present and in simulations of the future. This variability is related to changes in vegetation cover, soil moisture, and anthropogenic disturbance of the soil and vegetation surface, as well as to changes in climatic parameters such as wind speed and precipitation [Mahowald et al., 2002; Tegen et al., 2002; Werner et al., 2002; Luo et al., 2003]. Variability at interannual to decadal time scales makes longer-term trends hard to discern. Earlier views that anthropogenic soil disturbance was responsible for as much as half of the present-day dust flux have been revised downward in the last few years, and the contribution of dust from agricultural activities is now estimated to be <10% [Tegen et al., 2004]. Since dust affects precipitation through both radiative transfer and cloud microphysics, there is a feedback through soil moisture on dust emission. Dust exerts a negative forcing at the surface and a positive one in the atmosphere. The magnitudes of these forcings are dependent on the optical properties (mineralogy) of dust, which are still poorly known. Model calculations by Miller et al. [2004] suggest that increased dust loadings results in more rain in the Sahara, thus producing a negative feedback. In their model, which does not include the effects of dust on cloud physics, dust reduces global rainfall, thus lengthening the lifetime of aerosols, and increasing the burden of dust and other aerosols.

Dust fluxes and burdens in future decades will depend strongly on changes in land cover and climate, but reliable estimates cannot be made at present. One estimate suggests that dust fluxes may decrease with future land use and climate change, that is, by 10-60% by 2100, [Mahowald and Luo, 2003], while another suggests smaller changes (10-25%), with increases or decreases possible depending on which climate/land-cover model is used [Tegen et al., 2004]. Because of the strong dependence of dust mobilization on windspeed (increasing with a power of approximately 3-4), dust fluxes in future climates are highly sensitive to changes in the wind regimes in dust source areas.

For an analysis of the impact of dust aerosols on clouds and precipitation, the most important information would be the distribution of the number concentration of dust particles that can
act as CCN, giant CCN (GCCN), and IN. Unfortunately, the published studies on dust budgets do not provide this information. The issue is further complicated by the fact that dust particles may or may not contain soluble materials already at the time of emission, e.g., in the form of evaporation residues, or that they may become coated with sulfate, nitrate, or other soluble salts during atmospheric processing, especially in clouds [Levin et al., 1996; Yin et al., 2002; Trochkine et al., 2003]. Most of the dust models carry information about number size distributions, explicitly or implicitly through specification of the modal or bin characteristics of the mass/size distributions. The number maximum of the dust aerosol mode is generally well below 1 μm diameter. For example, Zender et al. [2003] show the number maximum at 0.6 μm diameter. Given the relatively high contact angles between water and silicates [Pruppacher and Klett, 1997], we would expect that most of these particles would not be able to act as CCN at commonly found supersaturations, unless they were coated with soluble material. If they are not activated as cloud droplets, a substantial fraction of dust particles can remain as interstitial aerosol in convective clouds and reach the middle and upper troposphere, where they can act as IN. If, on the other hand, large dust particles contain appreciable amounts of soluble material or are coated with soluble substances, they can play an important role as GCCN (see Chapter 4). In summary, we find that the source fluxes and burdens of dust aerosols, and the CCN- and IN-active fractions of this aerosol, depend strongly and interactively on atmospheric processes, especially on their behavior in clouds.

### 3.2.1 Sea Salt

Sea salt aerosols are generated by various physical processes, especially the bursting of air bubbles entrained during whitecap formation and the tearing of droplets from wave tops [Schulz et al., 2004]. The rates of particle production by both mechanisms are strongly dependent on wind speed, with production of spume droplets from wave tops occurring only at wind speeds >9 m s\(^{-1}\). Under moderate wind regimes, seasalt particle number concentrations are typically about 10 cm\(^{-3}\) or less, while under high wind conditions (>10 m s\(^{-1}\)), their concentration can rise to 50 cm\(^{-3}\) or more [O'Dowd et al., 1997; Shinozuka et al., 2004].

In ocean regions where wind speeds are high and/or other aerosol sources are weak (such as the winter-time remote oceans of the Southern Hemisphere) seasalt may be the dominant
contributor to CCN [O’Dowd and Smith, 1993; Murphy et al., 1998a; Quinn et al., 1998; Gong et al., 2002; Shinozuka et al., 2004]. By presenting a primary aerosol surface area, they can also inhibit new particle formation.

Sea salt particles cover a wide size range (ca. 0.05-10 μm diameter), with most of their mass in the supermicron range. The highest number concentration of seasalt particles, however, is in the submicron range, with a pronounced number mode at ca. 500 nm, and significant number concentrations still present at 100 nm [Campuzano-Jost et al., 2003]. Unfortunately, field observations of seasalt particles presently only reach down to ca. 100 nm [O’Dowd et al., 1997; Murphy et al., 1998a; Campuzano-Jost et al., 2003], but laboratory studies of seasalt aerosol production by Mårtensson et al. [2003] indicate that the maximum of the number distribution is actually somewhat below 100 nm (~50-80 nm). Very small sea-salt particles, with a number maximum around 20-30 nm, were found to be produced by coastal breaking waves [Clarke et al., 2003]. This is supported somewhat indirectly by measurements of particle flux from the sea surface [Nilsson et al., 2001] that also yield a mode at or below 100 nm, but in this study the composition of these particles could not be identified as seasalt.

A substantial source of very small seasalt particles could be of importance to the marine CCN budget even if they are below the critical diameter for droplet activation at the low supersaturations in marine stratus clouds, because they can serve as nuclei for the growth of sulfate particles from the oxidation of dimethyl sulfide, DMS which is produced from ocean biota [Charlson et al., 1987]. Direct production of sulfate particles by homogeneous nucleation is thought to be only rarely possible in the clean marine atmosphere [Pirjola et al., 2000]. On the other hand, larger seasalt particles also serve as a sink for gaseous H₂SO₄ molecules, thereby reducing the amount of sulfate available for the formation of accumulation mode particles [Gong and Barrie, 2003]. In summary, seasalt particles are only a minor fraction of the CCN population under most situations, except for the winter Southern Ocean, North Atlantic and North Pacific. There is evidence for the existence of fine and ultrafine seasalt particles, but their abundance and their direct or indirect role as CCN source remains unclear. More detailed studies on the size distribution, composition, and production rates of the submicron marine aerosol will be required to resolve these issues.
Several studies in the last few years have shown that “seasalt” aerosol actually contains a substantial amount of organic matter, consisting both of insoluble material (biological debris, microbes, etc.) and water soluble constituents [Novakov et al., 1997; Middlebrook et al., 1998; Cavalli et al., 2004; O'Dowd et al., 2004]. The fraction of organic components increases with decreasing particle size, and, in biologically active regions, may approach 90% in the size fraction around 100 nm [Cavalli et al., 2004]. The chemical and isotopic composition of this organic matter suggests a marine biogenic source [Turekian et al., 2003]. Organic materials and seasalt are present as internal mixtures, consistent with a production mechanism that involves fragmentation of organic-rich surface film layers during the bursting of air bubbles at the sea surface. Some evidence on dry particles suggests that the organics may be present as coatings on the surface of salt crystals [Tervahattu et al., 2002b]. In this case, they could interfere with the uptake of water by the particles [Saxena and Hildemann, 1996], and with the collision-coalescence process [Medina and Nenes, 2004], but these findings may not be applicable to the more liquid-rich seasalt particles at high (especially super-saturated) humidities [Abbatt et al., 2005].

Since sea salt particles are very efficient CCN, characterization of their surface production is of major importance for aerosol impacts on clouds. Estimates of the total sea salt flux from ocean to atmosphere vary over a wide range. The most recent estimate [Schulz et al., 2004] is 2690 Tg a$^{-1}$ in a mode centered at 2 µm diameter, plus 17,100 Tg a$^{-1}$ in a mode at 11 µm diameter. At first glance, this appears considerably higher than most previous ones, e.g., 3300 Tg a$^{-1}$ [Houghton et al., 2001], 6500 Tg a$^{-1}$ [Grini et al., 2002], or 5900 Tg a$^{-1}$ [Tegen et al., 1997]. A large part of the differences, however, can be attributed to the choice of the upper size cut-off. With increasing cutoff size, a large mass flux of very coarse and therefore very short-lived salt particles is included in the estimate.

Because most of the seasalt mass is in the coarse mode, while most of the seasalt particle number is in one or more submicron modes, the conventional estimates for the mass source flux of seasalt are of little relevance anyway for the discussion of seasalt particles as CCN, and it is necessary to analyze their number production rates and atmospheric distribution in a size-resolved model. This applies both to the submicron fraction, where most of the seasalt-CCN are found, and to the coarse fraction, which contains the giant CCN (see Chapter 4). There are several parameterizations available for the windspeed-dependent, size-resolved
emission flux of particles from the sea surface, most of which have been summarized in Schulz et al. [2004]. The best match with the observed size distributions \cite{O'Dowd et al., 1997} is obtained using the Vignati and Gong-Monahan schemes \cite{Vignati et al., 2001; Gong, 2003}, among which the Gong-Monahan scheme probably has the more faithful representation of the wind-speed dependence. Estimates of the number flux of seasalt particles have not been published so far, even though it must be assumed that they have been calculated in the various aerosol models. Below (section 2.2.4) we are presenting an estimate that was retrieved from the model runs that formed the basis of the papers by Stier et al. [2005; 2006]. Because of the strong dependence of seasalt particle production on wind speeds, the role of these particles in cloud and precipitation processes is expected to change with anthropogenically induced climate change.

3.2.3. Carbonaceous Aerosol

Carbonaceous material, consisting of organic compounds and near-elemental, sub-graphitic material (“soot”) represents an important fraction of the atmospheric aerosol. Only a minor part of this complex mixture, in the range 10 to 40%, has been identified at the molecular level \cite{Gelencsér, 2004; Decesari et al., 2006; Kanakidou et al., 2005}. The carbonaceous aerosol contains a large fraction of oligomers and highly polymeric matter from various origins. This unique chemical complexity is the cause of much of the prevailing uncertainties regarding the sources and properties of carbonaceous aerosol.

Carbonaceous materials represent a continuum from semi-volatile organic species (SVOC), chemically-identified particulate organic carbon (POC), polymeric POC, to black carbon (BC) or elemental carbon (EC). The BC or EC component is operationally defined as light-absorbing (“black”) or thermally refractory (“elemental”) carbon, and is generally thought to consist of aggregates of 20-50 nm granules with a graphite-like crystal structure and near-elemental composition with some surface functionalities \cite{Smith et al., 1989}. Such particles are also called “soot” particles (which term often includes the organic coating that is usually associated with these particles), and originate from various types of combustion \cite{Andreae and Gelencser, 2006}. Especially in biomass smoke, however, there are also other refractory organic compounds that operationally act like EC, and other light-absorbing organics that are determined as BC, so that at this time this aerosol component remains rather ill-defined. This
is reflected in large discrepancies between the results of OC and BC/EC measurements from different laboratories (up to factors of 10 for BC/EC), even when identical samples and nominally identical analytical protocols are used [Ten Brink et al., 2004]. The strong increase of the absorption with decreasing wavelength can sometimes be used to identify aerosols from biomass burning [Kirchstetter et al., 2004; Schmid et al., 2006].

As the most commonly used combustion-based chemical analyses provide carbon contents only, a conversion factor $k$ is required to account for organic functional groups in particulate organic matter (POM = $k \times$ POC). The $k$ factor varies with the origin and age of the particles over a wide range: 1.3 (close to sources) to $>2.0$ [Putaud et al., 2000; Turpin and Lim, 2001; Kiss et al., 2002; Russell, 2003].

In addition to primary carbonaceous aerosols, there is a substantial amount of secondary organic aerosols (SOA) formed from a large variety of anthropogenic and biogenic gaseous precursors [Seinfeld and Pankow, 2003]. As a result, organic aerosols contain variable amounts of primary (PPOM: Primary particulate organic matter; PPOC: Primary particulate organic carbon) and secondary (SOA) material. It has been shown that in source regions the formation of combustion-derived SOA occurs readily and could represent the same amount as PPOM [Cooke et al., 1999]. Once formed, carbonaceous aerosols are subject to continuous chemical modification in the atmosphere by photochemical and multiphase reactions [Maria et al., 2004]. This includes reactions with OH and other radicals from the gas phase [Molina et al., 2004; Zuberi et al., 2005], oligomerization with and without acid catalysis [Claeys et al., 2004; Gao et al., 2004; Kalberer et al., 2004], and formation of humic-like substances (HULIS) by photochemical polymerization of biogenic precursors [Gelencsér et al., 2002].

3.2.3.1 Primary Biogenic Aerosols

Primary biogenic aerosol particles (PBAP) may enter the atmosphere from a number of sources. Plants shed various types of debris (cuticular waxes, leaf fragments, etc.), and soil deflation releases a mixture of humic matter, plant decomposition products, fungi and microbes. Atmospheric aerosols over land and oceans contain a large variety of microbial particles (bacteria, fungi, viruses, algae, pollen, spores, etc.). These particles cover a wide size range from tens of microns down to a few tenths of microns. Morphological studies using light and electron microscopy have shown that PBAP make up a large fraction of the aerosol down to $\sim$0.2 $\mu$m at a variety of clean and polluted sites (e.g., the Amazon, Siberia,
and even urban Central Europe [Matthias-Maser and Jaenicke, 1995; Matthias-Maser et al., 2000; Graham et al., 2003a; Graham et al., 2003b; Posfai et al., 2003][Jaenicke, 2005]. Their number concentrations are in the range of tens to hundreds cm\(^{-3}\). Unfortunately, it becomes quite difficult to unequivocally identify particles in the submicron range as primary biogenic, as they often have little or no diagnostic morphological or bulk chemical signatures. Chemical tracers studies suggest, however, that primary biogenic material contributes a large fraction of the pristine aerosol even in this size range [Simoneit et al., 1990; Graham et al., 2003b]. Very small PBAP (20-50 nm) may also play a role as growth nuclei for CCN-sized particles over the remote oceans, through condensational growth by deposition of sulfate derived from the oxidation of DMS [Leck et al., 2002; Leck et al., 2004].

There is considerable evidence that primary biogenic particles may be able to act both as CCN and IN [Schnell and Vali, 1976a; Levin and Yankofsky, 1983; Diehl et al., 2001; Bauer et al., 2003]. This is consistent with their often relatively large size and soluble matter content, but surface properties are also thought to play a significant role [Bauer et al., 2003]. It is therefore possible, that PBAP represent a significant fraction of CCN and IN in the clean atmosphere, and therefore are of importance for the regulation of cloud microphysical properties under pristine or remote conditions, but not enough is known at this time to assess their role with any confidence. Since their atmospheric abundance may undergo large changes as a result of land use change, they deserve more scientific study.

Since many of these PBAP are of large size, they are of special interest because they can already be activated at very low supersaturations (<0.02%). Therefore, they will be the first to activate at cloudbase, and they may therefore represent a major source of GCCN. This might enable them to play an important role in precipitation formation under circumstances when high concentrations of pollutant CCN otherwise would suppress warm rain production [Yin et al., 2000a; Rudich et al., 2002]. To date, there are no reliable estimates on the rates of PBAP emissions. The estimates given in Table 3-1 have been calculated based on the commonly observed mass concentrations of 1-2 µg m\(^{-3}\) over remote, vegetated continental regions and an estimated atmospheric lifetime of 2 days for these particles.
3.2.3.2 Primary Anthropogenic Carbonaceous Aerosol

Black carbon and a fraction of PPOM and SOA originate from combustion sources. Combustion sources using fossil fuel are found in industry, power generation, traffic and residential heating, which take place to about 90% in the Northern hemisphere [Andres et al., 1996]. It has, however, been shown by carbon-14 analysis that even in cities the biogenic fraction may comprise about 60% of the carbonaceous aerosol, which in summer is mainly attributed to SOA from biogenic precursors, while in winter there is a substantial contribution from wood combustion [Szidat et al., 2006]. In the intertropical zone, biomass burning (open savanna, forest and agricultural waste fires) and biofuel use form the main sources of anthropogenic carbonaceous aerosols. Biomass burning from natural fires provides a major source of carbonaceous aerosols mainly in the mid- to high latitudes of the northern hemisphere [e.g., Lavoué et al., 2000]. SOA is also formed by the conversion of natural organic gases such as terpenes and, probably to a lesser extent, isoprene [Claeys et al., 2004]. From both biomass burning and fossil fuel sources, POM (PPOM + SOA) is always the predominant fraction of the carbonaceous aerosol. However, the chemical composition of the aerosol as represented by the BC/POM ratio is highly dependent on the type of sources: on average, the BC/POM ratio is higher for fossil fuel than for biomass burning sources. Also BC/POM ratio is linked to the type of burning: BC/POM is probably related to the temperature of the combustion process and thus for example higher for diesel emissions than for gasoline [Bond et al., 2004]. Biomass combustion aerosols emitted under flaming conditions will have a higher BC/POM ratio than smoldering aerosols, which are almost entirely of organic nature [Reid et al., 2005].

At present, lack of adequate data for BC and PPOC concentrations precludes the use of inverse modeling. To estimate the emissions of combustion carbonaceous aerosols, the only suitable method is at present the bottom-up approach. Emission estimates are based on two different data sets: one is related to fuel consumption and another to particulate emission factors.

Spatial distributions of fuel consumption (fossil fuel and biofuel) may be obtained from data provided by a variety of national and international organizations. They are complemented by regional inventories that make special efforts to account for regional source characteristics.
and may provide more detailed consumption maps [Streets et al., 2001; Reddy and Venkataraman, 2002; Streets et al., 2003; Schaap et al., 2004].

Mapping the amounts of burnt biomass is a difficult task. In previous studies [Hao and Liu, 1994; Hao et al., 1996] estimates of burnt areas for savanna and forest fires were based on statistical data. For emissions related to agricultural practices, the FAO (Food and Agriculture Organization of the United Nations) data are usually associated with assumptions of the burnt fraction [Yevich and Logan, 2003]. Qualitative improvements are now obtained from multi-year global distributions of fire pixel counts and burnt area given by satellite imagery, as these data allow taking into account the spatial and temporal variability of fires at different scales [Kajii et al., 2002; Generoso et al., 2003; van der Werf et al., 2003; Hoelzemann et al., 2004; Ito and Penner, 2004; Liouss et al., 2004; Michel et al., 2005].

The need to couple the two existing satellite tools (fire counts and burnt area) is illustrated by the work of Michel et al. [2005] on biomass burning inventory for Asia using SPOT (Satellite Pour l'Observation de la Terre) imagery for burnt areas. The fire spatial and geographical distribution they obtain is totally different than that of Streets et al. [2003] derived from fire pixel count maps. Provided that current problems with detection algorithms can be solved, the uncertainty of such products might be reduced to less than 15% for burnt areas. Uncertainty now resides primarily on the estimates of fire properties (biomass density and combustion efficiency) and further improvement will necessitate accounting for regional differences and relating in a common effort local measurements, remote sensing techniques and fuel models.

Emission factors for BC and PPOC (or PPOM) particles depend on the nature of the fuel used and on the physical conditions of the combustion processes. Thus for a given fuel, the type of use (industrial vs. domestic) and combustion technology will be of high importance. The level of development of each country and the regulating standards they apply are also crucial. For biomass burning, BC and PPOC emission factors have been reviewed based on a synthesis of the most recent experiments and existing papers [Andreae and Merlet, 2001].

For fossil fuels and biofuels, problems originate from the complexity, heterogeneity, and uneven coverage of available information. To cope with this difficulty, global inventories have been then constructed following a standard procedure that may apply to all activities in all countries and is flexible enough to be easily improved [Cooke et al., 1999; Junker and
Liousse, 2005]. The detailed sources were reduced to three main sectors of activities (traffic, industrial and domestic) and three levels of country developments (developed, semi-developed and developing) based on national gross income. This yields nine emission factors, each with a specific BC/PPOC ratio, for each fuel category for BC and primary POC. In an alternative approach, however, the exhaustive description of fuel combustion activities, technology divisions and emission controls (mostly for developed regions) has been preserved [Bond et al., 2004]. But numerous assumptions had to be made for regions where combustion technologies and emission controls were poorly known.

The few existing BC and PPOC emission inventories show significantly different budgets and different regional partitioning of the emissions. Most of them do not account for all sources (industrial versus biomass burning) or are restricted to one component of the aerosols (BC or primary OC only). Furthermore, the use of either PPOC or POM (including the SOA fraction) is often confusing.

These disagreements persist in the recent existing comprehensive fossil fuel and biofuel (BC and PPOC) global emission inventories by Bond et al. [2004] (B04), Liousse and coworkers [Liousse et al., 2004; Junker and Liousse, 2005] (L05) and Ito and Penner [2005] (I05) (Table 3-1). We find that differences sometimes exceeding a factor of 2 exist between the various estimates. The B04 and I05 emissions are significantly lower than the L05 values, although new sources were added that had not been taken into account before (such as off-road vehicle emissions and waste combustion). An important source of disagreement between these two studies probably arises from the choice of significantly different emission factors (EFs) (Table 3-2). I05 in most cases used the B04 EFs, which also supports differences in EFs as the main cause of discrepancies.
Table 3-1. Global emission estimates for black carbon and primary particulate organic carbon (Primary POC) from fossil fuel, biofuel, and open biomass burning.

<table>
<thead>
<tr>
<th></th>
<th>Black Carbon</th>
<th>Primary POC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tg C a⁻¹</td>
<td>Tg C a⁻¹</td>
</tr>
<tr>
<td>Fossil Fuel</td>
<td>Biofuel</td>
<td>Open Burning</td>
</tr>
<tr>
<td>Liousse et al. [2004]</td>
<td>5.9</td>
<td>2.9</td>
</tr>
<tr>
<td>Juncker &amp; Liousse [2005]</td>
<td>4.3¹</td>
<td>6.4</td>
</tr>
<tr>
<td>J05</td>
<td>6.4</td>
<td>8.7</td>
</tr>
<tr>
<td>Bond et al. [2004]</td>
<td>3.0</td>
<td>1.6</td>
</tr>
<tr>
<td>I05</td>
<td>2.4</td>
<td>6.5</td>
</tr>
<tr>
<td>Ito &amp; Penner [2005]</td>
<td>2.8</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>3.5</td>
<td>2.4⁴</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.3⁴</td>
</tr>
</tbody>
</table>

¹) Including savanna, open agricultural fires and tropical fires [Liousse et al., 1996b] and extratropical fires [Lavoué et al., 2000].
²) Primary particulate organic carbon only, including savanna, open agricultural fires and tropical fires [Liousse et al., 1996b] and extratropical fires [Lavoué et al., 2000].
⁴) Using a conversion factor of 1.3 for PPOC=>PPOM.

Table 3-2. Black carbon emission factors for the main fossil-fuel-related sources.

<table>
<thead>
<tr>
<th></th>
<th>Black Carbon Emission Factor [g C per kg fuel]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[Bond et al., 2004]</td>
</tr>
<tr>
<td>Developed Countries</td>
<td></td>
</tr>
<tr>
<td>Industrial, Coal</td>
<td>0.001-0.006</td>
</tr>
<tr>
<td>Industrial, Lignite</td>
<td>0.015-0.03</td>
</tr>
<tr>
<td>Traffic, Diesel</td>
<td>0.85</td>
</tr>
<tr>
<td>Traffic, Gasoline</td>
<td>0.04</td>
</tr>
<tr>
<td>Developing Countries</td>
<td></td>
</tr>
<tr>
<td>Industrial Coal</td>
<td>0.28-4.5 (1.12)</td>
</tr>
<tr>
<td>Industrial Lignite</td>
<td>0.09</td>
</tr>
<tr>
<td>Traffic Diesel</td>
<td>2-7</td>
</tr>
<tr>
<td>Traffic Gasoline</td>
<td>0.14-0.6</td>
</tr>
</tbody>
</table>

To assess the role of the choice of EFs in the discrepancy, a calculation was conducted for the year 1996, with the B04 and L05 sets of EFs, but using the same UN fuel database and the same methodology with the 9 BC EFs per fuel. A comparison of the results shows about...
10-fold larger emissions from lignite burning, 6-fold larger emissions from Diesel use, and about 25% greater emissions from hard coal burning when the L05 EFs are used instead of the B04 values. Recent combustion chamber experiments (not yet published) have shown that for Diesel emissions, the EF(BC) value of 2 for developed countries used by *Junker and Lioussse* (2005) would be in the highest part of the acceptable range, whereas on the other hand for coal industrial combustion, B04’s EF(BC) value of 0.0024 would be too low. There is indeed a critical need for further targeted experiments and for an effort towards consistency between the authors of these two studies to avoid the numerous existing or forthcoming discrepancies.

A way to test the validity of these inventories is to apply them in a global model and to compare the modeled values with existing BC and POM measurements. Reasonable agreement can be obtained for surface BC and PPOM concentrations using the L05 emissions implemented in the TM-3 model (3-D Transport Model; Velders et al., 1994). As the BC global burden estimated by B04 is almost half that of L05 it might be expected that comparisons between model and measurements would be less satisfactory. At this point, a recommendation for global inventories for fossil fuel and biofuel combustion emissions of BC and PPOC would be to use the *Bond and Lioussse* values respectively as the low and the high part of the range. Further improvements may come from simultaneous refinements in bottom-up estimates and combination with top-down approaches using distributions of emitted substances detected by remote sensing, such as in the recent study by *Ito and Penner* [2005].

Estimation of the source strength of the predominantly carbonaceous CCN from biomass and fossil fuel burning entails large uncertainties. In the case of biomass burning, the emission ratio of CN (Condensation Nuclei: particles larger than about 10 nm) to CO is a relatively robust parameter, with a characteristic value of about $25 \pm 10$ cm$^{-3}$ ppb$^{-1}$ [*Guyon et al.,* 2005], corresponding to about $2 \cdot 10^{13}$ particles emitted per gram CO. Of these particles, the fraction that can be activated at 1% SS (supersaturation) is very high, with most studies showing CCN/CN ratios of 60-100% for fresh or aged biomass smoke [*Radke et al.,* 1988; *Hudson et al.*, 1991; *Rogers et al.*, 1991; *Novakov and Corrigan*, 1996; *Ross et al.*, 2003]. Assuming an annual global emission of 700 Tg CO from biomass burning [*Andreae and Merlet*, 2001] and a CCN$_{1\%}$/CN ratio of 0.8 yields an estimate of $1.1 \cdot 10^{28}$ CCN$_{1\%}$ for the annual emission from biomass burning. An alternative calculation, based on CCN (at 1% SS) with an emission
factor of $2 \times 10^{15}$ per kg dry biomass, yields similar results. Estimates of CCN production from in-situ measurements yield production rates of $6 \times 10^{10}$ to $8 \times 10^{12}$ CCN per gram of material burned [Warner and Twomey, 1967; Eagan et al. 1974b].

Comparable emission factors are not available for fossil fuel combustion. Bond et al. [2002] have shown that residential coal burning produces $12 \cdot 10^{15}$ particles/kg for lignite (mostly $>100$ nm) and $14 \cdot 10^{15}$ for bituminous coal (mostly $<100$ nm). Given that residential burning tends to have considerably larger emission factors than industrial or power related combustion, the overall CCN emission factors for fossil fuel burning are likely to be considerably lower. Carbonaceous aerosol emissions are projected to decline over the 21st century to about two-thirds of their present values by 2050, but most likely with a significant rise in the BC/OC ratio in industrialized regions [Streets et al., 2004].

3.2.3.3 Secondary Organic Aerosols

As mentioned before, organic aerosols are comprised of primary and secondary particles. In the previous paragraphs, we have considered emission of primary particles (PPOC) only. The treatment of SOA in emission inventories of carbonaceous particles is still an open question. There is no evidence for the direct formation of SOA particles by nucleation of organics from the gas phase in the present-day atmosphere. SOA-rich particles have been shown to form by condensation of organic gases with low vapor pressures on pre-existing particles, which may be as small as freshly nucleated sulfate particles in the ultrafine size class [Kerminen et al., 2000; Kulmala et al., 2000; Kerminen, 2001; O'Dowd et al., 2002; Lihavainen et al., 2003; Maria et al., 2004; Zhang et al., 2004a]. The resulting particle may consist almost completely of organic material, and the presence of the initial sulfate nucleus may not be readily detectable any more. This is most likely the case in environments such as the Amazon Basin, where nucleation events have never been observed in spite of high concentrations of VOC and high rates of photochemical oxidation [Rissler et al., 2004]. Here, significant amounts of primary biogenic and other particles act as an efficient sink for condensable organics. On the other hand, in the presence of high concentrations of $\text{SO}_2$ and a high rate of $\text{H}_2\text{SO}_4$ production, new particle formation and growth by incorporation of organics occurs even at a high concentration of pre-existing aerosols, such as in the cities of Milan and Pittsburg [Baltensperger et al., 2002; Zhang et al., 2004a]. Once incorporated into the condensed phase, organics can be made less volatile by oxidation and polymerization.
reactions [Gelencsér et al., 2002; Limbeck et al., 2003; Claeys et al., 2004; Gao et al., 2004; Kalberer et al., 2004].

A number of studies has been conducted in recent years to assess the magnitude of SOA formation from biogenic and anthropogenic precursors [Liouss et al., 1996a; Griffin et al., 1999; Kanakidou et al., 2000; Chung and Seinfeld, 2002; Kanakidou et al., 2005; Tsigaridis et al., 2005]. In spite of these efforts, the range of estimates remains huge: 2.5-79 Tg yr\(^{-1}\) for biogenic SOA and 0.05-2.6 Tg yr\(^{-1}\) for SOA from anthropogenic precursors. Ignoring the more extreme results of sensitivity studies, we can use the values from the review of Kanakidou et al. [2005] as a currently best guess: 19, 15, and 2 Tg yr\(^{-1}\) from terpenes, oxygenated VOC (OVOC) and anthropogenic VOC, respectively. There is no published estimate of the number of CCN produced in the atmosphere as a result of SOA formation from biogenic or anthropogenic precursors. Recently, Heald et al. (2005) have pointed out a large discrepancy between model-predicted and observed concentrations of organic aerosols in the upper troposphere. They suggested that a substantial amount of SOA production, presumably from anthropogenic precursors, is required to explain the observed levels of organic aerosols in the free troposphere. If it is mixed downward, this aerosol could even make a significant contribution to the POM concentration in the lower troposphere.

### 3.2.4 Other Primary Anthropogenic Aerosols (Industrial Dust, etc.)

This category includes aerosols from transportation (e.g., tire and brake detritus, road dust, etc.), coal combustion (fly ash, etc.), cement manufacturing, metallurgical industries, and waste incineration, but excludes carbonaceous aerosols to avoid double counting. Primary anthropogenic aerosols are prominent in uncontrolled emissions from old industrial and energy technology, but are now being controlled fairly tightly, especially in developed countries. Growing industrialization, especially in Asia, was expected to lead to significant increases of primary anthropogenic aerosols in some regions. On the other hand, increased awareness of the negative health impacts of aerosols in developing countries, coupled with readily available emission control technology, may lead to less significant growth in these emissions than had been previously anticipated [Wolf and Hidy, 1997]. To our knowledge, there are no recently updated global emission estimates for this aerosol type, so that the very rough estimates given in IPCC-TAR had to be reused in Table 3-3. Furthermore, no data are
available on which a reliable estimate of the number source flux for this particle type could be based. The mass size distribution of these aerosols is dominated by the coarse mode, but a substantial submicron mode is also present, which contains significant amounts of soluble material [Kleeman and Cass, 1998]. Most of the primary particles from coal combustion are below 1 µm in diameter [Chen et al., 2004]. (A fraction of these particles is carbonaceous, and therefore is represented in the category of primary carbonaceous aerosols below.) In Beijing, for example, primary aerosols have been estimated to account for more than half of PM2.5 (Particulate Matter in the size class with diameters <2.5 µm) [Zhang et al., 2004b]. We conclude that primary anthropogenic particles could represent a significant source of CCN (and possibly IN), especially in polluted regions of the developing world.

### 3.2.5 Sulfates

Knowledge of the temporal and spatial details of anthropogenic and natural sulfur emissions is vital for estimating the atmospheric aerosol loadings of sulfate aerosols. The substantial growth of industrialization and energy production since the 1950s has resulted in increased burning of fossil fuels, causing extensive emissions of SO₂. Growing emissions are additionally driven by an increase in population. On the other hand local air quality concerns, structural change in energy supply and end-use, advanced technology and cleaning procedures have dramatically reduced air pollution, including SO₂ emissions, in some regions. Regional changes in emission pattern are additionally driven by a shift of strong polluting industry from the developed to the developing countries.
Table 3-3. Primary particle emissions and burdens for the year 2000.

<table>
<thead>
<tr>
<th></th>
<th>Mass Emission</th>
<th>Low</th>
<th>High</th>
<th>Mass Burden</th>
<th>Number Prod.</th>
<th>Number Burden</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tg a⁻¹</td>
<td>Tg</td>
<td>a⁻¹</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Carbonaceous aerosols</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organic Matter (0-2 µm)</td>
<td>95</td>
<td>40</td>
<td>150</td>
<td>1.2</td>
<td>-</td>
<td>310·10²⁴</td>
</tr>
<tr>
<td>Biomass burning</td>
<td>54</td>
<td>26</td>
<td>70</td>
<td>-</td>
<td>7·10²⁷</td>
<td>-</td>
</tr>
<tr>
<td>Fossil fuel</td>
<td>4</td>
<td>3</td>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Biogenic</td>
<td>35</td>
<td>15</td>
<td>70</td>
<td>0.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Black Carbon (0-2 µm)</td>
<td>10</td>
<td>8</td>
<td>14</td>
<td>0.1</td>
<td>-</td>
<td>270·10²⁴</td>
</tr>
<tr>
<td>Open burning &amp; biofuel</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Fossil fuel</td>
<td>4.5</td>
<td>3</td>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Industrial Dust, etc.</strong></td>
<td>100</td>
<td>40</td>
<td>130</td>
<td>1.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Sea Salt</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d&lt; 1 µm</td>
<td>180</td>
<td>60</td>
<td>500</td>
<td>3.5</td>
<td>7.4·10²⁶</td>
<td>-</td>
</tr>
<tr>
<td>d=1-16 µm</td>
<td>9940</td>
<td>3000</td>
<td>20,000</td>
<td>12</td>
<td>4.6·10²⁶</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>10,130</td>
<td>3000</td>
<td>20,000</td>
<td>15</td>
<td>1.2·10²⁷</td>
<td>27·10²⁴</td>
</tr>
<tr>
<td><strong>Mineral (Soil) Dust</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 µm</td>
<td>165</td>
<td>-</td>
<td>-</td>
<td>4.7</td>
<td>4.1·10²⁵</td>
<td>-</td>
</tr>
<tr>
<td>1-2.5 µm</td>
<td>496</td>
<td>-</td>
<td>-</td>
<td>12.5</td>
<td>9.6·10²⁵</td>
<td>-</td>
</tr>
<tr>
<td>2.5-10 µm</td>
<td>992</td>
<td>-</td>
<td>-</td>
<td>6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>1600</td>
<td>1000</td>
<td>2150</td>
<td>18±5</td>
<td>1.4·10²⁶</td>
<td>11·10²⁴</td>
</tr>
</tbody>
</table>

(a) Range reflects estimates reported in the literature. The actual range of uncertainty may encompass values both larger and smaller than those reported here. Values are based on the following publications: [Andreae, 1995; Andreae and Merlet, 2001; Guelle et al., 2001; Penner et al., 2001; Gong et al., 2002; Luo et al., 2003; Bond et al., 2004; Lioussse et al., 2004; Zender et al., 2004; Ito and Penner, 2005; Junker and Lioussse, 2005; Stier et al., 2005; Stier et al., 2006]

3.2.5.1 Sulfate Precursor Emissions

The most abundant sulfur gases in the atmosphere are carbonyl sulfide (OCS), DMS, sulfur dioxide (SO₂) and hydrogen sulfide (H₂S). They are all subject to oxidation by a variety of atmospheric species, including hydroxyl and nitrogen radicals, ozone, and excited oxygen atoms, and to photochemical decomposition [Seinfeld and Pandis, 1998]. The ultimate fate of these atmospheric sulfur compounds is the irreversible oxidation to sulfate (SO₄²⁻), which
forms particles mixed with water and ammonium (i.e., H$_2$SO$_4$, NH$_4$HSO$_4$, (NH$_4$)$_2$SO$_4$). A small fraction of DMS oxidizes to methanesulfonic acid (MSA), which condenses on existing particles. These atmospheric sulfur compounds have a variety of natural and anthropogenic sources, with anthropogenic fluxes dominating at present and into the foreseeable future [Dentener et al., 2005]. The most important natural sources of DMS and SO$_2$ are the marine biosphere and exhalation from volcanoes, respectively. Intra-annual variability and climate-induced changes of the source strength and distribution of natural sulfur sources are poorly known. Gabric et al. [2004] used model simulations of the period of equivalent CO$_2$ tripling (2080) to derive the changes in oceanic DMS production and flux to the atmosphere. They predicted a globally integrated DMS flux increase of 14%. The greatest perturbation to DMS flux is simulated at high latitudes in both hemispheres, with little change predicted in the tropics and sub-tropics. The largest change in annual integrated flux is simulated in the Southern Hemisphere between 50°S and 60°S. At this latitude, the DMS flux perturbation is most influenced by the GCM-simulated changes in the mixed layer depth. This indicates that future increases in stratification in the polar oceans will play a critical role in the DMS cycle and therefore the flux of sulfate aerosols in remote ocean regions.

Major uncertainties exist for the volcanic sulfur emissions and current estimates are probably too low [Textor et al., 2004]. Volcanic emissions are however released at mountain peaks mostly above the planetary boundary layer where they have extended life times and disproportionate effects on cloud properties.

The major anthropogenic sulfur source is SO$_2$ emissions from the burning of sulfur-containing fuels, mainly coal. Metal smelting and other industrial processes also release significant quantities of SO$_2$. Global anthropogenic sulfur emissions are estimated to range between 65 and 90 Tg S in 1990 [Houghton et al., 2001]. Reviews of most recent inventories indicate a most likely value of 75±10 Tg S in 1990 [Smith et al., 2001], and 71 Tg S in the AEROCOM inventory for 2000 [Dentener et al., 2005]. Smith et al. [2001] estimate that 56% of 1990 world sulfur emissions are from coal, 24% from oil, 15% from industrial processes and 3% from biomass burning. Emission from bunker fuel burned by the international cargo and passenger ship fleet represents a contribution of 4-7% to the global anthropogenic SO$_2$ emissions (3.0 Tg S in 1996 [Endresen et al., 2003]; 3.7 in 2001 [Endresen et al., 2005]. The majority of these ship emissions occur in the Northern Hemisphere.
Globally, emissions have been roughly constant from 1980 to the present. However, a significant shift has occurred in the spatial distribution of emissions. Many climate studies to date have used the emissions pattern from the GEIA inventory for 1985 and simply scaled these backwards and forwards in time. This is clearly an oversimplification, as the regional pattern of emissions has changed dramatically over the last 30 years [Massie et al., 2004]. Different inventories are similar at the global-mean level, but show marked differences at the regional level.

While 60% of global emissions in 1980 were from around the North Atlantic basin, this region contributed less than 40% of the global total by 1995 and will contribute even less in the future [Smith et al., 2001]. As a result of the implementation of international emissions agreements, SO$_2$ emissions decreased 32% from 1980 to 1990 in Europe, and 3% from 1980 to 1995 in North America. In 1990, the spatial pattern of emission shows that the US, the USSR, and China were the main sulfur dioxide emitters (i.e., approximately 50% of the total). Historical data for Asia are more problematic, but the trends are clearly in the other direction. The data indicate a roughly 60% increase in sulfur dioxide emissions from 1980 to 1990. The two main reasons for this are the region's rapid economic growth and its high dependence on coal. Streets and coworkers have estimated that China’s SO$_2$ emissions grew 3.6% per year from 1985 through 1996, and India’s SO$_2$ emissions grew at a rate of 4.9% per year during the 1990s [Streets et al., 2000; Streets and Waldhoff, 2000]. This trend is confirmed by satellite data (TOMS) of aerosol optical depth (AOD). Large increases in AOD between 1979 and 2000 are evident over the China coastal plain and the Ganges River basin in India. For instance AOD increased by 17% per decade during winter over the China coastal plain [Massie et al., 2004].

In contrast, Carmichael et al. [2002] report that since 1995 Asian SO$_2$ emissions have declined from 38.5 Tg in 1995 to 34.4 Tg in 2000, a decrease of 2.3% per year. This remarkable change is almost entirely due to a reduction in emissions in China, which emits about two-thirds of SO$_2$ from Asia. This has been brought about by several factors: a marked reduction in industrial coal use from the closure of small and inefficient plants, a slow-down in the Chinese economy, the improved efficiency of energy use, the closure of some high-sulfur coal mines, a general reform in industry and power generation, and a rising awareness of the dangers of air pollution.
3.2.5.2 Transformation and Sulfate Particle Formation

When emitted into the atmosphere, sulfur species undergo transport and chemical reactions to form compounds that condense on aerosol surfaces or nucleate. Wet deposition removes gases and particles in the atmosphere and deposits them on the Earth's surface by means of precipitation. With dry deposition, particles and gases reach land and water surfaces without precipitation. All the processes controlling dispersion, residence time in the atmosphere, aerosol load and composition vary with weather and climate. Since sulfur emissions themselves influence climate, and so change the transport and sink processes, one cannot accurately model the influences of anthropogenic emissions unless the climate influences are also modeled in a consistent way. While this clearly requires an appropriate coupled climate-sulfur-chemistry model, the primary driving force must be reliable, spatially-resolved SO$_2$ emissions [Smith et al., 2001].

The percentage of SO$_2$ transformed into sulfate, according to simulations with global chemistry models, falls between 51 and 56% (four models) with one outlier calculating an efficiency of 83% (due to low dry deposition flux of sulfur dioxide) [Berglen et al., 2004]. The COSAM (Comparison of Large Scale Sulfate Aerosol Models) model comparison [Barrie et al., 2001; Roelofs et al., 2001] reflecting the state of modeling in 1999 and the AEROCOM [Kinne et al., 2006; Textor et al., 2005] comparison performed in 2004 give a range of transformation rates between 45 and 67%, and between 50 and 81%, respectively. This transformation rate depends on transport, sink processes and chemical transformation. The latter will change due to changes in the emission of oxidants and oxidant precursors. Only gas phase oxidation by OH will lead to formation of new particles, whereas aqueous phase oxidation and catalytic oxidation by metals take place on existing particles, thereby changing their CCN properties. Gas phase sulfate formation contributes between 10 and 22% to the total oxidation of SO$_2$ in the COSAM comparison and between 14 and 42% in the more recent AEROCOM comparison. In addition, a southward shift of the sulfur emissions toward regions with more incoming solar radiation and hence more OH would increase the fraction of SO$_2$ oxidized in the gaseous phase [Berglen et al., 2004].

Secondary particle formation from gas-phase sulfuric acid dominates the nucleation mode and contributes significantly to the particle number concentrations although contribution to total aerosol mass is small. Model calculations show that nucleation of sulfate particles is
favored in regions with little available aerosol surface area, low temperatures, and high relative humidity. Thus, the maxima of the nucleation mode number concentration can be found in the upper tropical troposphere and in the remote regions of the Antarctic. The soluble Aitken mode numbers are dominated by particles growing from the nucleation in the Aitken size range. Accumulation mode soluble numbers are highest in the lower troposphere between 30° S and 60° N close to the sources from biomass burning and fossil fuel use. Increased levels can also be found in the upper troposphere, attributable to convective detrainment of particles and their precursors and evaporation of cloud droplets and ice crystals. Figure 3-1 shows the effect of the omission of SO₂ emissions from fossil fuel and biofuel use on the accumulation mode number concentration. Between 0° and 60°N sulfur dioxide from anthropogenic sources contributes about 25 to 30% of the total particle number burden.

Conventionally, additivity is assumed for aerosol emissions and for their climate impacts. However, due to the complex interactions between transport, aerosol and hydrological cycle, one cannot expect a linear relationship between emissions and aerosol load or aerosol effects. Graf et al. [1997] have shown that the efficiency by which sulfur precursors are transformed to sulfate depends on the source type. However, it depends additionally on the location and the season of the release. Barth and Church [1999] used the NCAR Community Climate Model (CCM3) to determine the contributions of southeast China to the global aerosol burden. Southeast China emitted 11.6% of the global anthropogenic sulfur emissions and contributed 9% to the global sulfate burden in their simulations, indicating a non-linear response between emission and burden. When anthropogenic sulfur emissions were doubled, the aerosol burden contributed by China increased by a factor of 2.2. Their simulations show a gas-phase oxidation rate from south-east-Asian emissions above average. Sulfate formed via this pathway is less susceptible to wet scavenging and thus contributes more efficiently to the atmospheric sulfate load than sulfate formed in clouds, which is already incorporated in cloud droplets and thus is more likely to undergo wet removal.

Stier et al. [2006] analyzed the response of the global aerosol system to changes in anthropogenic sulfur emissions. In a scenario without anthropogenic sulfur emissions (corresponding to a reduction by 55%), the sulfate burden decreases globally by almost 50% and the accumulation mode number concentration by 21%. However, not only the sulfate budget is affected, but also annual global-mean life-times of black carbon, particulate organic
matter, and dust increase by 8.9%, 2.5%, and 1.5%, microphysical aging times increase by 164%, 84%, and 66%, and annual zonal-mean column burdens in high northern latitudes increase by up to 70%, 20%, and 20%, respectively. This emphasizes that changes in one particular anthropogenic emission have impacts on the cycles of other aerosol components, resulting in changes of their life-times, column burdens, and microphysical aging times. These results suggest that the separate treatment of the different aerosol components and their precursors can lead to non-negligible errors in estimates of the aerosol climate effects.

Concerning future emissions of sulfur, the SRES (IPCC Special Report on Emissions Scenarios) scenarios portray similar emission dynamics – at various future dates (between 2020–2030 and 2070, depending on the scenario and its underlying storyline), global SO$_2$ emissions reach a maximum level and decline thereafter. By 2030 sulfur emissions range between 40 and 160 Tg S, by 2070 between 20 and 165 Tg S, and by 2100 between 10 and 95 Tg S. Emission trajectories of the SRES scenarios reflect a combined impact of different scenario driving forces (local air quality concerns, structural change in energy supply and end-use, etc.), which lead to a gradual decline in sulfur emissions in the second half of the 21$^{st}$ century.

3.2.6 Nitrates

The production of secondary aerosol nitrate is closely tied to the relative abundances of ammonium and sulfate. If ammonia is available in excess of the amount required to neutralize the stronger acid H$_2$SO$_4$, gaseous HNO$_3$ and NH$_3$ can enter the condensed phase, and their subsequent dissociation yields nitrate (NO$_3^-$) and ammonium (NH$_4^+$) ions. Submicron particles containing nitrate are potentially efficient CCN [Metzger et al., 2002a]. In the presence of acidic accumulation-mode sulfuric-acid containing aerosols, however, HNO$_3$ deposits on larger, alkaline mineral or salt particles [Dentener et al., 1996; Murphy and Thomson, 1997; Gard et al., 1998]. This deposition of soluble material on otherwise insoluble dust particles increases their CCN activity.

Until recently, nitrate has not been considered in assessments of the climatic effects of aerosols, and even current models often ignore the role of nitrate [e.g., Stier et al., 2005] or the sub-micron nitrate aerosol fraction [Derwent et al., 2003]. Andreae [1995] estimated that
the global burden of ammonium nitrate aerosol from natural and anthropogenic sources is 0.24 and 0.4 Tg (as NH$_4$NO$_3$), respectively, and that anthropogenic nitrates cause only 2% of the total direct forcing. Adams et al. [1999] obtained a value of only 0.17 Tg (as NO$_3^-$) for the global nitrate burden, which may be due to the fact that their model did not include nitrate deposition on sea salt aerosols. Their most recent estimate [Adams et al., 2001], predicts a nitrate burden of 0.38 Tg in the form of secondary ammonium nitrate/sulfate particles.

The importance of aerosol nitrate is likely to increase substantially over the next century. For example, the SRES A2 emissions scenario projects that NO$_x$ emissions will more than triple in that time period while SO$_2$ emissions decline slightly. Using this scenario, Adams et al. [2001] predict a nitrate burden of 1.8 Tg for 2100, compared to a sulfate burden of 2.2 Tg for the same year. In their model run for 2100, the total anthropogenic forcing (direct and indirect) associated with nitrate actually exceeds that caused by sulfate.

### 3.3 Aerosol Number Fluxes

The calculation of the aerosol impacts on clouds, precipitation and hence on climate, and their interaction with the hydrological cycle requires size-resolved knowledge of the global distribution of aerosol number concentrations. Because primary aerosol emissions result in a direct increase of the atmospheric particle number burden, whereas secondary aerosol production often results only in the addition of mass to existing particles, the fluxes of primary aerosols play a key role in the context of cloud physical effects [Adams and Seinfeld, 2003]. Recently, the aerosol size distribution has been added to the prognostic variable space of global microphysical aerosol models to account for these effects. In addition to the emission mass fluxes, these models require as input the primary aerosol number fluxes and therefore the knowledge of the emission size distribution.

Natural sea salt and mineral dust are predominantly emitted in the accumulation and coarse mode size range and their size distribution is mainly altered by size-dependent sink processes. A number of size-resolved emission parameterizations are available for sea salt and mineral dust (see above), and can be interactively calculated in numerical models. These parameterizations can also be used in future scenarios.
However, for primary emissions from combustion processes, the initial evolution of the size distribution is characterized by small-scale microphysical processes [Jacobson and Seinfeld, 2004] that cannot be resolved in regional and global aerosol models. Therefore, the sub-grid scale evolution of the size-distribution and mixing state shortly after emission should be parameterized in a way that is appropriate to the model scale. Up to now, such parameterizations are not available and the emission size-distribution is generally estimated from measurements made on “slightly” aged plumes. For the AEROCOM (Aerosol Model Intercomparison Initiative) aerosol model inter-comparison, an emission inventory has been provided [Dentener et al., 2005], based on mass flux estimates by Bond et al. [2004] for carbonaceous aerosols from fossil- and bio-fuel use, by Van der Werf et al. [2003] for vegetation fires, and by Cofala et al. [2005] for anthropogenic sulfur emissions (data and documentation available from ftp://ftp.ei.jrc.it/pub/Aerocom/). AEROCOM assumes that 2.5% of the total SO\(_2\) emissions are emitted in the form of primary sulfate to account for sub-grid scale sulfate formation at the sources. In addition, source-specific size distributions are proposed (Table 3-4) for freshly emitted primary aerosols, which are then subject to further microphysical processing in global aerosol models. For comparison, the parameters of a log-normal fit [Stier et al., 2006] of natural sea-salt [Schulz et al., 2004] and mineral dust [Tegen et al., 2002; Tegen et al., 2004] as used in the emission parameterizations, neglecting the super-coarse mode emissions are also given in Table 3-4.

The total emission number flux \( F_N \) of primary particles can be calculated from a given emission mass flux \( F_M \) and a log-normal emission size-distribution as follows:

\[
F_N = F_M \frac{3}{4\pi \rho} \left( r_{NMR} \exp(1.5 \ln^2(\sigma)) \right)^3
\]

where \( \rho \) is the density of the compound and \( r_{NMR} \) the number median radius and \( \sigma \) the standard deviation of the log-normal emission size-distribution.

Based on the AEROCOM anthropogenic emission inventory and the proposed emission size distributions, the total anthropogenic primary aerosol number flux can be estimated as listed in Table 3-4. This flux is highly dependent on the choice of the emission size distribution. For example, the usage of the same emission mass flux for primary sulfate, but with the significantly smaller emission size-distributions used by Adams and Seinfeld [2003] results in a 6\( \times 10^4 \) larger number flux than using the large “fly ash” AEROCOM size distribution.
recommendation. This underlines the necessity of adequate emission size distributions and for careful adaptation to the respective model scales.

Table 3-4. *Global annual anthropogenic primary aerosol number fluxes as recommended by the AEROCOM emission inventory. For comparison, values for the natural emissions of sea salt [Schulz et al., 2004] and mineral dust [Tegen et al., 2002; Tegen et al., 2004] are also given.*

<table>
<thead>
<tr>
<th>Source</th>
<th>Number Median Diameter (µm)</th>
<th>Standard Deviation</th>
<th>Total Mass Flux (Tg a⁻¹)</th>
<th>Total Primary Number Flux (a⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fossil-Fuel Burning (BC+POM)</td>
<td>0.030</td>
<td>1.8</td>
<td>6.4¹</td>
<td>4.8 x 10²⁸</td>
</tr>
<tr>
<td>Bio-Fuel Burning (BC+POM)</td>
<td>0.08</td>
<td>1.8</td>
<td>10.7¹</td>
<td>4.2 x 10²⁷</td>
</tr>
<tr>
<td>Vegetation Fires (BC+POM)</td>
<td>0.08</td>
<td>1.8</td>
<td>37.7²</td>
<td>1.5 x 10²⁸</td>
</tr>
<tr>
<td>Primary Sulfate (H₂SO₄)</td>
<td>1.0</td>
<td>2.0</td>
<td>4.3³</td>
<td>5.1 x 10²³</td>
</tr>
<tr>
<td>Sea Salt (Accumulation Mode)</td>
<td>≈0.4</td>
<td>varying with wind-speed</td>
<td>1.6</td>
<td>7.4 x 10²⁶</td>
</tr>
<tr>
<td>Sea Salt (Coarse Mode)</td>
<td>≈4.4</td>
<td>varying with wind-speed</td>
<td>2.0</td>
<td>4.6 x 10²⁶</td>
</tr>
<tr>
<td>Mineral Dust (Accumulation Mode)</td>
<td>0.4</td>
<td>1.6</td>
<td>7.5</td>
<td>4.1 x 10²⁵</td>
</tr>
<tr>
<td>Mineral Dust (Coarse Mode)</td>
<td>0.8</td>
<td>2.0</td>
<td>655</td>
<td>9.6 x 10²⁵</td>
</tr>
</tbody>
</table>

¹ [Bond et al., 2004]
² [van der Werf et al., 2003]
³ Industry, fossil- and bio-fuels from Cofala et al. [2005] and vegetation fires from van der Werf et al. [2003] with 2.5% of SO₂ emitted as primary sulfate.

3.4 Physical and Chemical Nature of Atmospheric Aerosols

Knowledge of the various primary and secondary aerosols sources, combined with the size and composition dependent aerosol sinks, in principle allows prediction of the chemical and physical characteristics of the atmospheric aerosol. The properties most relevant to the impact of aerosols on precipitation are the size distribution and composition. They determine the interaction of the aerosol with water or, more specifically, their activity as either CCN or IN.
Before delving into these aerosols attributes, we discuss the concept of different aerosol regimes. We shall use a series of aerosol “regimes”, as they may be found at various locales around the Earth: over the remote oceans and continents far from pollution sources (“remote marine” and “remote continental”, respectively), over regions influenced by industrial pollution or biomass burning (“polluted continental” and “polluted marine”), and over areas dominated by mineral dust aerosols. Earlier classifications have attempted to separate the atmosphere into “marine” (equated with low pollutant levels and low CCN concentrations) and “continental” (high pollutant levels and high CCN concentrations) regimes. This view has been challenged, as it has become apparent that unpolluted continental regions may be just as clean, and have just as low CCN concentrations as remote marine areas [Twomey et al., 1978; Squires and Twomey, 1966; Delene and Deshler, 2001; Hoppel et al. 1973; Komppula et al., 2005] [Roberts et al., 2001; Andreae et al., 2004]. Conversely, polluted marine areas have aerosol concentrations just as high as polluted continental regions [Hudson and Xie, 1999; Hudson and Yum, 2002][Andreae et al., 2000; Raes et al., 2000]. In this chapter, we will limit the discussion of the nature of aerosols to an analysis of generic properties within broad generic regimes; a further breakdown into regional characteristics is provided in Chapter 4.

3.4.1 Size Distributions

While many different types of distribution functions have been utilized over the years to represent the atmospheric aerosol, the discussion here is in terms of the ubiquitous lognormal function introduced and popularized by Junge and later Whitby and his colleagues [cf., Junge, 1953; Junge, 1963; Whitby et al., 1972; Whitby, 1978]. The distribution function itself can be expressed in terms of any of its moments, the most useful representation being dependent on the particular issue being addressed. For example, for exploring the dependence of cloud drop number concentration on aerosol concentration, the number size distribution (Figure 3-1) has the most utility, whereas if one is considering the chemical nature of the aerosol, the mass distribution is commonly employed.
3.4.1.1. The remote marine aerosol

This category includes both the seasalt aerosol and particulate oxidation products of biogenic DMS [Raemdonck et al., 1986; Charlson et al., 1987; Andreae and Crutzen, 1997; Bates et al., 1998; Andreae et al., 1999; Bates et al., 2000; Bates et al., 2001]. While there is substantial variability [cf, Heintzenberg and Covert, 2000], the most salient features are ubiquitous. Perhaps most characteristic, there is a prominent mode in the number distribution in the size range from about 0.1 to 0.2 µm diameter, which contains most of the DMS-derived non-seasalt sulfate (the sulfate which is present in marine aerosols in excess of what originates from the sulfate in seasalt). This mode corresponds to the accumulation mode and is virtually global in extent [e.g., Kim et al., 1995; Quinn et al., 1995; Weber et al., 1998; Bates et al., 2002]. Hoppel and his colleagues have argued persuasively that its location and prominence are attributable to cloud processing of aerosol [Hoppel et al., 1986; Hoppel et al.,

![Figure 3-1: Typical particle number size distribution. The shaded band represents the range of sizes activated as CCN at 0.3% SS, a typical median SS in clouds.](image_url)
A second, smaller diameter, mode commonly occurs around 30-50 nm, particularly in the most remote locales. It has been variously referred to, most frequently as the aged-nucleation or Aitken mode, and is thought to arise from in-situ particle nucleation associated with gas to particle conversion under conditions of low pre-existing aerosol surface area [Wiedensohler and Covert, 1996; Brechtel et al., 1998; Covert et al., 1998; Collins et al., 2000]. The concentration of particles in these submicron modes ranges from a few tens per cm$^3$ to several hundred per cm$^3$ with concentrations in the Northern Hemispheric oceans generally (though not always) appreciably higher than those in the Southern Hemisphere oceans. Finally, a larger, or coarse, size mode, the so-called sea salt mode, is nearly always present (~ 0.5-3 µm diameter). While few in number, the contribution of this mode to particle mass – and thus the volume distribution – is always very substantial and commonly dominant [Bates et al., 2002; Maring et al., 2003]. These particles are not only sea salt, but contain also mineral dust that can be transported over very long distances and can have a marked presence even in remote marine air [e.g., Andreae et al., 1986; e.g., Arimoto et al., 1997]. Non-seasalt sulfate is also always present in variable amounts in the coarse mode of the marine aerosol [Andreae et al., 1999; Andreae et al., 2000]. Furthermore, the impact of even small numbers of large particles to marine precipitation can be quite substantial [e.g., Feingold et al., 1999]. Some representative size distributions, taken from Bates et al. [2002], are shown in Figure 3-2 and illustrate the features just discussed.
Figure 3-2. Number (heavy line) and volume (light line) distributions at 55% RH for different air masses of origin during the first 2 ACE experiments. The vertical bars represent one standard deviation in mean number or volume in the size bin over the averaging period (23-114 hr). From Bates et al. [2002].

3.4.1.2 The remote continental aerosol

Because of the vast amounts of anthropogenic emissions of aerosol particles and their gaseous precursors, combined with efficient long-range transport, it is nowadays very difficult to find areas of the Earth that are not measurably impacted by pollutant aerosols. This applies especially to the continental regions of the northern hemisphere, where most of the human activities are concentrated.

Aerosol particles have typical atmospheric lifetimes of 3-10 days, which implies that even after 10-30 days (3 lifetimes) about 5% of the initial burden is still left. Given that airmasses can easily travel several 1000 km in 15 days, and that, at least in the northern hemisphere, few places are more than a few 1000 km from major pollution sources, there are really no places where we can expect to find pristine conditions in the northern hemisphere (see
Chapter 4). The remote continental aerosol nowadays consists of some natural material (dust, biogenic materials, etc.), mixed with pollution aerosols at varying levels of dilution. For example, a study of the European aerosol climatology shows that the average PM2.5 concentration ranges from some 30 $\mu$g m$^{-3}$ at urban sites to about 4 $\mu$g m$^{-3}$ at the cleanest site, Sevettijärvi in northern Finland [Putaud et al., 2004; Van Dingenen et al., 2004]. The composition of the aerosol at this remote site is, however, very similar to that at urban and regionally polluted sites, with a large fraction of non-seasalt sulfate, and about 4% of the combustion tracer black carbon. Pollution aerosols dominate even in most continental areas in the Southern Hemisphere, especially in the dry seasons, when biomass burning is widespread [Artaxo et al., 1988; Artaxo et al., 1994; Artaxo et al., 2002; Sinha et al., 2003]. Only under relatively rare circumstances do aerosol concentrations over the remote continents approach pristine conditions. This usually occurs when clean marine airmasses are advected over practically uninhabited continental regions. Examples are the rainy season in central Amazonia, some of the desert regions in Australia and parts of the boreal regions in Eurasia and North America. Unfortunately, there are very few studies that document these conditions carefully enough to be useful for a assessment of what aerosol concentrations (and composition) would be like in the absence of anthropogenic emissions. Size distributions from a variety of relatively unpolluted venues such as rural Canada [Leaitch and Isaac, 1991], the Southwestern U.S. and rural Midwestern U.S. [Kim et al., 1993a], and the Amazon Basin [Guyon et al., 2003a; Rissler et al., 2004; Krejci et al., 2005] and the high altitude site Jungfraujoch (3580 m asl) [Weingartner et al., 1999] are differentiable from more strongly polluted air and share some common characteristics. The mass accumulation mode of the remote continental aerosol tends to center in the size range from 0.1 to 0.4 $\mu$m diameter, thus showing more size variability than the remote marine aerosol, particularly on the upper bound, though tending to a slightly smaller modal size. In number concentration, the remote continental aerosol also shows more variability, but with concentrations usually appreciably higher than those of its marine counterpart, typically from a few hundred to a few thousand particles per cm$^3$ [Weingartner et al., 1999; Putaud et al., 2004; Van Dingenen et al., 2004]. This difference is the result of two factors – the presence of pollution aerosols and the frequent occurrence of large particle number concentrations in the nucleation and Aitken modes. In the few instances when nearly pristine conditions could be sampled, number concentrations in the continental tropics [Roberts et al., 2001; Rissler et al., 2004; Krejci et
al., 2005] as well as in the lower free troposphere during winter [Weingartner et al., 1999] were in the low hundreds, indistinguishable from marine values. From vegetated continental areas at northern mid-to-high latitudes, no published estimates of pristine levels exist, but concentrations are estimated to be in the tens to low hundreds per cm$^3$ (M. Kulmala, personal communication, 2005).

Besides the accumulation mode, separate Aitken or nucleation modes are frequently present, especially in the presence of low particle surface area and small, but significant concentrations of gaseous H$_2$SO$_4$ from SO$_2$ oxidation [Kulmala et al., 2000; Lihavainen et al., 2003]. Despite their high number concentrations, these particles are of little direct consequence for precipitation processes, as they do not have enough soluble mass to act as CCN. Only when they have grown to sizes approaching 100 nm, e.g., by condensation of terpene oxidation products [Leaitch et al., 1999][Kerminen et al., 2000], are they able to nucleate cloud droplets.

The supermicron, or coarse mode, of the remote continental aerosol is commonly centered in the range from 2 to 5 µm diameter (well away from desert sources of dust). The number concentration in this mode is more variable than its marine counterpart, but generally significantly lower, with typical concentrations in the range from 0.1 to 1 cm$^3$, as compared to 1 to 10 cm$^3$ in marine air. It consists of mineral and soil dust and a variety of primary biogenic particles, and dominates the aerosol mass size distribution under the cleanest conditions [Artaxo et al., 2002; Guyon et al., 2003b].

Because of the contrast between the size distributions of continental and marine remote aerosols, it must be noted that even the simple advection of continental air into the marine venue will alter the marine size distribution and chemical composition. This can be clearly seen in Figure 3-3, taken from Covert et al. [1998], which shows such impact for the Southern Ocean aerosol. For the submicron portion of the size distribution, the bi-modal structure characteristic of clean marine air is overwhelmed by a massive modal peak centered at 40-50 nm diameter. Similar changes have been reported for both the North and South Atlantic [Hoppel et al., 1990; Kim et al., 1995; Bates et al., 2000; Collins et al., 2000; Johnson et al., 2000], the Indian Ocean [Cantrell et al., 2000], the Pacific Ocean, [Moore et al. 2003], as well as by the review by Bates et al. [2002]. For the coarse particle mode,
advection of pollution generally has little effect unless soil dust is present [cf. Li-Jones and Prospero, 1998; cf. Bates et al., 2002, [Maring et al., 2003; Moore et al., 2003].

Figure 3-3. Differential aerosol number size distributions from Cape Grim during ACE 1. CGBS in the legend refers to Cape Grim Baseline Station. Curve e is an expanded scale for curve b to show an additional mode above 200 nm. Taken from Covert et al. [1998].

3.4.1.3 The polluted continental aerosol

Much of the early work on industrial air pollution was done in the highly polluted Los Angeles Basin. Studies there show typically a tri-modal number size distribution with prominent nucleation, accumulation and coarse modes, and associated modal diameters of 0.01-0.02, 0.1 and 1.0 µm, respectively [e.g., Heisler et al., 1973; Sverdrup et al., 1975; Whitby et al., 1975]. Later work confirmed this basic structure but showed that the nucleation and accumulation modes are not always fully resolved [e.g., John et al., 1990; Hering et al., 1997; Hughes et al., 1999]. In European pollution both the bi-modal submicron structure and a prominent though variable coarse mode have been commonly found [e.g., Birmili et al., 2001; Baltensperger et al., 2002; Petzold et al., 2002]. Number concentrations
vary in the range from thousands to tens of thousands per cm$^3$. The partition in number concentration between the nucleation and accumulation modes is highly variable with the larger concentration switching back and forth between the two as a function of time of day, as well as season and various other variables.

### 3.4.1.4 Biomass burning aerosol

Besides industrial pollution, biomass burning emissions are the other main pollution source on a global impact scale and occur mostly from the tropical regions [Crutzen and Andreae, 1990], especially Brazil and Southern Africa. Andreae et al. [1996], summarizing a number of studies in Africa, report a large accumulation in the number size distribution at $\sim 0.1-0.2$ µm diameter for in-plume measurements. Regional haze associated with biomass burning was shifted to slightly larger sizes. There is also a coarse mode in both the plume and haze aerosol, which is composed to a large extent of resuspended soil dust, with some additional coarse ash particles [Reid et al., 2005]. The number modal diameter is a bit unclear, in part because volume rather than number distributions are reported. Nevertheless, it would appear to be at least 1-2 µm in size. Le Canut et al. [1996] also report an accumulation mode centered at $\sim 0.15$ µm and a coarse mode centered at $\sim 3$ µm. More recently, from the SAFARI-2000 campaign, Haywood et al. [2003a; 2003b] reported an accumulation mode at a slightly larger volume modal diameter of 0.25 µm with a coarse mode at $\sim 10$ µm. In none of these studies was a distinguishable Aitken or nucleation mode reported.

Data on Brazilian biomass burning aerosol has recently increased sharply as a result of the SCAR-A and SCAR-B studies [Reid and Hobbs, 1998; Reid et al., 1998] and the Large Scale Biosphere-Atmosphere Experiment in Amazonia – Smoke, Aerosols, Clouds, Rainfall and Climate (LBA-SMOCC) [Andreae et al., 2004; Guyon et al., 2005]. Reid et al. [1998; 1998] have summarized work on both distinct fire plumes and the regional haze produced by such plumes. As in African emissions, there is no distinct nucleation mode in distinct plumes, but rather a single large accumulation number mode centered at 0.10-0.13 µm diameter, and a volume mode around 0.2-0.3 µm. In contrast, the regional haze associated with biomass burning showed an accumulation mode centered in the range 0.2-0.3 µm. In the fresh smoke, a coarse mode at $\sim 3$ µm was evident, while in the regional haze a significant coarse mode was not commonly present. Airborne measurements of fresh smoke from deforestation fires during the LBA-SMOCC campaign showed a single number mode near 0.10 µm, while in detrained smoke from pyrocumulus clouds the modal diameter had grown to about 0.13 µm.
At a ground site, where a mixture of fresh and aged smoke was present, the distribution was bimodal, with number modal diameters of 0.09±0.01 and 0.18±0.02 μm [Rissler et al., 2006]. A nucleation mode with a diameter near 10 nm appeared sporadically.

Irrespective of locale, the submicron aerosol of tropical biomass burning is usually present as a large accumulation mode with number median diameters in the range of 0.1 to 0.2 μm and volume median diameters of 0.2-0.3 μm [Reid et al., 2005], and with a distinct coarse mass mode centered around 4 μm diameter [e.g., Radke et al., 1991]. After some appreciable aging, the number mode shifts up to 0.2-0.3 μm. The coarse mode diminishes in aging smoke and is typically centered at a few micrometers in size. These characteristics contrast to some extent with the other major source of pollution, namely industrial emissions, where the accumulation mode and coarse mode are at somewhat smaller modal diameters and a distinct nucleation mode is frequently encountered.

### 3.4.2 Size-Dependent Chemical Composition and Solubility

#### 3.4.2.1 The marine aerosol

The chemical composition of marine aerosols varies substantially with geographic locale. A review by Heintzenberg et al. [2000] finds that anthropogenic pollution is ubiquitous over the oceans of the Northern Hemisphere due to long-range transport. The composition of the undisturbed remote marine aerosol is most easily assessed in data from the Southern Hemisphere. Sea salt is the single most common component of the aerosol with latitudinal maxima in both number concentration and mass in the mid-latitudes, where strong winds associated with baroclinic systems predominate. The number size distribution of the sea salt aerosol extends well down into the sub-micron range and indeed commonly peaks there [e.g., McInnes et al., 1997; O'Dowd et al., 1997; Kreidenweis et al., 1998; Murphy et al., 1998a; Campuzano-Jost et al., 2003]. It had long been thought that most remote marine sub-micron particles are primarily comprised of sulfate derived from the oxidation of dimethyl sulfide (DMS) emitted from the ocean surface and SO₂ from long-range transport of continental sources and ship emissions [Twomey, 1971; Charlson et al., 1987; Andreae et al., 1988; Andreae, 1990; McInnes et al., 1997; Andreae et al., 2000]. However, numerous recent studies have revealed a substantial organic presence even in remote marine air [Middlebrook et al., 1998; Murphy et al., 1998b], both in coarse mode and accumulation mode particles.
McInnes et al., 1997; Matthias-Maser et al., 1999; Guazzotti et al., 2001; Mayol-Bracero et al., 2001; Cavalli et al., 2004][O’Dowd et al., 2004]. The speciation of the organics is still not entirely resolved but there are clearly substantial amounts of both soluble and insoluble carbon present.

While sea salt contributions to the marine aerosol mass are governed primarily by wind speed [Lewis and Schwartz, 2004] (see Chapter 4), the contribution of organics and inorganics such as sulfate and nitrate are strongly a function of geography, induced largely by variability in the proportions of biogenic and anthropogenic emissions. The loadings of sulfate and carbon species increase dramatically in polluted marine air [Cachier et al., 1983; Rau and Khalil, 1993; Andreae et al., 2000; Heintzenberg and Covert, 2000; Putaud et al., 2000; Guazzotti et al., 2001]. However, the dominant species depends on the source of the pollution. For example, pollution from either Europe or North America advected over the North Atlantic increases sulfate much more than the various carbon species, resulting in a marked increase in the inorganic mass fraction [Andreae et al., 2000; Putaud et al., 2000]. In contrast, pollution from India transported over the Indian Ocean results in more organics and especially black carbon [Cantrell et al., 2000; Chowdhury et al., 2001; Mayol-Bracero et al., 2002a]. While both the carbon and sulfate species are predominately sub-micron in size, especially the elemental and other insoluble carbon species tend to be disproportionately concentrated at the smallest sizes [Sicre et al., 1990; Rivera-Carpio et al., 1996; Neusüß et al., 2000]. This is of considerable interest since a number of these organic species act as surfactants, forming layers on the aerosol surface that might alter their CCN and IN properties [Facchini et al., 2000; Charlson et al., 2001; Feingold and Chuang, 2002]. Indeed, numerous surfactants and even surfactant layers have been found on marine aerosols [Barger and Garrett, 1976; Gagosian et al., 1982; Mochida et al., 2002; Tervahattu et al., 2002a].

3.4.2.2 The remote continental aerosol

Although remote continental aerosols often have a considerable contribution from soil dust [e.g., Jaenicke, 1993; e.g., Eleftheriadis and Colbeck, 2001] a substantial portion of the mass, particularly the sub-micron mass, is composed of sulfate and organic matter [Malm et al., 1994; Shrestha et al., 1997; Andreae et al., 2002; Henning et al., 2003; Putaud et al., 2004]. While there are in fact land sources of natural sulfur compounds (e.g., volcanoes and terrestrial biota), their magnitude is rather small compared to anthropogenic sulfur emissions
Small but significant amounts of nitrate are present even in remote locales, at least occasionally in excess of what would be expected from natural sources [Talbot et al., 1988; Talbot et al., 1990; Malm et al., 1994; Putaud et al., 2004]. Much of the sulfate found in remote continental aerosols is likely of anthropogenic origin. The other major component of the remote continental aerosol is that of carbonaceous compounds, mostly composed of organic matter [e.g., Artaxo et al., 1988; Talbot et al., 1988; Artaxo et al., 1990; Talbot et al., 1990; e.g., Malm et al., 1994; Zappoli et al., 1999; Krivacsy et al., 2001; Artaxo et al., 2002; Guyon et al., 2003b; Putaud et al., 2004; Kanakidou et al., 2005]. This includes both primary biogenic aerosols and secondary organic aerosol from biogenic precursors. Indeed, in the most remote locales, the organic fraction can be as large or larger than that of sulfate [Artaxo et al., 2002; Graham et al., 2003a; Graham et al., 2003b; Guyon et al., 2003b]. There is a definite trend towards a larger soluble organic fraction as the aerosol approaches pristine conditions [Mazurek et al., 1997; Zappoli et al., 1999; Graham et al., 2003a]. While this might well be in part due to evolution of industrial pollution, several studies have shown that the natural organics are qualitatively different from polluted air [Mazurek et al., 1997; Graham et al., 2003a]. This aspect of the remote continental aerosol composition is of considerable importance when considering CCN activity. The few existing studies show, that the aerosol from the least polluted sites in the Amazon Basin, for example, is very efficient at nucleating cloud droplets [Roberts et al., 2001; Andreae et al., 2004; Rissler et al., 2004].

3.4.2.3 The polluted continental aerosol

Polluted continental aerosols are by far the most widely studied. In industrial pollution, the most clear-cut trend with higher pollution levels is the absolute and relative increase in the abundance of black carbon [Malm et al., 1994; Zappoli et al., 1999; Neusüss et al., 2002; Putaud et al., 2004], although the tendency is not universal [Molnar et al., 1999]. Organic carbon also increases substantially in absolute and relative terms but the latter trend is much less clear than for the black carbon. In part, this variability may be due to the fact that organic aerosol has a distinct biogenic source (primary particles and SOA) which produces a significant natural background, which is not the case for BC. The inorganic portion of the aerosol mass is dominated by sulfate and nitrate, accompanied by an extensive suite of metal cations [Putaud et al., 2004].
The mass size distribution of the organics and inorganic in aerosols is also of interest. In accord with the changes in the number size distribution discussed in the previous section, most of the mass of industrial aerosol is confined to the sub-micron size range, for both organic and inorganic species. However, the carbon species, and particularly elemental carbon, tend to be disproportionately concentrated at the smaller sizes examined, especially below approximately 0.1 µm diameter [Hering et al., 1997; Maenhaut et al., 2002; Neusüß et al., 2002; Jimenez et al., 2003]. Additionally there is a tendency for the insoluble organics to be at smaller sizes than the soluble species [cf. Neusüß et al., 2000]. However, it should be noted that most of the particles even in urban source areas are largely internally mixed [e.g., Noble and Prather, 1996].

The ratio of insoluble to soluble organics is generally higher in urban, industrial pollution than in background aerosols [e.g., Zappoli et al., 1999] [Turpin and Lim, 2001]. These characteristics of urban pollution may have implications for the formation of organic films on aerosols, with significant consequences for CCN properties. However, there has been relatively little recent work on this compared to numerous studies on marine aerosols. An illustrative example of some of the composition variability that can be seen in transitioning from background to polluted venues can be seen in Figure 3-4.

3.4.2.4 Biomass burning aerosol

For pollution arising from biomass burning, a somewhat different compositional picture arises. Much of the aerosol mass consists of carbonaceous species [Andreae et al., 1996; Cachier et al., 1996; Mayol-Bracero et al., 2002b; Sinha et al., 2003; Decesari et al., 2006; Fuzzi et al., 2005; Reid et al., 2005]. A sizable fraction of this is black carbon, somewhat higher than that in industrial pollution. A more significant difference from industrial emissions resides in the relatively high organic fraction of the emissions [Andreae and Merlet, 2001; Sinha et al., 2003]. The organic composition of the biomass smoke is quite distinct from that of industrial emissions, with a substantial component attributable to dehydrated sugars – essentially cellulose breakdown products such as levoglucosan [Graham et al., 2002; Mayol-Bracero et al., 2002b; Gao et al., 2003; Decesari et al., 2006]. However, from the standpoint of impact on precipitation, the key attribute of the pyrogenic organic emissions is their relatively high solubility, about half of the organic matter in these aerosols being water soluble [Reid et al., 2005; Decesari et al., 2006]. The inorganic fraction of the
Biomass aerosol is minor in terms of mass initially, but can become substantial with aging due to secondary production of species such as sulfate and nitrate [Reid et al., 1998]. In spite of their lower mass fraction, inorganic species can make a significant contribution to the solute content of pyrogenic aerosols [Roberts et al., 2002; Mircea et al., 2005].

Figure 3-4. Average mass balance at remote (a), rural (b) and polluted (c) venues in Western Europe. EC= elemental carbon, WINSOC= water soluble organic compounds, WINSIC= water insoluble inorganic carbon, WSOC=water soluble organic compounds, ND= not determined. (From Zappoli et al. [1999]).
3.4.2 Hygroscopic Growth, CCN and IN Activity

All the interactions of aerosol particles with water, including hygroscopic growth at RH <100% as well as CCN and IN activity, are derivative from the aerosol composition and size, at least in principle [Dusek et al., 2006]. However, our theoretical understanding of this linkage, even for the long-studied CCN activity, is still not entirely satisfactory, and our understanding of the IN linkage is at present rudimentary [Pruppacher and Klett, 1997] (for more detail on ice processes see Chapter 2).

3.4.3.1 Hygroscopic Growth

At ambient relative humidity, water represents a significant component of most aerosol particles. The water content of a given particle is a function of its composition and size, and of the ambient relative humidity. Inorganic ionic components provide a large fraction of the soluble content of most aerosols, while the carbonaceous components include both water-soluble and insoluble species.

In most climate models, a fraction of emitted carbonaceous species, particularly the soot particles, is considered to be hydrophobic. These hydrophobic carbonaceous aerosols are then converted into hydrophilic aerosols by photochemical processes, condensation and coagulation, usually in a matter of hours to a day [Riemer et al., 2004]. Once the carbonaceous aerosols are hydrophilic, they can act as CCN and thus are subject to in-cloud and below-cloud scavenging, whereas wet scavenging of hydrophobic carbonaceous aerosols is limited to impaction scavenging (see sub-section 4.7.2.1).

The equilibrium water content of inorganic species as functions of ambient relative humidity (equivalent to water activity for bulk solutions) is well known. Polynomial fits have been published for binary solutions [e.g., Tang and Munkelwitz, 1994] and models for rigorous calculations for multi-component mixtures are readily available [Kim et al., 1993b; Clegg et al., 1998; Nenes et al., 1998; Pilinis et al., 2000; Clegg et al., 2001; Metzger et al., 2002b]. If particles composed of soluble inorganics are completely dried and then exposed to increasing RH, they take up no water until exposed to a relative humidity corresponding to the saturated-solution water activity, at which point a droplet of saturated solution forms. The water content progressively increases as RH is increased beyond this point. The wet-to-dry volume ratio at a particular RH can vary substantially with composition, with sulfuric acid
(which is soluble in water in all proportions and thus has no deliquescence point) and sodium salts having relatively higher water contents than neutralized sulfates. As RH is decreased, particles generally do not dry spontaneously at the deliquescence RH but rather retain water to a lower RH. This hysteresis implies that most particles in the atmosphere, having been exposed to relatively high RH at some time, contain at least some water and rarely exist in a completely dry state.

Much less is known about the hygroscopic behavior of individual organic species and their mixtures found in atmospheric aerosols. Most of the constituents of organic aerosol matter remain unidentified, and any single compound rarely accounts for a substantial fraction of the mass. Laboratory studies have measured equilibrium water contents for dicarboxylic acids, sugars, humic and fulvic acids, model polymers, and several other species [Peng et al., 2001a; Peng et al., 2001b; Prenni et al., 2001; Choi and Chan, 2002; Chan and Chan, 2003; Wise et al., 2003; Brooks et al., 2004; Hansen and Beyer, 2004; Mikhailov et al., 2004; Chan et al., 2005; Svenningsson et al., 2006][Varathbangkul et al., 2006]. The semivolatile nature of many organic aerosol species makes them more difficult to study in a controlled laboratory setting, and techniques that are readily applied to nonvolatile salts must be modified [e.g., Choi and Chan, 2002]. Furthermore, many relevant organic compounds are only sparingly soluble, and deliquesce at RH > 95%. It is difficult to control RH at these high values and thus laboratory water uptake experiments on such species are lacking, except for observations of their behavior at water supersaturations [e.g., Cruz and Pandis, 1997; Corrigan and Novakov, 1999; Giebl et al., 2002; Raymond and Pandis, 2002; Hori et al., 2003; Kumar et al., 2003; Raymond and Pandis, 2003; Shantz et al., 2003; Bilde and Svenningsson, 2004; Broekhuizen et al., 2004].

It is generally adequate to model the density of multicomponent solutions assuming volume additivity [e.g., Brechtel and Kreidenweis, 2000a]. With respect to the water content of mixtures, the simplest treatment is the Zdanovskii-Stokes-Robinson (ZSR) assumption [Seinfeld and Pandis, 1998], which states that the total equilibrium water content at a specified relative humidity (RH) is the sum of the equilibrium water contents for each individual species at the same RH. This approach can capture non-idealities that arise in individual binary solutions, but more sophisticated techniques are needed to represent solute-solute interaction effects. In other words, the ZSR approximation assumes that there are no interactions between species that either enhance or diminish hygroscopic growth beyond that
expected for each component. Although some counterexamples have been documented in the literature [Tang, 1997], the ZSR assumption appears to be adequate in many cases [Malm and Kreidenweis, 1997; Prenni et al., 2003; Brooks et al., 2004]. However, the deliquescence RH of a mixed salt is always lower than that of any of the components in the mixture [Wexler and Seinfeld, 1991], an effect that is not captured with the ZSR treatment. Marcolli et al. [2004] demonstrate that mixtures of organic species deliquesce at very low RH and thus atmospheric organic aerosols, which are known to be mixtures of many soluble and insoluble species, are expected to retain water to very low RH.

Recently, several workers [Brechtel and Kreidenweis, 2000a; Brechtel and Kreidenweis, 2000b; Rissler et al., 2004; Koehler et al., 2006; Kreidenweis et al., 2005; Rissler et al., 2006; Svenningsson et al., 2006; Petters and Kreidenweis, 2006] have suggested that information on particle hygroscopic growth below 100% RH can be extrapolated to predict the critical supersaturations required for activation to cloud drops. This method involves using growth factors obtained from Humidified Tandem Differential Mobility Analyzers (HTDMA) and deducing composition-dependent water activities of aqueous solutions of pure materials or mixtures, or directly extrapolating HTDMA wet-to-dry volume ratios into the supersaturated regime. Thus far the methods have produced reasonable agreement with direct measurements of critical supersaturations for particles composed of inorganic [Kreidenweis et al., 2005] and organic [Koehler et al., 2005; Svenningsson et al., 2006; Petters and Kreidenweis, 2006] compounds and with field measurements on ambient aerosols [Rissler et al., 2004; Rissler et al., 2006], generally within the uncertainties of CCN measurements. When used to predict nucleated drop numbers in an adiabatic parcel model, even the least accurate extrapolations predict drop numbers within the range of uncertainties of those predicted using the most accurate thermodynamic models [Koehler et al., 2005]. The validity of this approach has not yet been examined for some secondary organic aerosol.

### 3.4.3.2 Cloud Condensation Nuclei (CCN)

CCN are a sub-set of the atmospheric aerosol upon which water vapor condenses to form cloud droplets at the supersaturations achieved in clouds (<1%). It follows from the Köhler equation (section 2.1.2.1) that the greater the particle mass and water solubility, the lower will be the supersaturation at which the particle can serve as a CCN, and the greater the
chance that it will influence the droplet number concentration in a cloud. For example, to serve as a CCN at 1% supersaturation, completely wettable but water insoluble particles need to be at least 200 nm in diameter, whereas soluble particles can serve as CCN at 1% supersaturation even if they are as small as 30 nm. Most CCN consist of a mixture of inorganic and organic components. Most inorganic species, including those normally found in the greatest abundance in the atmosphere (e.g. ammonium sulfate), are highly water soluble. On the other hand, the solubilities of organic species are widely varying [e.g. Saxena and Hildemann, 1996]. Moreover, some organics, such as condensed hydrocarbons, may not even be wettable, i.e. unable to adsorb water on the surface of the particle. Water-soluble organic carbon (WSOC) can influence aerosol activation by increasing the amount of soluble material [Shulmann et al., 1996; Cruz and Pandis, 1997].

Facchini et al. [1999] measured the surface tension of partially evaporated cloud water samples, and attributed the values lower than for pure water to dissolved organics. However, there is a limit to the extent to which surface tension can be reduced close to saturation because of dilution [Ervens et al., 2005]. Also, Lohmann and Leck [2005] have shown the presence of potentially surface-active particles from the ocean microlayer may account for an important fraction of CCN in a high Arctic aerosol. Another potential effect of organic material residing on the surface of the particle or solution droplet is to reduce the rate at which the water molecules condensing on the droplet stick to the surface [e.g. Cantrell et al., 2001; Feingold and Chuang, 2002].

Particle size is an important consideration in determining its effectiveness as a CCN [Fitzgerald, 1974; McFiggans et al., 2005; Dusek et al., 2006]. However, it needs to be pointed out that an insoluble but wettable particle of 100 nm radius will require a supersaturation of at least 1% in order to activate (see Chapter 2, Figure 2-5), whereas a 100 nm radius particle of ammonium sulfate will activate at a supersaturation more than an order of magnitude smaller (about 0.05%). A discussion of recent advances in knowledge of organics as CCN is given in the following section.

3.4.3.3 Carbonaceous Aerosols as CCN

The organic fraction of the fine aerosol mass is 20-50% at continental mid-latitudes [Saxena and Hildemann, 1996; Putaud et al., 2004], up to 90% in tropical forested areas [Talbot et al., 1990; Andreae and Crutzen, 1997; Roberts et al., 2001], and a significant fraction is
Thus, there is some expectation for the organic to influence CCN activity. In this section, measurements pertaining to the CCN activity of organic particles are summarized. In the subsequent section, results of “closure” experiments of ambient aerosols are used to examine the relative importance organic relative to inorganic components of aerosol particles.

Primary carbonaceous particles originating from fossil-fuel pollution may initially consist to a large extent of EC and water-insoluble substances (e.g., engine oil, etc.). At that stage, most are probably not efficient CCN [Hudson et al., 1991; Weingartner et al., 1997; Zuberi et al., 2005], but some combustion particles do exhibit CCN activity [Petzold et al., 2005]. Regardless of their initial composition, as they become partially oxidized and as soluble compounds (sulfates, nitrates, WSOC) deposit on them, they become CCN at realistic supersaturations [Tsigaridis and Kanakidou, 2003]. This process can occur on timescales of minutes to hours. Primary particles from biomass burning (pyrogenic particles) are already fairly efficient CCN near the source and become even more easily activated as they age [Warner and Twomey, 1967; Hobbs and Radke, 1969; Eagan et al., 1974b; Hallett et al., 1989; Radke et al., 1991; Rogers et al., 1991; Ross et al., 2003; Andreae et al., 2004].

Laboratory measurements of the CCN properties of SOA particles from monoterpene oxidation indicate a large variation in the CCN activity of the resulting compounds [VanReken et al., 2005; Hartz et al., 2005]. Field observations suggest that the condensation onto pre-existing particles of SOA from the oxidation of biogenic terpenoid compounds significantly enhances the CCN activity of those particles [Leaitch et al., 1999; Sun and Ariya, 2006]. Whether these SOA compounds form new particles from the gas phase or simply condense on existing particles remains unknown [e.g. Hoppel et al., 2001; Gao et al., 2001].

Other laboratory studies of activation of single-component aerosol particles at atmospherically relevant supersaturations have clearly shown that the solubility of the compounds is a key parameter for organic aerosol activation [Novakov and Corrigan, 1996; Giebl et al.,2002; Raymond and Pandis, 2002; Hori et al., 2003; Shantz et al., 2003; Broekhuizen et al., 2004] and that the lowering of the surface tension can dramatically reduce the activation diameter of a moderately soluble organic particle [Broekhuizen et al., 2004].
It has been shown that the ability of organic particles to activate can be substantially enhanced by adding small amounts of soluble compounds as inorganic salts or surface-active species [Hegg et al., 2001; Bilde and Svenningsson 2004; Broekhuizen et al., 2004; Abbatt et al. 2005]. Broekhuizen et al. [2004], Bilde and Svenningsson [2004] and Henning et al. [2005] also found that the initial particle phase played a significant role in the critical supersaturation of particles containing only slightly soluble substances. If the particle was initially totally liquid the measured critical supersaturations were lower. Marcolli et al. [2004] recently demonstrated that multicomponent organic solutions favored the liquid phase. For complex mixtures found in the atmosphere, it may be expected that aerosol comprising slightly soluble organics would be unlikely to contain undissolved material. Also, it is important to note that even relatively small amounts of inorganic material will allow the activation of organic particles [Ervens, et al., 2005; Lohmann, et al., 2004], though the abundance and nature of organics will affect both equilibrium water content and droplet growth kinetics. The recent work of Petters and Kreidenweis [2006] provides a method for connecting these results for model applications. Reviews of the carbonaceous aerosol and their roles in cloud droplet nucleation are found in Kanakidou et al. [2005] and McFiggans et al. [2005] respectively.

3.4.3.4 CCN Closure Studies

The classic description of the relationship between aerosol size, composition and CCN activity is the well-known Köhler theory (For a more detailed discussion see Chapter 2). If this theory is valid, then the CCN activity is specified by aerosol size and composition measurements, and direct measurement is unnecessary. This theory has been subject to laboratory trials in which the activation size of the particles of various sizes and compositions have been measured at supersaturated water environments in CCN chambers. Much of the early such work utilized inorganic salts and acids since these were thought to dominate ambient water soluble aerosol composition. While experiments with these particles lent support to the applicability of the theory [cf. Fitzgerald, 1973], more recent work on aerosol composition, which (as discussed above) attributes a very substantial role to carbonaceous species, led to questioning of the prognostic power of the Köhler theory [e.g., Shulman et al., 1996; Facchini et al., 1999; Charlson et al., 2001; Anttila and Kerminen, 2002]. Consequently, modified forms of the theory have been developed to take into account the lower water solubilities of many organics, their possible impact on haze particle surface
tension, and even the influence of soluble gases that may co-condense with water during particle growth [Laaksonen et al., 1998].

A further complication of the organics is their surfactant properties and the consequent possibility that they could form films that would hinder water condensation/evaporation and thereby impose a kinetic limit to the thermodynamic Köhler theory [Gill et al., 1983; Cruz and Pandis, 1998]. However, laboratory experiments to quantitatively assess this phenomenon – and even some detailed theoretical assessments - have yielded mixed results. Most investigators have concluded that such a limitation is at best small [Cruz and Pandis, 1998; Hegg et al., 2001; Anttila and Kerminen, 2002; Raymond and Pandis, 2003; Shantz et al., 2003; Broekhuizen et al., 2004]. Hygroscopic growth factors were well reproduced at the high altitude site Jungfraujoch using the chemical composition of the mixed particles and the Zdanovskii-Stokes-Robinson (ZSR) relation [Gysel et al., 2002; Weingartner et al., 2002]. The temporal variability of the hygroscopic growth factors was mainly a result of varying fractions of organic/inorganic mass, and the monomodal distribution of the hygroscopic growth factor confirmed the presence of an internally mixed aerosol. Nevertheless, since there is some laboratory support for such limitation and evidence for organic films has indeed been found on ambient aerosols [Mochida et al., 2002; Tervahattu et al., 2002a], the possibility must be considered when analyzing atmospheric data.

Evidence for a role of the organic aerosol as CCN has been found through the investigation of apparent changes in CCN efficiency and statistical correlations with organics [e.g. Bigg, 1986; Hudson and Da, 1996; Matsumoto et al., 1997]. In a landmark study, Novakov and Penner [1993] and later Rivera-Carpio et al. [1996] combined CCN spectral measurements with size resolved composition measurements to show that both the sulfate and the carbonaceous compounds contributed to the CCN activity.

An effective method for investigating the CCN activity of ambient aerosols is the CCN closure experiment [e.g. Bigg, 1986; Liu et al., 1996; Covert et al., 1998; Chuang et al., 2000; Snider and Brenguier, 2000; Cantrell et al., 2001; VanReken et al., 2003; Mircea et al., 2005; Rissler et al., 2006; Broekhuizen et al., 2006; Ervens et al., 2007]. In such an experiment, the aerosol size distribution and chemical composition as a function of size are measured simultaneously with the CCN activity. The linkage between the variables is the Köhler theory, and the closure constitutes a test of the theory and provides some level of
confidence in our knowledge of the system. Ideal closure experiments would include the measurement of the individual composition of particles. Such experiments may not yet be possible, and instead there is a spectrum of closure experiments with varying degrees of measurement completeness. However, even the most rudimentary of these types of experiments have shed light on the CCN activity of atmospheric aerosols.

Closure studies have taken three different approaches. In the first, the aerosol size distribution and CCN activation spectrum have been measured and a certain soluble inorganic composition (commonly ammonium sulfate) is assumed. Comparison of predicted (based on Köhler theory) and measured CCN concentrations then test this assumption [Bigg, 1986; Cantrell et al., 2000; Snider and Brenguier, 2000; VanReken et al., 2003]. In some instances [e.g., VanReken et al., 2003], reasonable closure has been achieved but in most instances, particularly in more polluted air, the predicted concentrations have substantially exceeded the measured CCN.

In the second more constrained approach, the CCN activation spectrum, the aerosol size distribution and the size dependent aerosol hygroscopicity have been simultaneously measured [cf. Brechtel and Kreidenweis, 2000]. In a practical sense, it is a more complete approach than utilizing size-dependent composition data since the latter must then be coupled with characteristic thermodynamic data to derive the activity of the aerosol chemical species. Recent studies that have utilized this approach [Covert et al., 1998; Zhou et al., 2001; Dusek et al., 2003; Rissler et al., 2004; Ervens et al., 2007] found modest discrepancies between predicted and measured CCN concentrations. While the differences were generally not large, they were systematic with the predicted concentrations nearly always in excess of the measurements. The Covert et al. [1998] study showed that the percentage magnitude of the overprediction varies directly with the extent of anthropogenic impact, consistent with the earlier work of Hudson and Da [1996].

The third type of closure study involves particle size, size-resolved composition and CCN activity measurements. Examples of this are Chuang et al. [2000], Cantrell et al. [2001] Mircea et al. [2005] and Broekhuizen et al., [2006]. Chuang et al. [2000] overpredicted the measured CCN concentrations by as much as an order of magnitude, in part because the CCN activity and compositional measurements were not co-located and in part due to problems
with the measurements. The other studies were able to achieve closure within 10-20% with a tendency to overpredict the CCN concentrations.

*Cantrell et al.* [2001], *Broekhuizen et al.* [2006] and *Ervens et al.* [2007] all obtained the best closure assuming no role of the organic components in the particles, despite the presence of significant organic mass fractions (20-90%). *Cantrell et al.* [2001] attributed their overprediction to the possibility that the some of the organic components in the particles inhibited water uptake, while *Broekhuizen et al.* [2006] suggested limited resolution of the model and the description of the mixing states could explain the slight bias in their closure.

In a comprehensive analysis of size-dependent chemical composition, hygroscopic growth, and CCN activity, *Mircea et al.* [2005] showed that knowledge of aerosol WCOC composition in terms of classes of compounds and of their molecular weights and acidic properties were necessary to predict CCN activation.

A key question of the closure studies is how are the various aerosol species are mixed? While some studies have not found this to be a critical assumption [e.g. *Liu et al.*, 1996], later studies, have demonstrated a strong sensitivity of the activity of either organic aerosols or inorganic aerosols to trace quantities of each other [e.g. *Roberts et al.*, 2002; *Bilde and Svenningsson*, 2004; *Broekhuizen et al.*, 2004; *Abbatt et al.*, 2005; *Mircea et al.*, 2005; *Broekhuizen et al.*, 2006].

The results of these closure experiments suggest a widely varying role for the organic compounds in the water uptake by the atmospheric aerosol. Studies in or near urban regions are able to explain most, but not all, of the CCN activity using the inorganic chemistry. In the case of a strong biomass burning influence [*Mircea et al.*, 2005], there appears to be a strong role for the organic in determining CCN activity. And, as discussed in section 2.3.4.3, SOA particles from biogenic sources have at least some importance for water uptake. Recently, *Petters and Kreidenweis* [2006] have introduced a pseudo-empirical approach for modeling the CCN activation of multi-component aerosols that eliminates many of the thermodynamic details necessary for the description of water activity, and may provide a more streamlined approach to the representation of CCN activation in models. Still, the chemical components, concentrations, solubilities, and surface tensions of solution droplets that form in various environments are not well documented, and knowledge of the effects of these parameters on
droplet activation individually and in combination is very limited [Kanakidou et al., 2005; McFiggans et al., 2005].

3.4.3.5 Source-related global distribution of CCN

Global microphysical aerosol models are required to globally predict the evolution of the size-distribution and the contribution of specific aerosol sources to the number of radiatively important aerosols and cloud condensation nuclei. In a study with the global microphysical aerosol-climate model ECHAM5-HAM, Stier et al. [2005] demonstrated that the global CCN number concentration is non-linearly related to the emission mass flux. While primary carbonaceous aerosol emissions disproportionately contribute to the CCN number, anthropogenic SO$_2$ emissions less than proportionally contribute to the number of CCN, because a significant fraction of the produced sulfuric acid condenses on pre-existing aerosols. The co-emission of anthropogenic carbonaceous and sulfur emissions produces generally less CCN in the vicinity of the source regions than would be obtained by summing up the CCN produced by the individual emissions considered separately. Therefore, in contrast to results from traditional bulk mass aerosol models, the sum of the change of CCN due to individual emission changes is different from the total CCN change due to several combined emission changes.

As a measure of the contribution of specific aerosol sources to the potentially CCN-relevant size range, one can define the difference in CCN due to omission of the specific aerosol source with respect to a simulation with the complete set of emission. In Figure 3-6 a) the annual zonal mean distribution of hygroscopic accumulation and coarse mode particles (an approximation of the CCN concentration) for the year 2000 is depicted, computed by the ECHAM5-HAM aerosol model [Stier et al., 2005]. Further shown are the differences with respect to this reference scenario due to the omission of b) carbonaceous emissions from fossil-fuel use and industry, of c) sulfur emission from fossil-fuel use, industry, and bio-fuels, and of d) total anthropogenic emissions.

The zonal mean distribution of CCN (a) shows a pronounced maximum in the lower to mid-troposphere (1000 hPa – 500 hPa) in the region of strongest anthropogenic emissions between 30°S and 50°N. The omission of carbonaceous emissions from fossil fuel use and industry (b) reduces the zonal mean number of CCN in the lower troposphere between 20°N and 60°N by up to 30 %. Fossil-fuel use and industry contribute (with 6.4 Tg) only
10% of the total carbonaceous mass emissions and a significantly smaller fraction of the total anthropogenic aerosol mass source. In the lower to mid-troposphere, sulfur emissions from fossil fuels, industry, and biofuels (c) contribute between 30% and 50% of the total number of CCN between 20°N and 60°N. Higher up in the NH mid-to-upper troposphere (750-250 hPa), they dominate the CCN numbers. The impact of anthropogenic sulfur emissions on southern hemispheric CCN in the lower troposphere is negligible. Total anthropogenic emissions (d), including fossil-fuel use, industry, vegetation fires and biofuel use, largely control the low to mid-troposphere CCN numbers between 30°S and 70°N contributing to 50%-80% of total CCN numbers. Only the CCN at high southern latitudes are predominantly controlled by natural emissions.

Figure 3-6. a) Total zonal annual-mean cloud condensation nuclei (cm$^{-3}$ STP{1013hPa, 273.15K}) as simulated by the ECHAM5-HAM aerosol model. CCN are defined here as hygroscopic accumulation and coarse mode aerosol particles. Change in CCN due to the omission of b) carbonaceous emissions from fossil-fuels and industry, c) sulfur emission from fossil-fuels, industry, and bio-fuels, d) all anthropogenic emissions including fossil-fuels, industry, bio-fuels, and vegetation fires (adapted from Stier et al.[2005,2006]).
3.4.3.6 IN Activity

Very little is still known about the IN activity. The situation circa 2000 has been well summarized in the 3rd IPCC report [Houghton et al., 2001], but, unfortunately this field has seen relatively little activity in the last 30 years. Such issues as the precise mode of action of aerosol acting as IN (e.g., contact, immersion, deposition, etc. nuclei), a lack of consistency in measurement techniques, in-situ modification of particle IN efficiency and secondary ice production all contribute to the uncertainty in IN activity of atmospheric aerosols.

Nevertheless, a few facts have emerged. Firstly, there is no question that particles acting as IN are relatively rare compared to CCN. IN concentrations typically are measured at a few per liter or less, thus constituting no more than a few tenths of a percent of the normal total aerosol population. In part this is no doubt due to the limited chemical components that are known to act as IN. Best known are crustally-derived minerals, for which there is a large and still growing body of corroborative data [Kumai, 1976; Hagen et al., 1995; Heintzenberg et al., 1996; Castro et al., 1998; Kreidenweis et al., 1998; DeMott et al., 2003a; Cziczo et al., 2004; Sassen, 2005].

Recently, experiments have been performed in a large-volume (84 m$^3$) chamber called AIDA (Aerosol Interaction and Dynamics in the Atmosphere), which can be operated at variable pressures and temperatures, between 30°C and –90°C, and at high relative humidities. A unique feature of AIDA is the ability to supersaturate the gas phase with respect to ice and liquid water by quasi-adiabatic cooling, thereby realistically simulating cooling rates of air parcels in the atmosphere with updraft velocities between 0.5 and 6 m s$^{-1}$. Mohler et al. [2006] investigated the deposition mode ice nucleation efficiency of various dust aerosols (Arizona test dust (ATD) and two dust samples from the Takla Makan desert in Asia (AD1) and Sahara (SD2)) at cirrus cloud temperatures between 196 K and 223 K. Deposition ice nucleation was most efficient on ATD particles with ice-active particle fractions of about 0.6 and 0.8 at an ice saturation ratio $S_i<1.15$ and temperatures of 223 K and 209 K, respectively. The desert dust samples SD2 and AD1 showed a significantly lower fraction of active deposition nuclei, about 0.25 at 223 K and $S_i<1.35$. Field et al. (2006), using similar samples, reported that at temperatures warmer than -40°C, droplets were formed before ice crystals formed, and there was generally no deposition nucleation observed. At temperatures colder than -40°C, dust samples exhibited dual nucleation events that were observed during the
same expansion experiment. The primary nucleation event occurred at ice saturation ratios of 1.1 to 1.3, and is likely to be a deposition nucleation mode. The secondary nucleation event occurred at ice saturation ratios between 1.35 and 1.5, and probably was a condensation mode, as there was some soluble material in the dust samples. For all samples, the ice activated aerosol fraction could be approximated by an exponential equation as function of $S_i$. This formulation of ice activation spectra may be used to calculate the formation rate of ice crystals in models, if the number concentration of dust particles is known. More experimental work is needed to quantify the variability of the ice activation spectra as function of the temperature and dust particle properties.

There is also evidence suggesting carbonaceous aerosols [e.g., Karcher et al., 1996; Petzold et al., 1998; Strom and Ohlsson, 1998] and various biogenic particles [see the review by Szyrmer and Zawadzki, 1997] to be IN. Nevertheless, it is commonplace for the concentration of IN to be far below the concentration of particles containing any or all of these various favorable components. Differing degrees of IN deactivation due to pollutant deposition on IN surfaces or cloud processing have been invoked as an explanation for this phenomenon but clear evidence of this is wanting. Furthermore, it is also likely that favorable ice-forming substrates could be deposited as organic films on previously inactive particles. The mix of such in situ processes could well explain the long-puzzling variability in the IN fluxes from pollution sources [e.g., Hobbs and Locatelli, 1969; Braham and Spyers-Duran, 1974a]. Once again, however, evidence of the impact of such secondary processes is sparse at best. For further discussion of IN activity, see Chapters 2 and 6.

3.5 Recommendations for Future Observations and Studies

- Emission inventories have been conventionally limited to estimates of mass emission fluxes for bulk aerosol components. To be useful in cloud and precipitation studies, size-resolved information about the numbers of particles emitted is required for the various primary aerosol sources, especially seasalt, mineral dust, pyrogenic aerosol and primary biogenic aerosol.
- Emission information must be provided in forms that are suitable for dynamic emission inventories. It is much more useful to have emission algorithms that permit calculation of an
emission flux from model variables than a static array of emissions valid for one particular time.

- The abundance and rate of production of submicron seasalt particles is still under dispute. Since these particles can play a significant role in cloud microphysics and precipitation in remote marine regions, observational constraints to their source parameterizations are urgently needed.

- Primary biogenic aerosols probably play a significant role as cloud and ice nuclei at present in very remote regions, and must have been a key control on cloud processes in the pre-anthropogenic atmosphere. Their characteristics and rates of emission need to be investigated.

- Due to the complexity of carbonaceous aerosol composition and emission processes, a critical concern is to define an adequate classification scheme for these aerosols based on observable characteristics.

- Building a reliable inventory for carbonaceous aerosols (BC, PPOC and SOA) remains a challenge because of uncertainties in emission factors and activity estimates. Emission parameterizations must be provided that include size distributions and number fluxes.

- The injection height of emissions also differs strongly from one aerosol source type to another, and even shows pronounced variability within some types of aerosol source. This applies especially to biomass burning and vegetation fires. Any changes in injection height due to shifts in agricultural or technical practices will have important consequences on the lifetime and fate of particles.

- The rates of sulfate aerosol production are at least as dependent on the accurate representation of the processes leading from SO$_2$ to sulfate aerosol as they are on emission estimates of SO$_2$. Accurate knowledge of these processes and their correct parameterization in models is essential.

- Aerosol nitrate is expected to increase in climatic importance in the 21$^{st}$ century. Accurate measurement techniques and reliable modeling approaches for this component must be established.

- In order to assess the human impact on cloud physics, we need to know the aerosol, CCN, GCCN and IN distributions in the pre-anthropogenic atmosphere. The sources, characteristics, and fluxes of natural aerosol types must be investigated by carefully conducted field and laboratory studies.
• The ability of aerosol particles to act as CCN, GCCN and IN should be determined as a function of particle size, composition and supersaturation. Uncertainties are greatest for IN because of the different processes in which that ice can be nucleated. Measurements on cloud-active particles should be conducted in the context of closure studies.
4.1 Introduction

Because the time for air parcels to circle the Earth on winds in the troposphere is of the same order of magnitude as the residence time of atmospheric aerosols, there is no location on the globe that is not influenced by aerosol sources. Once released into the atmosphere from primary production or produced via gas-to-particle conversion (for source details, see Chapter 3), aerosols are subject to many processes that affect their global distribution, chemical and physical properties, and hence their influence on climate, weather, human health and ecosystems. They are dispersed in the atmosphere through processes of advection, convection and turbulence. They are also transformed and removed by physical and chemical processes involving clouds, precipitation as well as processes occurring in cloud-free air. During the life cycle of an aerosol in the atmosphere, gas to particle conversion and mixing of aerosols from different sources changes the chemical, physical and optical properties of the original aerosols.

4.2 Geographically Distinct Aerosol Regimes

Hemispheric-scale differences in land area (39% of the northern hemisphere is covered by land vs. 19% in the southern hemisphere) and in human population have led to large inter-hemispheric differences in primary aerosol source strengths (Chapter 3). This involves both natural and anthropogenic aerosols, but the greatest differences are seen for fossil fuel and aircraft emissions. According to data summarized in Chapter 5 of the IPCC 2001 assessment report [IPCC, 2001], more than 90% of the aircraft emissions and ~99% of the fossil fuel emissions are emitted into the atmosphere over the northern hemisphere. More than 80% of the mineral dust is produced in the northern hemisphere, according to Jickells et al. [2005], and less than 20% of the dust flux into the oceans occurs in the southern hemisphere. In contrast, biomass burning emissions are more evenly distributed between hemispheres, with a split of nearly 50%:50%. Sea salt is the only primary aerosol produced in greater quantities in the southern hemisphere (Chapter 3).
Consider the global atmospheric aerosol distribution from the perspective of the global distribution of mean annual aerosol optical depth (AOD, Figure 4-1) obtained from a combination of satellite observations and verified using a large surface-based network of aerosol optical depth remote sensing instruments (Chapter 5). There are some biases in this distribution due to the dependence of satellite observations on sunlight. For instance, satellites cannot observe the polar regions in the dark winter half of the year, thus missing Arctic haze pollution \cite{Barrie1986, Barrie1990, Christensen1997}, and they are unable to measure beneath clouds, thus possibly under-sampling the atmosphere where clouds are most frequent (Fig. 1-5a) such as in the roaring forties belt between 40 and 60 S of the southern oceans. In addition, the distribution of AOD does not completely correlate to that of aerosol mass, because the fine aerosol particles have a stronger optical effect (mass extinction coefficient), than larger particles. Therefore coarse sea salt and dust aerosols cannot be readily seen on this figure (see section 4.2.6). Nevertheless, the geographical distribution of AOD in Figure 4-1 helps us to organize a description of the global aerosol and aids in understanding their influence on weather, climate, air quality and, in particular, precipitation reaching the Earth’s surface. Seasonal counterparts to this figure can be found in Chapter 1 (Fig. 1-8).

Figure 4-1: A best estimate of the global distribution of annual average tropospheric aerosol optical depth (AOD) compiled by combining data from six satellites (operating for limited periods between 1979 and 2004). Observations for a region were selected using ground-based AOD observations as guidance (courtesy of S. Kinne MPI, Hamburg, Germany). **AFR**: Africa, **SAS**: South Asia, **SEA**: Southeast Asia, **EAW**: East Asia and Western North America, **NAE**: Eastern North America, North Atlantic and Europe, **SAM**: South America, **MAR**: Marine and **AH**: Arctic Haze
For the purposes of discussion, we divided the global aerosol into 8 geographically distinct regimes each involving particles of one or more aerosol types (e.g. sea salt, marine biogenic, soil dust, biomass burning, anthropogenic sulfates and organics).

4.2.1 Africa (AFR)

The aerosol regime in Africa is a good example of a primary (soil dust) aerosol mixing with primary and secondary biomass burning aerosols. The soil dust component covers the Saharan desert source region and extends thousands of kilometers downwind. It affects not only the surrounding oceans but also other continents such as the Americas, Europe, the Middle East and south Asia. Transport across the Atlantic can affect visibility and deposition of nutrients in the Caribbean. There are outstanding questions regarding its role in influencing hurricane development over the tropical Atlantic [Evan, et al., 2006]. For instance, does it affect sea surface temperature, which is one variable that is widely recognized as important to hurricane development? The effects of soil dust on cloud and precipitation formation processes are only beginning to be understood. There is evidence that a proper treatment of Saharan dust can improve the forecast of winds such as the easterly jet over West Africa in numerical weather predictions [Tompkins, et al., 2005].

Emissions from biomass burning sources in tropical and southern Africa drift over the tropical Atlantic into tropical South America as well as into the Indian Ocean off the southeast tip of Africa. A so-called “river of smoke” (Fig. 1.8), observed during the SAFARI 2000 field campaign flowed from northwest to southeast over the subcontinent, and caused heavy haze and reduced visibility over Botswana and South Africa [Swap, et al., 2003]. The satellite image from the MODIS satellite in Figure 4-2 from Kaufman et al. [2005] separates coarse Saharan dust and fine bio-mass burning aerosols in central and west Africa. Superimposed on these major sources are lesser but still quite significant sources of fossil fuel aerosol resulting from human activities in the region.

The African monsoon occurs adjacent to the Saharan soil dust sources and is co-located with biomass and air pollution sources. The role of aerosol radiative forcing in the dynamics of this monsoon is thought to be significant, but is only beginning to be investigated [Caminade, et al., 2006].
Figure 4-2: MODIS aerosol composites for 10 June 2004. Blue represents clean conditions, aerosol optical thickness <0.1, and green and red show higher optical thickness corresponding to the coarse (green) and fine (red) modes. The fine fraction (y axis) varies from green for fine fraction of zero to red for fine fraction of one. Therefore pure dust is green, and pure smoke or pollution is red [Kaufman, et al., 2005].

4.2.2 Eastern North America, North Atlantic, Europe (NAE) and Arctic Haze (AH)

Widespread fossil fuel combustion in eastern North America leads to a North American plume that moves eastward across the North Atlantic where it mixes with sea salt (see below) and enters Europe receiving fresh fossil fuel combustion aerosols and precursors [Methven, et al., 2006]. Not evident in Figure 4-1 because of a blind spot of aerosol satellite observations in the polar troposphere is the reverse flow from Eurasia to North America over the Arctic from November to May that constitutes the Arctic haze [Barrie, 1986; Barrie, 1990; Christensen, 1997]. European pollution also drifts eastwards in mid latitudes joining pollution from regions of the former Soviet Union and soil dust. The surface concentrations of aerosols in both North America and Europe from ground monitoring stations are discussed in Section 4.3.2.

4.2.3 South Asia (SAS) and Southeast Asia (SEA)

The aerosol regime in south Asia including the Arabian Peninsula and Indian Sub-continent is a result of soil dust, pollution and biomass burning aerosols. Saharan soil dust advected from the west affects the Indian Sub-continent, and this is combined with local sources of
dust. In addition, the Indian sub-continent has pollution from fossil fuel and biomass burning that leads to a highly absorbing aerosol [Ramanathan, et al., 2005]. It is advected out of the region affecting clouds and precipitation in the Himalayas in summer and over the Indian Ocean in winter. There is documented evidence of transport as far as the Mediterranean basin [Lelieveld, et al., 2002].

In the Atmospheric Brown Cloud (ABC) study, the presence of highly absorbing aerosol over the Indian Ocean was observed as far south as the Maldives. It had the effect of reducing solar radiation reaching the surface by as much as 20 W m$^{-2}$ [Ramanathan, et al., 2005] and increasing the stability of the atmosphere through redistribution of solar energy from the surface to soot-bearing aerosols aloft. This affects the evaporation of water vapor and hence the cloud regime. In addition, aerosol dust and pollution is thought to affect the Asian monsoon through direct radiative forcing of changes in regional circulation [Lau, et al., 2005] and indirectly through aerosol impacts on clouds and precipitation formation [Chylek, et al., 2006]. The latter which is the focus of this review, is more difficult to quantify due to the non-linearity of processes involved (Chapter 6 and 7). In any case, there are documented changes in precipitation in India over the last half century that raises the issue of direct aerosol pollution effects as well as other climate change effects that modify the monsoon dynamics. For instance, river basins in central India (Sabarmati, Mahi, Narmada, Tapi, Godavari and Mahanadi) have been experiencing reduced rainfall from about the 1960s, while in others—Indus from 1954, Ganga from 1993, Brahmaputra from 1988, Krishna from 1953 and Cauvery from 1929—there has been an upward trend in rainfall. The recent trend in rainfall across India is believed to be due to global warming [Singh, et al., 2005]. Whether changing aerosol pollution in the region has played a role remains an open question.

Another aerosol source region that is prominent in Figure 4-1 is south-east Asia centered on Indonesia and Borneo. Biomass burning aerosols coupled with air pollution from mega-cities are the main components of this aerosol regime. During the 1997 forest fires in Indonesia, approximately 42 Tg of aerosol mass were emitted into the atmosphere, and resulted in a -0.32 Wm$^{-2}$ forcing averaged over the 4 months of the fire season [Davison, et al., 2004]. This is more distinctly seen on Figure 1-10.
Another major aerosol regime is located in eastern Asia, mainly including China, Korea and Japan and the downwind region of the north Pacific as far as western North America. Soil dust from the Gobi and Taklimakan deserts and anthropogenic pollution from industry and fossil fuel burning are the main aerosol types. The soil dust source is seasonal, peaking in February to May and originating from the western and northern parts of the region. The mineral dust advects in the northerly and westerly winds over populated areas, often mixing with pollution. It also is lofted by convection into strong westerly winds of the middle troposphere (2-7 km) that periodically transport it in a matter of 5 to 10 days to North America [Jaffe, et al., 1999] and even into the Arctic [Sassen, 2005; Sirois and Barrie, 1999; Welch, et al., 1991]. Air pollution in the region is persistent throughout the year and has increased greatly in the latter half of the 20th century largely as a result of the economic growth in China. The southern part of this region is affected by monsoonal precipitation in summer and, as in Africa and south Asia, the effects of aerosols are poorly understood. Observations have shown that regional precipitation has changed substantially in China between 1951 and 2000 [Zhai, et al., 2005]. Simultaneously, there is a decreasing trend in the total global radiation reaching the surface over most industrial regions [Che, et al., 2005] reflecting an increase in aerosol loadings (details in section 4.3.4). Whether these changes are linked remains an open question.

Natural biomass burning aerosols originating from fires ignited by lightning in the boreal region of Canada, Russia and Alaska are highly seasonal but climatologically persistent (see Fig. 1-8a). Occurring mostly in summer, the burning products sometimes loft to great heights and can even produce their own thunderstorm clouds into which pollution is mixed, scavenged and dispersed in the outflow regions. They can also drift long distances near the ground affecting air quality and human health. These aerosols appear most prominently in summer season in the northern hemisphere. A good example was an event during the summer of 2002, when over 360 fires occurred in the James Bay region of Canada. Sparked by a combination of lightning and dry conditions, these fires consumed hundreds of square kilometers of forest (8871 km²) and created a huge plume of smoke (see Figure 4-3) that affected air quality in much of eastern Canada and, as far south as Washington, DC, in eastern United States.
Once airborne, boreal biomass burning aerosols can circle the hemisphere in 15-20 days in the strong westerly winds of the middle to upper troposphere. They are much longer lived than lower tropospheric aerosols. Their radiative and microphysical effects on clouds are poorly understood. Since northern boreal regions tend to receive most precipitation in the summer period these aerosols are potentially very important.

**Figure 4-3.** In this satellite image taken on July 7, 2002, large fires in northern Quebec are identified as red spots, clouds are white and smoke has a brownish tint. MODIS image from NASA’s Terra and Aqua satellites. Where the smoke plume enters the United States over the state lines of New York and Vermont, it measures almost 300 km wide. The smoke affected air quality and visibility in major cities including Montreal, New York, Philadelphia, Baltimore and Washington, (D.C. Source: Canada Center for Remote Sensing.)

### 4.4.5 South America (SAM)

South American aerosols originate mainly from biomass burning in the Amazon and savannah regions as well as from pollution emitted from mega-cities. Large urban areas such as Sao Paulo, Buenos Aires, Rio de Janeiro, Santiago and others are large sources of aerosol particles. These aerosols, when released in the tropical Amazon basin, drift westwards towards the Andes. When released further south, they drift over the south Atlantic on westerly winds. Recently, the influence of smoke aerosols on the microphysical properties of clouds and convection has been extensively reported (see section 4.3.5).
The biomass burning occurs during the dry season that extends from June to November \cite{Artaxo, et al., 2002}. Huge plumes of particles extending for 3 to 6 million square kilometers with AODs at 500 nm of 2 to 4 are constantly observed. This has strong effects on net energy input to atmosphere and its distribution in the vertical over large areas of South America \cite{Eck, et al., 2003; Procopio, et al., 2004}. In addition, the aerosol pollution affects microphysics and occurrence of clouds \cite{Andreae, et al., 2004; Koren, et al., 2004}. The influence of smoke aerosols on the microphysical properties of clouds and convection has been extensively reported. One simulation result suggests that biomass burning aerosols, peaking shortly before the onset of the South American monsoon, may influence the development of the monsoon \cite{Liu, et al., 2005b}. Monsoon precipitation is part of a complex weather and climate system that affects large portions of southeastern South America. The long range transport is well documented \cite{Andreae, et al., 2001}. Deposition of particles occurs over most of South America area, with important climatic and environmental effects.

4.2.6 Marine Aerosols (MAR)

Marine aerosols consisting of sea salt often mixed with marine biogenic sulfur are another prominent aerosol optical depth maximum (Figure 4-1) in the Earth’s atmosphere particularly over the high wind “roaring 40’s” belt between 40 and 60 S and in the north Pacific and north Atlantic in winter. The AOD-signal of marine aerosol is relatively weak in comparison to the marine aerosol mass present in this region since sea salt aerosols tend to have larger size than pollution aerosols and are thus less effective scatters of light per mass unit. A global model of the distribution of the coarse and fine fraction of sea salt (Figure 4-4) shows clearly that they are prominent in and downwind of windy regions of the Earth’s oceans \cite{Gong, et al., 2002}. The fine particle distribution of sea salt (Figure 4-4b) shows that advection deep into continental regions can occur in some places. For instance, the penetration of marine aerosols into Amazonia is observed in precipitation 1000 km inland where it dominates the concentrations of major inorganic ions \cite{Andreae, et al., 1990}. Coarse marine sea salt particles (Figure 4-3a) provide giant cloud condensation nuclei that can dramatically alter the effect of pollution aerosols on clouds.
Figure 4-4. Predicted global distributions of atmospheric sea-salt Na⁺ mixing ratios in January in the atmospheric surface layer (0-50 m) for two size ranges: (a) Super-micron and (b) Sub-micron [Gong, et al., 2002].

4.3 Aerosol Surface Observations and Regional Characteristics

This section reviews selected observational studies that characterize atmospheric aerosols and their chemical compositions in various regions. Chapter 5 covers additional aspects of in situ and satellite measurements of aerosols.

4.3.1 Global Monitoring Networks

The Global Atmospheric Watch (GAW) program of the World Meteorological Organization consists of 24 Global observatories and approximately 200 Regional stations that measure the
composition of the atmosphere including aerosols, their gaseous precursors and precipitation chemistry. Global observatories host facilities for a broad range of atmospheric observations and research and are often located in remote locations (Figure 4-5) and Regional stations measure at least one of the six GAW target variable groups and complete the global coverage. All observations are made according to GAW measurement guidelines [GAW, 2003]. Measurement of five core aerosol variables is recommended for all GAW stations: fine and coarse mass, fine and coarse aerosol composition, in situ aerosol scattering, in situ aerosol absorption and aerosol optical depth. In addition, a comprehensive list of optional observations is offered including Lidar profiling, number size distributions, and cloud condensation nuclei. The GAW Global observatories Cape Grim, Australia and Mace Head, Ireland have unique long term records of cloud condensation nuclei which are directly related to cloud formation and initial cloud radiative properties.

**Figure 4-5.** The network of the WMO Global Atmosphere Watch (GAW) global observatories.

Recently the GAW program initiated an international effort to identify contributors to long term AOD observations globally and steps needed in quality assurance and coordination to merge them into a global AOD network [GAW, 2004]. Such observations will form the surface-based component of an integrated global aerosol system. With satellite observations, they will yield global products such as the AOD map in Figure 4-1 and be used in predictive weather and climate models. The surface AOD network in March 2004 [GAW, 2004] is
comprised of approximately 90 stations operated by about five different networks (Figure 4-6).

![Global AOD Network Long-term Sites](image)

### Latitudinal distribution
- Polar regions: 4
- Midlatitude North: 50
- Tropics: 26
- Midlatitude South: 10
- Total: 90

### Major data gaps
- Africa, Asia, India, Polar region, Oceans

**Figure 4-6.** The configuration of the long term surface-based aerosol optical depth network as of March 2004 [GAW, 2004] that will be an important component of an integrated global aerosol observation system. It is comprise of a large number of international and national networks.

In North America, several networks have been operational since late seventies and early eighties. They were initiated mostly in response to acid deposition and visibility impairment problems. One of the goals of the Canadian Air and Precipitation Monitoring Network (CAPMon) and the U.S. Clean Air Status and Trends Network (CASTnet) is a description of the aerosol chemical composition across North America. This is greatly augmented in the United States by the IMPROVE (Interagency Monitoring of Protected Visual Environments) network, which is operated by the U.S. Parks Service and the Environmental Protection Agency. It started with stations in most US National parks but later expanded into national coverage including some urban areas. IMPROVE focuses on the monitoring of aerosol chemical properties and their relationship to visibility impairment [Malm, et al., 2002]. It has a comprehensive aerosol chemical characterization network.

Since 1977, the European Monitoring and Evaluation Program (EMEP) has been routinely measuring the major ion composition of aerosols at stations mostly across western and northern Europe [Kahnert and Tarrasón, 2003; Tørseth, 2004; Tørseth, et al., 2002]. In
addition, there is an extensive regulatory network for aerosol mass in Europe operated to monitor compliance with air quality standards. The stations tend to be in urban areas. Aerosol mass and chemical composition monitoring networks have more recently been developed in Asia, including projects underway for the Air Pollution in the Mega-cities of Asia Program and Asiairnet (http://www.asiairnet.org/).

4.3.2 Observations in North America

Figure 4-7 shows the isopleths of the integrated aerosol mass for d<10 µm (PM$_{10}$) and for d<2.5 µm (PM$_{2.5}$) over the three-year period from March 1996 through February 1999 from the IMPROVE network. Higher PM$_{10}$ concentrations are found in the eastern United States. The highest concentration was observed in Edwin B. Forsythe National Wildlife Refuge in New Jersey at 23 µg/m$^3$, followed by the Southeast region, Sequoia National Park, and the Mid South, which experienced concentrations >15 µg/m$^3$. Outside of southern California and the Northern Rockies, the lowest PM$_{10}$ concentrations occurred in the western United States where the concentration of PM$_{10}$ was less than 8.0 µg m$^{-3}$. The spatial variations in PM$_{2.5}$ mimic those observed for PM$_{10}$ (Figure 4-7b) although the spatial gradients tend to be smaller compared with PM$_{10}$; this trend is consistent with the known longer residence time of fine aerosol particles in the atmosphere.

IMPROVE has provided data for aerosol sulfate and nitrate, organic carbon (OC), light-absorbing black carbon (BC), and mineral dust; that is, all of the major aerosol components except sea salt, are measured [Malm, et al., 2004]. Results on fine aerosol (<2.5 um) sulfate from the U.S. are summarized by the isopleths plot in Figure 4-8. The average sulfate component of the fine aerosol measured over the three-year period March 1996 through February 1999 is shown. Since sulfate is a major component of fine aerosol mass, its spatial distribution and gradient across the United States is similar to that observed for total fine aerosol mass as shown in Figure 4-7b.
Figure 4-7. Isopleths of the average mass concentrations (in $\mu$g m$^{-3}$) from the IMPROVE Network (a) PM$_{10}$ and (b) PM$_{2.5}$. Since IMPROVE stations are all in the rural areas, hot spots in big cities are not specifically represented in these plots [Malm, et al., 2004].

Sulfate concentrations (Figure 4-8) typically average $\sim$5 $\mu$g m$^{-3}$ over a broad region of the central eastern United States, amounting to 50 to 60% of the aerosol mass. Sulfate tends to peak in the summer, presumably due to photochemical processing of SO$_2$ from coal-burning sources. Sulfate decreases towards the west, both in terms of absolute concentration and the percent contribution to reconstructed fine mass (RCFM). Organic matter also exhibits its highest concentrations in the eastern U.S., averaging 2 to 3 $\mu$g m$^{-3}$, with some enhancements
in the Southeast due to emissions from fires and secondary particles formed from volatile biogenic organic compounds. Particulate organic matter (POM) accounts for the highest percentage (40 to 50%) of the RCFM in the Northwest and 20 to 30% of the mass over much of the country, and the concentrations there are highest during April to September when fires are most common. Fine soil particles are highest in April, especially in the southwestern/south-central states, roughly from south-central Texas to northern Utah, where the annual average fine soil concentrations are 0.5 to 2 µg m⁻³ and make up 30 to 50% of the RCFM. A secondary maximum in soil dust occurs in August over most of the western US, including the Northern Great Plains and Mid-South. The large-scale spatial and temporal patterns in dust loads observed at the IMPROVE surface sites led Malm et al. [2004] to suggest that large-range transport contributes significantly to the fine particle soil mass. While vertical variations in soil dust concentrations from long-range transport are known and may have influenced the IMPROVE results, the transport of dust from Africa [Perry, et al., 1997] and Asia [VanCuren and Cahill, 2002] to North America has been demonstrated independently.

![Figure 4-8. Average fine sulfate aerosol concentrations (in µg/m³) for each site in the IMPROVE Network, excluding Washington, D.C. [Malm, et al., 2004].](image)

High concentrations of fine particle nitrate (2 µg m⁻³, 20% of the RCFM) were measured south of Lake Michigan, and similarly high concentrations were found in southern California, but there the percent fine nitrate was 20 to 40% of the RCFM due to the relatively low sulfate and organics loadings. Nitrate is highest in the winter when conditions favor its formation. Low concentrations of nitrate are seen in the northwestern US.
Data for particulate matter in Canada have been compiled by Environment Canada into the National Chemical/Particulate Matter Database, with the purpose of investigating spatial and temporal trends in particulate matter, particular as PM relates to climate change, visibility, human health, and acid deposition. Data from the National Air Pollution Surveillance (NAPS) network in Canada are reported annually and they are periodically used to summarize trends, most recently from 1990 to 2001 (Environment Canada, 2004). Analyses of the NAPS data showed 89% of the PM$_{10}$ and 72% of the PM$_{2.5}$ mass was from open sources, including forest fires, and several sources, i.e., agricultural tilling, wind erosion, construction, and road dusts, that involved the generation of mineral particles.

The highest PM$_{2.5}$ concentrations occurred in southern Ontario where the peak levels occurred in June. Prairie sites showed variable levels of PM$_{2.5}$, with peaks tending to occur in August, while in Vancouver, the PM$_{2.5}$ loadings were of the same magnitude but fewer variables than those over the prairies. PM$_{10}$ concentrations were highest at the prairie sites, most likely due to the presence of mineral aerosol [Brook, et al., 1997]. The chemical composition also varies geographically in Canada (Figure 4-9); For example, the organic fraction in PM$_{2.5}$ decreases from 37.6% at Abbotsford on the west coast, to 25.8% at Egbert in Ontario, to 20.4% at St. Andrews in the east while the sulfate fraction increases from 16.9%, 25.6%, 27.0%, to 31.5%, respectively [Vet, et al., 2001]. This fractional difference is primarily due to higher PM$_{2.5}$ concentrations from higher emissions of NOx and SOx in the east and transports of PM and precursors from up-wind regions.

### 4.3.3 Observations in Europe

PM$_{10}$ and PM$_{2.5}$ measurements of the EMEP Monitoring Network shown in Figure 4-10a indicate that the PM$_{10}$ concentrations are the highest in Spain, where annual mean concentrations are typically above 20 µg m$^{-3}$ but can be as high as >40 µg m$^{-3}$. Data from Switzerland and Germany indicate more moderate levels of PM$_{10}$ between 10 and 20 µg m$^{-3}$. Data are lacking in large parts of Europe, in particular the eastern and southeastern parts of the continent. The geographical distribution of annual mean concentrations of PM$_{2.5}$ in 2002, as measured at EMEP sites, is presented in Figure 4-10b [Torseth, 2004]. Due to the limited number of monitoring stations, it is difficult to obtain a generalized spatial distribution pattern for the PM$_{2.5}$ mass. However, the available data indicate that PM$_{2.5}$ was generally
between 5 to 15 µg m$^{-3}$ in most of Spain, Switzerland, and Germany, with higher values observed in northern Italy and eastern Austria. Additional modeling work showed the highest PM$_{2.5}$ would be expected in the Netherlands [Tørseth, 2004].

**Figure 4-9.** Fractional composition of Rural PM$_{2.5}$ measured at Abbotsford (49.0° N, 122.3° W, BC), Ester (51.6° N, 110.2° W, AB), Egbert (44.2° N, 79.8° W, ON) and St. Andrews (45.9° N, 67.1° W, NB) of Canada. The size of the circle represents the magnitude of the total PM$_{2.5}$ concentration in that station [Brook, et al., 1997].
Figure 4-10. (a) Annual averages of PM$_{10}$ concentrations at EMEP sites in 2000 [Kahnert and Tarrasón, 2003] and (b) PM$_{2.5}$ in 2002 at selected monitoring stations in Europe [Tørseth, 2004].

In the future a growing network of continuous PM$_{2.5}$ observations, presently generated by national agencies, will provide European air quality observations for regulatory monitoring, thus increasing the capability of studying links between aerosols and precipitation.

Measurement results of aerosol sulfate in Europe, from the EMEP particulate matter network, can be seen in Figure 4-11 [Kahnert and Tarrasón, 2003]. Similar to North America, aerosol sulfate concentrations in Europe coincide with anthropogenic SO$_2$ emissions, a consequence of coal as the primary energy source. Sulfate concentrations are highest in southern Europe, mostly at sites in Spain and Italy, and in central Europe, mostly in the eastern part of Germany, Poland, Austria and Hungary, where they can be higher than 1.35 µg S m$^{-3}$ (or 4.05 µg SO$_4^{2-}$ m$^{-3}$) on an annual average (Figure 4-11). In Europe, sulfur dioxide emissions increased from the post war years and peaked mainly in the 1960s and 1970s. Following emission control measures, gradual reductions took place in most countries and especially in Western Europe in the 1980s.
Figure 4-11. The annual averages of sulfate mass (µg S m\(^{-3}\)) concentrations for 2000 (a) and 2001 (b) measured at EMEP sites [Kahnert and Tarrasón, 2003].

The Putaud et al. [2003] compilation summarized data from thirty-four sites which were classified as follows: natural background, rural background, near city background, urban background, free troposphere and kerbside based on distances from large pollution sources, vehicular traffic, etc. The background PM\(_{10}\) and PM\(_{2.5}\) mass loadings at those sites averaged 7.0 ± 4.1 µg m\(^{-3}\) and 4.8 ± 2.4 µg m\(^{-3}\), the latter being 5 to 10 times lower than the maximum PM\(_{2.5}\). The persistent background was said to be a result of anthropogenic substances, such as black carbon, transported long distances in addition to particles originating from natural sources. The regional background was concluded to be a major influence on PM concentrations in cities, and this background was at least partially responsible for the observed variability in the PM masses within the various categories of sites. The PM mass data, which showed PM\(_{10}\) concentrations occasionally exceeding 40 µg m\(^{-3}\), were consistent with measurements reported in EMP/CCC 5/2001 of PM\(_{10}\) reaching 50 µg m\(^{-3}\) in Western Europe and up to 70 µg m\(^{-3}\) in Eastern Europe (Figure 4-12). Interestingly, during polluted periods, the increases in aerosol mass were mainly due to PM\(_{2.5}\).
Figure 4-12. Annual average PM$_{10}$ concentrations in different European countries for various site types (units: $\mu g/m^3$) during 1999, from EMEP/CCC-Report 5/2001. The x-axis indicates the country and the number on the bar represents the number of sites under that category.

Indeed, the PM$_{2.5}$ masses at a subset of thirteen sites in the Putaud paper were found to be fairly similar, varying by roughly 4-fold, while in comparison, the particle number concentrations were more variable among this subset of data, differing by a factor of ten. The scales of these differences illustrate the limitations of mass-based data for problems more appropriately addressed by considering particle numbers, as is the case for most issues concerning aerosol/cloud interactions. Particle size distributions reviewed by Putaud et al. showed diurnal influences from motor vehicle traffic and photochemistry, with influences from the regional background on particles about 100 nm in size. These particle size data showed that even the clean sites were affected by pollution aerosol and that during transport particle numbers are reduced by dilution and coagulation, the latter process removing smaller particles from the aerosol population as they collide with large ones.

Further analyses presented in the Putaud et al. compilation showed that the PM$_{2.5}$ masses were linearly related to particle number concentrations at the clean sites, suggesting that the small particles in clean air mostly originate from distant sources. This relationship did not hold at the more polluted sites, however, and consistent with what was noted above, the numbers of particles with diameters (d) greater than 10 nm tended to be more variable than
PM$_{2.5}$ mass. Furthermore, 70 to 80% of the particles had d < 100 nm, and compositional analyses for size-separated samples indicated that these particles were mainly carbonaceous.

Seasonality was evident in the mass loadings at all but one of the polluted sites, that is, both PM fractions were highest in winter. However, this trend was not seen at the natural or rural background sites. The elevated PM loadings in winter were thought to be driven primarily by meteorology, mainly large-scale stratification, but seasonality in emissions and long-range transport patterns and in the chemical processing of condensable species, such as ammonium nitrate and organics, at the lower winter temperatures were offered as additional explanations for the higher wintertime PM.

Chemical analyses showed that ~70% or more of the PM$_{10}$ and PM$_{2.5}$ mass at the sites in the Putaud et al. compilation could be accounted for by the chemical measurements and some simple assumptions relative to elements not measured, such as non-carbon atoms in organic compounds (Table 4-1). In comparison, a larger percentage of the mass, 28 to 46%, could not be identified in the coarse fraction (PM$_{10}$ - PM$_{2.5}$). At a few sites, the PM$_{10}$ and PM$_{2.5}$ fractions were similar in composition, but more often, the two fractions differed chemically.

Table 4-1. “Grand average” composition (%) of the annual mean contributions to PM$_{10}$, PM$_{10}$-PM$_{2.5}$ and PM$_{2.5}$ adapted from EMEP/CCC-Report 5/2001.

<table>
<thead>
<tr>
<th></th>
<th>PM$_{10}$</th>
<th></th>
<th></th>
<th>PM$_{2.5}$</th>
<th></th>
<th></th>
<th>PM$<em>{10}$-PM$</em>{2.5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Natural &amp; Rural Background</td>
<td>Near-city &amp; Urban Background</td>
<td>Kerbside</td>
<td>Natural &amp; Rural Background</td>
<td>Near-city &amp; Urban Background</td>
<td>Kerbside</td>
<td>Natural &amp; Rural Background</td>
</tr>
<tr>
<td>Black Carbon</td>
<td>6</td>
<td>5</td>
<td>13</td>
<td>8</td>
<td>8</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>Organic Matter</td>
<td>16</td>
<td>20</td>
<td>22</td>
<td>23</td>
<td>22</td>
<td>29</td>
<td>4</td>
</tr>
<tr>
<td>Nitrate</td>
<td>7</td>
<td>15</td>
<td>10</td>
<td>5</td>
<td>17</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>Ammonium</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>8</td>
<td>10</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Sulfate</td>
<td>19</td>
<td>13</td>
<td>10</td>
<td>29</td>
<td>17</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Sea salt</td>
<td>8</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Mineral dust</td>
<td>10</td>
<td>9</td>
<td>19</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>Unknown (*)</td>
<td>28</td>
<td>27</td>
<td>19</td>
<td>20</td>
<td>17</td>
<td>15</td>
<td>40</td>
</tr>
</tbody>
</table>

(*) calculated only when all the main aerosol components were measured.

At most of the sites in the Putaud et al. review, sulfate and OC were the main components of PM$_{2.5}$. Sulfate mainly occurred in the PM$_{2.5}$ size fraction, as it often does elsewhere,
accounting for 12 to 30% of the PM$_{2.5}$ mass. The EMEP/CCC-Report 5/2001 indicates that sulfate concentrations tend to be high in central and southern Europe, approaching 1 µg m$^{-3}$, with lower values ~0.4 µg m$^{-3}$ in northern Europe and Scandinavia. Annex 4 of the Putaud et al. [2003] compilation shows non-seasalt sulfate concentrations of 5 to 6 µg m$^{-3}$ in urban regions and 1 to 2 µg m$^{-3}$ in natural and rural areas. In comparison with the IMPROVE values cited above, the concentration and the percent contribution of sulfate to aerosol mass in Europe appears to be slightly less than for the eastern US and more comparable to mid-western and western states.

Carbonaceous particles were the most abundant component of PM$_{2.5}$ on a mass basis at the near-city and urban background and kerbside sites in the Putaud et al. paper, reaching 5 to 10 µg m$^{-3}$ at the most heavily impacted sites. With further reference to carbonaceous materials, Noone et al. [2003] drew attention to the uncertainties in the composition and sources of organic compounds, and they noted that studies current at that time were beginning to show that aerosol organics might play a more important role in the multi-phase atmospheric system than previously thought. These authors highlighted the potential importance of water-soluble organic compounds (WSOC) for cloud formation and cloud properties, citing a study by Saxena and Hildeman [1996] showing that WSOC accounted for 20-70% of the total aerosol carbon.

Overall, mineral dust was the main contributor (20 to 40%) to the coarse particle fraction at all three general classes of the Putaud et al. sites, with seasalt accounting for 5-20% of the coarse particles (Table 4-1). Nitrate and black carbon contributed roughly equal proportions to the PM$_{10}$ and PM$_{2.5}$ fractions, and when the PM$_{10}$ concentrations exceeded 50 µg m$^{-3}$, nitrate and organic matter were the main contributors to the aerosol mass.

In addition to the recognition of WSOC as a source of CCN, studies conducted for EUROTRAC-2 showed that road traffic is an important a source for nanometer-sized particles [Noone et al. 2003 and references therein]. Even though these particles are rapidly lost during transport, the processes involved became better understood as a result of research conducted for the program. Another general conclusion drawn from the EUROTRAC studies was that the algorithms for describing particle emissions from surface processes are not
universal but rather there are differences between the situations in Europe compared with North America.

4.3.4 Observations in Asia

Asia is an immense and diverse source of aerosol particles and trace gases. Mineral dust particles mix with pollutants from various industrial and energy-related sources, biomass burning and transport exhaust to produce an extraordinarily complex regional aerosol mix. In the late winter and early spring of 2001, two major field campaigns were mounted to study the atmosphere over Asia and the western Pacific: TRACE-P (TRAnsport and Chemical Evolution over the Pacific) and ACE-Asia, the third in a series of Aerosol Characterization Experiments (ACE).

Several findings from these programs concerning Asian-aerosol outflow are directly relevant to this aerosol/precipitation review. First, Jordan et al. [2003], Seinfeld et al. [2004], and others have presented data showing that gas-phase species mix with Asian dust particles, making the dust particles more hygroscopic, and presumably affecting the formation of clouds over regional scales. In particular, the Seinfeld et al. [2004] studies highlighted the fact that the structure of the atmosphere is highly complex and that layers of dust, pollution, and biomass burning products can be transported long distances as distinct entities or they can be mixed together. Data obtained with a single-particle mass spectrometer by K. Prather and colleagues aboard the research vessel Ronald H. Brown and reported in Arimoto et al. [2006] indicate that the processes involved in the mixing of sulfate, nitrate, and chloride with dust are competitive or selective in nature. These uptake and mixing reactions not only affect the physical properties of the mineral dust but also affect the number and size distributions of the ensemble aerosol population, especially nitrate-containing but also sulfate-containing particles. In TRACE-P, Kittaka et al. [2004] showed that the mass of sulfate in ice clouds was much less than in liquid clouds and that sulfate particles could be released into the atmosphere following the oxidation of SO₂. Their study showed a close and direct linkage between meteorology and aerosols, that is, accurate predictions of clouds and precipitation were needed to successfully model the atmospheric sulfate budget.

ACE-Asia and TRACE-P also have provided information on new particle formation, particle distributions, Asia outflow, and how areas downstream could be affected by long range
transport, all topics relevant to this review. For example, a study by McNaughton et al. [2004] showed evidence of secondary aerosol particle formation over synoptic scales. These particles formed heterogeneously in a postfrontal situation when dry continental air mixed with maritime air; that is, nucleation occurred when mixing led to supersaturations of condensable species. Comparisons of model results with aircraft observations showed that in areas where dust was concentrated, it composed up to 90% of the super-micrometer particles, and even though dust impacts sometimes extended above 6 km, the influences were strongest in the boundary layer [Tsai, et al., 2004]. These authors also showed correlations between submicron aerosol mass loadings and ethyne, an indicator of combustion sources, and from this they concluded that anthropogenic emissions were the major source for the submicron particles. Similar conclusions have been reached by many others; indeed Clarke et al. [2004] suggest that even though interactions sometimes do occur, the coarse mode and fine mode particles can at times function as nearly independent phenomena, with coarse-mode dust often nearly non-hygroscopic and the fine pollution aerosol much more hygroscopic.

Aircraft studies for TRACE-P and ACE-Asia have shown that aerosol properties vary strongly with altitude, and comparisons with ground-based measurements show that boundary layer measurements generally are not good indicators of the aerosol properties above the surface. More specifically, the aerosol population in the boundary layer (below ~2000 m) is in many ways distinctly different from that in the free troposphere. Pollution aerosol is largely confined to the BL, while dust is distributed much more evenly throughout the column [Clarke, et al., 2004].

One drawback from the above analysis is that these campaigns were conducted over the marine atmosphere outside the mainland China downwind of the major pollution sources. Within the mainland continent, observational data are very limited. Recently, monitoring stations have been established in China to quantify air pollution levels and to better define the problems. Table 4-2 summaries the typical range of aerosol concentrations observed in China from several monitoring stations. Data from the newly developed stations (X.Y. Zhang, Center for Atmosphere Watch and Services, CMA, personal communication) as well as from field campaigns [Zhang, et al., 2005] show elevated aerosol concentrations. The total suspended particle (TSP) concentrations generally range from 50 to 100 µg m$^{-3}$. 

147
Table 4-2. Aerosol concentrations observed at Chinese stations (unit: µg m\(^{-3}\))

<table>
<thead>
<tr>
<th>Station</th>
<th>Lat/Lon</th>
<th>TSP</th>
<th>Sulfate</th>
<th>Nitrate</th>
<th>BC</th>
<th>OC</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin An</td>
<td>30.3°N, 119.73°E</td>
<td>50–70</td>
<td>15–18</td>
<td>2–4</td>
<td>~2.0</td>
<td>10–12</td>
<td>2004</td>
</tr>
<tr>
<td>Zhenbaitai</td>
<td>38.3°N, 109.7°E</td>
<td>45–50</td>
<td>4–9</td>
<td>~2.0</td>
<td>2–5</td>
<td>9–17</td>
<td>2003–2004</td>
</tr>
<tr>
<td>Lian Yun</td>
<td>34.7°N, 119.3°E</td>
<td>100</td>
<td></td>
<td>3–5</td>
<td>~7.0</td>
<td>10–14</td>
<td>2003</td>
</tr>
<tr>
<td>Tong Liao</td>
<td>43.2°N, 122.2°E</td>
<td>78</td>
<td></td>
<td>~2.0</td>
<td>~7.0</td>
<td></td>
<td>2003</td>
</tr>
</tbody>
</table>

High concentrations were also found for other aerosol species shown in Table 4-2. One notable example is the elevated organic carbon content at all those sites. Increased aerosol concentrations over China are reflected in long term decreases in the total global radiation reaching the surface [Che, et al., 2005]. Over the latter half of the 20th century, there have been significant decreases in global radiation (-4.5 W m\(^{-2}\) per decade), direct radiation (-6.6 W m\(^{-2}\) per decade), clearness index (-1.1% per decade), and the percentage of possible sunshine duration (-1.28% per decade). This is at least partially attributable to the increase in the aerosol loading in China. Changes in precipitation at the ground in China from 1951-2000 have been observed in the same areas where the total global radiation decreased [Zhai et al., 2005]. Even though there is little trend in the country-wide precipitation for China, there are distinct regional and seasonal patterns. Annual total precipitation has significantly decreased over southern Northeast China, North China, and over the Sichuan Basin but significantly increased in western China, the Yangtze River valley and the southeastern coast. There are many factors influencing the precipitation variations. Nevertheless, coherent changes in the radiative fluxes and precipitation patterns no doubt contain valuable information on the link between the concentrations of aerosols and patterns of precipitation in China. This needs to be investigated in more detail using the extensive monitoring network observations and numerical models.

4.3.5 Southern Hemisphere Observations

PM\(_{10}\) and PM\(_{2.5}\) have been determined gravimetrically, and more recently, with the use of TEOM monitors at a network of Australian monitoring stations that are mostly located near
the major urban centers along the east and southeast coast. Annual median PM$_{10}$ concentrations are typically 15 to 20 µg m$^{-3}$ in coastal locations in New South Wales, Victoria, Queensland, and West Australia, but appear to be higher in Tasmania while the annual median of PM$_{2.5}$ concentrations range from 5-10 µg m$^{-3}$ [DEH, 2004].

Figure 4-13 presents long term aerosol concentrations in the center of the biomass burning region in South America, showing that concentrations of PM$_{10}$ particles exceeds 300 µg/m$^3$ every year [Artaxo, et al., 2002]. These high aerosol loadings also shows in the AERONET sun-photometer network, presented in Figure 4-14 with a time series of AOD in the central Amazonia. AOD at 500 nm peaks at higher than 3, corresponding to a very high atmospheric aerosol loading. Areas covered by high aerosol concentrations occur not only in Amazonia but also in a large part of the rest of the continent. Figure 4-15 shows a MODIS image of smoke covering much of South America, with AOD approximately 3. The high concentrations of aerosols and the large geographical extent show clearly that smoke is a major driver of weather, climate and air quality in South America.

![Alta Floresta Aerosol Mass Concentration 1992-2000](image)

**Figure 4-13.** Long term aerosol concentrations in the center of the biomass burning region in South America, showing that every year concentrations of PM$_{10}$ particles exceeds 300 µg/m$^3$ [Artaxo, et al., 2002].

The forested areas of South America are also source of natural biogenic aerosol, both primary and secondary, the latter produced from precursors such as volatile organic compounds. Dry areas of Northern Chile and Southern Argentina are significant sources of dust. In addition, emissions from large urban areas such as Sao Paulo, Buenos Aires,
Santiago, Quito, Lima, Rio de Janeiro and others strongly contribute to the overall regional aerosol burden.

Figure 4-14. Time series of aerosol optical depth obtained by an AERONET station in central Amazonia, showing AOD values higher than 3 at 500 nm.

Figure 4-15. MODIS image of smoke from biomass burning covering a large fraction of South America. Image taken in September 25, 2002.
4.4 Vertical Profiles

4.4.1 Vertical Profiles of Aerosols – Surface-based Network

The vertical distribution of atmospheric aerosols are an important property that determines their life time and long range transport potentials as well as their climatic impact. However, due to limitations in measuring equipment and technology, information on the vertical distribution of aerosols over the globe is hardly available not to mention the lack of data on size distribution and chemical compositions as a function of height. Ground-based Lidar networks for aerosols have been established in various locations (Chapter 5). Since these networks are in an early stage of operational monitoring, a global data set for the vertical profiles of aerosols is not readily available. However, some regional characteristics of aerosol vertical profiles can be deduced from the networks.

Figure 4-16 shows the mean aerosol Lidar extinction profiles for northern and southern European stations in 2000 [Matthias, et al., 2004] and for a US station representative of the inner continental background in North America during 2000 [Ferrare, et al., 2006]. These profiles are compared with the global aerosol model outputs from AeroCom. The model results agree reasonably well with the observations. Both observations and models show a distinct aerosol layer at higher altitudes for the southern European stations, which is not present at the northern stations. The simulations attribute this layer to the presence of dust. Aerosol concentrations in the planetary boundary layer are higher at the northern stations; the simulations explain this by increased sulfate concentrations. The aerosol profiles at the US station are similar to the northern European profiles. The average aerosol extinction profiles simulated by the AeroCom models (Section 4.5.2) typically show good agreement with the observations above about 2 km, but are less good in the planetary boundary layer.
Figure 4-16. (a) Comparison of the mean aerosol extinction profiles from Lidar (at 355 nm) with the LMDzT INCA model for northern and southern EARLINET stations in 2000. The observed extinction profile is indicated by the solid blue line. The all-year-average modeled profiles are plotted with a solid red line. The modeled contributions from individual species are shown as well in this figure, as indicated in the legend (adapted from S. Guibert et al. personal communication); (b) Average aerosol extinction profiles over the ARM Southern Great Plains (SGP) CRF (36.62 N, 97.5 W, 317 m height) site during 2000-2001 compared to AeroCom models (see section 4.5.2). Figure taken from [Ferrare, et al., 2006].

Figure 4-17 shows monthly mean extinction coefficient profiles observed at the AD-Net Lidar network site in Sri Samrong, Thailand (17.1° N, 100.0° E) during dry seasons in 2001-2004 [Shimizu, et al., 2006]. The boundary layer height increases from January to April, and the extinction due to aerosols also increases. A marked increase of aerosol extinction occurs from November to February during the burning season. Year-to-year differences are seen in the aerosol extinction and seasonal variation. The aerosol extinction in 2003 was lower than other years. This result was consistent with the co-located skyradiometer data.
Figure 4-17. Monthly mean extinction coefficient profiles in Sri Samrong, Thailand during dry seasons in 2001-2004 [Shimizu, et al., 2006].

One of the most useful applications of Lidar data is to validate aerosol transport models. For this purpose, event-to-event comparisons using a network of continuously operated Lidars is effective. Continuity of the observations and suitable locations of the observation sites (depending on the scale of phenomena) are very important. Aerosol type classification methods using depolarization ratio, wavelength dependence, and the Lidar ratio (if available) are useful for the comparison. Figure 4-18 shows an example of Asian dust transport captured by the Lidars in Beijing, Suwon (37.1° N, 127.0° E), Nagasaki (32.8° N, 129.9° E), and Tsukuba (36.1° N, 140.1° E). Vertical structures and relative magnitudes of the dust storm were clearly observed at those stations. Downwind of the dust source regions, measuring the clouds and aerosol together using a polarization-Lidar, Sassen [2005] detected the long-range transported dust aerosols in 2004 spring at Alaska; these particles served as ice nuclei (IN) in the atmosphere and caused ice clouds to form at modest super-saturations
and temperatures, 4.5 and 6.5 km above sea level. The aerosol profiles peaked at a much higher height above the ground in Alaska than those observed at East Asian stations, reflecting the evolution of dust long-range transport.

![Figure 4-18. Mie-Lidar Dust Extinction Coefficient (S<sub>i</sub>=50) at 4 locations in China and Japan, April 2002.](image)

### 4.4.2 Vertical Profiles of Aerosols – Space-borne

Unlike Lidars that profile the atmosphere above, at a single geographic location, or airborne Lidars that measure above and below aircrafts on regional scales along flight trajectories, spaceborne Lidar allows the study of aerosol transport in regions that are difficult or impossible to explore by other means. As such spaceborne Lidar may be the only means to capture the vertically profile of aerosols in those regions that is essential in understanding the global transport perspective. The era of spaceborne backscatter Lidars started with the successful launch of LITE (Lidar In-Space Technology Experiment) in September, 1994 onboard the Space Shuttle Discovery. There were two purposes to the LITE mission: evaluate technological requirements of a spaceborne Lidar [Couch, et al., 1991] and evaluate the science capability [McCormick, et al., 1993]. LITE convincingly demonstrated the value of spaceborne Lidar in retrieving the vertical structures of clouds [Omar and Gardner, 2001; Winker and Trepte, 1998] and aerosols on a global scale.
Several examples of the global transport of aerosols were evident from the LITE data. Aerosols from natural sources, such as Saharan dust was measured on several orbits by LITE. Although it has been known for quite some time that large quantities of Saharan dust are transported across the Atlantic towards the Caribbean, the unique capabilities of a spaceborne Lidar proved ideal for tracking and quantifying the magnitude of these events [Powell, et al., 1996]. LITE showed these enormous plumes would stretch 100’s of kilometers, reaching altitudes of greater than 5 km (Figure 4-19).

Figure 4-19. LITE orbit showing Saharan dust plume over the Atlantic. Significant structure (as the raw signal counts at 532 nm wavelength) is shown within the plume as it reaches altitudes of 5 km (courtesy of NASA Langley Research Center).

LITE also proved invaluable at pinpointing source regions of anthropogenic aerosols: from small local sources and combined net regional impacts, to large scale activity such as biomass burning. Observations from LITE showed biomass burning in South America extending hundreds of kilometers from the source region. The depth and complexity of these plumes are nicely captured by the Lidar making it possible to apportion their impact and quantify their contribution to long-range transport. Using multiple orbits as LITE processes around the globe made possible the tracking of regional transport. A good example would be the measurement of anthropogenic aerosols leaving the Eastern United States and riding the “gulfstream highway” towards Europe [Hoff, et al., 2001]. An example of a LITE validation track from the Canadian validation component is given in Figure 4-20. This represents an orbital overpass along the Californian coastline while aircraft measurements along the track...
were being made below. The most spectacular feature from the LITE data is the identification of the urban aerosol plumes from San Francisco and Los Angeles (the cities locations are obvious by the position of the plume origin) flowing out along the coastline for several hundred kilometers [Strawbridge and Hoff, 1996]. What is truly remarkable is the 60 s trip for a spaceborne Lidar to go from San Francisco to Los Angeles. Other Lidar data from various aircraft and ground sites provided validation of the measurements, thus increasing confidence in the LITE’s performance.

![Image of LITE data](image)

**Figure 4-20.** A two minute snapshot from LITE along California’s coast showing the long-range transport of pollutants from the Los Angeles and San Francisco areas (courtesy of NASA Langley Research Center).

Among the most significant contributors to the radiative budget are the aerosols (and water vapor) found in the planetary boundary layer (PBL). Lidars have been used for several years to determine PBL height [Melfi, et al., 1985; Piironen and Eloranta, 1995; Strawbridge and Snyder, 2004] because the large gradient in aerosol concentration that occurs between the top of the PBL and the free troposphere.

The LITE technology provided the backbone for more recent satellite-based Lidars; e.g. CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation) [Winker, et al., 2002] which will provide 4-D distributions of aerosol backscattering in the atmosphere after its launch in April 2006 (See Chapter 5 for more details).
4.5 Transport of Atmospheric Aerosols

Transport is responsible for the spread of aerosols from their source locations to downwind destinations caused by dispersion and convection related to turbulence and by advection on winds. Dispersion is controlled by small-scale turbulent processes that depend on atmospheric stability and advection by the large-scale general circulation patterns.

4.5.1 Characteristics of Global Aerosol Transport

Long-range horizontal transport is important, for example, for dust aerosols being transported from Africa and Asia to North America [Gong, et al., 2003b; Gong, et al., 2006; Husar, et al., 2001; Prospero, 1999]. The spatial and vertical distribution of Asian aerosols was a main focus of the 2001 ACE-Asia study [Huebert, et al., 2003] as discussed in 4.3.4. In terms of aerosol-cloud interactions, long-range transport is important for the transformation of aerosols with consequences for their ability to act as CCN and IN [Levin, et al., 1996]. A study by DeMott et al. [2003a] in which African dust aerosols were sampled in the east part of the US, concluded that these particles were very effective IN.

Pathways for trans-Pacific and trans-Atlantic transport differ in the following ways. For trans-Pacific transport, the highest concentrations in the outflow are in the boundary layer (0-2 km), but the strongest outflow fluxes are in the lower free troposphere (2-5 km) and reflect episodic lifting of pollution over central and eastern China ahead of eastward moving cold fronts. This frontal lifting, followed by westerly transport in the lower free troposphere, is the principal process responsible for export of both anthropogenic and biomass burning pollution from Asia [Bey, et al., 2001]. In contrast, trans-Atlantic transport of North American pollutants takes place in the lower troposphere year round, and transport in the middle and upper troposphere is also important in summer [Li, et al., 2002]. Compared to trans-Pacific transport of Asian pollution to North America, which takes place mainly in the free troposphere followed by subsidence [Jacob, et al., 1999; Yienger, et al., 2000], the relatively short distance between North America and Europe and the prevailing westerly flow extending down to the surface favor transport in the boundary layer. Over eastern Asia, in contrast, westerly winds do not extend down to the surface, so that lifting to the free troposphere is a prerequisite for trans-Pacific transport [Bey, et al., 2001].
Long-range transport to the Arctic is also of importance, because the background atmosphere, especially in summer, is very clean. *Girard and Curry* [2001] point out that the concentration of IN, aerosol number concentrations, the slope of the aerosol size distribution, and the aerosol solubility, may impact substantially the cloud phase and total water content. A mesoscale model simulation by *Lohmann et al.* [2003] showed that the addition of anthropogenic aerosols to the background aerosol in the Arctic not only impacts liquid water clouds and drizzle, but can also alter snowfall rates. In this study, depending on the assumed snow crystal shape, which determines the accretion rate, the total amount of precipitation reaching the surface after 7 hours of simulation can be larger or smaller than the polluted case compared to the clean background.

The transport of atmospheric aerosols from their source regions is also complicated by the seasonal variation of transport patterns and removal processes [*Barrie*, 1990]. The strong seasonality of Arctic aerosols is a function of their atmospheric lifetime which is controlled by the variations in transport and removal processes [*Barrie*, 1986]. Because of stable thermal stratification in the atmospheric surface boundary layer and less than 10-mm liquid H₂O/month of frozen precipitation in the cold half of the year from October to May, aerosol residence times are much longer in winter (~3 to 7 weeks) than in summer (~3 to 7 days). Therefore, the seasonality of any aerosol species in the Arctic is more likely dominated by factors such meteorology, transport and removal processes specific to the major transport pathways and less to seasonal variations of emissions in the mid-latitude source regions of these species.

In addition, inter-continental transport is also modulated by climate fluctuations. By studying the transport of Asian dust to North America for the last 44 years with a numerical model, *Gong et al.* [2006] identified shifts in the inter-continental pathways for dust aerosols related to the El Niño Southern Oscillation. In that study, an analysis of the variability of the 44-year zonal mass transport flux (i.e. concentrations multiplied by westerly/easterly wind speeds) was conducted for eight typical El Niño and eight La Niña years. The zonal mass transport flux indicates the major transport routes and direction of Asian dust across the Pacific. Most trans-Pacific transport of Asian dust aerosol occurs in the middle troposphere between 2 and 5 km [*Zhao, et al.*, 2006]. Figure 4-21 shows the averaged dust transport flux (filled contours) for El Niño (Figure 4-21a) and La Niña (Figure 4-21b) years integrated from 3 to
10 km. On the same plot, eight year anomalies (dashed contour lines) from the 44 year averaged values are also superimposed.

(a) El Niño

(b) La Niña

**Figure 4-21:** The zonal transport fluxes ($\mu$g.m$^{-2}$.s$^{-1}$) of Asian dust in spring of (a) 8 El Niño years (1966, 73, 83, 87, 92, 98 and 2003) and (b) 8 La Niña years (1965, 71, 74, 76, 86, 89, 99 and 2000). The contour lines are for the anomalies relative to the 44-year mean of Asian dust aerosol in spring [Zhao, et al., 2006].

A sharp difference between zonal transport during El Niño and La Niña years is most clearly seen in the center of the transport path. During El Niño years, trans-Pacific transport is centered at 45° N while during La Niña years it is around 40° N. The zonal transport fluxes of Asian dust aerosol through 125° E in spring for the same El Niño and La Niña years indicate a shift of the main out-flow point in Asia from 45° N to 40° N during La Ninas. In contrast, in the eastern Pacific (140° W) where the outflow reaches the coast of North America, peak transport moves southward from 45° N to 40° N during El Nino years and to lower heights (5000 m during El Niño versus 6500 m during La Niña years).
Interannual variability of African dust transport over the north tropical Atlantic has been investigated by Chiapello et al. [2005] who analyzed in situ surface concentration data from Barbados beginning 1966, along with the Total Ozone Mapping Spectrometer (TOMS) and Meteosat dust optical thickness (DOT) records covering the last two decades. Their analysis shows a large regional impact of drought conditions in the Sahel on dust emissions and transport both in winter and in summer.

4.5.2 Numerical Modeling Studies - AEROCOM Results

As many quantities of aerosols including their transformation and spatial distribution on the global scale cannot completely be covered by observations, numerical models are used to integrate the current knowledge and to better understand aerosol processes. AeroCom (http://nansen.ipsl.jussieu.fr/AEROCOM/), an initiative to compare aerosol modules in global models among each other and to observations, started in 2003 to systematically identify and quantify uncertain elements in global aerosol simulations. An analysis of the simulated aerosol life cycles for 16 global models shows large diversities especially in spatial distributions, aerosol composition, and water uptake [Textor, et al., 2006]. Results from this study are summarized below.

The zonally averaged vertical concentrations of the total (component-combined) aerosol mass for the AeroCom models are shown in Figure 4-22. The aerosol concentrations are dominated in mass by the natural aerosols: dust contributes 50% to 80% to mass in all models except two models where sea-salt is dominant. All models show two maxima, the one in the northern hemisphere resulting from dust and that in the southern hemisphere from sea-salt in the ‘roaring forties’ of the southern oceans.
Figure 4-22. Zonally and annually averaged concentration of total aerosol in [µg m⁻³], the model name and the global average are given on top of the plots. Please note, we use a non-linear color scale. The white shading of lowest layer above ground in some models indicates that no data have been available in this layer. (Figure adapted from Textor et al. [2006])

Models agree least on the vertical transport of sea-salt, followed by dust, particulate organic matter, black carbon and sulfate. Of the total aerosol mass, 14% is placed above 5 km height and this varies from 5% to 26%. The mass composition close to the surface, i.e., below 1 km,
is dominated by sea-salt in eight models and by dust in seven models. In this layer contributions to the total mass vary for sea-salt between 20% and 80% and for dust between 15% and 70%. All models have a dust mass maximum in the upper PBL between 1 and 2.5 km. Vertical transport is relatively weak for the natural species. The sea-salt mass decreases strongest with height. The all-models-average mass fractions of sea-salt and dust above 5km are 9% (varying from 0.001% to 25%), and 14% (varying from 3% to 28%), respectively. The annual global average mass fractions above 5 km height are indicated in Figure 4-23. This height has been chosen as representative for aerosols above clouds where they have increased life times due to the absence of wet removal processes. The ranges of vertical transport, i.e., the differences for each model between the species with the largest and the smallest mass fractions above 5 km, respectively, are indicated by the gray shadings in Figure 4-23. These ranges, which indicate the degree of similarity of the vertical transport among the species within a given model, differ among models; the all-models-average range is 25%, varying from 10% to 45%.

![Figure 4-23](image)

**Figure 4-23.** Global, annual average mass fractions in [% ] of total mass above 5 km altitude for the AeroCom models. The gray shadings frame the range for each model [Textor, et al., 2006].
The vertical transports of black carbon and particulate organic matter are more similar to each other than to the other aerosol components in most models. This is because both are mainly contained in the fine mode in contrast to dust and sea salt, they are less soluble than fine sulfate, and they have similar source regions. The two species are well dispersed in the vertical with the maximum situated in the upper PBL between 1 and 2.5 km in all models. The all-models-average mass fractions above 5 km are 22% (varying from 6% to 38%) and 21% (varying from 6% to 40%) for black carbon and particulate organic matter, respectively. In eight models, the vertical transport is stronger for black carbon than for particulate organic matter, in six models the situation is reversed, and in two models similar.

Sulfate is present at the highest altitudes of all species in most models. The all-models-average sulfate mass fraction above 5 km is 33% (varying from 4% to 47%). The main sulfate burden is situated in the upper PBL between 1-2.5 km in nine models, in the lower free troposphere between 2.5 and 5 km in three models, in the upper free troposphere between 5 and 10 km in one model, and above 10 km in the tropopause region in three models. In general, the sulfate-contribution to the total mass composition becomes increasingly important with height due to the removal of dust and sea-salt, and due to chemical sulfate-production at greater altitudes within the atmosphere. AeroCom predictions confirm the results of an extensive analysis of sulfur cycle simulations in large scale atmospheric models which was performed in the COSAM exercise [Barrie, et al., 2001; Lohmann, et al., 2001; Roelofs, et al., 2001]. The uncertainty in predicting the global sulfate distribution is related to vertical mixing of emitted sulfur species from the planetary boundary layer into the free troposphere. In addition, cloud physics and cloud distributions play a major role as they influence cloud-related processes, i.e., wet deposition and the aqueous oxidation of SO$_2$ to sulfate. Furthermore, model agreement is low for the dry deposition of SO$_2$. The chemical production of sulfate at high altitudes explains why its mass fractions at higher altitudes are larger than those of dust, although these components have similar atmospheric residence times.

A similar diagnostic as for the vertical transport can be performed for the horizontal distribution of aerosol. The mass fractions in Polar Regions can serve as an indicator for the horizontal transport, i.e., for meridional long-range transport, because Polar Regions are far from the aerosol sources. We show the mass fractions of aerosols south of 80° S and north of 80° N in Figure 4-24. The all-models-averages of the mass fractions at the poles in relation
to global aerosol are 2.4% (varying from 0.3% to 8%) for total mass of five aerosol types, 6% (varying from 0.2% to 10%) for sulfate, 4.2% (varying from 0.4% to 10%) for black carbon, 3.3% (varying from 0.2% to 18%) for sea-salt, 3.3% (varying from 0.3% to 7%) for particulate organic matter, and 1.5% (varying from 0.03% to 4%) for dust, respectively. The AeroCom results show major differences in aerosol transport and more observations especially in remote regions and vertical profiles are needed to constrain aerosol distributions.

![Polar (80 degree) fraction for Load](image)

**Figure 4-24.** Global, annual average mass fractions in [%] of total mass in polar regions (south of 80°S and north of 80°N) for all AeroCom models. The gray shadings frame the range for each model [Textor, et al., 2006].

Within the AeroCom model intercomparison, simulated aerosol parameters (optical depth, concentrations, and extinction coefficients) have been compared to various observational data sets (see AeroCom web interface for individual comparisons). This analysis did however not allow for a simple ranking of models, because model performance depends strongly on the parameter, region and process under investigation. A model that performs well for one dataset may show inconsistent results for another.
Models perform reasonably well when compared to the component-combined aerosol optical depth (AOD) [Kinne, et al., 2006]. The annual global mean of the simulated AODs are with 0.11 to 0.14 at the lower end of global averages suggested by AERONET (~ 0.135) and satellite composites (~0.15). More detailed comparisons, however, revealed that larger differences in regional distributions and in compositional mixture remain.

Aerosol size and its micro-physics are also simulated by some sophisticated models within the AeroCom. The results show a large diversity in the fine aerosol fraction (radius smaller than 0.5 microns), especially for dust and sea salt [Textor, et al., 2006], indicating a large variations in aerosol size predictions. Therefore, verification of the resulting size distribution is still in an early stage. A comparison of the simulated global annual mean composition of ambient aerosol has also shown major disagreement among models, especially concerning the water content.

The global aerosol models compare less well to measured aerosol surface concentrations than to AOD, in terms of both spatial and temporal correlation and bias. Results indicate that these models overestimate the surface sulfate concentration by about 40% (correlation coefficient of 0.6) and underestimate POM concentrations by about 15% when compared to station data from the US (IMPROVE network) and Europe (EMEP network) [S. Guibert, et al., personal communication]. Simulated black carbon concentrations match the level of concentration as observed, but the comparison is just based on the IMPROVE network so far. Mean correlation between modeled and observed monthly mean concentrations is smaller than for sulfate with 0.49 and 0.58 for BC and POM, respectively.

For dust, the data from the 14 remote sites in the AEROCE network have been used as available from the University of Miami Aerosol Group for the years 1983 to 1998. Observed dust concentrations are slightly underestimated by the models (12%), but average correlation is quite good (r=0.73). However, the limited number of dust measurements would not allow to judge on the performance of the dust modeling in AeroCom. This statement is also true for the other aerosol components, and more data and more detailed analysis are necessary to allow for reliable statements about individual and ensemble model performance.
4.5.3 Numerical Modeling Studies - DMIP Results

An intercomparison study involving eight mineral dust/emission models over Asia (DMIP) has been evaluated [Uno, et al., 2006] in order to quantify model differences resulting from the variety of implemented dust emission schemes, model resolutions, different deposition parameterizations and numerical methods. The models were run for two major dust storms developed in spring 2002 over East Asia. The model results were evaluated against in-situ concentration measurements, satellite aerosol index data, Lidar observations and synoptic visibility data in the region.

Figure 4-25. Surface dust concentration (Log $\mu g/m^3$) valid for 7 April 2002 at 03 UTC, as predicted by eight dust models participated in DMIP. The lower-right corner map shows observations [Uno, et al., 2006]

Figure 4-25 shows that all models produce surface concentration patterns that in general well coincide with regions of observed maximum concentrations. However, the concentration levels from models are quite different. For example, differences of about one order of magnitude appear over the downwind Beijing region. Similar variations also occur in the
source areas. A large variability between models exists in the amounts of dust emitted into the atmosphere as well. During one of simulated dust storm, models produced total dust emissions ranging from 27 to 336 Tg.

Figure 4-26. Time-height cross sections of dust concentration in Beijing during the March 2002 dust episode, as simulated by eight models; Lidar data is also shown for comparison (the lower-right corner map) [Uno, et al., 2006]

Within the DMIP exercise, simulated vertical concentration was compared with the Lidar measurements available in the region. As Figure 4-26 indicates, the most of the models shows increased concentrations in the second half of the shown period of March 2002 when a major dust transport approached Beijing. However, there are significant diversities in details in the vertical-time structure of models. Most of the models overestimate the vertical extent of concentration, some of them retaining the dust too elevated.

This study demonstrated that although general agreement of models is achieved in reproducing major geographic extent and position of maxima of the surface concentration, they experience large differences in concentration values and in emission fluxes. Variations in vertical dust structure are also obvious. There are numerous possible reasons for the diversities such as different parameterizations and numeric solutions incorporated in such
dust models. The differences are also due to a serious lack of relevant dust measurements that may improve both model dust initialization and model verification.

4.6 Physical and Chemical Transformations

Even though there is an increasing evidence that aerosol particles are predominantly a conglomerate of different internally mixed chemical substances (i.e., each particle is composed of more than one chemical species) [Cziczo, et al., 2004; Kojima, et al., 2004; Murphy and Thomson, 1997]. However, the knowledge about the composition of aerosols on the global scale is still incomplete and most GCMs (global climate models) still treat aerosols as external mixtures (i.e., each particle is composed of only one chemical species) in terms of their optical properties (e.g., [Lohmann, et al., 1999; Textor, et al., 2006]). Advanced aerosol modules in some GCMs have been expanded to include aerosol internal mixtures (e.g., [Ghan, et al., 2001; Gong, et al., 2003a]).

After aerosol particles are formed, they undergo chemical and physical transformations as they move through the atmosphere. Collision and coagulation due to Brownian motions is most important for the nucleation and Aitken-mode aerosol particles. Larger particles can collide if they have a substantial difference in gravitational settling velocity and phoretic forces. Also, vapors can condense on pre-existing aerosol particles, which is especially important for converting aerosol particles from being externally mixed to being internally mixed. Chemical transformation refers to chemical reactions on the surface or in the volume of aerosol particles that can increase aerosol mass, and/or change the ability of aerosols to act as CCN or IN. In-cloud oxidation and cloud processing can also produce new aerosol mass and modify the size distribution. All these processes make aerosols more hygroscopic. This section reviews the processes modifying aerosols in the atmosphere and assesses their implications to the cloud and precipitation formations.

4.6.1 Aerosol Microphysics - Nucleation, Condensation and Coagulation

When condensable species such as H$_2$SO$_4$ (g) are formed in the gas phase by chemical reactions, they participate in two competing processes: nucleation and condensation on pre-existing particles depending on the particle size, vapor concentrations and source rates, and
concentration, as well as relative humidity and temperature. It is apparent that nucleation and condensation can occur at the same time. Condensation is responsible for the growth of atmospheric particles and for changing the aerosol chemical composition. When an aerosol exists in a supersaturated vapor environment, the vapor condenses on the particle population, resulting in a change of the particle size distribution. The supersaturation is usually produced by cooling of saturated vapors via adiabatic expansion (atmospheric clouds) or by gas-phase chemical reactions yielding products of low vapor pressure (e.g., in photochemical smog or in oxidation of biogenic hydrocarbons). In the case of soluble particles, condensation can take place even in unsaturated conditions [Hinds, 1982]. On the other hand, higher saturation ratios may be needed for condensational growth of very small particles because the attractive forces between surface molecules are smaller due to sharply curved surface so that vapors more easily evaporate (Kelvin effect).

The condensational growth of aerosol particles is accompanied by simultaneous mass and energy transport. Analytical expressions for the condensational growth rate of a single particle have been presented by a number of researchers [Kulmala, 1993; Kulmala, et al., 1989; Mason, 1971]. Once the growth rate of a single particle is known, the change in particle size distribution can be obtained by solving the condensation equation [Park and Lee, 2000; Seinfeld and Pandis, 1998]. The condensational growth rate of a single particle is proportional to particle surface area (free-molecule regime) or to particle diameter (continuum regime). Its implication is that the particle size distribution becomes narrower as a result of condensational growth.

Nucleation events in highly polluted regions can produce CCN [Laaksonen, et al., 2005]. 2.5 years of continuous nucleation observations from San Pietro Capofiume, Italy (44°39′N, 11°37′W), showed that nucleation events occur frequently in the Po Valley region, even though the region is rather polluted with high pre-existing particle concentrations. The nucleation events are often very intensive, and the newly formed particles can grow to sizes as large as 100–200 nm in diameter within a few hours and thus constitute an important source of CCN.

Coagulation is a process whereby aerosol particles come into contact because of their relative motion and then stick together to form larger particles. During the coagulation process, the number of particles decreases and the average particle size increases while the total mass of
particles is conserved. High coagulation rates among disparate particles have another important implication, which is the change of mixing state of particles. Particles originated from different sources often have different size and chemical compositions. Therefore, selected coagulation between unequal-sized particles will convert externally to internally mixed aerosols. This results in changes of optical properties and hygroscopic growth characteristics of the particles.

To simulate the change in the particle size distribution due to simultaneous nucleation, coagulation and condensation, it is necessary to resort to numerical methods. Among the most widely used numerical methods for aerosol microphysics is the modal dynamics method [Whitby and McMurry, 1997], which expresses the particle size distribution as a sum of two or three log-normal distribution functions, and the sectional method [Gelbard, et al., 1980], in which the entire particle size range is divided into a finite number of sections [Gong, et al., 2003a].

4.6.2 Chemical Modification by Non-reactive Heterogeneous Uptake

For the adsorption of semivolatile organics onto aerosols, particularly for persistent organic pollutants (POPs), the partitioning of the organics between the gas and aerosol phases is often modeled using the Junge-Pankow partitioning theory [Junge, 1977; Pankow, 1987; Pankow, 1994a; 1994b] under non-saturated vapor pressure conditions. The incorporation of organic species into the particle phase increases with pre-existing aerosol mass and is generally favored at lower temperatures [Pankow and Bidleman, 1991]. This temperature dependency is likely a contributing factor to the observed increase in the percentage of aerosol mass attributed to organic compounds with height in the troposphere [Jaffe, et al., 2005].

Some of oxygenated compounds can be directly emitted, particularly from bio-mass burning sources. However, most hydrophilic organic compounds are often formed in the atmosphere through gas phase photochemical oxidation of less hydrophilic precursors, for example, as hydrocarbons are oxidized to alcohols, ketones, aldehydes, and finally acids, as the oxidation progresses toward a breakdown of the entire molecule. The end results are that the originally hydrophobic aerosols become more hygroscopic and more effective as CCNs.
4.6.3 Chemical Modification by Reactive Heterogeneous Uptake

There has been a great deal of attention given recently to heterogeneous and multiphase reactions occurring on the surfaces and within tropospheric particles [Ravishankara and Longfellow, 1999]. For example, increased relative humidity enhances the reaction of CaCO₃ in dust with HNO₃ [Goodman, et al., 2000] and the reacted particles have been shown to be more hygroscopic than the un-reacted dust and are thus better CCN [Krueger, et al., 2003; Laskin, et al., 2005]. There is also evidence that SO₂ can be adsorbed onto mineral particles and oxidized to sulfate, creating mixed particles that may have complex hygroscopic behavior depending upon which cations are associated with the sulfate [Usher, et al., 2003]. Dust is also frequently associated with organic matter, although the mechanisms creating such mixtures and their effects on dust hygroscopicity are not well understood.

Deviations from model predictions by the Junge-Pankow partitioning theory, which is based on equilibrium are often observed for organic gases and organic particulate matter in the atmosphere [Jang and Kamens, 2001a; 2001b]. The main reason for the deviation is that the absorbed product can further react within the host aerosols, thereby significantly increasing the capacity of the aerosols for continued uptake of the organic gases. Many aldehydes form polymers in aqueous solution under acid catalysis. Such polymerization has been proposed to occur in atmospheric aerosols [Gao, 2004; Jang, et al., 2003; Jang and Kamens, 2001a; Kalberer, et al., 2004; Noziere and Riemer, 2003; Olson and Hoffmann, 1989; Tobias and P.J. Ziemann, 2000]. Recent laboratory studies have shown that indeed for certain aldehydes, polymerization can occur at significant reaction rates [Liggio, et al., 2005a; 2005b]. Liggio et al. [2005b] observed the process of polymerization of glyoxal, a dialdehyde commonly encountered in the atmosphere from both anthropogenic and biogenic precursors, on a time scale of one to three hours in acidic aerosols and estimated that in a typical urban-suburban atmospheric environment, the glyoxal uptake alone would be sufficient to account for up to 10% of aerosol organic contents. Oligomers up to trimers were observed under conditions that can be directly applied to the atmosphere in less than one hour. This polymerization leads to a significant growth in the size and mass of the aerosols. It has been demonstrated that freshly-formed secondary organic aerosols from anthropogenic and biogenic precursors in a smog chamber continued to undergo oligomerization to higher-molecular-weight compounds over 28 h, and the particles exhibited increasing thermal stability with aging [Baltensperger, et al., 2005]. Interestingly, although the hygroscopicity of the particles (at
increased during the first few hours, it remained virtually unchanged during the remainder of the experiment and was similar to that observed for humic-like substances (HULIs) that have been identified in atmospheric aerosols [Baltensperger, et al., 2005; Brooks, et al., 2004].

4.6.4 Black Carbon Aging Processes

Arguably the most complex constituent of the tropospheric aerosol is the carbonaceous fraction, which includes both soot and myriad organic species. Many studies have shown that soot, although initially quite hydrophobic unless formed during combustion of high-sulfur containing fuels [Popovitcheva, et al., 2001], becomes more hydrophilic as it ages through interaction with atmospheric gases such as SO₂, sulfuric acid, and nitric acid, and through oxidation [Zuberi, et al., 2005] in addition to non-reactive uptake and condensation/coagulation processes as discussed above. Chand et al. [2005] measured the hygroscopic growth of particles produced from burning of biomass in the laboratory, and noted that the growth was less than observed for aged smoke particles from similar origins. It has also been demonstrated that acidic particles catalyze the formation of secondary organic aerosol species from the gas phase on their surfaces [Gao, 2004; Jang, et al., 2002], a mechanism whereby an acidic inorganic particle could become internally-mixed with organic species.

The aging process is dominated by the condensation of sulfuric acid with the time scale of about 8 hours near the source region in summer during the day, while in winter ammonium nitrate becomes more important at the same time scale [Riemer, et al., 2004]. Equations of first order decay rate as a function of height and time of the day can be derived from these data. Such equations have been implemented into the Canadian climate model [Croft, et al., 2005] where the BC aging is assumed to be a first order decay and the aging e-folding time is parameterized as the coagulation of insoluble BC particles with soluble particles, the condensation of sulfuric acid and organic compounds, and the oxidation of a BC coating by ozone. A modal aerosol module coupled with GCM ECHAM5 model was developed by Stier et al. [2005], using detailed microphysical approach for aerosol modeling through condensation and coagulation [Vignati, et al., 2004]. The impact of BC aging processes from different schemes differs substantially on the removal and activation processes of BC
aerosols [Croft, et al., 2005]. Liu et al. [2005a] suggest, based on a modeling study that included a detailed representation of aerosol microphysics, that most carbonaceous aerosols in the troposphere are internally mixed with sulfate, nitrate, and secondary organics and are thus more hygroscopic and have significantly shorter lifetimes than predicted using typical hydrophobic-to-hydrophilic conversion rates. Clearly, more work is needed before atmospheric aerosol aging processes and their consequences for particle hygroscopicity and CCN activity are adequately represented in chemical cycle and climate models.

4.6.5 Aerosol Thermodynamics

The gas-particle partitioning and solubility of organic and inorganic species is critically important to our ability to accurately predict the water uptake of aerosols and clouds/precipitation formation. Nenes et al. [1998] compared the equilibrium composition and partitioning of aerosols as calculated by different thermodynamic models – including their ISORROPIA, as well as the SCAPE model of Kim et al. [1993b] and SEQUILIB of Pilinis and Seinfeld [1987] – for representative urban, marine, remote continental, and non-urban continental mass loadings taken from Heintzenberg [1989] and Fitzgerald [1991]. These thermodynamic models consider the inorganic aerosol system of sulfate, nitrate, ammonium, chloride, sodium, potassium, calcium, magnesium, carbonate, and water. They generally agree within several percent except at high relative humidities, where differences in concentrations of particular ionic species differed by as much as 30%. The discrepancies in ionic content relate to the varying treatments especially of bisulfate, equilibrium constants, thermodynamic data, water content, and deliquescence. Using the coupled IMAGES (Intermediate Model of Global Evolution of Species) -SCAPE2 model, Rodriguez and Dabdub [2004] found that nitrate concentrations predicted compare better with observations than those calculated with IMAGES alone. They also find that over polluted continental regions, the presence of sea-salt and dust aerosol potentially increases the formation of aerosol sulfate by 20–80%. Aerosol nitrate formation is enhanced by 14–60%, whereas ammonium formation is decreased by 20–60%.

Organic aerosols and condensable organic species are observed in all environments, and they are the dominant aerosol-phase species in some. Organic species also no doubt contribute to the composition of many primarily inorganic particles [Abdul-Razzak and Ghan, 2000]. However, the state of the science is such that the great majority of the organics present in
aerosol are unidentified. For those that have been identified, the lack of measurements of solubilities and their dependence on inorganic solute concentration is imposing, not to mention the lack of a suitable unified thermodynamic formulation for mixed inorganic/organic aerosol [Clegg and Seinfeld, 2004; Cruz and Pandis, 1997; Saxena, et al., 1995; Turpin, et al., 2000]. The impact of the chemical compositions on aerosol activation is discussed in Chapter 7.

4.7 Removals of Atmospheric Aerosols

The average residence times of atmospheric aerosols are on the order of a few days to about two weeks, depending on their size and location. Aerosols are removed from the atmosphere by wet and dry deposition. Wet deposition removes aerosols from the atmosphere in precipitation with increasing efficiency for soluble aerosols. In the absence of precipitation, smaller aerosols are deposited to the Earth surface within turbulent eddies. Larger aerosols settle out of the atmosphere very quickly under gravity, and some surfaces are more efficient at capturing aerosol than others. Both wet and dry removal processes need to be parameterized in global climate models and which results in rather large uncertainties in budgets. We will first examine some removal pathways before looking at the differences in atmospheric residence for different aerosol components.

4.7.1 Size-dependent Dry Deposition of Aerosols

The process of particle dry deposition includes several mechanisms. Transported from the free atmosphere down to the viscous sublayer that envelops all surfaces is the first process. This happens due to passive transport within turbulent eddies, or for larger particles through sedimentation. They then are transported across the viscous sublayer by Brownian diffusion, phoretic effects, inertial impaction, interception or sedimentation. Once across the viscous sublayer, particles will then interact with the surface, either sticking to it or bouncing off [Ruijgrok, et al., 1995]. A small fraction of very fine particles may also diffuse into leaf stomata. After collected by leaf surface, water-soluble particles may be gradually removed by wetting and insoluble particles may become embedded in epicuticular waxes. Thus, the characteristics of the atmosphere, the nature of the surface, and the physical/chemical properties of the atmospheric aerosols all affect dry deposition [Ruijgrok, et al., 1995; Wesely and Hicks, 2000; Zufall and Davidson, 1998]. Although the dry deposition process is a
slower process compared to wet deposition, it is a continuous event happening anytime over any surfaces. The accumulated amount of dry deposition can be much more important for some aerosol species/sizes [Ginoux, *et al.*, 2001], especially for the coarse aerosol types, sea salt and dust. It has been estimated that dry deposition contributes about 2/3 of the total deposition flux as shown for the AeroCom models [Textor, *et al.*, 2006].

Most measurements on aerosol flux made between 1970-1990’s are made with bulk sampling due to the limitations in aerosol measurement techniques [Nicholson, 1988; Sehmel, 1980; Sievering, 1989]. The development of more sophisticated measurement techniques in 1990’s made it possible to measure size-resolved aerosol physical and chemical properties. The number of studies making direct aerosol fluxes using eddy correlation methods also greatly increased [Gallagher, *et al.*, 2002]. Based on the measurements, dry deposition models were developed to calculate the dry deposition velocity $V_d$ as a function of particle size (e.g., [Giorgi, 1986; Nho-Kim, *et al.*, 2004; Zhang, *et al.*, 2001]. Estimates from several earlier models [Ruijgrok, *et al.*, 1995] revealed that these models differ from each other greatly and the largest uncertainty is for the 0.1-1.0 $\mu$m particle size range, which is not important for mass fluxes but is important in terms of number concentrations.

Dry deposition velocity ($V_d$) generally increases with particle size for particles larger than 2 $\mu$m because the collection efficiency by interception and impaction increases with particle size. The collection efficiency is a parameter defined as the ratio of the total number of aerosols reaching the surface to the total number of aerosols actually sticking to it. Particles larger than 2 $\mu$m tend to have a deposition velocity as high as a few cm s$^{-1}$ over rough forests (e.g., [Hofschreuder, *et al.*, 1997]). $V_d$ also increases with the decrease of particle size for particles smaller than 0.1 $\mu$m since the collection efficiency by Brownian diffusion is larger for smaller particles. Available measurements show that $V_d$ of ultra-fine particles (< 0.01 $\mu$m) at field conditions can be higher than 5 cm s$^{-1}$ [Schery and Whittlestone, 1995]. Deposition velocities for particles in the size range of 0.1-2 $\mu$m are lowest and can vary by several orders of magnitude. Many earlier laboratory experiments suggest that $V_d$ for this size range particles should be on the order of 0.01 cm s$^{-1}$ or less, over both smooth and rough surfaces [Nicholson, 1988]. Higher values have been obtained in many recent field studies investigating some trace species considered to be representative of particles in this size range. Field observations for sulfate and other sub-micron particles showed that the dry deposition
velocities are one to two orders of magnitude higher than in some earlier studies ([Gallagher, et al., 1997; Gallagher, et al., 2002], and references therein), especially for sub-micron particles, over rough vegetated surfaces.

The particular particle size for which the minimum \( V_d \) is predicted to occur is decided by the relative magnitude of the collection efficiencies by Brownian diffusion and interception. Since the collection efficiencies by Brownian diffusion and interception are different over different canopies, it follows that the minimum \( V_d \) will appear at different particle sizes over different canopies. Most measurements show that the minimum should be located at particle sizes around 0.1-0.3 \( \mu m \); however, some individual field studies show a minimum \( V_d \) located at particle sizes close to 1 \( \mu m \). In order to locate the exact size for minimum \( V_d \), more measurement studies on size-resolved \( V_d \) over different canopies are needed. Meteorological conditions and surface characteristics need to be measured simultaneously in order to relate the dry deposition process to different controlling factors.

### 4.7.2 Wet Deposition

Wet deposition is usually split into two categories, in-cloud and below-cloud (precipitation) scavenging. In-cloud scavenging includes contributions from both nucleation and impaction scavenging while below-cloud scavenging only includes contributions from impaction scavenging.

#### 4.7.2.1 Below-cloud scavenging

For practical considerations, the below-cloud scavenging of aerosols is typically represented by a scavenging coefficient \( S_c \) in aerosol mass continuity equations in large scale models. Most \( S_c \) parameterizations currently used in large-scale atmospheric models are treated as a function of known precipitation properties (e.g., \( S_c = aP^b \), with PR being the precipitation rate). Such a simple representation causes large uncertainties as shown in a comparison of model results from many different large-scale models [Rasch, et al., 2000]. \( S_c \) values for particles ranging from 0.001 to 20 \( \mu m \) this size range of particles differs over three to five orders of magnitude. Since most aerosol number is associated with small particles while most mass is associated with large particles, \( S_c \) is expected to be different for the bulk number concentration and the bulk mass concentration. However, \( S_c \) will be the same for the number and mass concentrations of a specific particle size. With the development of size-resolved aerosol modules within large-scale models, new parameterizations of \( S_c \) as a function of
aerosol size have been developed (e.g., [Gong, et al., 2003a; Jung, et al., 2003]). Particles ranging from 0.001 to 20 µm have to be included in climate and air-quality models and $S_c$ values for this size range can differ by three to five orders of magnitude.

Jylhä [2000] derived the scavenging coefficients from radar reflectivity based on the fact that both are functions of the hydrometeor size distribution. Although many studies show an increase of the scavenging efficiency with precipitation intensity, a measurement study by [Chate and Pranesha, 2004] actually obtained the largest $S_c$ values for submicron particles when the precipitation intensity was weakest. This implies that precipitation intensity alone may not be able to characterize $S_c$ for all particle sizes. Other precipitation properties such as total surface area of hydrometeors may be needed in parameterizing $S_c$. For example, [Zhang, et al., 2004] showed that $S_c$ for any size particles has a strong dependence on total droplet surface area. This is especially important for in-cloud impaction scavenging since the cloud droplets are much smaller compared to the rain drops below the cloud, and thus the total droplet surface area inside the cloud-layer is much larger. For the same reason, including the droplet surface area in the parameterization of $S_c$ is more important for weaker precipitation than for strong precipitation.

$S_c$ is also a function of the collection efficiency ($E_c$), a parameter defined as the ratio of the total number of aerosols collected by falling drops to the total number of aerosols within the swept volume of falling drops. Usually, $E_c$ is much smaller than 1 because particles can follow the streamlines around the drop. It can however be much larger than one under certain conditions, e.g., for charged particles. For small aerosol particles with little inertia Brownian diffusion is important to bring them into contact with the drop thus increasing their collection efficiency. The Brownian diffusion of particles decreases rapidly as particle size increases and this mechanism is most important for particles of diameters smaller than 0.1 µm. Large aerosols experience inertial impaction because their inertia prevents them from following the streamlines around falling droplets. Inertial impaction increases with the increase of particle size and is important for particles larger than 2 µm. Interception and inertial impaction are closely related, with interception occurring as a result of a particle’s size irrespective of its mass while inertial impaction occurs due to its mass and not its size. For particles within the size range of 0.1-2 µm, commonly referred to as ‘Greenfield gap’ in the literature [Greenfield, 1957], neither the Brownian diffusion nor the inertial impaction plays an
effective role. Thus the collection efficiency is smallest in this size range [Pruppacher and Klett, 1997; Slinn, 1984]. The exact depth, width and position of this minimum depend on the properties of the aerosols and hydrometeors, and on ambient conditions. Within the Greenfield gap, thermophoresis dominates over diffusiophoresis, so that in the case of single-phase hydrometeors, phoretic scavenging becomes more efficient with decreasing relative humidity of air (with respect to either water or ice) [Pruppacher and Klett, 1997].

Many experimental and theoretical studies have been dedicated to estimating $S_c$ and $E_c$ values under a variety of conditions (e.g., [Chate, 2005; Facchini, et al., 1999; Jung, et al., 2003; Laakso, et al., 2003; Zhang, et al., 2004]). $S_c$ and $E_c$ for very small and very large particles estimated from measurement studies generally agree well with theoretical studies. However, significant discrepancies between observations and theoretical estimates have been found for sub-micron particles, with measured values being one to three orders of magnitude higher than predicted values [Chate, 2005; Davenport and Peters, 1978; Laakso, et al., 2003; Radke, et al., 1980; Schumann, 1989; Volken and Schumann, 1993]. This is because the magnitudes of the $S_c$ measured for this size range of particles are very low (e.g., $10^{-6} – 10^{-3}$ s$^{-1}$), and any collection mechanisms that were not treated properly in the theoretical studies would cause large uncertainties such as those discussed above.

Collection of aerosols by snow crystals is more complicated than by liquid drops due to the irregular shape of ice elements. Large uncertainties exist in the collection efficiency between experimental results and theoretical predictions, especially for submicron particles. The differences between collection efficiencies by rimed and unrimed snow crystals seem to be small ([Murakami, et al., 1985]; and references therein), although earlier laboratory studies showed an enhancement of collection by rimed snow crystals. Collection efficiencies depend on the shape of the particles and generally increase with the temperature. Charged snow crystals and holes associated with them may also increase the collection efficiencies (see a review by [Jennings, 1998]).

Considering the scarcity of field measurement data on below-cloud scavenging, detailed cloud microphysics models such as the one described in Zhang et al. [2004a] are needed to develop/verify new parameterizations. More field measurements on size-resolved $S_c$ under different precipitation intensities are also needed to improve existing parameterizations and to develop new parameterizations.
4.7.2.2 In-cloud scavenging

In-cloud scavenging includes both nucleation scavenging and post-nucleation processes including Brownian diffusion, impaction scavenging and phoretic effects. Very few measurements are available, and most studies are based on numerical simulations. Although impaction scavenging inside a cloud layer removes little aerosol mass, it removes a substantial number of small particles if the cloud lasts for more than a few hours [Zhang, et al., 2004]. The mechanisms of in-cloud impaction scavenging are the same as below-cloud scavenging. The differences are that cloud droplets are much smaller than raindrops and only very small aerosols (i.e., < 0.1 µm) exist as interstitial aerosols within a cloud-layer. Thus, carefully parameterizing the collection efficiency of small aerosol particles by small droplets is very important for in-cloud impaction scavenging if the interstitial aerosol number concentration inside the cloud layer is a concern. In this section, nucleation scavenging is only briefly discussed since some discussions have been presented in Chapter 7.

Earlier studies used a simple power law to link the concentration of CCN to the aerosol number concentration \( C \) under a specific supersaturation \( S \): \( N_{CCN} = CS^k \) (where \( k \) is an empirical constant and varies with the origins of the aerosols, thus implicitly accounting for the influence of aerosol chemical composition). Although the power law approach was refined in several different ways, it still has large uncertainties due to the large variability of \( C \) and \( k \) with different aerosol source types [Hegg and Hobbs, 1992]. Other forms of parameterizations using pre-assumed aerosol size distribution for predicting droplet concentration through the nucleation process were made available more recently (e.g., Abdul-Razzak and Ghan, 2002 and references therein).

Cloud droplet nucleation from aerosols cannot be properly represented in models with modal representations of the size distributions. Abdul-Razzak and Ghan [2002] extended their earlier parameterizations, which assume single or multiple log-normal aerosol distributions, to sectional models. The parameterization performs reasonably well when compared to detailed numerical simulations in a parcel-model framework except for weak updraft conditions. Nenes and Seinfeld [2003] developed a parameterization which includes two steps to predict the number concentration of nucleated drops. The first step involves the representation of the aerosol number and the chemical composition distribution with respect to size and the calculation of the number concentration of droplets that can potentially form
at a certain level of supersaturation. In the second step, the CCN spectrum is included within the framework of an adiabatic cloud parcel model, with a specified updraft velocity, to compute the maximum supersaturation of the cloud parcel. The parameterization seems to perform better than earlier parameterizations; however, more computer time is needed. Evaluations of the above parameterizations using field measurements have not yet been performed.

The influence of organic aerosols on CCN concentration has also gained some attention (see discussions in Chapter 3). With modifications, several parameterizations for sectional aerosol models described above can also handle the nucleation of organic aerosols. Aerosols can also be scavenged by serving as ice nuclei as discussed in Chapter 3.

Note that in-cloud scavenging may occur in non-precipitating clouds. If these clouds evaporate, the scavenged aerosols will be returned to the free atmosphere and ultimately no removal from the atmosphere occurs via wet deposition. However, the aerosol properties might be changed, and this process is referred to as ‘cloud processing’.

### 4.7.3 Residence Times of Various Aerosols

The residence time $\tau$ cannot be measured globally, but has to be obtained from model simulations. It is independent of model differences in emissions strengths and reflects the integral of all simulated aerosol properties and processes. The residence time is controlled by the aerosol properties as it depends on particle size and solubility, but also the model-specific transports and parameterizations of aerosol processes. In addition, it reflects the spatial distributions of aerosols, particularly as they relate to relative humidity, precipitation, and surface properties.

Some common features in deposition patterns have been found using global simulation results [Ginoux, et al., 2001; Han, et al., 2004; Satake, et al., 2004; Zhao, et al., 2003]. Large particles usually deposit rapidly close to their sources (e.g., within the continent). On the other hand, small particles, especially submicron particles, can be transported far away from their sources (e.g., Asia to North America, Africa to Atlantic Ocean).
Figure 4-27 shows the simulated residence times from the AeroCom models [Textor, et al., 2006]. Sea salt has the shortest $\tau$ of about half a day, followed by sulfate and dust with about four days, and particulate organic matter and black carbon with about six and seven days, respectively. The largest model diversities in the residence times are found for the coarse aerosol types, in particular for sea salt.

The residence time also varies with geographic locations. The size dependence of sea-salt residence time was investigated by Gong et al. [2002] for diameter ranges from 0.01 $\mu$m to 40.96 $\mu$m. For two size ranges of sea-salt aerosols, the residence time follows a similar pattern; that is, $\tau$ is larger on either side of the equator than at middle or high latitudes. There are two subtropical anticyclones in the Southern Hemisphere. One is in the Pacific Ocean and another in the Indian Ocean. Three climatologically persistent subtropical anticyclonic regions prevail in the Northern Hemisphere. Two are in the Pacific Ocean and one in the Atlantic Ocean. These anticyclones are regions of long aerosol residence time. They are located on either side of the ITCZ where tropical cyclones with strong precipitation
scavenging of aerosols are frequent. The removal of sea-salt is less in the anti-cyclonic regions resulting in a residence time peak.

4.8 Recommendations

To better understand the transport, transformation and removal of atmospheric aerosols, the following activities are recommended:

- Development of high resolution atmospheric models with explicit treatment of aerosols, clouds and precipitation formation processes that effectively utilize observations from surface-based networks, aircraft and satellites through advanced data assimilation techniques.
- Identification and analysis of systematic long term measurements of both aerosol distributions and precipitation rates in regions where great changes in both aerosols and precipitation have been observed in order to find statistically significant correlation between them.
- Field measurements of dry deposition fluxes for specific particle sizes over different natural surfaces (e.g., using eddy-correlation method above canopies rather than using surrogated surfaces) are needed in order to better understand this process and to improve existing models or develop new models. Meteorological conditions and surface characteristics need to be measured simultaneously in order to relate the dry deposition process to different controlling factors.
- Development of a suitable parameterization of aging and transformation processes and incorporation into weather and climate models to account for: (i) the secondary organic aerosol fraction, (ii) the hygroscopic nature of particles and (iii) modifications of aerosol optical properties by physical and chemical transformation processes.
- Studies of cloud droplet activation as a function of chemical compositions of aerosols and cloud dynamics such as vertical velocity and turbulent mixing.
CHAPTER 5: IN SITU AND REMOTE SENSING TECHNIQUES FOR MEASURING AEROSOLS, CLOUDS AND PRECIPITATION

Lead authors: Didier Tanré, Paulo Artaxo, Sandra Yuter, Yoram Kaufman

5.0 Introduction

In this chapter, we review the methods presently available and expected in the near future to measure the effect of aerosols on clouds and precipitation, and the reasons why it is difficult to assess this connection.

The short lifetime of clouds and aerosols, the large variability of cloud properties, and the variability in the aerosol chemical and physical properties, makes it difficult to decipher the impact of aerosols on cloud properties and on the onset of precipitation. An array of satellite instruments, and surface networks that routinely measure aerosol, clouds and precipitation are designed to overcome some of these obstacles by generating large statistics of precipitation, cloud properties and the nearby aerosols. The aerosol properties include measures of the column concentrations (expressed by the aerosol optical depth) and size (given by distribution of the aerosol in 2 to 3 size modes or measurement of the Angstrom coefficient). A major remaining obstacle, that is still hard to overcome, relates to the precise measurement of precipitation and its collocation with cloud and aerosol measurements.

The measurement tools of aerosols and precipitation, therefore, include in situ measurements from the ground or from the air, remote sensing from the surface or from space and indirect assessments using other attributes. The use of measurements to deduce the effect of aerosols on precipitation includes the following groups: long term statistics of precipitation with assessment of aerosol based on location and time of the measurements [Warner, 1968; Givati and Rosenfeld, 2004]; individual case studies [Rosenfeld, 2000].

5.1 Measuring Aerosol Properties From The Ground

It is not easy to collect and analyze atmospheric aerosols with high precision and with high reproducibility and specificity [McMurry, 2000]. A number of factors make this task difficult. The dynamic range of aerosol particles in the atmosphere is ~1 nm to 100 µm, and they occur in large variety of shapes and different chemical compounds. In particular, our
knowledge of the composition of organic compounds is rudimentary and a significant fraction of the aerosol mass is semi-volatile. In this section issues of aerosol sampling, followed by the physical and chemical characterization of aerosols are discussed. We will focus on the properties that are relevant to aerosol-cloud interactions. More information is found in the GAW Measurement Guide compiled by the WMO/GAW Scientific Advisory Group for Aerosols [WMO, 2003]. A full comprehensive description of the methods of aerosol sampling analysis can be found in Hinds, 1999, Baron and Willeke, 2001 and Heard, 2006.

5.1.1 Ground Based Aerosol Sampling and Mass Analysis

Particle mass is one of the most commonly measured aerosol properties. Sampling artifacts are of concern for some chemical species, particularly those exhibiting evaporative losses of semi-volatile organic compounds [Baltensperger et al., 2001]. The very broad range of particle sizes makes unbiased sampling and analysis of aerosol properties difficult. Aerosol particles may be divided into three general size ranges: $PM_{10}$ (particles smaller than 10 $\mu$m aerodynamic diameter), $PM_{2.5}$ (<2.5 $\mu$m aerodynamic diameter), and $PM_{1.0}$ (<1 $\mu$m). The ideal aerosol sampling inlet would draw in 100% of the particles in a specified size range and transport them all without modification to a detector or collector. A broad range of instruments has been developed over the past 10 years to measure aerosol mass concentration in certain size ranges [Artaxo et al., 2002]. The gravimetric method is the most common, in which a filter (Teflon, Nuclepore, quartz fiber, nylon, etc.) is exposed to a flow of air loaded with particles from ~1 to 24 hours, after conditioning the aerosol sample in a temperature and humidity controlled environment, and the mass of aerosol deposited onto the filter is weighed using a microbalance. This method provides high accuracy, but is very labor intensive and cannot be done in real time. Sampling artifacts can affect aerosol concentrations in special situations where semi-volatile compounds are present in significant amounts, either increasing the collected aerosol mass or decreasing the mass of already collected aerosol particles. New real time $PM_{10}$, $PM_{2.5}$ or $PM_{1.0}$ instruments, such as the Tapered Element Oscillating Microbalance (TEOM) instrument measure the gravimetric mass of the particles with high time resolution (<30 minutes) and are now widely used in air pollution monitoring networks. In its original configuration the air had to be heated to 50°C or more, which resulted in some loss of semi-volatile organic and inorganic compounds. New versions of this
instrument correct for the particle losses in real time, and provide gravimetric mass with good accuracy.

Aerosol monitors using beta attenuation (also called beta gauges) are widely used for real time aerosol mass measurements in urban areas, although they can have large errors due to the different beta radiation attenuation coefficients needed for different aerosol elemental composition and size. New instruments derive PM$_{10}$ or PM$_{2.5}$ using light scattering from small lasers, but they require careful calibration and are prone to errors when aerosol types or size distributions change. Some aerosol samplers, such as the Dichotomous Partisol-Plus (Model 2025 Sequential Air Sampler; Patashnick et al., 2001), collect aerosols in two size ranges: PM$_{2.5}$ and PM$_{10}$ ($2.5<d_p<10$ µm) on two Teflon filters that are suitable for gravimetric, ionic and trace element analysis. Stacked Filter Units (SFU) collect aerosols in the same two size ranges on sequential Nuclepore filters for subsequent gravimetric and elemental analysis [Hopke et al., 1997]. High volume samplers can also be used to collect large amounts of PM$_{10}$ and PM$_{2.5}$ for analysis of various organic and inorganic compounds when large amount of material is required [Artaxo, 2002].

5.1.2 Aerosol Size Distributions and Particle Counting

Measurement of aerosol size distribution for the whole size range (from 1-2 nanometers to several hundred micrometers) requires different approaches and different instruments. Cascade impactors separate particles in several size ranges for subsequent analysis. For example, the MOUDI (Microorifice Uniform Deposit Impactor) cascade impactor acts as a 12-stage impactor when the inlet stage and final filter are included, with the stages having 50% cut-points ranging from 0.056 to 18 µm in aerodynamic diameter [Marple et al., 1991]. New versions, such as nano-MOUDI, can size and collect particles down to 10 nm. Many other types of cascade impactors, such as the Berner impactor, are widely used to collect aerosols. Condensation particle counters (CPCs) grow particles in a highly supersaturated environment, allowing them to be detected optically. Particles down to 3 nm can be detected and counted, and very low particle number concentrations can be measured. There are CPCs that use butanol as the condensing vapor and some new instruments are water-based. However, size distribution measurements are not possible using a CPC alone.
Particle number size distributions can be measured with a scanning mobility particle spectrometer (SMPS), which is an adaptation of the differential mobility particle sizer (DMPS). Both methods can measure aerosol size distributions from 3–800 nm with a time resolution of a few minutes. They include a differential mobility analyzer (DMA) to separate particles according to their mobility, which is a function of size and a condensation particle counter (CPC) to count them. Coarse mode particles can be sized optically with the Aerodynamic Particle Sizer (APS) spectrometer or with an optical particle counters (OPC). The latter instrument measures the intensity of light scattered by individual particles as they traverse a tightly focused beam of light.

At relative humidity >70–80%, water can comprise a large fraction of the fine particle mass. Water uptake capacity of aerosol particles is critically important in determining the CCN activities and optical properties. The water uptake capacity can be measured with instruments such as the Humidified Tandem Differential Mobility Analyzer (HTDMA), which measures the hygroscopic diameter growth of aerosol particles at different relative humidity (RH) [McMurry and Stolzenburg, 1989]. The HTDMA works with two DMA working in tandem, where the first selects a particular narrow size range in a dry air flow that is then conditioned to a certain relative humidity, while the second scans the size range to analyze the ratio of the wet and dry diameter (growth factor) of the particles. In general, HTDMA can measure growth rates for relative humidity up to 80%, but new methods can measure the growth factor up to 97% relative humidity. These instruments can separate hygroscopic from hydrophobic particles based on their water affinity, a critical property for determining their ability to act as Cloud Condensation Nuclei (CCN).

5.1.3 Individual Particle Analysis

Particles collected on filters can be analyzed by scanning electron microscopy (SEM), to provide size, morphology and elemental composition. Basically, a finely focused high energy electron beam scans the particle, and secondary electrons from the sampled particle are collected to form an image of the particle [Van Grieken et al., 1991]. X-rays emitted by the interaction between the high energy electron beam and the sample are analyzed by a Si(Li) X-ray detector to obtain elemental composition information [Artaxo et al., 1992]. Scanning electron microscopy can be used to measure particles >0.1 μm in diameter, with a
mass detection limit for elements of ~0.5%. The measurement and quantification of the elemental mass is complicated by particle matrix effects. Environmental SEMs allow analysis of particles under near atmospheric conditions and different humidity and temperature conditions, thus also allowing the measurements of volatile particles. Automated SEM can measure and quantify thousands of particles on a filter in a few hours. Transmission electron microscopy is used to observe particles less than 0.1 \( \mu \text{m} \) diameter. Optical microscopy can be used to observe and count large particles (larger than 0.5 \( \mu \text{m} \)). Laser Microprobe Mass Analysis (LAMMA) measures organic and inorganic ions by the use of a laser beam to vaporize particles that are then analyzed in a time of flight mass spectrometer. Measurement of the elemental composition of individual particles is a valuable addition to bulk measurements, especially when the effects of aerosols on clouds are of interest, because of the potential to measure the internal versus external mixture of the particles [Van Grieken et al., 1991]. This is because the presence of internally mixed particles, containing some soluble material in them, can change CCN efficiencies in the atmosphere. For example, dust particles that are normally not efficient CCN can become good CCN if some soluble material accumulates on their surfaces [Levin et al., 1996].

5.1.4 Measurements of Optical Properties of Aerosols - Light Absorption and Scattering

Quantification of light scattering and absorption by aerosols is needed to assess the climatic impacts of aerosols [e.g., Ogren, 1995, Ramanathan et al., 2001]. The aerosol single scattering albedo (\( \omega_0 \)) is a critically important aerosol parameter that expresses the ratio between scattering and extinction of radiation and can be derived from simultaneous measurements of scattering and absorption coefficients. To measure light scattering, integrating nephelometers measure the amount of light scattered by a sample volume of air loaded with particles. Nephelometers of various types are used, some with a single wavelength (generally 500–550 nm), while others use three wavelengths that can provide useful information on aerosol size distributions and Angstrom coefficients with high time resolution and sensitivity. Instruments are available that measure both forward and backward scattering coefficients.

Measurements of light absorption by aerosols are notoriously difficult. Several instruments are available to measure the light absorption coefficient, which is related to the so-called
“black carbon” or "elemental carbon". The Aethalometer and the Particle Soot Photometer (PSAP) are two widely used instruments that measure aerosol absorption on a filter. New versions, operating at several wavelengths, allow measurement of the absorption coefficient as a function of wavelength. These instruments suffer from multiple scattering processes in the filter substrate, resulting in high uncertainty of the data. The Multi Angle Absorption Photometer (MAAP) uses transmitted as well as scattered light for the quantification of the absorption coefficient, resulting in a more accurate reading [Petzold and Schönlinner, 2004]. Absorption coefficients are most commonly determined from measurements of particles collected on filters using integrating spheres or photometers. Photoacoustic spectroscopy measures the absorption coefficients of suspended particles in real time, and can obtain a reliable measurement of aerosol absorption; however, their sensitivity still needs to be improved. Individual light-absorbing particles can also be measured by incandescence [Baumgardner et al., 2005]. Aerosol absorption coefficients can be inferred from mass measurements, using real time elemental carbon measurements that can nominally separate the “organic” from the “elemental” carbon component using a temperature/vaporization profile. The method reveals large differences. It is difficult to distinguish unambiguously between elemental, black or organic carbon based only on the temperature profile because the elemental composition affects this separation. Commercially available quasi-real time organic and elemental carbon monitors are available, but the methods are far from being widely accepted, and intercomparison between measurements is problematic. This area requires important advances in terms of measurement methods and instrumentation.

5.1.5 Elemental Composition of Aerosol Particles

Aerosol source apportionment methods use elemental composition information to infer and quantify aerosol sources [Artaxo, 2002]. The use of factor analysis and chemical mass balance methods requires the knowledge of the elemental composition of aerosols with as much detail as possible [Artaxo et al., 1999]. Varieties of techniques are used for the determination of chemical composition of bulk or size-segregated samples. X-Ray Fluorescence (XRF) uses an X-ray beam to provide information on aerosol elemental composition. The newest instruments use polarized X-rays for excellent detection limits. Detection of light elements (Z<20) requires analysis in vacuum. About 15-20 elements can be routinely measured by XRF in a non-destructive analysis lasting about 5 minutes. Synchrotron radiation sources provide the best detection limits, but are a rather expensive
analysis [Bukowiecki et al., 2005]. Particle Induced X-ray Emission (PIXE) uses a beam of protons with energy between 2 and 3 MeV and X-ray spectrometry to measure routinely the concentration of trace elements [Artaxo et al., 1993]. The detection limits are ~0.1–10 ppm or a few ng/m³. For optimum results using PIXE analysis, samples should be collected on the thinnest possible backing materials composed exclusively from low atomic number elements, such as carbon, hydrogen, oxygen or nitrogen. Polycarbonate Track Etched (PCTE) and Teflon filters are often used, as are impaction surfaces such as Kapton and Mylar [Johansson and Campbell, 1988]. Blank values should be carefully taken into account for reliable measurements of trace elements.

Multi-element analysis by means of inductively coupled plasma-mass spectrometry (ICP-MS) provides excellent detection limits, in the range of parts per trillion (ppt). The aerosol sample is extracted from filters and mixed with water and nitric acid and is injected into the plasma torch via a nebulizer. The dissociated elements are detected by mass spectrometry for a very fast and precise quantitative determination. In INAA (Instrumental Neutron Activation Analysis) spectrometry a flux of neutrons in a nuclear reactor is used to produce unstable nuclei and gamma-ray spectrometry to measure trace elements with high accuracy and very low detection limits. Ion chromatography (IC) is commonly used to measure nitrate, sulfate and a broad range of ionic compounds, including organic components. Ion chromatography has also recently been adapted for semi-continuous measurements of particulate and gaseous species on commercial instruments. Detection limits and precision are excellent, but calibration and blank control must be followed closely.

Real time sulfate and nitrate measurements are now possible with commercial instruments. Flame photometric detectors (FPD) measure sulfates by detecting 394 nm light produced by excited-state S₂ molecules formed when sulfur compounds are burned in a hydrogen-rich flame. A few minutes time resolution is possible, with 1 µg/m³ detection limit for sulfate. Particulate nitrate concentrations can be measured with a single stage impactor that is 95% efficient for particles >0.1 µm aerodynamic diameter and the collected particles are analyzed by flash vaporization using a chemilluminescent NOₓ analyzer.
5.1.6 Real-Time Aerosol Mass Spectrometry

Recently new approaches have been developed to measured aerosol composition in real time, either in single particles or bulk composition (see review by Sullivan and Prather, 2005). Particle analysis by laser mass spectrometry (PALMS and similar instruments) makes it possible to determine the ionic composition of single particles [e.g. Prather et al., 1994, Noble and Prather, 2000]. Aerosol particles are drawn by vacuum into the instrument. A fraction of these particles is detected by a visible light laser (532 nm), the scattered light from which triggers an excimer laser (193 nm) that ablates and ionizes individual particles. A complete positive or negative mass spectrum is obtained on particles by using a time-of-flight mass spectrometer.

A different approach to mass spectrometry was adopted in the Aerodyne Aerosol Mass Spectrometer (AMS) that measures properties of an ensemble of particles. The AMS electron impact ionization (EI) and quadrupole mass spectrometry (QMS) applies thermal desorption for aerosol characterization. The AMS consists of three main parts: an aerosol inlet, a particle sizing chamber, and a particle composition detection section. The different sections are separated by small apertures and are differentially pumped. The aerosols are sampled through an inlet that focuses them into a narrow beam. Size-dependent particle velocities created by expansion into vacuum are used to determine particle size through a particle time-of-flight measurement. The focused particle beam is modulated by a rotating wheel chopper. Time-resolved particle detection, after a known flight distance, gives the particle velocity from which the particle aerodynamic diameter is obtained. Detection is performed by directing the particle beam onto a resistively heated, roughened surface under high vacuum (~10^{-7} Torr). Upon impact, the volatile and semi-volatile components in/on the particles flash vaporize. The vaporization source is integrally coupled to an electron impact ionizer at the entrance of a quadrupole mass spectrometer. When the quadrupole is tuned to a representative mass, bursts of ions are produced that are averaged to give a size-resolved mass distribution [e.g., Jayne et al., 2000, Jimenez et al., 2003]. Refractory component are not measured by this technique.
5.1.7 The Counterflow Virtual Impactor (CVI)

The counterflow virtual impactor (CVI) has been utilized both in aircraft experiments and on the ground to study aerosol/cloud interactions, cloud physics, and cloud impacts on climate [Ogren et al., 1985; Noone et al., 2003]. At the CVI inlet tip, cloud droplets or ice crystals larger than a minimum aerodynamic diameter are separated from the smaller interstitial aerosol and impacted into a dry carrier gas. This separation is possible via a counterflow stream of gas out the CVI tip, which assures that only larger particles (cloud droplets or ice crystals) are sampled. The non-volatile residual nuclei remaining after droplet evaporation and water vapor removal are sampled downstream of the inlet with selected instruments. These may include a condensation particle counter, an optical particle counter, and various chemical techniques. Inclusion of a water vapor sensor allows determination of bulk cloud condensed water content (Strom et al., 1997, Twohy et al., 1997). CVI’s have also been applied to interface cloud activation instruments to particle composition measuring instruments (e.g., Cziczo et al. 2004).

5.1.8 Cloud Condensation Nucleus Measurements

A fraction of the particles in the atmosphere can serve as cloud condensation nuclei (CCN) for cloud droplet formation. The fraction depends on the size, composition and surface properties of the particles, and on the water supersaturation of the air [Roberts and Nenes, 2005]. CCN are linked to cloud microstructures and precipitation formation although this last link in the chain is still under discussion and is the main aim of this assessment. In addition, the role of the CCN in climate feedback remains largely unknown. Measurements of CCN are very challenging [e.g., Twomey, 1960] because they depend on supersaturation, which is a function of both size and chemical composition, although recent measurements in Germany by Dusek et al. [2006] show that size is the most important parameter in determining CCN efficiency. Significant improvements in measurement techniques are needed in this area, and even basic concepts need improvements.

A few instruments are available for measuring the CCN number concentration (or activation spectrum), including those based on the static thermal gradient diffusion cloud chamber (SDCC), continuous flow spectrometers (CFC), and other principles [Nenes et al., 2001]. The
SDCC consists of two parallel plates, held at different temperatures with their facing surfaces wetted, the lower plate being several degrees colder than the upper plate. By varying the temperature difference between the plates, it is possible to produce maximum supersaturations in the chamber that range from a few tenths of 1% to a few percent, which are similar to the inferred supersaturations that activate droplets in clouds [Wieland, 1956]. Small water droplets form on those particles that act as CCN at the peak supersaturation in the chamber. The concentration of these droplets can be determined by photographing a known volume of the chamber and counting the number of droplets visible in the photograph, or by measuring the intensity of light scattered from the droplets [Radke and Hobbs, 1969]. By repeating the above procedure with different temperature gradients in the chamber, the CCN activation as a function of supersaturation can be determined. Figure 5-1 shows the CCN Spectra for different types of aerosol particles in Amazonia, expressed as CCN efficiency defined as the ratio of CCN to CN (condensation nuclei).

Several CCN instruments were developed and provided CCN spectra and concentration through several experimental designs. The Desert Research Institute (DRI) CCN spectrometer (Hudson, 1989) measures nuclei that activate at critical supersaturations down to 0.002%. This instrument continuously provides more than 100 channels of CCN and large nuclei resolution with less than a few seconds time resolution. The Wyoming CCN counter (Delene and Deschler, 2000) is a static thermal-gradient diffusion chamber, taking a grab sample every 30s. Both the top plate and the bottom plate of the chamber are covered with water-saturated blotter paper.
Figure 5-1. CCN Spectra for different types of aerosol particles in Amazonia. Green ocean measurements refer to natural biogenic conditions during the wet season. Most of the CCN are biogenic particles in the wet season in Amazonia. Fresh and aged smoke refers to conditions during heavy biomass burning emissions. (Andreae, personal communication)

Recently, a commercially available cylindrical continuous flow thermal gradient diffusion chamber has been built [Roberts and Nenes, 2005], employing a novel technique of generating a supersaturation by establishing a constant streamwise temperature gradient so that the difference in water vapor and thermal diffusivity yields a quasi-uniform centerline supersaturation. The droplet spectrum at the outlet of the growth column is detected using a standard optical particle counter (OPC). This design allows for atmospheric CCN measurements with high time resolution for supersaturations in the range from ~0.07-3% achieved by varying the flow rates and temperature gradient along the growth column.

In the dynamic CCN spectrometer, the aerosol sample is exposed to a varying supersaturation field and the CCN spectrum is obtained by analyzing the outlet droplet size distribution. Time resolution for CCN spectrometers can vary from 1–30s. Although these two types of counters provide some measurement of the CCN concentrations and their dependence on
supersaturation, they both suffer from deficiencies, among which is the inability to correctly measure giant CCN that fall out in the chambers before they can be measured.

Giant CCN (GCCN) are wettable particles greater than about 5 µm in diameter that can initiate collision and coalescence in a very efficient way, and thus efficiently produce precipitation embryos. Their effect on cloud droplet spectra can be critical. Due to their large size and low number concentration, it is difficult to measure GCCN. Giant aerosol particles can be measured with large uncertainties from aircraft by optical scattering sensors (PMS FSSP, PCASP), because their sample volumes are very small (about 1 m³ in 12h of flight). Since typical concentrations of these particles are a few hundreds per cubic meter, the sampling statistics by these probes is poor. Another method of sampling GCCN is to sample air by an inlet tube into an aircraft using a low turbulence inlet [Huebert et al., 2003]. This approach is good for particles up to 7 µm in diameter so many ultra-giant particles are not sampled. A third method of sampling GCCN is to sample them by impaction on a glass slide. With slide sample areas of 1 cm², 1 m³ of air can be sampled in 1-2 minutes of aircraft time [Levin et al., 2005]. The disadvantage of this approach is that tedious, labor-intensive microscopic analysis of each slide is required. However, new methods that use computer-controlled counting systems can automatically size the particles. It is also possible to set up the system in a humidity-controlled environment where the hygroscopic growth of particles can be examined.

5.1.9 Ice Nucleus Measurements

A significant fraction of supercooled clouds contains ice crystals, while some clouds (e.g., cirrus) are generally completely glaciated. The presence of ice in clouds is very important for the development of precipitation. It is also an important component of the aerosol indirect effect on climate. At temperatures warmer than about -40°C, ice can form by heterogeneous nucleation on particles called ice nuclei (IN) (see Chapters 2 and 6). For clouds with temperatures below about -40°C, ice can form both by heterogeneous nucleation on insoluble particles and by homogeneous nucleation. In spite of the importance of IN in clouds, there is a great deal of uncertainty in measurements of their concentrations and even in their characterization. Heterogeneous ice nucleation can take place via a number of processes (see Chapter 2): immersion freezing, contact freezing, condensation-freezing and deposition. In
many of the instruments that measure IN activity or spectrum, ice may be formed by one or more of these processes, some of which depend on the supersaturation of the air as well as on temperature, without being able to differentiate between them. For example, in expansion chambers cooling is produced by compressing the air and then suddenly expanding it. In case of mixing chambers, cooling is produced by refrigeration. Moist air is then introduced into a cloud but the amount of supersaturation is not well controlled.

A traditional way of measuring IN activity was to collect aerosols from air samples of about 300 liter on Nuclepore filters. Filters are then inserted into an IN chamber made up of two parallel plates held at different temperatures below the 0°C. Supersaturation with respect to ice is reached just above the sample. The concentration of ice crystals formed at each temperature is recorded through a downward looking microscope or a video camera. The number of crystals detected as a function of temperature is then plotted and a best-fit curve is obtained [e.g., Gagin, 1975].

Another method for measuring the IN activity spectrum is via drop freezing measurements. This uses an isolated chamber, the lower horizontal plate of which is temperature controlled. Aerosol samples are collected on a filter and washed into a beaker with a known quantity of water. About 100 drops from this water are placed on the cold stage. The number of drops that freeze at each temperature as the plate is cooled down, is recorded. Knowing the volume of air sampled and the amount of water used in the washing process, a qualitative estimate of immersion freezing spectra can be obtained [e.g., Vali, 1985]. A more recent development is the continuous-flow diffusion chamber (CFDC), which measures ice nucleation by in situ aerosol particles under defined temperature (between 0°C and −80°C) and humidity conditions [Rogers et al., 2001]. The instrument focuses on an aerosol particle stream in the central portion (approximately 10% of volume) of a vertically downward-oriented laminar flow between two ice-coated cylindrical walls held at different temperatures. The inter-wall temperature gradient and vapor field expose the aerosol stream to constant temperature and relative humidity conditions for a period (5–7 s), enough to initiate ice nucleation and for growth to occur. A hydrophobic material replaces ice on the warmer wall in the lower one third of the CFDC, which helps to evaporate liquid particles, including cloud droplets, leaving the ice crystals to pass through. The ice crystals that grow in the CFDC are detected and counted optically. By changing the supersaturation conditions, different modes of ice
nucleation (e.g., deposition, condensation-freezing, immersion-freezing and homogeneous-freezing) can be detected [DeMott et al., 2003a]. It should be noted that none of these techniques simulate contact nucleation, since the droplets must remain in the chamber for a long enough time to scavenge the aerosol particles acting as IN. A few prototype devices have been built to simulate contact nucleation [Meyers et al., 1992] but none has been implemented for field measurements.

5.2 Airborne Instrumentation for Characterizing Aerosols and Clouds

5.2.1 Introduction

There are important weakness and technical difficulties in measuring aerosol and cloud properties in airborne platforms. In situ measurements of aerosol vertical profiles necessitate the use of airborne platforms with instrumentation suitable to fast and accurate analysis. Isokinetic inlets are required when using airplanes to correctly represent the natural conditions in still atmosphere. The adequacy of the inlet and the necessary high time resolution make measurements of aerosol from aircraft particularly challenging. Aircraft used for atmospheric measurements typically fly at speeds from ~40 equal to greater than 200 m s$^{-1}$, while the flow speed through filter samplers or aerosol counting/sizing instrumentation is typically less than 3 m s$^{-1}$. Variations in pitch, roll and yaw, together with turbulence generated at high aircraft speeds, make the conditions for aircraft sampling very difficult and with large associated uncertainties, especially for sampling of particles $>0.5$ µm [Huebert et al., 2003]. In addition, many airplanes are pressurized and sampling of aerosols inside the cabin requires special care to prevent cabin air from contaminating the samples.

The aerosols samples collected on board the airplane are then analyzed by one of the techniques mentioned in Section 5.1. Instruments capable of conducting real time measurements with time resolution better than a few minutes are desirable. In spite of these constrains, aerosol mass spectrometers, nephelometers, PSAP (Particle Soot Photometer), CPCs (Condensation Particle Counters), SMPS (Scanning Mobility Particle Sizer), DMPS (Differential Mobility Particle Sizer), OPCs (Optical Particle Counter), and high time resolution instruments have provided critically important information in recent experiments. Bulk aerosol composition is still a challenge in aircraft measurements, since large samples require high volume pumping over long time periods, something that is not always feasible.
aboard aircraft. New real time instruments providing aerosol composition such as PILS (particle-into-liquid sampler) is helping to advance the knowledge on the chemical composition of aerosols and CCNs in airborne experiments.

Distortion of airflow around the aircraft fuselage can enhance or deplete certain sizes of particles, generating errors that may be larger than those inherent in the measurement sensor itself [King, 1984; Twothy and Rogers, 1993]. However, instruments located on aircraft wings can measure particle size distributions in real time. The PMS FSSP-100 (for measuring cloud droplets) and the PMS-PCASP-100 (for measuring aerosol particles) are examples of such instruments. Cloud droplets and ice particles can also be collected, counted, sized and identified using airborne platforms.

5.2.2 Cloud Physics Aircraft Instrumentation

Particle Measuring Systems (PMS), have manufactured a set of aerosol and cloud droplet probes based on single droplet light scattering and photodetector imaging [Knollenberg, 1981]. These instruments came to dominate airborne cloud microphysics observations for more than two decades. However, these instruments are no longer manufactured by, or supported by PMS. The major instruments developed and still used today are the PMS PCASP-100X (Aerosol sizes 0.1 - 3.0 \( \mu m \)), PMS FSSP-100 (Cloud particle sizes 2 – 47 \( \mu m \)), PMS OAP-2D2-C (Cloud and drizzle particle sizes 25 – 800 \( \mu m \)) and the PMS OAP-2D2-P (Precipitation particle sizes 200 – 6400 \( \mu m \)).

Particle and droplet instruments made by Measurement Technologies (DMT) also rely on light scattering and occultation of photodiode arrays. The Cloud Droplet Probe (CDP) provides a size distribution similar to the FSSP-100. CDP can measure cloud particles in the range of 1–50\( \mu m \). The Cloud Aerosol Precipitation Spectrometer (CAPS) provides a large number of particle measurements, including cloud droplet and particle aerosol.

5.2.3 Liquid water meters

Cloud liquid water is a very fundamental property of clouds. It can be measured in-flight by the Johnson and Williams (JW), King liquid water probes (King et al., 1978) and the DMT Liquid Water Content (LWC-100) Sensor. The temperature of the sensor is maintained constant. For a given air flow past the wire, the higher the liquid water content,
the greater the current required to maintain a fixed wire temperature. Liquid water content at 100Hz is determined as a function of current through the probe and the true air speed. The range of measurements is 0–3 g m\(^{-3}\) at 5–50 µm droplet diameter. The very fast responding and versatile Gerber Scientific optical cloud water sensor (the Particulate Volume Monitor PVM-100) is an optical cloud microphysics probe designed to characterize the liquid water content (LWC), droplet surface area (PSA), and droplet effective radius (Re) of the smallest droplets in a cloud. The PVM makes these measurements optically on a cloud volume of \(\sim 10 \text{ cm}^3\), such that the measurements are independent of air speed. The accuracy of the PVM is estimated to be better than 10% for droplet spectra with volume medium diameter (VMD) smaller than 30 µm, with a precision on the order of 0.002 g m\(^{-3}\) [Gerber et al., 1994]. Another instrument that is gaining wide acceptance on many research aircraft is the Nevzorov total water content/liquid water content probe [Korolev et al., 1998]. This is a constant-temperature, hot-wire probe designed for aircraft measurements of the ice and liquid water content of clouds. It consists of two separate sensors for measurements of cloud liquid and total (ice plus liquid) water content.

Recently, there have been intercomparison studies using the major LWC probes in wind tunnels (e.g. Strapp et al., 2003; Twohy et al., 2003; Wendisch et al., 2002). Such studies are necessary to document the differences between different types of probes.

### 5.2.4 Cloud Particle Shapes

2D probes can measure cloud particle shape in real time. The cloud particles enter a sensitive volume where they occult a laser beam casting a shadow on a photodiode array. Because the size range of cloud particles is several orders of magnitude, and the size of the arrays in the early 1980’s was rather small, several 2D instruments had to be flown to sample the full range of cloud particles likely present in precipitating clouds. These digital images provided a substantial quantitative advance, especially as recognition analysis allowed separation of the images by crystal habit and particle dimensions. However, the relatively small number of pixels in the array detector limited the resolution of these devices. It was not always possible to decide whether the imaged particle was liquid or frozen. The cloud imaging probe (CIP from DMT) is a new design based on older optical imaging probes. Improvements include a
solid-state diode laser, on-board Pitot tube, an on-board Digital Signal Processor, and a
synchronous RS-422 data channel providing statistics and compression of image data. The
Stratton Park Engineering Company (SPEC) Cloud Particle Imager (CPI) (Lawson et al.
1995; Lawson et al., 2001) produces digital, near photographic quality, images of the cloud
particle at aircraft speeds. The Cloud Integrating Nephelometer model CIN-100 is designed
for aircraft mounting and measures the optical extinction coefficient and asymmetry
parameter. It has a large range of detection, spanning from 5 to 3000 micrometer (Gerber et
al., 2000). The Small Ice Detector, (SID) developed for the Meteorological Research Flight at
Farnborough employs a set of six high sensitivity detectors arranged about an illumination
beam (from an Nd:YAG laser) to assess particle shape, and a further two detectors to provide
trigger and particle size signals. The instrument is designed to classify cloud particles in the
range from ~2 to ~50 µm diameter, and is intended to provide, for the first time, an ability to
discriminate between micrometer sized ice crystals and super-cooled water droplets (Hirst et
al. 2001).

5.2.5 Remaining issues

Measurement of aerosols and cloud properties using airborne platforms is difficult.
Maintaining isokinetic sampling conditions over a large range of aircraft speeds is not easy,
and efficient collection of large particles remains a challenge. The time resolution of aerosol
measurements must be improved in order to better describe the rapid changes in aerosol
population over different heights. Better instruments are needed to measure CCN and ice,
including the capability to distinguish between drops and very small ice crystals (< 15 µm).
These are vital in order to improved characterization of CCN and ice within vertical profiles
under a range of environmental conditions. Despite attempts at instrument de-icing, icing can
still be a problem in supercooled clouds as even light icing will distort the sample flow,
resulting in particle size sorting, shadowing and possible particle shattering. Additionally,
several droplet instruments have significant “dead time” after a particle is observed, before
they are ready to measure the next particles. At high droplet concentrations, this can result in
significant errors. Light scattering instruments are calibrated with spherical non-absorbing
particles. Hence, the calibration is not strictly valid for absorbing and non-spherical particles.
The airflow surrounding any instrument attached to the exterior surfaces of an aircraft will be
changed as the aircraft maneuvers. Therefore, mounting locations should be selected to minimize these effects.

Despite satisfaction with recent instrument developments, the path forward to the next generation of airborne instruments remains uncertain. We must now consider measuring clouds ranging from the local to the global scale. For example, on the smallest scales the interplay between aerosols, water vapor supersaturation, and condensation on CCN, turbulence, coalescence, electrical charge separation and the production of precipitation remains observationally shrouded. To address such complex systems it will be necessary to observe down to the inter-droplet spacing. While in-cloud measurements of temperature and humidity using radiation transfer and tunable diode laser technology show promise, much more work is needed in this area. Tomorrow’s instrumented research aircraft will need, in addition to in situ instruments, a compliment of remote sensing tools that permit a 3-D perspective of cloud structures and motions. Microwave imager’s, such as NCAR’s ELDORA Doppler radar, proportioned to much shorter wavelengths, is a powerfully research tool. Such miniaturized 3-D imagers can then be used in real time to depict the optimal locations to employ in situ observing tools. Characterizing the 3-D nature of clouds is particularly critical in radiation transfer calculations [e.g., Kasyanov, 1999].

5.3 Surface-Based Remote Sensing of Aerosols and Clouds

5.3.1 Measurement of Aerosol Optical Depth

Aerosol optical depth (AOD) is a quantitative measure of the extinction of radiation by aerosol scattering and absorption along a path, often between a point of observation and the top of the atmosphere. Column AOD is a measure of the integrated aerosol load through the atmosphere, and is an important parameter for evaluating aerosol amount, variability, and direct radiative forcing. AOD can be determined from the ground by measuring the spectral transmission of solar radiation through the atmosphere, using rather simple and relatively inexpensive instruments pointed directly at the sun called sun photometers or filter radiometers. Routine ground-based AOD observations are of utmost importance for the calibration and validation of AOD retrievals from satellites. In addition, they are needed to correct for aerosol effects in the retrieval of other satellite products. The Ångström exponent,
that gives an indication of the column integrated aerosol size distribution, can be derived from simultaneous AOD measurements at several wavelengths.

AOD is not directly measured, but is retrieved from observations of atmospheric spectral transmission. The solar irradiance $I$ at a given wavelength can be expressed as $I = I_0 \exp(-m \tau)$ with $I_0$ the extraterrestrial (top-of-the-atmosphere) solar irradiance of the sun, $m$ the air mass and $\tau$ the total optical depth. The air mass equals 1 for a vertical path and is roughly proportional to $1/\cos z$ with $z$ the zenith angle of the sun during the observation. The total optical depth $\tau$ at a given wavelength results from several components such as scattering by gas molecules, $\tau_R$ (Rayleigh scattering), extinction by aerosol particles, $\tau_A$, absorption of trace gases, $\tau_G$, such as ozone, and possible cloud contamination. Thus, the AOD can be obtained from the total optical depth by subtracting modeled estimates of the other components $\tau_A = \tau - \tau_R - \tau_G$. Because AOD is essentially a difference between two large numbers, it is very sensitive to calibration errors and to a minor degree also to the methods chosen to model the other components [WMO 2003].

The current state of long-term observational networks for AOD was reviewed by a group of international experts gathered at a meeting in Davos, Switzerland in March 2004 [WMO, 2005]. There are a number of international and national networks that have extensive records of AOD measurements including AERONET (which encompasses AEROCAN, PHOTONS, AEROSIBNET, BSRN, GAW-PFR, and SKYNET). National networks are supported by Australia (Bureau of Meteorology), China, Finland (Finnish Meteorological Institute), Germany (DWD, the German Weather Service), Japan (JMA), Netherlands (KNMI, including Surinam), Russia (Sibrad), and USA (networks including ARM, SURFRAD, CMDL, USDA). Figure 5-2 presents the Global AOD network of long-term sites. These comprise 90 stations with a continuous record for the past 4 years and temporal data for at least 50% of the day (In any given year many more stations may be operating if short term campaign observations are taken into account. For instance, AERONET reported 110 stations operating during 50% of 2003). Half of the stations fall within the AERONET project, while the other half is maintained mainly by WMO Members and SKYNET. Hemispheric coverage corresponds roughly to the landmass distribution, ($\frac{1}{3}$SH, $\frac{2}{3}$NH), with Australia, Europe and North America accounting for more than 50% of stations. Major gaps exist in Africa, India, Latin America and the Polar Regions.
Figure 5-2. Global AOD network of long-term sites, i.e., with a continuous record for the past 4 years and temporal data coverage of more than 50%.

The AERONET and the GAW-PFR networks have particularly important roles in worldwide aerosol monitoring. AERONET is by far the largest network. During the workshop in Davos [WMO, 2005], the GAW-PFR network was assigned a special role in the harmonization of worldwide AOD data in the process of forming a federation of diverse networks coordinated under the WMO/GAW umbrella. These two networks will therefore be described in more detail below.

5.3.1.1. The AERONET Network for Monitoring Aerosol and Cloud Properties

The AERONET network of ~200 well calibrated sunphotometers and sky radiance radiometers distributed around the world [Holben et al., 1998, 2001], measures the aerosol optical properties in key locations that are representative of a wide diversity of aerosol conditions. Both direct sun and diffuse sky directional radiance measurements taken throughout the daylight hours are utilized to characterize total column integrated aerosol optical properties. The spectral aerosol optical depths are measured from 340 to 1020 nm in 7 wavelengths for the standard CIMEL sun/sky radiometer and also to 1640 nm in the extended wavelength CIMELs. The AERONET spectral aerosol optical depth (AOD) measurements are highly accurate (~0.015 uncertainty; Eck et al., 1999) and are used to validate satellite retrievals of AOD from all the current satellite sensor systems, including MODIS, TOMS,
SeaWiFS, MISR, and AVHRR [Hsu et al., 1999; Remer et al., 2002a; Chu et al., 2002; Diner et al., 2002; Kahn et al., 2005; Abdou et al., 2005; Torres et al., 2002]. In addition to validation, the aerosol size distributions and single scattering albedo retrievals from AERONET are utilized in algorithm development for some of these sensors, such as MODIS and TOMS [Remer et al., 2002b].

The CIMEL sky radiance almucantar measurements (measurements along a conical surface with an angle equal to the sun zenith angle) at 440, 675, 870, and 1020 nm in conjunction with the direct sun measured $\tau_a$ at these same wavelengths are used to retrieve aerosol size distributions following the methodology of Dubovik and King [2000]. Almucantar sky radiance measurements are made at optical airmasses of 4, 3, and 2 in the morning and afternoon, and once per hour in between. The option of either spherical or spheroid particle shape may be assumed in the retrievals, depending on whether coarse mode dust aerosol is present in the aerosol mixtures [Dubovik et al., 2002]. Sensitivity studies performed by Dubovik et al. [2000] were used to analyze the perturbations of the inversion resulting from random errors, possible instrument offsets and known uncertainties in the atmospheric radiation model. Simultaneous retrievals of aerosol single scattering albedo are also made with this algorithm and the sensitivity analysis shows that these retrievals have an uncertainty of $\sim 0.03$ for both desert dust and biomass burning aerosols when $\tau_{a440}\geq 0.5$ [Dubovik et al., 2000].

The use of the AERONET network of ground-based radiometers to obtain average size distribution integrated over the whole atmosphere has been proven to be highly effective. Figure 5-3 taken from Dubovik et al. [2002] is an example of the averaged optical properties of different types of tropospheric aerosol.
Figure 5-3. The averaged optical properties of different types of tropospheric aerosol retrieved from the worldwide-multi-year AERONET network of ground-based radiometers. Urban/industrial, biomass burning and desert dust aerosols are shown for AOD(440nm) = 0.7. Oceanic aerosol is shown for AOD(440nm) = 0.15 since oceanic background aerosol loading does not often exceed 0.15. The figure shows the single scattering albedo (top), the size distribution (bottom) and values of the real refractive index (n) and the Angstrom exponent (α), estimated using optical depth at two wavelengths 440 and 870 nm. (After Dubovik et al., [2002]).

The aerosol size distributions from all of the biomass burning sites are dominated by accumulation mode particles, however the sub-micron particle size is somewhat smaller in the savanna and cerrado sites due perhaps to the phase (flaming or smoldering) of combustion, but also possibly due to differences in aerosol aging, fire intensity, and ambient relative humidity and temperature.

Based on AERONET data, recent analysis of the effects of aerosols on cloud cover, (Kaufman and Koren, 2006) show a decrease in cloud cover with increasing column integrated aerosol absorption (Figure 5.4) and an increase in cloud cover with increasing non-absorbing aerosols. This work shows that the effect of aerosol on clouds depends on the
composition and radiative properties of the aerosols. This result is consistent with studies that showed that biomass burning aerosols strongly inhibit cloud formation (Andreae et al., 2004, Koren et al., 2004).

AERONET’s distribution across the globe with standardized instrumentation and processing algorithms provides similar characterizations, aerosol climatologies and aerosol models for marine, urban/industrial, aeolian dust, biogenic, arctic and mixtures of these aerosol types. Polarization observation capabilities along with non-spherical modeling of scattering, show great potential for more universal characterization of aerosols. AERONET data in combination with other web-based data sets such as back trajectories, lidar profiles, satellite and chemical transport and global climate models can be used to address issues of radiative forcing, public health and long term environmental impacts.
Figure 5.4: Regional analyses of the AERONET data for the effect of aerosols on cloud cover. Fraction of cloud cover per unit of AOD \(\Delta f_{\text{cl}} / \Delta l_{\text{int}}\) was plotted as a function of absorbing aerosol amount. Each point represents an analysis of more than 3000 measurements from a given location and 2 calendar months averaged over 3 to 5 years. Green symbols indicate continental sites, blue symbols indicate marine sites, and red symbols indicate biomass-burning sites. Error bars are printed for three representative points and indicate the average uncertainties in the least-squares fit used for individual points and the estimated error in tabs. Adapted from Kaufman and Koren [2006].

5.3.1.2. The GAW Precision Filter Radiometer (PFR) Network

A new generation of sunphotometers called 'Precision Filter Radiometers' (PFR) was designed at the World Optical Depth Research and Calibration Center (WORCC) in Davos with emphasis on instrumental stability and ruggedness in field use. A series of 34
instruments were built by PMOD/WRC for deployment in the trial network and for use as traveling standards (GAW, 2003). The remaining instruments were sold to different research groups and national weather services; PFR instruments are now in operation at 19 stations ranging from the tropics to beyond the polar circle and from sea level to sites above 3,500 m. The PFR instrument is presently not available commercially. Measurements are taken continuously at 1-minute intervals. The ASCII datasets are self-contained and suitable for further evaluation into AOD results or Langley calibration runs either at the stations or at a centralized data center. During the trial phase, data are transmitted monthly to WORCC for evaluation, quality control and archiving. Instantaneous values of multi spectral AOD are calculated using exoatmospheric calibration coefficients and individual air masses for Rayleigh scattering, ozone absorption and aerosol components. Ångström exponents $\alpha$ are calculated by log-log regression of AOD from 368 to 862nm.

In the GAW-PFR network, a new calibration method was tested as a possible alternative to classic Langley extrapolations at high altitude stations. A reference instrument has obtained its exoatmospheric values in October 1998 during a stratospheric balloon flight at 40km height, and its radiometric stability has been monitored since then by repeated spectral comparisons with a silicon trap detector that is traceable to a standard at the Physikalisch-Technische Bundesanstalt in Berlin [Wehrli, 2000]. Additional reference instruments were calibrated by Langley extrapolations at Jungfraujoch and Mauna Loa. All field PFR radiometers are calibrated at WORCC by comparison with 1 or 2 reference instruments before deployment in the trial network. In-situ Langley calibrations are routinely made at the high altitude or arid sites; for other sites, parallel measurements with a traveling standard, exchange of instruments or re-calibration at Davos are used.

GAW-PFR stations are co-located with other networks at several sites. A first comparison during 2000 at Mauna Loa with AERONET individual AOD results showed excellent agreement within 0.005 OD. In 2001, the PFR at Bratt's Lake participated in a 3-month, multi-network comparison [McArthur et al., 2003] where average differences smaller than 0.01 OD per observation were found for direct-pointing instruments at most wavelengths.
5.3.2. Aerosol vertical profiles retrieved by Lidar

Lidars are useful tools for measuring vertical profiles of aerosols. There are various lidar techniques for measuring aerosols, air pollutants, water vapor, temperature, wind velocity, etc. The lidar methods used for aerosol observations are Mie-scattering lidar, Raman lidar, and high-spectral resolution lidar (HSRL). Mie scattering lidar is the simplest lidar technique, which detects elastic backscattering from aerosols. Mie lidars, however, have limitation in quantitative measurement because the lidar equation cannot be solved without an assumption on aerosol optical characteristics or some additional constraint such as independent optical depth measurement, for example, with a sunphotometer. Usually, the lidar equation is solved with an assumption on the extinction-to-backscatter ratio, or lidar ratio (or S1). An additional feature of Mie scattering lidar is the depolarization ratio measurement, which detects polarization components parallel and perpendicular to the polarization direction of the transmitted laser. The depolarization ratio is a good index of non-sphericity of the scatterers. It is useful for detecting mineral dust particles and distinguishing water and ice clouds. Mie lidar measurements at multiple wavelengths also provide useful information for characterizing (or classifying types of) aerosols.

Raman lidars and HSRL can measure extinction and backscattering coefficients independently. Consequently, the profile of the lidar ratio is also obtained with these methods. Raman lidars for aerosol measurements utilize Raman scattering of atmospheric molecules such as nitrogen. Because the profile of atmospheric molecules is known from radiosonde observations and models, the extinction coefficient profile of aerosols is derived from the Raman scattering signals. Using the Mie-scattering signals simultaneously measured with the Raman signals, the backscattering coefficient is also determined [Whiteman, 2003]. Raman lidars can be extended for multi-wavelength measurements and for water vapor profile measurement. It is also possible to have rotational Raman channels for measuring temperature profiles. A weak point of Raman lidar method is the low sensitivity. Raman lidars, generally, require larger laser pulse energy and larger telescope size. HSRLs use a similar method but utilize Rayleigh scattering of atmospheric molecules [Liu et al., 1999]. HSRLs have a higher sensitivity, because Rayleigh scattering cross sections of molecules are much larger than Raman scattering cross section. HSRLs are
technically more difficult, because they need very high-resolution filters for separating spectra of Mie scattering and Rayleigh scattering.

Currently, several Lidar networks are operated in the world (Figure 5-5), which include the Micro-Pulse Lidar Network of NASA (MPLNET), the European Aerosol Research Lidar Network (EARLINET), the Commonwealth of Independent States Lidar Network (CIS-LiNet) and the Asian Dust Network (AD-Net).

The Micro-Pulse Lidar Network (MPLNET) operated by NASA is the network of the micro-pulse lidar (MPL) co-located with the sun/sky photometer sites in the NASA Aerosol Robotic Network (AERONET). MPL is a single-wavelength (523 nm) eye-safe lidar capable of unattended operation. At present, MPLs are operated at 22 sites around the world. The combined lidar and sunphotometer measurements are able to produce quantitative aerosol and cloud products, such as optical depth, sky radiance, vertical structure, and extinction profiles (http://mplnet.gsfc.nasa.gov/).

The Regional East Atmospheric Lidar Mesonet (REALM) is the network of lidar research groups on the east coast of the United States (http://alg.umbc.edu/REALM/).

The European Aerosol Research Lidar Network, EARLINET established in 2000 is a highly instrumented lidar network (http://www.earlinet.org/). The primary purpose of EARLINET is to build a quantitative comprehensive statistical database of the horizontal, vertical, and temporal distribution of aerosols on a continental scale. Also, additional measurements are performed to address important processes, for example, dust events, forest fire events, and volcanic eruptions.
The Commonwealth of Independent States Lidar Network (CIS-LiNet) established in December 2004 is the lidar network in CIS regions (http://www.cis-linet.basnet.by). The basic concept of the network is similar to EARLINET.

The Asian Dust Network (AD-Net) (http://www-lidar.nies.go.jp/AD-Net/index.html) is a voluntary network of lidar researchers in the East Asia region for exchanging data and information. Various groups with various instrumentations are involved in AD-Net. There are multi-wavelength Raman lidars in Tokyo and Gwangju, Korea, and a high-spectral-resolution lidar in Tsukuba, Japan. The National Institute for Environmental Studies (NIES) Lidar Network (http://www-lidar.nies.go.jp/), which is a part of AD-net, is a network of
automated two-wavelength polarization lidars (backscattering at 1064 nm and 532 nm, and depolarization at 532 nm). Currently, 13 lidars (Japan (8), Korea (1), China (1), and Thailand (1)) including those of cooperative sites, are operated continuously, year-round. Many of the lidars are co-located with the Skyradiometer Network (SKYNET) (http://atmos.cr.chiba-u.ac.jp/)

5.3.3. Surface Based Remote Sensing of Clouds

AERONET also measures the cloud optical depth, the most important of all cloud optical properties and vital for any cloud-radiation parameterization. Its impact on radiative fluxes and therefore climate is matched only by cloud fraction. Cloud optical depth is also a fundamental optical variable in any cloud-resolving model. Progress on cloud-radiation parameterization may halt because without good optical depths, the radiation calculations could be wrong.

There are ambiguities in definitions of cloud fraction. Visual cloud fraction is a nice thing to know but climate models need a *radiative* cloud fraction, and not a visual quantity. By radiatively effective cloud fraction we mean the one that forces plane-parallel radiative transfer calculations to give the same radiation as the measured ones. Stephens [1978] showed that radiatively effective cloud fraction is the cloud property actually required in GCM cloud parameterizations and it is always different from a simple visual one. The AERONET network has begun monitoring *cloud* optical properties, such as cloud optical depth and (effective) cloud fraction, using AERONET “idle time” inappropriate for aerosol study. When the sun is blocked by clouds, a radiometer instead of going “to sleep” is forced to look straight up and measure zenith radiance. Several radiometers have now been equipped with this “cloud mode”.

There have been two major problems with inferring cloud optical depth from measurements of zenith radiance: (i) lack of a unique relationship to determine cloud optical depth; and (ii) a strong influence of 3D cloud structure on measured radiance. To solve these problems, a new method that exploits the huge jump in vegetated surface reflectance across 700 nm wavelength has been proposed [Marshak et al. 2004]. The idea was to use zenith radiance measurements in two narrow spectral bands on each side of the jump. In the RED spectral
region (~670 nm), the chlorophyll in green leaf absorbs 90-95% of solar radiation; thus, the vegetation albedo is low. In contrast, in the NIR spectral region (~870 nm), a green leaf reflects 90% of incident radiation; thus its albedo is high at this wavelength. If AERONET measurements are combined with surface reflectance properties, then zenith radiances at RED and NIR wavelengths can be used to determine the overlying cloud properties. This is illustrated in Figure 5-6 using surface data. The left panel shows a calculated set of curves for one particular choice of RED and NIR surface reflectances, and for a range of possible cloud optical depths $\tau$ and cloud fractions $N_c$. Also plotted are three groups of 10 data-points measured by an AERONET radiometer on July 28, 2002 at the DOE/ARM site in Oklahoma at three different times. The plot shows that the 3 clusters of AERONET measurements correspond to three different pairs of cloud properties – ($N_c = 0.9; \tau = 28$), ($N_c = 0.8; \tau = 22$), and ($N_c = 0.4; \tau = 12$). The right panel shows an image from DOE/ARM’s “all sky camera.” Note that $N_c = 0.8$ is not a horizontal cloud fraction that would be seen in the surface camera, or the satellite image, but a “radiatively effective” fraction that accounts for the horizontal inhomogeneity of the clouds. For details, see Marshak et al. [2004].

![Figure 5-6](image)

**Figure 5-6.** Left panel: DISORT calculated values of RED and NIR radiances for a wide range of optical depth $\tau$ and cloud fraction $N_c$ for SZA = 62°±3°, surface albedos $\rho_{\text{RED}}=0.092$ and $\rho_{\text{NIR}}=0.289$. When $N_c$ is constant and $\tau$ is varying, the set of calculated values define the cloud fraction isolines. When $\tau$ is constant and $A_c$ is varying, the set of values define the optical depth isolines. The red dots are data points from AERONET measurements at the ARM site on July 28, 2002 and show three different clusters of cloud properties. Measurements were taken around 13:45, 13:58, and 14:11 UT, respectively. Right panel: Total Sky Image taken at 14:00 UT with SZA=62.75°.
Recently, surface-based remote sensors have been used to examine aerosol effects on clouds. Although these methods can only be applied at a limited number of locations, they yield high temporal resolution data. They offer some interesting insights, and are complementary to the global satellite view. Satellite-borne probes measure cloud properties in cloudy pixels and aerosol properties in adjacent, cloud-free pixels, with the implicit assumption that these aerosols are representative of the aerosols entering cloud base.

5.4 Satellite Measurements of Aerosol and Clouds

5.4.1 The Large Scale Picture of Aerosol Distribution

Radiant energy reflected and emitted by the Earth’s surface and atmosphere, carries with it a signature of the atmospheric and surface properties. By measuring sunlight transmitted directly or diffusively to the surface through the aerosol and cloud layers, surface based networks of radiometers can derive aerosol and cloud properties. By measuring the reflected light’s spectral, angular and/or polarization properties, satellite sensors can also quantify several atmospheric and surface properties (polarization is the degree of organization of the electric field of the scattered solar radiation into a given direction).

To get a perspective on our remote sensing capability from satellites and surface radiometers, we start with our experience with the human eye. Our eye is sensitive to a narrow range of the solar spectrum, with receptors in the blue, green, and red. We get depth perception using the eyes’ slightly different angles of observation. By analogy, aerosol remote sensing started using a single wavelength and single angle of observation [Fraser et al., 1984; Husar et al., 1997], corresponding to the viewing ability of a color blind person using a single eye (e.g. METEOSAT, and GOES). The TOMS instruments, flown since 1978 have two channels sensitive to ultraviolet (UV) light that were discovered to be excellent for observations [J. Herman et al., 1997] of elevated smoke or dust layers above scattering atmosphere. A two-channel technique was used for the AVHRR with simultaneous analysis of aerosol and clouds [Nakajima et al., 2001; Sekiguchi et al., 2003; Mishchenko et al., 2003]. These instruments supply a long term series of aerosol and clouds properties, currently 20 years long – see Figure 5-7 [Torres et al., 2002], that can be used to assess regional changes in the aerosol concentration; but calibration and characterization of the instrument performance is difficult [Vermote et al., 1995], and the TOMS analysis requires an assumption of the height
of the aerosol layer [Torres et al., 2002]. Note that the geostationary satellites, METEOSAT and GOES, have an advantage over the AVHRR and TOMS of being able to generate long term series with the full diurnal cycle, albeit at a single wavelength. This approach has not yet been completely explored.

![Diagram showing aerosol optical depth from the TOMS satellite UV measurements. Top: for the globe as a function of latitude [Torres et al., 2002], Bottom time series for India [Massie et al., 2004]](image)

**Figure 5-7.** 20 year time series of aerosol optical depth from the TOMS satellite UV measurements. Top: for the globe as a function of latitude [Torres et al., 2002], Bottom time series for India [Massie et al., 2004]

The first remote sensing instrument designed for aerosol measurements, POLDER (Polarization and Directionality of the Earth Reflectances, [Deschamps et al., 1996]), uses a combination of spectral channels in a range wider than the human vision (0.44-0.91 μm). The instrument is a wide-angle camera that observes the same target on the Earth at 14 different angles. POLDER also measures light polarization (that we cannot see) to detect fine aerosols over land, taking advantage of the difference between the small and spectrally neutral polarized light reflected from the Earth’s surface and the spectrally decreasing polarized light.
reflected by fine aerosols. Two instruments, MODIS (MODeRate resolution Imaging Spectroradiometer) and MISR (Multi-angle Imaging SpectroRadiometer), on the Terra satellite [King et al., 1999] have been measuring global aerosol distributions and properties since 2000. Over the ocean, MODIS uses the aerosol spectral signature in a wide range (0.47 to 2.1 µm) to distinguish small particles, which in high concentrations are typically associated with anthropogenic pollution or smoke, from coarse particles usually identified as natural sea-salt or dust [Tanrè et al., 1997; Kaufman et al., 2002a]. This information is used to derive the fraction of aerosol optical depth that is due to anthropogenic sources (21±7% over the oceans) [Kaufman et al., 2005]. Over land, MODIS uses the 2.1 µm channel to observe the surface cover properties, to estimate surface reflectance at visible wavelengths, and to derive the aerosol optical depth [Kaufman et al., 1997; Remer et al., 2005] from the residual reflectance at the top of the atmosphere. The brightness of the earth-atmosphere system at the top of the atmosphere is affected by the angle at which the light is reflected by the surface and atmosphere. MISR [Diner et al., 1998] takes advantage of this fact by detecting the reflected light at different viewing angles (nadir to 70° forward and backward) along the satellite’s track in a narrower spectral range (0.44-0.87 µm), thus being able to separate the aerosol signal from that of the surface, and to derive information about particle size and shape [Martonchik et al., 1998; Kahn et al., 2001]. A mixed approach using two view directions but a wider spectral range (0.55-1.65 µm) is used by ATSR [Veefkind et al., 1998] to derive the aerosol concentration and type. These instruments, POLDER, MODIS, MISR and ATSR have controlled calibration and characterization that results in known and smaller errors than the previous generation of instruments [King et al., 1999].

The realization that aerosols affect surface temperature, Earth’s radiation budget and precipitation patterns, creates a demand for more informative spaceborne observations. This is now achieved with the completion of the so-called A-Train in April 2006 (see Figure 5-8) that is designed to measure aerosol, clouds and precipitation, with PARASOL (Polarization and Anisotropy of Reflectances for Atmospheric Sciences coupled with Observations from a Lidar [POLDER-like instrument launched Dec. 2004]) and MODIS for passive aerosol and cloud observations in the visible to IR, the OMI instrument for aerosol observations in the UV to visible, the CALIPSO lidar for aerosol profiling in two spectral channels and cloud top height detection, and CloudSat radar for profiling cloud liquid water and structure.
In parallel with the development of aerosol and cloud measurements, the Tropical Rainfall Measuring Mission (TRMM) [Kummerow et al., 2000] was launched in 1997 to measure tropical-subtropical rainfall and in so doing, acquiring the first accurate, representative, and consistent ocean climatology of precipitation. TRMM carries the first spaceborne precipitation radar (PR) producing new measurements of precipitation’s vertical structure (~250 m resolution) and several additional instruments to conduct new observations of the Earth’s assortment of convection, frontal zones, precipitating storms, and tropical cyclones.

The set of measurements described above covers information from aerosol distribution and properties to cloud properties and precipitation. The satellite data are evaluated and complemented by ground-based remote sensing of aerosols (e.g., AERONET and MPL networks – Holben et al., 1998), aerosol effects on clouds (at the ARM sites – Feingold et al., 2003a; Kim et al., 2003; Garrett et al., 2004) and precipitation [Kummerow et al., 2000]. The data are used to evaluate aerosol transport models [e.g. Chin et al., 2002] and as input to climate models that study the effects of aerosol on clouds and precipitation [e.g. Menon et al., 2002]. Statistics have been acquired of the aerosol effect on cloud microphysics and

**Figure 5-8.** The A train: a constellation of satellites that jointly measure aerosol, clouds and precipitation properties.
Some issues relating passive remote sensing aerosol retrievals in the vicinity of clouds have been raised [e.g., Coakley et al., 2005]. It is recognized that the cloud properties (cloud droplet size distribution, extended cloud fields or broken clouds) being studied depend on the sensor resolution and on the method of investigation; which in turn affects the strength of the aerosol impact on clouds [Rosenfeld and Feingold, 2003].

The accuracy of satellite retrievals is higher over the dark ocean surface than over the variable land albedo. Densely vegetated land areas, e.g., the Amazon Forest, are dark and relatively uniform and are therefore some of the best targets for measurements of the aerosol impacts over the land. Surface conditions affect both the aerosol and cloud retrievals and the measurements of liquid water content from TRMM. Therefore, the statistical analysis of aerosol-cloud interaction can be done globally over the oceans and over some continental regions, in particular the dark forests. The increase in cloud fraction and cloud development with increase in aerosol loading is in agreement with the notion that increase in CCN loading modifies cloud microphysics and inhibits precipitation (see chapter 6 for more details), although the changes in cloud microphysics due to increase in aerosol on inhibition of precipitation has not yet been statistically established. New instruments are being designed or proposed for future launches with several critical innovations:

- Spaceborne instruments that combine a wide spectral range with a wide viewing angle range and with polarization [e.g., the Glory mission, Mishchenko et al., 1997; Advanced PARASOL, (Tanrê, private communication); Next-generation MISR (Diner, private communication)]. Such instruments are expected to improve derivation of fine and coarse aerosol concentrations, their sizes, shapes and refractive index. Refractive index, an optical property, is sensitive to the aerosol composition and water uptake [Chowdhary et al., 2002].
- Derivation of aerosol absorption over the land [Kaufman et al., 2001] and over the oceans by measuring the aerosol spectral attenuation of the bright glint [Kaufman et al., 2002b].
- Measuring the vertical distribution of cloud microphysics using slant angle spectroradiometers [Martins et al., private communication].
- New strategies and instruments to measure precipitation from space [Smith et al., 2003].
It is recognized that there is a need for coordination of the aerosol, clouds and precipitation measurements and coordination with in situ and network measurements as well as development of models that can integrate large data sets [Diner et al., 2004; Seinfeld et al., 2004].

In the last decade, we have witnessed a major step in the amount and quality of satellite and surface-based remotely sensed data of aerosol, clouds, and precipitation. We are probably at the peak of research benefits from this new array of satellite and surface measurements. Specific plans for the next 2 decades are presently limited in scope. More attention is needed to ensure that the degree of sophistication in the aerosol, clouds, and precipitation measurements and the degree of coordination among them continues to increase.

In the following we describe in more detail some of the satellite instruments used to measure the effect of aerosol on clouds and precipitation. One of the advantages of the multi-instrument system to measure clouds and aerosol is the ability to intercompare the results. Figure 5-9 shows an example of such intercomparison for 4 satellite observations of the pollution over the Indian subcontinent.
Figure 5-9. Intercomparison for 4 satellite observations (MODIS, POLDER, ATSR, and MISR) of the pollution over the Indian subcontinent. Top left - population density per km², top right - POLDER measurements of the Aerosol Index that is similar to the fine aerosol optical depth, center right - ATSR optical depth and Angström exponent for 1999, bottom right - MODIS aerosol optical depth for 2001 through 2003, left bottom [DiGirolamo et al., 2004] - MISR optical depth averaged over 2001-2004.

5.4.2 Remote sensing of clouds and aerosols by the POLDER Instrument

The POLDER instrument [Deschamps et al., 1994] measures the solar radiance reflected by the Earth’s surface and atmosphere in the visible and near infrared (440 to 910 nm). The two main characteristics of POLDER are the ability to measure the linear polarization of the radiance in three spectral bands, and to acquire the directional variation of the reflected radiance: the instrument concept is based on a wide field of view lens and a bi-dimensional Charge-Coupled Device CCD that provides an instantaneous field of view of ±43° along-track and ±51° cross-track. As the instrument flies over the target, up to fourteen views are acquired; these can be composited to infer the directional signature of the reflectance. This signature provides information on the surface, aerosol, and cloud characteristics. A limitation of POLDER is the rather coarse spatial resolution of about 6 km, which affects the ability to account for scene heterogeneity.
The POLDER instrument flew onboard the ADEOS 1 and 2 platforms in 1996-1997 and 2003, respectively. Unfortunately, due to the failure of the satellite solar panels, the measurement time series are limited to respectively 8 and 7 months. Another very similar instrument was launched in December 2004 onboard the microsatellite PARASOL to be part of the A-train. Algorithms have been developed to process the measurements in terms of aerosols [Deuzé et al., 2001; Herman et al., 2005] and cloud characteristics [Buriez et al., 1997, Parol et al., 1999].

5.4.2.1 Remote sensing of Aerosol Properties
Over ocean, the satellite algorithm [Herman et al., 2005] assumes spherical or non-spherical particles, non-absorbing particles and the size distribution follows a combination of two log-normal aerosol size distributions in the accumulation and coarse modes respectively. When the geometrical conditions are optimal, the shape (spherical or not) of the particles is derived. The refractive index retrieval is next attempted from the polarization measurements. If the real part of the refractive index of the coarse mode is retrieved when spherical particles are present (close to 1.35, indicating hydrated particles), the derivation of the refractive index of the accumulation mode is very tentative. Comparisons of AOD with ground based AERONET measurements show excellent agreement, with typical RMS errors on the order of 0.05, including errors due to cloud cover, with no significant bias (Goloub et al., 1999). The fine-mode optical depth can also be compared to AERONET measurements, albeit with some uncertainty in the aerosol radius cutoff. Statistical results indicate a low bias of 0.02 with a standard deviation of 0.02.

The retrieval of aerosol properties over land surfaces is based on polarized reflectance measurements. Polarized reflectance of land surfaces is small and fairly constant, although it does have a very strong directional signature [Nadal and Bréon, 1999]. Scattering by submicron (accumulation mode) aerosol particles generates highly polarized light [Deuzé et al., 2001], which makes the polarized satellite radiances more sensitive to the presence of aerosols than the total radiances. On the other hand, larger aerosol particles, such as desert dust, do not polarize sunlight and are therefore hardly detected from polarization measurements. Results of the optical depth of the accumulation mode derived from PARASOL are reported on Figure 5-10 for September 2005.
Figure 5-10. POLDER aerosol monthly optical depth of the accumulation mode for Sep. 2005. The accumulation fraction of the aerosol loading is derived with different methods over land and ocean. The continuity at the land/sea boundaries observed in most regions is a good indicator of the quality of the inversions. Note the aerosols emitted from South America, Equatorial Africa and Asia.

Over land, the evaluation of POLDER retrievals is made against the fine mode optical depth derived from AERONET measurements. The results show no significant bias and an RMS error on the order of 0.04 when dust-loaded atmospheres are excluded (i.e., a validation in regions affected by biomass burning or pollution aerosols).

5.4.2.3 Remote sensing of Cloud Properties

POLDER provides an innovative method to derive the droplet radius. It is based on the analysis of cloud bows (Figure 5-11) that can be observed in the polarized radiance measurements [Bréon and Goloub, 1998]. The method is very precise (error of <0.3 µm) with no known causes for error or biases, but it is applicable only in very specific conditions, i.e., extended cloud fields (150 x 150 km²) with narrow size distributions at cloud top. The requirements are most often fulfilled over stratocumulus cloud fields on the Eastern part of oceanic basins, as well as over the Antarctic Ocean [Bréon and Colzy, 2000]. A comparison with MODIS estimates of cloud droplet radius show a high correlation over the oceans, but poor correlation over land surfaces. There is a bias in the two estimates that is not understood [Bréon and Doutriaux-Boucher, 2005]. The CDR from POLDER has been used together with aerosol estimates to assess their effect on cloud microphysics [Bréon et al., 2002].
The polarization capabilities of POLDER also permit an unambiguous identification of the cloud top phase. The method is based on the presence or the absence of a “cloud bow” for a scattering angle of 140° [Riédi et al., 2000].

The POLDER instruments provide two means of evaluating the cloud top pressure. One method is based on the measurement of the polarized radiance generated by molecular scattering (in the blue part of the spectrum where molecular scattering is very efficient), that is decreasing as a function of the cloud top pressure. The statistics show a negligible bias and a standard deviation of 65 hPa. The other method is based on the ratio of two radiance measurements centered on the oxygen A-band at 763 nm [Vanbauce et al., 1998]. One channel is 10 nm wide while the other is 40 nm wide with the same central wavelength. As a consequence, the ratio of the narrow and wide channels provides an indication of the oxygen absorption between the reflector and the satellite, which is a direct proxy of the reflector pressure. A comparison with lidar measurements of cloud profiles has checked the validity of the two methods [Vanbauce et al., 2003].

**Figure 5-11.** Polarized reflectance from a cloud field as a function of the scattering angle in the three polarized channels of POLDER (blue - 443, green - 665 and red - 865 nm). The measurements over a 150x150 km² area have been averaged in 0.3 scattering angle bins. The figure shows the average (diamond) and the standard deviation (crosses) together with the result of a best fit (line). In this case, the best fit indicates a droplet effective radius of 8 µm.
5.4.3  Remote Sensing of Aerosols and clouds by MODIS

The MODerate resolution Imaging Spectroradiometer (MODIS) began collecting data in February 2000 from the Terra (10:30 am and 10:30 pm equatorial crossing time) spacecraft and June 2002 from the Aqua spacecraft (1:30 pm and 1:30 am equatorial crossing time). Special emphasis is given to on-board calibration facilities, lunar observations, and detailed analysis of the calibration time series on the ground [Barnes et al., 1998]. MODIS provides measurements of the spectral radiances from 0.41 to 14 µm in 36 spectral bands daily over most of the globe. MODIS measurements are used to characterize the global aerosol, clouds, water vapor, fires, vegetation, sea surface temperature, and ocean color. The aerosol characteristics are derived over the oceans [Tanré et al., 1997] and land [Kaufman et al., 1997] using independent algorithms.

5.4.3.1 The Aerosol Algorithm

Over oceans, the MODIS aerosol algorithm uses the measured 500 m resolution radiances from six MODIS bands (550-2100 nm) to retrieve aerosol information. Specifically, in cloud-free, glint-free ocean scenes [Martins et al., 2002], MODIS retrieves at a 10 km resolution: the aerosol optical depth at 550 nm, \( \tau_{550} \), the fraction of \( \tau \) contributed by the fine (sub-micron size) mode aerosol, \( f \), and the effective radius of the aerosol, \( r_{eff} \) [Tanré et al., 1997]. Over land, the 2.1 µm channel is used to estimate the surface properties [Remer et al., 2005] and the 0.47 and 0.66 µm channels are used to derive \( f \) and \( \tau \). Aggregation of the MODIS aerosol information from the 500 m pixels to the 10 km product allows for rigorous cloud screening, avoiding data gaps, and still generates large enough statistics for a stable and accurate product. The MODIS-derived aerosol properties have been validated using AERONET [Remer et al., 2005]. In agreement with theoretical error analysis [Tanré et al., 1997; Kaufman et al., 1997], the aerosol optical depth is derived with an error of \( \Delta \tau_{550} = \pm 0.03 \pm 0.05 \tau \) over the oceans and \( \Delta \tau_{550} = \pm 0.05 \pm 0.15 \tau \) over the land. Over the oceans, the errors were found to be mostly random with very little bias remaining for large statistics of data. For aerosol dominated by dust, a bias of about +10% was observed. Figure 5-12 shows example of the MODIS observations over the globe. [MODIS monthly data can be accessed interactively from http://lake.nascom.nasa.gov/movas/. Information and images on the MODIS atmospheric products can be found at http://modis-atmos.gsfc.nasa.gov/.
5.4.3.2 Remote Sensing Measurements of Cloud Properties

The MODIS cloud product [Platnick et al., 2003; King et al., 2003] combines infrared and visible techniques to determine the physical, radiative, and microphysical properties of clouds. Cloud optical depth and effective radius are derived globally using six visible and near-infrared bands at 1-km spatial resolution. Cloud-top properties, including cloud-top temperature, cloud-top pressure, and effective emissivity, are derived using the infrared split window and longwave CO2 absorption bands (both day and night) at 5-km spatial resolution. Cloud thermodynamic phase is computed at 5-km resolution using a two-band algorithm that includes an additional thermal band at 8.55 µm, and also at 1-km resolution using a different technique based on results from the cloud mask tests followed by a bispectral threshold test, shortwave infrared tests, and finally cloud top temperature [King et al., 2004]. Finally, the cloud product contains a cirrus reflectance product at a visible wavelength for use in removing cirrus scattering effects from the land surface, and it utilizes an additional band at 1.38 µm. Hence, the cloud product contains many different cloud properties derived from 14 bands in total, and the file size is different during the night (only cloud-top properties and thermodynamic phase at 5-km resolution) than during the day (when additional 1-km resolution products are included).

Figure 5-13 [King et al., 2003] shows an example of the cloud optical depth, cloud top pressure, and effective radius for a daytime granule of Terra-MODIS data over the Western Pacific Ocean near the Kamchatka Peninsula on August 10, 2001 at 0025 UTC. The true color image in Figure 5-13 (a) shows extensive cloud cover over the Sea of Okhotsk, including mid-level and upper-level ice clouds, whereas the Bering Sea to the east of the Peninsula contains extensive marine stratocumulus clouds with numerous ship tracks in the southeastern portion of the image. Figure 5-13(b) shows cloud optical depth, Figure 5-13(c) shows cloud-top pressure, and Figure 5-13 (d) shows cloud effective radius, where different color scales for water and ice clouds were used in Figure 5-13(b) and (d). The optically thick marine stratocumulus to the east of the peninsula is identified as water clouds with optical depths (0.65 µm) up to 25. The optically thick ice clouds over the Sea of Okhotsk and around the southern portion of the Kamchatka Peninsula have cloud optical depths approaching 40.
Figure 5-12. MODIS aerosol monthly composites for 2001. Each composite is for the 15th of each month ± 5 days to find enough cloud free regions. Data for June are not shown since no MODIS data were available during the middle of the month. The color bar is located instead. The color bar was constructed so that blue represents clean conditions, aerosol optical depth <0.1, green and red show higher optical depth corresponding to the coarse (green) and fine (red) modes. Therefore, pure dust is shown as green and pure smoke or pollution red. Note that the color of the aerosol emitted from Africa changes from mixed red and green in January through April, to green in July-August. Biomass burning occurs in the Sahel during January-March, and moves to Southern Africa for July-August, when it is separated from the dust flow.
Figure 5-13. Cloud properties over the western Pacific Ocean off the Kamchatka Peninsula on August 10, 2001. Panel (a) is a true color composite of one MODIS granule, showing marine stratocumulus clouds with ship tracks as well as upper level ice clouds. Panels (b) and (d) show the cloud optical depth and effective radius derived from all cloudy pixels, where a separate color bar is used to denote clouds processed as ice and water clouds. Panel (c) shows the cloud-top pressure for all clouds in this scene, (from King et al. [2003]).
Ship tracks are not easily identified in the cloud optical depth image shown here, but they result in reduced effective radii in the microphysical retrievals shown in Figure 5-13(d). For this scene, the thermal infrared and decision tree algorithms for deriving cloud thermodynamic phase, discussed by Platnick et al. [2003], are in quite good agreement. Finally, the cloud-top pressure for the cloud-filled pixels, shown in Figure 5-13(c), clearly show that the water clouds lie predominantly between 700 and 850 hPa, whereas the optically thick ice clouds over the Sea of Okhotsk lie at altitudes above the 500 hPa level.

5.4.4 *Measuring Aerosols With the Multi-angle Imaging Spectro Radiometer (MISR)*

The Multi-angle Imaging SpectroRadiometer (MISR) instrument on Terra [Diner et al., 2002] observes the Earth globally at nine angles, ranging from 70° forward to 70° backward along the spacecraft track. MISR aerosol, cloud, and surface products are available through the NASA Langley Atmospheric Sciences Data Center (http://eosweb.larc.nasa.gov); further information about the MISR Mission can be found at http://www-misr.jpl.nasa.gov.

Images are obtained in four spectral bands (446, 558, 672, and 866 nm) with a ground spatial sampling of 275 m - 1.1 km. Such data offers several advantages for aerosol retrieval. Oblique viewing provides high sensitivity to aerosol optical depth. Over major aerosol source regions such as deserts and urban areas, where the high surface reflectance presents a challenge, MISR algorithms make use of the systematically changing ratio of surface to atmospheric radiance with view angle to separate the surface from atmospheric signals [Martonchik et al., 1998; 2002]. The resulting optical depths show good agreement with values derived from AERONET [Kahn et al., 2005; Abdou et al., 2005; Martonchik et al., 2004; Diner et al., 2001].

Over water, MISR’s scattering angle coverage (~ 60°-160° in mid-latitudes) helps distinguish particle size and shape. Three-to-five particle size groupings between about 0.1 and 2.5 microns are identified in the data [Kahn et al., 2001; 2005], and recent work indicates that MISR is able to separate different mineral dust shape classes [Kalashnikova et al., 2005; Kalashnikova and Kahn, 2006]. Although detailed aerosol type validation opportunities are rare, retrievals from MISR data acquired during the CRYSTAL-FACE campaign identified
three distinct aerosol modes within the column: maritime-type particles, cirrus, and Saharan dust, in proportions that agree with independent data to within about 20% [Kahn et al., 2001]. On another occasion, retrievals over Galveston Bay, near Houston, TX, distinguished two aerosol modes, one with an effective radius near 0.1 µm and the other with an effective radius near 0.6 µm, showing sensitivity to differences between >~0.95 and <~0.88 in aerosol single scattering albedo, in agreement with field measurements upwind in the city, and consistent with pre-launch sensitivity expectations. Over water, MISR also generally has multiple glint-free cameras, making aerosol retrievals possible even when nadir-viewing instruments are glint-contaminated.

MISR routinely obtains the heights of reflecting surfaces such as cloud and aerosol plume tops stereoscopically, provided there is some spatial contrast so image pattern matching can be performed. This is often the case for aerosols near volcanic, forest fire, and dust source regions, where distinct plume features are present. The technique retrieves the altitude at which the reflectance contrast is greatest typically near the top of the plume. MISR algorithms perform pattern matching globally in an automated manner [Moroney et al., 2002], with 1.1 km horizontal resolution and about 0.5 km sensitivity in the vertical. An example for a wildfire is shown in Figure 5-14. The stereo retrieval is successful in distinguishing the heights of two separate plumes associated with the fire complex.

5.4.5 Meteosat Second Generation – MSG

The Meteosat series of geostationary satellites, at 0° longitude, are ideally located to follow the export of dust from the Saharan and Sahel dry areas and the biomass burning activity in Sahel and central Africa, two major sources of aerosols to the atmosphere. This opportunity has been exploited since the early 1990s to quantify the export of dust over the Atlantic and the Mediterranean [Jankoviak and Tanré, 1992; Dulac et al., 1992; Moulin et al., 1997]. Likewise, the thermal infrared channel of Meteosat was used to locate the main source areas of dust in Northern Africa through their impact on the apparent temperature at midday during dust storms and a corresponding Infrared Difference Dust Index (IDDI) has been defined [Legrand and N’Doumé, 2001].
Figure 5-14. MISR nadir camera image (left), 70° camera image (middle), and standard stereo height retrieval product (right) for the B&B fire complex in Oregon, acquired September 4, 2003. The smoke snakes northward from several blazes located near the lower left.

The first generation of Meteosat satellites carried only three wide spectral bands respectively in the visible band [0.45 to 1.0 µm], in the water vapor absorption band [5.7 to 7.1 µm] and in the thermal infrared (window) band [10.5 to 12.5 µm], so it is not possible, for instance, to distinguish between aerosol types like dust or biomass burning aerosols over ocean. This limitation does not apply to the Meteosat Second Generation (MSG) that carries three channels in the solar spectral range, from 0.6 to 1.6 µm, and two in the thermal infrared (window) band, 10.8 and 12.0 µm. It is expected that this operational satellite will permit an excellent monitoring of the aerosol sources and transport over the African continent. One major advantage of the geostationary satellites, compared to the polar orbiters, is their ability to provide images of a given area with high temporal sampling, every 15 min for MSG, which permits the monitoring of aerosol transport and impact on clouds and precipitation in different parts of the diurnal cycle.
5.4.6 Future missions for cloud and aerosol remote sensing.

The aerosol effect on the cloud albedo can be detected and quantified from space by means of long-term global measurements of the change in the number concentration of aerosol particles and the associated change in the cloud albedo and lifetime. The National Polar-orbiting Operational Environmental Satellite System (NPOESS) is being developed jointly by the National Ocean and Atmospheric Administration (NOAA), the Department of Defense and NASA. The satellite system is planned for long term monitoring of the Earth and atmosphere, from the Visible/Infrared Imager/Radiometer Suite, VIIRS, a multi-spectral imager somewhat similar to MODIS.

Most CCN are smaller than the particles that can be detected with current passive remote sensing. The aerosol size spectra retrieval together with information on the chemical composition could be a good proxy for CCN estimation [Mishchenko et al., 1997]. Although the chemical composition of aerosols cannot be determined directly from remote sensing measurements, the spectral refractive index and shape of aerosols provides strong constraints on aerosol composition. Other measurable manifestations of the indirect effect include modifications to the cloud droplet size distribution and number concentration and changing liquid water path [King et al., 1995; Schwartz et al., 1995, Brenguier et al., 2000, 2003]. As previously described, single radiance-only measurements provide estimates of column optical thickness, Ångström exponent, fine mode fraction, and effective particle size of aerosols as well as of optical depth, effective radius and phase (ice or water) of cloud droplets. Directional (MISR) and polarized radiances (POLDER) can better constrain the inversion to provide additional information on aerosol composition, single scattering albedo and shape, which also reduces the uncertainties on the retrieved aerosol and cloud parameters. This is of importance to understand the different type of aerosols that play a role in modifying cloud microphysics. It is also essential that the cloud properties do not have biases that depend on the type of cloud, or season. Extension to the short infrared spectral (such as MODIS) of the present directional and polarization capabilities will provide a detailed estimate of aerosol and cloud properties in which context the type of correlative studies currently being used to investigate the effect of aerosols on clouds can be better understood and evaluated.

In 2008, the National Aeronautics and Space Administration (NASA) Glory mission will deploy a high precision multi-spectral photopolarimeter called the Aerosol Polarimetry
Sensor (APS). Since these measurements will only be available along the spacecraft ground track, other satellites are needed to obtain such data at a global scale on a daily basis. Applications of satellite data are constrained by the spatial resolution of the sensor. A spatial resolution of 500m-1000m is required for an efficient screening between aerosol plumes and clouds.

5.4.9 Aerosol and Clouds Profiling From Space

5.4.9.1 The GLAS Lidar Mission

The Geoscience Laser Altimeter System (GLAS) was launched in 2003. It is the sole scientific instrument on the Ice, Cloud, and land Elevation Satellite (ICESat). GLAS combines a precision surface lidar with a sensitive dual wavelength cloud and aerosol lidar. GLAS has 3 lasers, operating one at a time, that emit infrared and visible laser pulses at 1064 and 532 nm wavelengths. On orbit, GLAS continuously emits laser pulses at a rate of 40 per second from the Earth facing (nadir) side of ICESat. GLAS measures precisely how long it takes for photons in a laser pulse to pass through the atmosphere to the Earth, reflect, and travel back to GLAS. Halving the total travel time, and applying corrections for the speed of light through the atmosphere, the distance from ICESat to the laser footprint on Earth can be calculated. ICESat collects data for calculating its position in space by using onboard GPS (Global Positioning System) receivers augmented by a network of ground GPS receivers and satellite laser ranging stations. The angle of the laser beam relative to stars is measured precisely by GLAS with star-tracking cameras on the zenith side of ICESat. As ICESat orbits, GLAS takes data along ground tracks defined by the sequence of laser spots. GLAS produces a series of approximately 70 m (230 ft) diameter spots that are separated by nearly 170 m (560 ft) along track. Though ICESat main mission is to measure polar ice, it also provides detailed information on the vertical distribution of clouds and aerosols illuminated by its laser spots (Figure 5-15).
5.4.9.2. The Calipso Satellite Mission

The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite, has been launched in April 2006. It continues the GLAS aerosol and cloud measurements but with emphasis on the atmospheric studies and in coordination with other measurements on the A-train set of satellites. CALIPSO combines an active lidar instrument with passive infrared and visible imagers to probe the vertical structure and properties of thin clouds and aerosols over the globe (Winker et al 2003). It is a joint U.S. (NASA) and French (Centre National d'Etudes Spatiales/CNES) satellite mission that will fly as part of the A-train with an expected 3 year lifetime.

Retrieving the spatial and optical properties of clouds and aerosols from satellite lidar is confronted by a number of difficulties that are not faced in the analysis of ground-based lidars. Among these are the very large distance from the target, the high speed at which the satellite traverses the ground track, and the resulting low signal-to-noise ratios [Winker et al., 2004]. To retrieve the aerosol extinction profile, the appropriate lidar ratio needs to be used in the optical analyses of aerosol layers. Field measurements of aerosol lidar ratios show large variability from 15 sr to 120 sr and cannot be derived from the backscatter lidar. Lidar-

Figure 5-15. GLAS profiles of aerosols on the region of the Himalayas. Image at (http://icesat.gsfc.nasa.gov/images/movies/india_noclouds_dates.mpeg).
only measurements of lidar ratios require a Raman lidar or a high spectral resolution lidar (HSRL). CALIPSO therefore determines the lidar ratios using a model-matching scheme, with the help of geophysical (e.g., latitude, longitude), and temporal (season) information. Six aerosol types are used, four of these (polluted continental, biomass burning, desert dust, and polluted dust) are derived from a comprehensive cluster analysis applied to AERONET data gathered from numerous sites around the globe [Winker et al., 2004]. Alternatively, the MODIS measurements can be used to constrain the inversion of the lidar data [Kaufman et al., 2003]. The combined lidar and spectral measurements can be used to sense the vertical distributions of aerosol concentrations and properties. Figure 5-16 shows the results of inversions of combined aircraft lidar and MODIS data to study the properties of smoke off the southwest coast of Southern Africa. The inversion derives profiles of the aerosol extinction due to fine and coarse particles. Comparisons with three sets of airborne in situ measurements show excellent agreement of the aerosol extinction profiles; however, the lidar inversion derives smaller spectral dependence of the extinction than the in situ measurements.

![Figure 5-16. Extinction profiles derived from the lidar & MODIS data (black), in situ nephelometer measurements (blue, light and dark green) converted to ambient extinction, airborne sunphotometer measured extinction profile (red), all for wavelength of 0.55 µm. The corrected in situ data reveal the smoke and boundary layer marine aerosol, shown in the inverted lidar-MODIS data.](image-url)
5.4.9.3. The Cloudsat Satellite Mission

CloudSat (Stephens et al. 2002), launched in April 2006 with CALIPSO, provides from space the first global survey of the vertical structure of cloud systems through 94 GHz radar reflectivity measurements. The unique feature of this radar lies in its ability to observe jointly most of the cloud condensate and precipitation within its nadir field of view and its ability to provide profiles of these properties with a vertical resolution of 500 m. The primary product is the level-1B calibrated, range-resolved radar reflectivities and the essential level-2 products are the cloud profile properties derived from these radar data. It flies in on-orbit formation with the Aqua, PARASOL and CALIPSO satellites, providing a unique, multi-satellite observing system (The A-Train) particularly suited for studying the atmospheric processes of the hydrological cycle.

5.5 In-Situ and Remote Sensing Of Precipitation

Precipitation is a discontinuous geophysical field. It is raining over only ~4% of the earth’s surface at any one time. Historical records indicate that rainfall has been measured for at least the last 2500 years. Simple rain gauges are based on catching rainfall in a bowl and measuring the contents at regular time intervals. Recent technological advances in rainfall measurement have focused on standardizing measurement in non-recording gauges, devising self-recording gauges, and on remote sensing methods.

The spatial variability of precipitation and its intermittency in time make areal rainfall difficult to measure accurately. The value of precipitation measurements depends on how well the measurement system is suited to the desired application. There is not a unique answer regarding the relative accuracy and precision of precipitation measurements since the magnitude of errors are dependent on the spatial and time scales of interest, the storm structure, and the instrument system making the measurement. There is also currently no consensus on how to report rainfall measurement uncertainties which makes comparison among different measurement systems and studies difficult.

5.5.1 In-Situ Point Measurements of Precipitation

Rain gauge measurements are usually underestimates of the true precipitation at the location of the instrument [Groisman and Legates, 1994]. Wetting on the internal walls of the gauge
and evaporation of liquid within the gauge decrease measured precipitation. Wind-related undercatch is usually the largest individual error source and is particularly large for snow. For objects such as precipitation gauges mounted above the surface, wind deviates around the object yielding a pattern of turbulence and accelerated and decelerated flow [Oke, 1987]. The divergence of flow acts to divert lighter particles such as small drops and snow, away from the top orifice of the gauge, which decreases the precipitation falling on the gauge in comparison with what it would be at the surface if the instrument was not present [Folland, 1988; Nešpor and Sevruk, 1999]. For a given particle size, undercatch increases with increasing wind speed [Sevruk, 1982]. Wind effects can be lessened by applying a wind correction based on the rain rate and wind speed [e.g., Sevruk, 1982; Hasse et al., 1998], by deploying a wind shield around the instrument, by modifying the shape of the instrument [Folland, 1988], and by placing the instrument in an area of low wind. Wind undercatch can be eliminated by placing the instrument in a pit so that the gauge orifice is at the ground surface. For typical rainfall events in the central plains of the USA, undercatch by above-ground tipping bucket and weighing gauges compared to a pit gauge is 4-5% [Duchon and Essenber, 2001]. For an extreme event in the same region with very heavy rain (200 mm/hr) and high winds (12.5 m s⁻¹) undercatch was 15% [Duchon and Essenber, 2001]. A separate study comparing measurements among above-ground gauges found that losses related to wind, wetting, and the common practice of counting trace precipitation (<0.127 mm) as 0 mm lead to average underestimates in rainfall of 20% for Barrow, Alaska [Yang et al., 1998].

The uncertainty in estimation of liquid water equivalent of snowfall is a function of wind speed, air temperature, whether a wind shield is present, and the type of wind shield and gauge [Goodison, 1978; Groisman et al., 1991]. Errors have been estimated to be as low as 10% at wind speeds ≤ 5.5 m s⁻¹ for Canadian Nipher shielded snow gauges [Goodison, 1978] to several hundred percent for unshielded 8” non-recording gauges historically used by the US weather service [Yang et al., 1998].

In situ rainfall measurements at sea vary in quality as a function of wind speed, instrument type and exposure, and to what degree the ship or buoy distorts the wind flow in addition to the distortion created by the instrument itself [Yuter and Parker, 2001].
5.5.1.1 Basic Types of In Situ Rain Instruments

Non-recording or manual rain gauges based on the bowl concept have hundreds of different designs and are still in widespread use. These instruments catch precipitation for recording by hand at regular time intervals. Measurement is performed either with the liquid in place or by pouring the contents of the gauge into a graduated cylinder. These methods yield slightly different estimates for the same catch. Depending on the length of time and the ambient environment, precipitation can be lost to evaporation before it is measured.

Different types of self-recording *in situ* rain instruments measure rainfall by myriad methods. Instruments that catch rainfall and direct it to a measuring mechanism employ a bucket on a pivot that tips when it is full (tipping bucket), the weight of collected water, the capacitance of the water within a container, and the conversion of collected water back into drops of fixed size and the counting of the formed drops [Strangeways, 1996]. Raindrop-induced optical scintillation is also used to estimate rain rate [Wang et al., 1979]. Hot-plate sensors estimate rain and snow rates using cooling due to melting and evaporation of precipitation [Rasmussen et al. 2002].

Disdrometers measure precipitation droplet size distributions by counting the number of drops within each of several size categories over a time interval. Disdrometers are a ground-based analog to aircraft precipitation particle probes (Section 5.2.2) since both yield size distributions of liquid and snow particles large enough to fall (≥ ~0.2 mm diameter). The range of particles sizes detected and the distribution of bin sizes vary among different instruments. The minimum particle size detected can also vary as a function of rain rate, and with the ambient wind and noise environment of the instrument [Joss and Gori, 1976, List, 1988, Kruger and Krajewski, 2002, Yuter et al. 2005]. The size distributions are used to estimate bulk properties of the raindrop distribution including median diameter, rain rate, liquid water content, and equivalent radar reflectivity [Joss and Gori, 1976]. Some disdrometers use the kinetic energy of a raindrop impact to count natural drops in specified size ranges [Joss and Waldvogel, 1967]. Other disdrometers use optical methods including occultation and image processing [Löffler-Mang and Joss, 2000; Fiser et al., 2002; Kruger and Krajewski, 2002; Barthazy et al., 2004]. Some optical disdrometers can measure particle size and velocity simultaneously and have the capability of distinguishing among particle

Vertically-pointing Doppler radar and profiler measurements are used to retrieve the drop spectra in a profile above the instrument [Gossard et al., 1990; Thomson and List 1995; Peters et al. 2002; Williams, 2002, Nissen et al., 2005]. Various methods are used to account for vertical air motion and turbulence, which if left uncorrected will reduce the quality of the retrieved DSD [Joss and Dyer, 1972]. Methods to retrieve precipitation rate based on line-integrated attenuation between microwave transmitters and receivers, including cell phone towers, are under development (e.g. Messer et al. 2006).

Cloud radar (mm-wavelength) measurements attenuate in rain but can provide information on precipitation properties within shallow clouds [Wang and Geerts. 2003; Bretherton et al. 2004; Li et al. 2004; Stevens et al. 2005]. A combination of cloud radar and backscatter lidar measurements can be used to estimate mean particle size and to retrieve liquid water flux in stratocumulus clouds [e.g. O’Connor et al. 2005]

Self-recording precipitation measurement instruments can be ranked by their precision and latency. The smaller the detectable accumulation of rainfall, the better the instrument is at sensing low rain rates and the start and end of rainfall events. Latency is the time delay between when the rain passes through the plane of the sensing orifice and when it is registered by the instrument. Instruments such as impact or optical sensors have latency equal to their measurement time interval. Instruments with funnels to direct rainfall into the measuring mechanism have latency greater than or equal to their measurement time interval. For instruments with funnels, the percentage of rainfall lost to instrument wetting and the latency increases with decreasing rain rate. At low rain rates, it may take minutes to hours for sufficient rain to accumulate and be detected by an instrument with coarse precision. The impact of precision and latency on rainfall measurement is illustrated in Figure 5-17 which shows 1 min rainfall data collected by an impact disdrometer [Joss and Waldvogel, 1967] with precision < 0.01 mm, a drop-forming rain gauge [Hasse et al., 1998] with precision of 0.1 mm, and two tipping bucket rain gauges with a precision of 0.254 mm. While all four time-series show the same general temporal variation in precipitation, the 1 min rain rates, storm accumulations, and storm durations differ among the instruments. The highest precision instrument yields the highest accumulation and the longest storm duration, while
the coarser precision instruments yield smaller accumulations and shorter storm durations. The latency at low rain rates is visible as gaps in the tipping bucket data, and to a lesser degree the drop-forming gauge data, when the disdrometer shows light rainfall. Relative errors between rain gauges and remote sensing estimates are addressed later under remote sensing systems.

5.5.1.2 Tipping Bucket Rain gauge Networks

Tipping buckets rain gauges are one of the least expensive automated gauges available, and are often deployed in operational rain gauge networks. These instruments have sampling limitations related to their coarse precision (0.1 mm to 0.254 mm increments of rainfall). Habib et al., [2001] characterized sampling errors in 0.254 mm (0.01 in) tipping bucket gauges commonly used in the USA. They simulated tipping bucket data from a high resolution data sample obtained during summer in Iowa. Their sample was dominated by intense convective storms, typically with significant temporal variability and intermittency, and had few stratiform events. The simulation of the tipping bucket data permitted isolation of the sampling error from other instrumentation errors such as calibration and malfunctions. Sampling errors decreased with increasing time scales and were primarily related to difficulty in capturing small temporal features of the rainfall time series. For measurements at individual tipping bucket rain gauges, sampling errors were ~50% for 1 min., 30% for 5 min. and 15% for 15 min. time scales.
Figure 5-17: Time series of recorded 1-min rain rates obtained from 1230-1630 UTC on 27 August 2003 from four instruments within 10 m of each other on Kwajalein Island in the Republic of the Marshall Islands. Instruments are (a) Distromet Inc. RD-69/ADA-90 impact disdrometer, (b) Eigenbrodt Inc. SRM-450 drop-form rain gauge, and (c)-(d) Qualimetrics Inc. 6011-A series 0.254 mm tipping bucket rain gauges.

Additionally, tipping bucket gauges are often badly sited relative to obstructions and poorly maintained [Strangeways, 1996]. Steiner et al. [1999] highlighted problems in operational tipping bucket rain gauge networks by comparing nearby gauges to each other and to a radar-derived rainfall map. Over a two year study period, Steiner et al. [1999] reported that none of the 30 gauges in the network covering an area of 21.4 km$^2$ worked 100% of the time, but that for 80% of the storms investigated ~70% of the gauges worked reasonably well. Tipping bucket rain gauges were prone to malfunction due to mineral particulate and biological detritus accumulation such as dust, blown grass, pine needles, spider webs, dead insects, and bird droppings. Detritus malfunctions include partial or complete blockage of the gauge funnel leading to underestimates of rainfall.

5.5.1.3 Point to Point Spatial Variations of Precipitation

Thiessen [1911] was one of the first to discuss large differences in measured rainfall among nearby gauges. Huff and Shipp [1969] used statistical correlation methods to describe spatial variability and to define sampling requirements for rain gauge networks based on data obtained during four summer seasons. Spatial correlations decay with distance at rates that
vary with rain type and synoptic pattern. Spatial correlations at a given scale decrease with increasing spatial variability. For thunderstorms, they found $r^2=0.98$ at 1.6 km distance and 0.88 at 9.7 km distance compared to $r^2=1.00$ at 1.6 km and 0.98 at 9.7 km during passage of low pressure centers [Huff and Ship, 1969]. Spatial correlation increased for increasing storm durations up to 12 hr and then decreased [Huff and Ship, 1969]. Morin et al. [2003] obtained similar results based on analysis of 15 summer airmass thunderstorms observed in southeastern Arizona. To measure at least 75% of the explained variance between sampling points within summer storms, Huff and Ship’s recommended rain gauge separations of 0.5 km for 1 min rain rates and 12 km for storm totals. In comparison, typical rain gauge densities in operational networks, such as the one in California, have gauge separations of ~10 km near population centers and ~20 km in rural areas. The Global Precipitation Climatology Project of the World Climate Research Program has estimated that between 5-20 rain gauges are required per 2.5° latitude box to meet a 10% criterion for relative sampling error for monthly precipitation [Rudolf et al., 1994]. This recommended rain gauge density is met in most industrialized areas but is not met over the majority of the earth’s land surface Figure 5-18.

Figure 5-18: Modified number of rain gauges used in GPCP monthly rainfall 2.5 deg x 2.5 deg grid product. The modified number includes the gauges within an individual grid box and 0.25 gauges for each surrounding grid box with at least one gauge. Plot courtesy of G. Huffman, NASA.
5.5.1.4 Point to Point Temporal Variation

The intermittence of rainfall describes its temporal variation between two points and provides information on the frequency of high spatial variability rainfall. Krajewski et al., [2003] examined the probability of the concurrent occurrence of rainfall accumulation and rainy time fraction for rainy periods exceeding 0.5 mm/hr at two points separated by 1 km distance for five regions: central Florida, Oklahoma, and Iowa City, Iowa in the United States, the Island of Rota (near Guam in the west Pacific), and Ji Paraná, Rondonia, in Brazil. The fraction of total time raining was small, less than 3% in Florida, Oklahoma, Iowa, and the West Pacific, and less than 6% in Brazil. The joint occurrence of rainfall accumulations 1 km apart greater than or equal to 0.5 mm/hr for 5 min was lowest in the west Pacific location near Guam (79%) and slightly higher in Brazil (85%) and Florida (85%). Widespread precipitating clouds were more common in the mid-latitude locations resulting in higher joint probabilities in Iowa (89%) and Oklahoma (89%) compared to the tropical sites. Joint occurrence of precipitation at gauges 1 km apart increased with increasing time interval.

5.5.1.5 Representativeness Error in Areal Precipitation Estimates

A good estimate of local precipitation (± 10%) can be made at the location of a gauge when it is properly sited, calibrated, and maintained. However, most applications of rainfall mapping require areal estimates of precipitation, frequently at a hydrologic basin scale. The transformation of a point measurement or set of point measurements into an areal estimate requires taking into account the temporal and spatial variability of precipitation. Multi-gauge and gauge-radar studies of populations of storms can provide information on the typical distribution of temporal and spatial variability characteristics. Usually, the actual variability of rainfall around a gauge is unknown and yields an error source associated with transforming between spatial scales that is independent of the errors associated with the instrument itself.

Representativeness error occurs when observations at the native scale of the sensor are used to represent data at a different spatial scale [Tustison et al., 2001]. For example, use of rain gauge data, a measurement made at a scale less than 1 m², to represent areal average rainfall for an area >> 1 m² introduces a representativeness error. Analogously, use of satellite data with a sensor spatial resolution of 20 km x 30 km to represent rainfall at 5 km x 5 km scale also introduces a representativeness error as does use of operational radar data at 1 km x 1 km scale to represent still finer scale precipitation variability. Representativeness error is
independent of measurement errors at the native scale such as calibration errors. Representativeness error between gauge measurements and areal estimates is dependent on the difference in spatial scales, the smoothness of the actual rainfall field, rain gauge spacing, and the method used to transform between the scales. A numerical experiment designed to isolate representativeness errors from other error sources found that representativeness error ranged from 25% to 95% of the hourly conditional mean rainfall and up to 50% of the hourly areal average rainfall for scales between 5 to 50 km [Tustison et al., 2001].

5.5.2 Radar Estimation of Surface Rainfall

Weather radar [Battan, 1973; Doviak and Zrnic, 1993; Meischner, 2004, Rinehart, 2004] can provide estimates of precipitating storm location, size, and intensity over large areas and at higher spatial resolutions and time resolutions than operational rain gauge networks. Ground-based scanning precipitation radars are used in short-term weather and flood forecasting, and to estimate the distribution and amount of cumulative rainfall over a region. The weather services of many countries have networks of operational radars that monitor precipitation near population centers, as well as in remote areas.

5.5.2.1 Methods and Sources of Error

Currently, radar reflectivity is the primary input to algorithms used to estimate surface rain rates. Both radar reflectivity (Z) and rain rate (R) are moments of the raindrop size distribution and can be related by a power-law of the form \( Z = aR^b \) called the Z-R relation. Sources of error in radar-derived estimates of rainfall [Austin, 1987; Joss et al., 1998; Krajewski and Smith, 2002; Yuter, 2002] include: non-meteorological echoes such as ground clutter, sea clutter, insects, birds, and anomalous propagation; instrument noise, absolute calibration, beam blockage, changes in the vertical profile of precipitation from the height of the radar beam to the surface, signal enhancement by melting particles or hail, attenuation, presence of downdrafts, inhomogeneities in rainfall within the radar resolution volume, and variations in the drop size distributions.

Inhomogenities within the radar resolution volume arise when precipitation comes in "sheets" of no more than a few hundred of meters [Thompson and List, 1999]. Figure 5-19 shows an example of the inhomogeneities in precipitation seen by a vertically pointing Doppler radar with 1° beam width (17m wide at 1 km height increasing to 174 m wide at 10 km height). It shows heavy precipitation features < 200 m in size. The large mismatch in
spatial scales between these fine-scale features producing the precipitation and the average reflectivity within a several km wide resolution volume obtained by a scanning radar beam yields an incorrect representation of precipitation and potentially significant errors [Thompson and List, 1999].

Figure 5-19  Vertically pointing Doppler radar scans through a passing hailstorm. The colored regions represent the total vertical velocity. The thick black line shows the contour of the 30 dBZ data. The letters A, B, and C mark the west-east locations of the high negative velocity streaks (hail) extending up to 6.5 km in the vertical direction. These hatched areas represent aliased data that have been corrected by hand. Figure 2b with different geometrical aspect ratios: Power spectra can be viewed as a measure of particle concentration for different particle speeds. They are superimposed to show the high spatial variability of precipitation. The axes for each spectrum span from -16 to 16 m s⁻¹ horizontally and -30 to 40 dBZ (m s⁻¹⁻¹) vertically. The origins of the power spectrum axes [0 m s⁻¹, -20 dBZ (m s⁻¹⁻¹)] are positioned with respect to the background velocity data at the locations where each of the spectra was measured. Noisy data above 10 km are caused by low signal quality (from Thomson and List, [1999]).
Different operational and research radar groups producing rain maps have different methods to mitigate these errors. Most apply some method of quality control to remove non-meteorological echo as an initial step. Some rain-map products use a mean field bias adjustment [Amitai et al., 2002, Gjertsen et al., 2004] to address multiple error sources simultaneously while others address individual sources of error independently [Joss et al., 1998; Zawadzki and Bellon, 2003; Houze et al., 2004]. The use of polarimetric radar variables for quantitative rainfall estimation mitigates some error sources such as attenuation and the presence of hail. Hybrid methods using a combination of reflectivity and polarimetric variables as input are being developed and evaluated [e.g., Bringi and Chandrasekar, 2001; Illingworth, 2004] and are planned for operational deployment in several countries later this decade.

5.5.2.2 Quantification of Errors in Radar-Rainfall Estimates

The weather service of Switzerland, MeteoSwiss, has a comprehensive program to quantify errors in surface precipitation estimates and, based on those empirical results, to refine quantitative rainfall products from their radar network [Joss et al., 1998; Germann and Joss, 2004]. Precipitation estimation in mountainous terrain has additional challenges compared to flat terrain. Gauge observations in mountains have large uncertainties associated with wind, snow-drift, and small-scale spatial variability. Radar observations are prone to both strong ground clutter and beam shielding by mountains. The historical data analysis and algorithm refinements by MeteoSwiss (Table 5-1) illustrate several important points. A particular algorithm refinement may improve some statistics but degrade others. To improve the statistics across the board, a combination of refinements was needed including: rigorous elimination of clutter, correction for the vertical profile of precipitation between the altitude of radar measurement and the ground surface, hardware calibration to insure reproducibility of measurements, and long-term adjustment to independent data to minimize bias. These data also provide information on the residual relative error for daily rainfall between rain gauge and the average of nine 1 km$^2$ radar pixels centered over the gauge when primary sources of error in radar rainfall mapping have been minimized. For 2004 storms with daily rainfall $> 0.3$ mm, probability of detection is 0.89 and false alarm rate is 0.14. For 2004 storms with daily gauge rainfall $\geq 0.3$ mm, average bias for over all of Switzerland (41,290 km$^2$) is -11% and scatter is 3 dB(R). The scatter refers to the spread of daily radar/gauge ratios when pooling all rainy days and gauge stations over all of Switzerland together. For example, 3.0 dB(R) scatter in summer 2004 means that the radar-derived estimate of daily precipitation at
a gauge location is within a factor of 2 of the gauge estimate for 68% of the rainy days, for the remaining 32% the uncertainty is larger. Uncertainties decrease closer to the radar. Within 70 km from a radar, scatter decreases to a factor of 1.4. For winter of 2004, average bias over Switzerland is -3% and scatter is 3.9 dB(R) or a factor of 2.5. These relative errors are a result of the superposition of gauge and radar uncertainties and hence the values in Table 5.1 are overestimates of the true uncertainty of the radar estimates [Germann et al., 2006].

Steiner et al., [1999] examined relative errors of 1 km² radar-derived rainfall compared to tipping bucket rain gauges within a network of 30 rain gauges within the Goodwin Creek watershed (21.4 km²) in Mississippi, USA. For heavy rain events with rainfall accumulations >10 mm, the root mean square errors of storm total rainfall between radar and rain gauges were 10% to 40% when using the subset of rain gauges passing quality control and up to 80% when all gauges in the network were included. For storms with rainfall accumulations <10 mm, uncertainties are expected to be higher.

Table 5-1. Relative errors between MeteoSwiss RAIN product for daily radar-derived rainfall and daily rain gauge totals for summer season in Switzerland. The radar data are averaged over nine 1 km² pixels centered on 58 gauge locations including 19 gauges used for long-term adjustment. Bias and scatter statistics based on events with daily gauge rainfall ≥ 0.3 mm. Probability of Detection (POD, perfect score = 1) and False Alarm Rate (FAR, perfect score = 0) based on events with daily rainfall > 0.3 mm. Adapted from Germann et al. [2006].

<table>
<thead>
<tr>
<th>Year</th>
<th>Bias (dB)</th>
<th>Bias (%)</th>
<th>Scatter (dB)</th>
<th>Scatter (factor)</th>
<th>POD</th>
<th>FAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>-2.4</td>
<td>-43</td>
<td>4.0</td>
<td>2.5</td>
<td>0.84</td>
<td>0.34</td>
</tr>
<tr>
<td>1998</td>
<td>-1.2</td>
<td>-23</td>
<td>4.0</td>
<td>2.5</td>
<td>0.87</td>
<td>0.30</td>
</tr>
<tr>
<td>1999</td>
<td>-4.2</td>
<td>62</td>
<td>4.8</td>
<td>3.0</td>
<td>0.73</td>
<td>0.08</td>
</tr>
<tr>
<td>2000</td>
<td>-3.2</td>
<td>52</td>
<td>4.4</td>
<td>2.8</td>
<td>0.74</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Changes in operational clutter elimination

<table>
<thead>
<tr>
<th>Year</th>
<th>Bias (dB)</th>
<th>Bias (%)</th>
<th>Scatter (dB)</th>
<th>Scatter (factor)</th>
<th>POD</th>
<th>FAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>-4.1</td>
<td>61</td>
<td>3.0</td>
<td>2.0</td>
<td>0.75</td>
<td>0.08</td>
</tr>
<tr>
<td>2003</td>
<td>0.2</td>
<td>5</td>
<td>3.4</td>
<td>2.2</td>
<td>0.89</td>
<td>0.18</td>
</tr>
<tr>
<td>2004</td>
<td>-0.5</td>
<td>-11</td>
<td>3.0</td>
<td>2.0</td>
<td>0.89</td>
<td>0.14</td>
</tr>
</tbody>
</table>
Amitai et al. [2002] found average relative errors of 20% between monthly tipping bucket rain gauge accumulations and 4 km$^2$ radar-derived rainfall accumulations over the gauge locations for August-September 1998 in Melbourne, FL. The 15 gauges were part of a dense rain gauge network of 10 km$^2$ area within 35 km range of the radar. They found that the average difference between the radar estimate over the gauge and the gauge accumulation was the same order as the differences between point rainfall measurements within the same 4 km$^2$ radar pixel.

Houze et al. [2004] used data from two rainy seasons at Kwajalein Atoll, Republic of the Marshall Islands, to estimate the magnitude of several independent components of radar-derived rain-map uncertainty under tropical oceanic conditions. Underlying assumptions regarding radar calibration, vertical profile correction, the Z-R relation, and missing data were varied within empirically determined bounds and their impact was assessed on the resulting monthly areal mean rainfall accumulation based on 4 km$^2$ pixels averaged over a ~7 x 10$^4$ km$^2$ area (between 17 km and 150 km range from the radar). The largest individual source of uncertainty was the absolute calibration of the radar, which was estimated at ±30% based on the practical limit of ±2 dB(Z) on reflectivity calibration using several methods in the Kwajalein setting. The next largest source of uncertainty (±10 - 15%) was associated with the correction for the vertical profile of precipitation from the altitude of the radar beam to surface. Since the radar beam altitude varies with range, this correction is largest for pixels at farthest ranges. Uncertainties in the Z-R relation based on disdrometer data accounted for only ±5% uncertainty and small number of data gaps when the radar was not operational had associated uncertainty of < ±2%. A total uncertainty in monthly areal average mean rainfall of 50% was estimated by adding the individual error sources.

### 5.5.2 Satellite Estimation of Surface Rainfall

Weather satellites in geosynchronous orbit (36,000 km altitude) and low earth orbit (typically 250-1000 km altitude) use remote sensing measurements to estimate precipitation at the earth’s surface. The primary methods are infra-red (IR), passive microwave, and radar (active microwave). None of these methods measure rain rates directly. Rather, the satellite observed physical characteristics of precipitating clouds are used as inputs to algorithms of varying complexity to retrieve surface rainfall.
5.5.2.1 Satellite Precipitation Retrieval Algorithms

Methods to estimate precipitation from geosynchronous IR satellite data are based on empirical relations between area fraction of cold cloudiness and areal average rainfall. These methods decompose the areal average rainfall into a rainfall area estimated by satellite and a mean conditional rain rate estimated by independent measurements. In a series of studies [Arkin, 1979; Richards and Arkin, 1981; Arkin and Meisner, 1987], Arkin and his collaborators refined empirical relations between IR cloud temperatures and areal-average surface rainfall accumulation estimated from an array of four shipborne radars using data sets collected in the eastern tropical Atlantic. The best correlations ($r^2=0.8$) were found for the largest spatial scale ($2.5^\circ \times 2.5^\circ$) and temporal scale (1 day) they examined. Correlations between rainfall area and IR cold cloudiness for deep tropical oceanic convection are very poor ($r^2=0.05$) for instantaneous data at 240 km scale [Yuter and Houze, 1998]. The GOES Precipitation Index (GPI) algorithm [Arkin and Meisner, 1987] defines areal average precipitation as the product of the mean fractional cloudiness in a $2.5^\circ \times 2.5^\circ$ box lower than an IR temperature threshold, the time interval in hours, and a mean conditional rain rate in mm/hr. Based on data from the tropical eastern Atlantic and South America, they recommended a temperature threshold of 235K, and a mean conditional rain rate of 3 mm/hr for the tropics. A temperature threshold of 220K was recommended for the extratropics based on comparisons over the continental USA. Comparisons between GPI estimates and independent estimates over land using rain gauges and radar show that GPI reproduces the occurrence of deep convective rainfall well but is less successful at estimating the intensity of rainfall [Arkin and Meisner, 1987; Janowiak, 1992].

Multispectral rainfall algorithms designed for geosynchronous satellites [e.g. Ba and Gruber, 2001] often follow a similar methodology to GPI by decomposing rainfall into the probability of detectable rain and a mean rainfall rate but utilize IR as well as visible, near infrared, and water vapor channels. Multispectral methods exhibit improved detection of rainfall from lower altitude (higher temperature) cloud tops as compared to IR-only methods. A limitation of these algorithms is that the visible channel is available only during the day. Additionally, more fractional cloudiness is not always correlated with more precipitation. Subtropical marine stratocumulus clouds typically yield more drizzle from regions with broken cloud compared to regions of unbroken clouds that have higher areal average albedo [Comstock et al., 2005, Stevens et al., 2005].
Satellite passive microwave precipitation estimation techniques utilize the modification of surface upwelling radiation by precipitation [Spencer et al., 1983, Wilheit, 1986]. Microwave brightness temperatures are the result of the column scattering and absorption properties of precipitation, cloud, and water vapor. Observed brightness temperatures are dependent on the particular hydrometeor types and their mixing ratios within the column and the frequency of the sensor. Between 22 and 60 GHz both scattering and absorption are important and either may dominate depending on the specific situation. Below 22 GHz, absorption, primarily by liquid water, dominates. Above 60 GHz, ice scattering dominates. Scattering by precipitation is detectable over both water and land. Surface water has low emissivity at microwave frequencies and thus presents a cold background, which makes the small increase in brightness temperature associated with absorption by precipitation detectable. Land has larger and variable emissivities and presents a much warmer background against which the absorption signal is difficult to detect unambiguously [Wilheit, 1986]. As a result, passive microwave precipitation retrievals weight the input frequency channels differently over oceans and land.

Passive microwave algorithms address the problem of relating brightness temperatures (column measurements) to surface rainfall (the flux at the bottom of the column) using some combination of “empirical” and “physically-based” methodologies [Smith et al., 1998]. Empirical algorithms [e.g., Liu and Curry, 1992; Ferriday and Avery, 1994] derive a statistical regression between measured brightness temperatures at several frequencies and ground-based measurements of rainfall (radar and/or rain gauge network). Physically-based algorithms [e.g., Kummerow, 1998; Bauer, 2001] use cloud models to simulate the three-dimensional water vapor, cloud, and precipitation properties of storms, which are in turn used as input to forward radiative transfer calculations to estimate brightness temperatures. The cloud model and radiative transfer calculations yield a database associating sets of brightness temperatures with different precipitation structures and surface rainfall rates.

Error sources in both empirical and physically-based passive microwave estimates [Tesmer and Wilheit, 1998; Bauer et al., 1999; Harris and Fofoula-Georgiou, 2001; Olson et al., 2001] are related to the three-dimensional precipitation structure and include uncertainties in the depth of the rain layer, mixing ratio profiles of cloud water, water vapor, and super-cooled precipitable water, the radiative characteristics of mixed phase layers such as wet
aggregates in the melting layer, the natural variability of the profile of particle size distributions, and inhomogeneities of the rainfall pattern at scales smaller than the sensor field of view.

Intercomparison of different passive microwave precipitation retrieval algorithms using the same input data set revealed a bias uncertainty between estimates of ±30% [Smith et al., 1998]. Differences among algorithms were partitioned into those associated with rainfall detection (rain area) and those associated with the conversion from brightness temperature to rain rate (rain intensity). For oceanic regions, detection differences mainly affected light rainfall regions that did not contribute significantly to area average rain rates. Over the ocean, the largest contribution to relative error among algorithms was the weighting of the input frequency channels, which impacts how intense rain rates were calculated and the maximum allowable rain rate. In contrast, over land rainfall detection was the dominant factor in producing differences among the algorithms.

The Tropical Rainfall Measuring Mission (TRMM) satellite, a collaborative effort between the U. S. and Japanese space agencies, has as one of its sensors the first spaceborne precipitation radar (PR) [Kummerow et al., 1998]. The PR operates at Kᵤ-band (2.3 cm) and yields reflectivity profiles at satellite-relative orientations from nadir (0°) to ±17°, yielding earth-relative vertical resolutions of 250 m to 1500 m. Unlike operational weather radars operating at S-band (10 cm) and C-band (5 cm), the Kᵤ-band signal is strongly attenuated by rainfall. The signal attenuation is cumulative downward and is largest near the earth’s surface. An attenuation correction must be applied to the near-surface reflectivity (Z) values before a meaningful transformation to surface rainfall (R) can be made. There are currently several methods to estimate total path-integrated attenuation and to distribute the attenuation correction in the vertical [Meneghini and Kozu, 1990; Iguchi et al., 2000]. Independent evaluation and error diagnosis of attenuation correction methods applied to satellite radar data has proven to be difficult. Error sources in the estimation of surface precipitation from an attenuated reflectivity profile include many of the same error sources in passive microwave algorithms. The set of passive microwave sensors on the TRMM satellite are referred to as the TRMM Microwave Imager (TMI). Zonal averages of areal average precipitation derived from TRMM TMI and PR at native sensor resolution exhibit relative differences of 20-40% [Kummerow et al., 2000].
Multi-satellite products attempt to combine the strengths of different types of satellite sensors to mitigate their individual weaknesses. Several multi-satellite products use variations of the GPI algorithm [Janowiak and Xie, 1999; Huffman et al., 2001; Kuligowski, 2002. IR cloud top temperatures from geosynchronous satellites are used to determine rain areas at high temporal resolution. The IR temperature threshold and average conditional rain rate are adjusted for designated time and spatial scales using some combination of global passive microwave data and rain gauge data over land. Comparisons between dense gauge networks over land to these types of multi-satellite products show similar results to those products using IR data alone, they perform better in deep convection typical of the warm season than in shallower convection typical of the cool season, and tend to perform poorly over complex terrain regardless of season [Arkin and Meisner, 1987; Huffman et al., 2001].

Global satellite observations, rain gauge-derived areal estimates over land, and numerical model outputs are quality controlled and merged to create monthly analyses as part of the Global Precipitation Climatology Project (GPCP) of the Deutscher Wetter Dienst/WMO/WCRP Global Precipitation Climatology Centre [Huffman et al., 1997, Adler et al., 2003] and the NOAA Climate Predication Center’s Merged Analysis of Precipitation (CMAP) [Xie and Arkin, 1997]. GPCP is an element of the Global Energy and Water Cycle Experiment of the World Climate Research Program (WCRP). The GPCP and CMAP groups use different methods to combine rainfall data from multiple sources. Products include global maps of monthly precipitation on 2.5° by 2.5° and 1.0° by 1.0° latitude/longitude grids and monthly precipitation anomaly compared to a long-term average (see Chapter 1, Figure. 1-6 and Table 1.1). The value of merged products is based on the assumption that the combination of several estimates with unknown errors is closer to the truth than any single estimate. Analogous to ensemble numerical forecast modeling, this assumption holds if the individual estimates are distributed around the true value.

5.5.2.2 Uncertainties Associated With Satellite Temporal Sampling

Common to all satellite precipitation estimates are uncertainties related to estimating temporal average precipitation (e.g., daily, 5-day, monthly) from the nearly instantaneous samples for a given area obtained by satellite. Depending on storm duration, structure, and stage of evolution, a single snapshot of a storm has varying value as a representative sample of the storm temporal and spatial average. Steiner et al. [2003] utilized a set of operational
radar network composites over the United States central plains during summer as a baseline
data set to investigate errors associated with different satellite temporal sampling schemes.
The radar composite had a spatial resolution of 2 km and a temporal resolution of 15 min.
These data were sub-sampled at 1-, 3-, 6-, 8-, and 12-h intervals and averaged over 100-,
200-, and 500-km spatial scale domains. The uncertainty associated with discrete temporal
sampling varies directly with sampling time interval and inversely with rainfall rate and
space and time domain size.

Although summer-time rainfall in the United States central plains has different characteristics
than rainfall in other seasons and at other locations, the uncertainty values in the Steiner et al.
[2003] study provide guidance on the magnitude of errors due only to satellite temporal
sampling. Based on Steiner et al.’s calculations for a 100 km scale domain with areal average
rainfall of 0.1 mm hr\(^{-1}\), median rainfall sampling uncertainty for daily rainfall is 10% for 1
hour, 43% for 3 hour, 92% for 6 hour, and 154% for 12 hour sampling time intervals. The
uncertainties decrease for 500 km scale to 3% for 1 hour, 11% for 3 hour, 31% for 6 hour and
61% for 12 hour. Current geosynchronous satellites sample IR full-disk images at 1 to 3 hour
intervals. The temporal sampling of individual low-earth orbit satellites such as TRMM is
dependent on the details of the orbit and latitude. TRMM Precipitation Radar revisit times
over an area of 120 km radius vary from 78 hours at the equator to 17 hours at 35 deg
latitude. A set of low-earth orbit satellites with similar passive microwave sensors can be
utilized as a group to obtain temporal sampling intervals of ~6 hour in the tropics (30°S to
30°N) decreasing to ~3 hour at latitudes above 60° [Turk et al., 2003].

Since satellite algorithm retrieval errors are strongly dependent on the details of the three-
dimensional precipitation structure, these errors vary with regional and seasonal changes in
precipitation structure. It would be useful to isolate errors in the satellite retrieval algorithm
from errors related to temporal sampling, differences in spatial scale, and differences in
minimum detectable rain rate that are a function of the sensitivity of the sensor. In practice,
this is difficult since the different types of sensors on the current generation of satellites were
not designed to have matching spatial resolutions or sensitivity.

5.5.2.3 Future Satellite Precipitation Measurements

The proposed joint US and Japanese space agency Global Precipitation Measurement (GPM)
Mission builds on the success of TRMM to address the critical need for a more
comprehensive global precipitation measuring program. The GPM Mission consists of a Core satellite similar to TRMM and a constellation of internationally contributed multichannel passive microwave satellites. The GPM Core satellite will be used as a calibration reference for moderate to heavy rainfall for the constellation satellites. Additionally, the replacement of the single frequency Ku radar used on TRMM with a dual frequency radar (Ku/Ka) with improved sensitivity on the GPM Core satellite provides the ability to estimate bulk properties of the rain drop size distribution and to improve the attenuation-corrected vertical profile of reflectivity and rain rate using differential attenuation methods. An additional satellite, EGPM, proposed to the European space agency will carry a multichannel radiometer and Ka-band radar. The higher sensitivity of the Ka-band radar on EGPM (~5 dBZ) compared to the GPM Core satellite Ka radar (~12 dBZ) and higher orbit inclination for EGPM (90 deg) versus GPM (65 deg) are designed to optimize observations of light rainfall and light to moderate snowfall. Analogous to GPM Core satellite, the EGPM satellite will be used as a calibration reference for light rainfall and snow for the passive microwave constellation satellites. The full set of satellites associated with the GPM mission--GPM core, EGPM, and the constellation satellites will be used together to improve the frequency and quality of global precipitation estimates (Figure 5-20).

![Figure 5-20](image-url)

**Figure 5-20.** Revisit times of satellite passive microwave radiometers. (left) Current sampling by TRMM, Aqua, and DMSP F13, F14 and F15 satellites yields < 3h sampling over 35% of the globe and < 4 h sampling over 57% of the globe. (right) Best-case scenario sampling in GPM-era based on planned orbits of GPM-core, NASA-90°, GCOM-W, EGPM, and 3 NPOESS satellites would yield < 3 hr sampling over 96% of the globe. (NASA, Goddard Space Flight Center).
5.5.3 Discrepancies Among Different Methodologies

One of the most extensively analyzed data sets of recent years is from the Tropical Ocean Global Atmosphere Coupled Ocean Atmosphere Response Experiment (TOGA COARE) [Webster and Lukas, 1992], which took place in the western Pacific warm pool region from November 1992 to February 1993. The TOGA COARE intensive flux array (IFA) had an area of \( \sim 2.25 \times 10^5 \) km\(^2\). Areal average daily rainfall estimates for the IFA (Table 5-2) were obtained using many different methods for estimating surface rainfall [Johnson and Ciesielski, 2000; Weller et al., 2004]. Values varied from 5.4 mm/day for the shipborne radars to 9.3 mm/day for a merged gauge/satellite method. Discrepancies are a function of differences in subregion examined, temporal sampling, and spatial resolution of the rainfall estimates [Weller et al., 2004]. These differences make it difficult to isolate errors associated with specific techniques or to determine the best estimate. For example, Johnson and Ciesielski [2000] showed that part of the reason the ship radar data estimate was lower than the other estimates was that the ships were not present for some of the heavier precipitation events during the study period, and that one of the two radar ships was stationed in a precipitation minimum.

Table 5-2. TOGA COARE intensive flux array intensive observation period estimates of areal average daily rainfall using different techniques. “Sat/mixed” [Curry et al., 1999] utilized a combination of passive microwave brightness temperatures and visible and IR radiances from geosynchronous and low-earth orbiting satellites. “CMAP” refers to the CPC Merged Analysis of Precipitation of rain gauge, IR and microwave satellite estimates [Xie and Arkin, 1997]. Adapted from Johnson and Ciesielski [2000] and Weller et al. [2004].

<table>
<thead>
<tr>
<th>Method</th>
<th>Rainfall estimate (mm/day)</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmospheric moisture budget multiquadratic technique</td>
<td>8.2</td>
<td>Johnson and Ciesielski [2000]</td>
</tr>
<tr>
<td>Atmospheric moisture budget-Barnes scheme</td>
<td>5.6</td>
<td>Johnson and Ciesielski [2000]</td>
</tr>
<tr>
<td>Ocean freshwater budget (near ship)</td>
<td>8.0</td>
<td>Feng et al. [2000]</td>
</tr>
<tr>
<td>Shipborne radar (2 km resolution within 145 km of ships)</td>
<td>5.4</td>
<td>Short et al. [1997]</td>
</tr>
<tr>
<td>Sat/Mixed</td>
<td>8.3</td>
<td>Curry et al. [1999]</td>
</tr>
<tr>
<td>CMAP</td>
<td>9.3</td>
<td>Johnson and Ciesielski [2000]</td>
</tr>
<tr>
<td>NCEP T62 reanalysis (210 km resolution)</td>
<td>8.4</td>
<td>Johnson and Ciesielski [2000]</td>
</tr>
<tr>
<td>ECMWF T106 reanalysis (125 km resolution)</td>
<td>6.7</td>
<td>Johnson and Ciesielski [2000]</td>
</tr>
</tbody>
</table>
Attempts to inter-compare monthly global estimates of precipitation [e.g., Adler et al., 2001] have suffered similar problems to the TOGA COARE rainfall comparisons. The impacts of differences in temporal sampling and spatial scales of the input sensor data are not usually removed. As a result, the factor of 2 to 3 differences among different satellite retrievals and ground-based observations [Adler et al., 2001] are difficult to diagnose in terms of specific algorithm or measurement system deficiencies. Estimates tend to cluster based on input data type and hence similar temporal and spatial sampling [Ebert et al., 1996]. Inclusion of all available operational data sets into blended or merged products limits opportunities for comparison to truly independent estimates.

The International Precipitation Working Group (IPWG) was established in 2001 as a permanent working group of the Coordination Group for Meteorological Satellites (CGMS) and is co-sponsored by CGMS and the WMO. The IPWG validation/intercomparison project provides standardized web-accessible intercomparisons over Australia, the continental US and Europe between daily precipitation estimates from rain gauge and radar networks and a variety of operational and semi-operational satellite algorithms and numerical weather prediction models [Ebert, 2004, Ebert et al., 2007]. Example daily statistics aggregated over two 3 month periods between the 0.25° operational rain gauge analysis for Australia [Weymouth et al., 1999] and a GPI algorithm, a merged multi-satellite algorithm, and the ECMWF forecast are shown in Table 5-3. Performance of both satellite algorithms is better during the December-February period corresponding to the wet season in tropical northern Australia as compared to the dryer June-August period. No one method performs best across all statistics for both the dry and wet seasons.
Table 5-3. Aggregate daily statistics from IPWG intercomparison from 1 Dec 2003 – 29 Feb 2004 and 1 June 2004-31 August 2004. Statistics shown were compiled over land areas based on estimates of the daily accumulation from a GPI algorithm, a NASA TRMM 1 hourly multi-satellite product (3B41RT, Huffman et al. [2003]), and the ECMWF 36 hour forecast for 24 hr precipitation accumulation compared to the Australian national rain gauge network. Spatial scale of intercomparison is 0.25° for all products. Data courtesy of E. Ebert.

<table>
<thead>
<tr>
<th></th>
<th>GPI</th>
<th>TRMM</th>
<th>ECMWF</th>
<th>GPI</th>
<th>TRMM</th>
<th>ECMWF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dec – Feb</td>
<td>Dec – Feb</td>
<td>Dec – Feb</td>
<td>June – Aug</td>
<td>June-August</td>
<td>June-August</td>
</tr>
<tr>
<td>Ratio of Estimated to Observed Average Volume</td>
<td>1.18</td>
<td>1.08</td>
<td>0.74</td>
<td>0.38</td>
<td>0.55</td>
<td>0.75</td>
</tr>
<tr>
<td>Bias Score</td>
<td>0.97</td>
<td>0.69</td>
<td>0.91</td>
<td>0.22</td>
<td>0.31</td>
<td>0.80</td>
</tr>
<tr>
<td>POD</td>
<td>0.71</td>
<td>0.56</td>
<td>0.73</td>
<td>0.13</td>
<td>0.16</td>
<td>0.69</td>
</tr>
<tr>
<td>FAR</td>
<td>0.26</td>
<td>0.19</td>
<td>0.19</td>
<td>0.41</td>
<td>0.49</td>
<td>0.13</td>
</tr>
<tr>
<td>Mean Absolute Error</td>
<td>4.0</td>
<td>3.9</td>
<td>2.8</td>
<td>0.66</td>
<td>0.89</td>
<td>0.39</td>
</tr>
<tr>
<td>(mm/d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMS Error (mm/d)</td>
<td>9.0</td>
<td>10.7</td>
<td>7.5</td>
<td>2.6</td>
<td>4.0</td>
<td>1.4</td>
</tr>
</tbody>
</table>

5.5.4 Testability of the hypothesis that aerosols impact precipitation

Uncertainties in precipitation measurement vary with the spatial scale, temporal scale, and storm type that are considered. For areal precipitation estimates derived from all types of sensors, storm total rainfall estimates from precipitating clouds covering larger areas with low spatial variability, of longer duration, and with moderate to heavy precipitation intensities generally have smaller uncertainties than estimates from short-lived, isolated showers or light precipitation near the detection threshold of the measurement instrumentation. The few quantitative estimates of relative rainfall estimation error in the literature focus either on large areas and long time periods, such as a month or more, or on the subset of storms with moderate to heavy precipitation. These studies indicate that many of the independent sources of error have substantial biases and scatter. Representativeness error is estimated at 25-95% of hourly conditional mean rainfall and up to 50% for hourly areal average rainfall [Tustison et al., 2001]. The daily radar/rain gauge ratio for storms with
at least 0.3 mm daily precipitation is only within a factor of 2 for 68% of summer days in Switzerland but these large daily uncertainties average to a summer seasonal bias of -11% [Germann et al., 2006]. Comparison of radar derived rainfall over gauges (1 km$^2$) for storm totals with rainfall accumulations > 10 mm had root mean square errors from 10-40% when using the subset of gauges passing quality control, and up to 80% when all gauges in the network were used [Steiner et al., 1999]. Errors for storms with lighter precipitation were not quantified but are expected to be higher. Houze et al. [2004] estimated $\pm$50% uncertainties in radar-derived monthly areal average rainrates over $7 \times 10^4$ km$^2$ tropical oceanic regions. Satellite temporal sampling errors, which are independent of the precipitation retrieval method, increase rapidly as temporal sampling decreases and range from 10% for 1 hourly sampling to 154% for 12 hourly sampling of daily areal average rainfall over a 100 km scale domain [Steiner et al., 2003]. These values are from studies that have better control of error sources, such as calibration and instrument maintenance, than is typical worldwide and hence represent lower limits for rainfall measurement uncertainty. Uncertainties in snowfall measurement are larger than those for rain.

For long time scales, the impact of aerosols on precipitation is difficult to isolate from other sources of precipitation variability such as the location of storm tracks, variations in surface moisture and heat fluxes, urban boundary layer dynamics-thermodynamics, and interannual cycles. Analysis of time scales of storm duration or shorter and spatial scales sufficiently small to distinguish regions under clearly observed aerosol plumes ($\leq 50$ km scale) would help to separate the impact of aerosols from other factors. At these short time and spatial scales, errors in precipitation measurement will typically be a factor of 2 or more. Given the large magnitude of these uncertainties, it will be infeasible to discern the impact of aerosols that modify rainfall by less than a factor of 2.

5.6 Conclusions

Even though we clearly see in measurements and in simulations the strong effect that aerosol particles have in cloud microphysics and development, we are not sure what is the magnitude or direction of the aerosol impact on precipitation and how it varies with meteorological conditions. The reasons behind this difficulty are the result of:

- The limited accuracy of the measurements:
Quantitative radar measurements of precipitation can be off by up to a factor of 2 for daily rainfall and spatial scales \( \leq 50 \) km.

Measurements of the cloud droplet size from satellites are biased by up to 30% for convective clouds, due to departure of the clouds from the plane parallel model. However, we can still measure the relative change of cloud properties with change in aerosol concentrations.

Aerosol column optical depth measured from satellites is not accurate enough: accuracy is \( \pm 20\% \) over the oceans, less accurate over vegetated land, deserts or urban areas. Ground based remote sensing of aerosol can have a high precision of \( \pm 0.02 \) in units of aerosol optical depth, but with limited spatial coverage.

Translation of the optical thickness into the column aerosol concentration and into CCN is difficult due to variation in the aerosol size distribution and chemical composition, the vertical distribution of the aerosols and thermodynamic conditions, and the lack of satellite passive remote sensing sensitivity to most CCN-sized particles.

In situ aerosol measurements from aircraft provide more physical and especially chemical detail. However, the aircraft may influence the aerosol and cloud properties in some cases, and interpretation of the measurements must take account of limitations in the sampling of all aerosol and cloud droplets sizes. Surface based measurements offer highly accurate column aerosol optical depth and physical as well as chemical information about boundary layer particle properties, but the property measurements may not represent the total column, especially at cloud altitudes needed to study aerosol-cloud interactions.

- **Small spatial and temporal coverage for highly variable parameters:**
  
  High spatial and temporal variability of precipitation and aerosols make it difficult to measure coincidental aerosol, clouds and precipitation. Therefore, we are not aware of studies that include direct statistical correlations of aerosol and precipitation.

  Cloud properties have a diurnal cycle not well represented by satellite observations. Geostationary satellite measure the diurnal cycle but with limited ability to derive the cloud droplet size and aerosol amount and properties.

- **Cause and effect:**
  
  A considerable number of studies have revealed a correlation between higher aerosol concentrations and higher cloud droplet concentrations in stratocumulus clouds. However, the impact of aerosol-induced changes in droplet concentrations upon drizzle rates is not well understood as different studies have yielded different results [Wood 2005]. A confounding
factor in determining causation is that once precipitation forms, aerosols are removed. It is still difficult to assess whether the changes in the precipitation patterns are primarily a result of the aerosol impact or whether it is changing of the local meteorological conditions including boundary layer properties that are the primary influences on changes in precipitation patterns near industrialized regions. Recent work [Ackerman et al. 2004, Wood, 2007] shows that precipitation from stratocumulus clouds is strongly correlated with cloud thickness and that aerosol-induced changes in the cloud droplet spectra may lead to thicker or thinner clouds depending on the large scale meteorological conditions. These feedbacks on precipitation, both positive and negative, further complicate the unraveling of the relation between aerosols and precipitation. Even the most informative measurements so far on the effect of aerosols on precipitation do not include simultaneous quantitative measurements of aerosols, cloud properties, precipitation and the full set of meteorological parameters.

Field campaigns involving coordinated surface, aircraft, and spacecraft measurements, along with aerosol transport and cloud models, could aim at better understanding the physical mechanisms that must be represented in models, and at identifying key indicators in satellite observations that could be used to extrapolate the more detailed field measurements to regional and global scales.
6.1 Introduction

The impact of aerosols on cloud macrophysical (cloud extent, cloud thickness and liquid water path), cloud microphysical (droplet and ice crystals concentrations and size distributions), and precipitation has received a great deal of attention for over 50 years. The pioneering work of Gunn and Phillips [1957], Squires [1958], Squires and Twomey [1961], Warner [1968] and Warner and Twomey [1967], to mention just a few, pointed out that high concentrations of cloud condensation nuclei (CCN) from anthropogenic sources, such as from industrial pollution and from burning of sugarcane, can increase cloud droplet number concentration (CDNC), hence increasing cloud microphysical stability and potentially reducing precipitation efficiency. Although the connection between increased CCN and increased CDNC has been supported by many in situ observations, the impact of CCN on precipitation is less well established.

One of the obstacles in assessing such a cause and effect relationship is the fact that different aerosol types generally correspond to different air masses, hence to different vertical profiles of moisture and stability. The sensitivity of cloud and precipitation to these meteorological parameters is particularly high in the boundary layer. In convective clouds the ability to produce precipitation embryos depends to a large extend on the maximum LWC, hence on the cloud depth, because the adiabatic liquid water content increases with altitude above cloud base (a rate of about 2 gm$^{-3}$ of condensed water per km of ascent at 20 C). In subtropical boundary layer clouds, the altitude of the cloud base, hence the cloud depth and the maximum LWC slightly below cloud top, depend on the total water mixing ratio of the mixed layer. For example, in a boundary layer with a total water mixing ratio of 20 gm$^{-3}$, topped by a 200 m thick cloud, a very small drying of the boundary layer (1% decrease of the total water mixing ratio, at constant temperature), leads to a decrease of the cloud layer thickness by a factor of 2, which corresponds to a factor of 4 reduction of the liquid water path (LWP). Since the precipitation rate in boundary layer clouds roughly scales with the LWP and the inverse of CDNC [Pawloswka et al., 2003, Comstock et al., 2004, Van Zanten et al., 2005], a 1% drying of the boundary layer might have the same impact on precipitation
as a factor of four increase in CDNC due to pollution. The same argument can be made with a slight warming of the boundary layer, since a 0.6°C increase of the temperature, at constant humidity, leads to a reduction of the cloud depth by about 100 m.

Deep clouds which often produce more rain than boundary clouds, are less sensitive to small variations in the boundary layer moisture and temperature, but are highly sensitive to the stability profile and to the low level convergence of sensible and latent heat.

The following discussion illustrates how observational studies try to progressively overcome the obstacles to understanding by addressing each element of the long chain of physical processes connecting aerosols to precipitation.

### 6.2 Warm Clouds

#### 6.2.1 Pollution aerosol impacts on CCN properties and CDNC

There are numerous observational evidences that the CCN concentration is increased in polluted air masses and that their hygroscopic properties are modified.

Considerable effort on aerosol sources was expended in Australia in the 70s and 80s that paints a very clear picture on anthropogenic influence. Five years of automated CCN measurements at a rural site in New South Wales showed convincingly that cities and towns were copious sources of CCN, with levels of CCN qualitatively in proportion to size of town (Figure 6-1) [Twomey et al., 1978].
Subsequently airborne methods for determining CN fluxes from individual anthropogenic sources were developed and applied to a very large, isolated, metal smelting complex at Mt Isa, then to towns and cities around Australia, demonstrating existence of well defined "pollution tracks" of aerosol CN and inferred CCN extending up to 1800 km from the source [Bigg and Turvey, 1978; Bigg et al., 1978; Ayers et al., 1979; Ayers et al., 1982; Carras and Williams, 1981; Williams et al., 1981; Carras et al., 1988]. The extremely well defined scaling of town CN flux with population was shown theoretically to have a sound physical basis [Manton and Ayers, 1982] (Figure 6-2).
Radke and Hobbs [1976] measured CCN (S=0.2%) concentrations of 1000–3500 cm$^{-3}$ along the East Coast of the USA from Virginia to Long Island, compared to 100 cm$^{-3}$ of CCN in clean maritime air and 300 cm$^{-3}$ in continental air. They concluded that the global anthropogenic production rate of CCN might be comparable to the natural production rate.

Mineral dust particles are known to be efficient ice nuclei [Pruppacher and Klett, 1997] but not very efficient CCN. However, recent in situ measurements of dust from the Sahara over the Mediterranean [Levin et al., 1996], and of Asian dust over China and Japan during ACE-Asia [Trochkine et al., 2003], show that dust particles can get covered by sulfate from pollution during their transport through the atmosphere, which may increase their effectiveness as CCN. More recently, Levin et al. [2005] showed that low-lying dust storms over the sea could lead to sea-salt coating on a large fraction of the dust particles. These types of coating could modify the particles into effective giant CCN (GCCN) as well as efficient ice nuclei. Indeed, sampling in clouds affected by particles coated with sulfate revealed the presence of relatively large drops close to cloud base (Figure 6-3).
At this point one may ask whether CDNC increases proportionally to the aerosol concentration. From collocated airborne measurements of condensation nuclei (CN), CCN and droplet concentrations, Leaitch et al. [1992], Martin et al. [1994], Taylor and McHaffie [1994], Gultepe et al. [1996], Chuang et al. [2000], Snider and Brenguier [2000], Ramanathan et al. [2001a] among others, have shown that regardless of location, increases in aerosol concentration lead to CDNC increases. However, these trends appear to taper off when the aerosol concentration increases, as shown in the compilation of these diverse results [Ramanathan et al., 2001b]. The origin of the large discrepancies in this compilation is two fold: First, CCN are only a fraction of the CN population and that fraction varies significantly with the size distribution and chemical composition of the aerosol. Second, the fraction of available CCN that are actually activated in a cloud increases with the intensity of the updraft at cloud base (see Chapter 2). Stratocumulus clouds, with updraft speed of less than 1 ms$^{-1}$ nucleate a few of the available CCN particles (Atlantic Stratocumulus Transition Experiment (ASTEX) and the Aerosol Characterization Experiment-2 (ACE-2) case studies in Figure 6-4), while deeper convective clouds can develop updraft intensities of more than 10 ms$^{-1}$, hence activating most of the available CCN particles (North Atlantic case study in Figure 6-4). The ratio of CDNC to the total aerosol concentration thus varies from 70 % for North Atlantic convective systems to about 10 % in ASTEX and ACE-2 sub-tropical stratocumuli.
Note also that the total aerosol concentration varies significantly depending on which instrument is used to measure the concentration (sizes bigger than 2.5, 10, 20 nm) because, during nucleation events, the relative concentration of nanometric particles can be very high, while such particles can hardly be activated in clouds.

![Figure 6-4: Aircraft data illustrating the CDNC increase with the increase in aerosol number concentration. The thick red line is obtained from a composite theoretical parameterization that fits the INDOEX aircraft data for the Arabian Sea [Ramanathan et al., 2001b].](image)

Observational studies of the CCN activation process therefore require a comprehensive characterization of the particle size distribution, chemical composition, mixing state and surface coating (see Chapters 3 and 4 for more details), all contributing to the activation properties of the observed aerosols [McFiggans et al., 2006]. In actual clouds however, CCN activation is not uniform because the vertical velocity that governs the kinematics of the activation process varies in space and time, hence does the resulting CDNC. Moreover, the droplet number concentration after activation can be significantly reduced by additional processes that are not directly linked to aerosol properties, such as entrainment-mixing, coalescence and precipitation scavenging.

A closure study on the CCN activation process has been performed during the ACE-2 experiment. Aerosol particle size distributions and chemical compositions were measured at the ground and on board research aircraft flying in the sub-cloud layer. The probability density function (PDF) of vertical velocity was also measured by an aircraft flying at the
cloud base [Guibert et al., 2003]. With these input parameters, a PDF of predicted CDNC was derived from the measured PDF of vertical velocity using different versions of an activation model. The actual values of droplet number concentration were measured in the cloud layer after selection of the samples that were not affected by mixing or drizzle scavenging [Pawlowska and Brenguier, 2000]. The conditionally sampled and the predicted PDFs of CDNC are compared in Figure 6-5, for 5 of the ACE-2 cases [Snider et al., 2003]. Some predictions are in good agreement with the observations, but the closure experiment is inconclusive because information were lacking on the size resolved chemical composition and mixing state to better constrain the activation model.
Figure 6-5: CDNC values predicted by a CCN activation model for 5 ACE2 cases (colors correspond to different initialization modes). Error bars indicate the width (1 standard deviation) of the predicted CDNC probability distribution function corresponding to the measured frequency distribution of the vertical velocity at cloud base. Horizontal lines show the average and the average plus or minus 1 standard deviation of the conditionally sampled CDNC probability density function. [Snider et al., 2003].
6.2.2 Aerosol Impact on Cloud Microphysics

6.2.2.1 Deep Convective Clouds

Warner and Twomey [1967] and Warner [1968] reported that, through their activity as CCN, the release of smoke particles results in large CDNC increases, and consequent reduction in the sizes of cloud droplets (Figure 6-6), which may impede the coalescence process of rain formation.

![Figure 6-6](image-url)

**Figure 6-6:** Droplet concentration versus liquid water content of clouds affected by smoke from sugarcane fires in Australia. Note the CDNC increase from 140 to 510 cm\(^{-3}\) (red) and the slight decrease in liquid water content from 0.14 to 0.12 g m\(^{-3}\), in the clouds affected by smoke compared to those affected by clean on-shore airflow (black). Adapted with modifications from Warner and Twomey [1967].

Recently, Andreae et al. [2004] reported on results from airborne and ground observations of cloud and precipitation development in the Amazon region. They divided their observations into four types of clouds. 1) Maritime shallow clouds over the ocean, which had low CCN concentrations and few cloud drops, including a few large ones. These clouds did not grow to
be tall but developed precipitation very effectively, as can be seen by the widening of the drop size distribution with height. These were termed “Blue Ocean”. 2) The name “Green Ocean” was given to clouds that developed inland in the western Amazon in unpolluted conditions, mainly in the rainy season. These conditions were characterized by low CCN concentrations, due to washout by rain, and therefore resembled maritime clouds. 3) “Smoky Clouds” were those that were affected by high concentrations of CCN from smoke emitted from older fires. These particles remained in the atmosphere for a long time due to the lack of rain in the dry season. The authors reported that the high CCN concentrations produced high concentrations of cloud droplets, hence smaller sizes and narrower size distributions, thus reducing the efficiency for growth by collection. Therefore, these drops continued to grow by condensation and were surmised to reach higher altitudes and lower temperatures where ice could form. Since these clouds were deep, they could produce lightning, hail and heavy rain. 4) “Pyro-Clouds” were clouds fed directly with smoke and heat from biomass fires. The high concentrations of CCN in the smoke produced large concentrations of small drops, thus limiting the warm rain production (see also Fromm et al., 2006). Although, some of the large ash particles in the smoke could produce large drops and enhance warm rain development, no raindrops were actually measured by the rain probe in the growing towers of these clouds. Figure 6-7, that summarizes the drop size distributions in the four types of Amazonian clouds listed above, illustrates how that the evolution of the droplet spectrum with altitude varies significantly depending on the type of aerosol.

6.2.2.2 Boundary Layer Clouds

Garrett and Hobbs [1995] studied the effects of long-range transport of particles from polluted and unpolluted sources on the characteristics of shallow, non-raining stratus clouds around the Azores. The aerosols originating from continental sources (Europe) had a monomodal particle size distribution at about 0.05 µm, while the clean maritime air contained a bimodal particle size distribution with modes at 0.02 and 0.08 µm. The clouds affected by the polluted air had 160% higher concentrations of droplets, and a 27% smaller effective radius, than the clouds affected by the clean air (Figure 6-8).
Figure 6-7: The mass drop size distributions at different altitudes in a number of typical Amazonian clouds. Note the similarity and the wide drop size spectra in both the “blue” and “green ocean” clouds, compared to the narrow drop size spectra in the “smoky” clouds. Also note the large aerosol sizes in the ash and the resultant second peak in the drop size distribution in the “pyro-clouds.” [Andreae et al., 2004].

Figure 6-8: Cloud droplet size distribution in clean (case 1) and polluted (case 2) warm clouds over the Azores Islands. Note the higher concentrations and the smaller sizes of the drops in the polluted clouds. Adapted from Garrett and Hobbs [1995].
Relying on Cape Grim observations of the CCN seasonal variability (see Sec. 6.2.1.1), the Southern Ocean Cloud Experiment (SOCEX) was designed to investigate whether these coherent seasonal cycles extend to the microphysical characteristics of the maritime cloudy boundary layer. The venue of the experiment was the atmosphere just west of Cape Grim, which is located at the north-western tip of Tasmania, about 300 km south of the Australian continent. When the wind direction is from the southwest (Baseline sector, between 190° and 280°) the air stream has a typical fetch of 4000 to 5000 km over the Southern Ocean and is virtually devoid of anthropogenic influences.

Two experiments were carried out: SOCEX I, in July 1993 at the minimum of the seasonal cycles in DMS flux and CCN concentration; and SOCEX II in February 1995 at the maximum of the seasonal cycle. The results [Boers et al., 1996; Boers et al., 1998] from the experiments are summarized in Figure 6-9, that shows the average droplet concentration and the average effective radius as a function of pressure difference with respect to cloud base, for the summer experiment (circles) and winter experiment (pluses). Only droplet properties (FSSP data) are shown, so the results exclude precipitation-size droplets. The seasonal differences are large and should be judged as a firm indication of the seasonality in the microphysical structure of clouds over the Southern Ocean. Near cloud base there is almost a factor of three differences in droplet concentration, which ranges from 32 cm\(^{-3}\) in the winter to 94 cm\(^{-3}\) in the summer.

These results are almost identical to the predictions by Boers et al. [1994] who found, using a parcel model of adiabatic ascent, that the cloud base droplet concentration in unpolluted conditions over the Southern Ocean would range from 35 cm\(^{-3}\) in winter to 92 cm\(^{-3}\) in summer.
Figure 6-9: In situ measured vertical profiles of droplet effective radius (left) and CDNC (right), for the SOCEX summer experiment (circles) and winter experiment (pluses). The ordinate is the pressure difference with respect to cloud base.

The ACE-2 CLOUDY-COLUMN experiment in 1997 provided a very comprehensive data set on the impact of anthropogenic aerosol on the microphysics of stratocumulus [Brenguier et al., 2000a]. Figure 6-10 compares vertical profiles of CDNC, mean volume droplet radius and liquid water mixing ratio for the most pristine case (26 June, CDNC = 51 cm$^{-3}$), and the most polluted one (9 July, CDNC = 256 cm$^{-3}$). Frequency distributions of these parameters are based on more than 400 km sampling along a 60 km square track, and stratified over 30 m thick layers from cloud base to cloud top. The liquid water mixing ratio increases in average with altitude, as predicted by adiabatic convective models (Chapters 2 and 7). CDNC reaches its maximum value slightly above cloud base because in the first 30 to 80 meters, all CCN are not fully activated and some droplets are too small to be detected by the airborne droplet spectrometer. Close to cloud top CDNC is substantially reduced by cloud-top entrainment-mixing. CDNC values that characterizes CCN activation are derived by selecting samples with quasi-adiabatic LWC (>0.9 LWC$_{ad}$), and no precipitating particles (the white vertical lines in the top row). The figure shows that the mean volume radius (second row) increases as predicted by the condensational growth theory using the CDNC value $N_{\text{act}}$ as defined above. It also shows that this diagnostic of the mean volume radius is still valid at cloud top despite the significant reduction of LWC following mixing with the
overlying dry air. This feature is commonly referred to in the literature as inhomogeneous mixing (see Chapter 2) [Latham and Reed, 1977; Baker and Latham, 1979; Baker et al., 1980, Burnet and Brenguier, 2007).

Figure 6-10: Frequency distributions of N, r, and q in five sub-layers, represented by their 5% percentiles, for the 26 June and 9 July ACE-2 cases. Statistics limited to cloud samples (N>20 cm$^{-3}$). The mean value in each sub-layer is indicated by a white circle. The vertical line in the top row corresponds to the characteristic CDNC value N$_{act}$ (51 cm$^{-3}$ for 26 June, and 256 cm$^{-3}$ for 9 July). In the next two rows the lines indicate the adiabatic droplet size prediction, with N=N$_{act}$, in the r$_v$ distribution, and the adiabatic LWC in the q$_l$ distribution. From Brenguier et al. [2003a].

6.2.3 Aerosol Impact on Cloud Radiative Properties

By modifying the microphysics of clouds, aerosols also affect their radiative properties hence allowing detection of their impact from satellite remote sensing measurements of the reflected solar and infrared emitted radiances. Most of the observational studies of the aerosol impact on precipitation in fact rely on the interpretation of cloud radiative property changes that must therefore be precisely evaluated. In the following section and more
comprehensively in Chapter 5, the issues of interpretation of cloud radiative properties and the changes due to aerosols are discussed. The limitation of using remote sensing methods for evaluation of precipitation amounts is also outlined.

6.2.3.1 Ship Tracks

Ship tracks are narrow lines of perturbed regions in marine stratiform clouds caused by emissions from ships. Ship tracks appear brighter in satellite imagery, particularly at a wavelength of 3.7 \( \mu \text{m} \) [Coakley et al., 1987]. Ship tracks (Figure 6-11) are often as long as 300 km or more and ~9 km wide [Durkee et al., 2000, 2001]. They typically form in relatively shallow boundary layers between 300-750 m deep. They do not form in boundary layers thicker than 800 m [Durkee et al., 2000]. Ship tracks are of considerable interest because they provide the most direct evidence that shallow, clean clouds can be modified by pollution sources that emit large concentrations of CCN.

Figure 6-11: Ship tracks off France. From bands 1, 4 and 3 of MODIS on the Aqua Satellite. (Courtesy of NASA.)

Hobbs and Garrett [2000] reported that various types of ships, burning both low-grade marine fuel oil (MFO) and U.S. Navy distillate fuel (NDF), emitted from \( \sim 4 \times 10^{15} \) to \( 2 \times 10^{16} \) total particles per kg of fuel burned (\( \sim 4 \times 10^{15} \) – \( 1.5 \times 10^{16} \) particles per second). Diesel-powered ships burning MFO emitted particles with a larger mode radius (\( \sim 0.03 \) – 0.05 \( \mu \text{m} \)) and larger maximum sizes than those powered by steam turbines using U.S. NDF. Consequently, for particles with similar chemical composition, those emitted by diesel ships burning MFO will serve as CCN at lower supersaturations, and will therefore be more likely
to produce ship tracks, than the particles emitted by ships powered by steam turbines using U.S. NDF.

The prevailing hypothesis is thus that the ship's trails appear brighter on satellite imagery because the effluent from the ships is rich in CCN particles. The more numerous CCN particles create larger concentrations of cloud droplets, which then reflect more solar energy than the surrounding clouds [King et al., 1993]. Aircraft observations in ship tracks and surrounding clouds [Radke et al., 1989, Ferek et al., 1998] confirm that ship tracks exhibit higher droplet concentrations, smaller droplet sizes, and higher liquid water content than surrounding clouds. In some cases, ship tracks appear to be thicker than the ambient cloud [Ackerman et al., 2000]. Radke et al. [1989] found that the concentration of drizzle drops (droplets of diameter $\geq 200$ µm in ship tracks was only $\sim 10\%$ of that in surrounding clouds. This supports the laboratory results of Gunn and Phillips [1957] and Albrecht's [1989] hypothesis that a CDNC increase should reduce drizzle production. Another hypothesis is that the ship track clouds exhibit higher liquid water content because heat and moisture emissions from ships create deeper and wetter (hence brighter) clouds [Porch et al., 1990]. This hypothesis is supported by the observation of ship tracks in an apparently cloud-free sky [Conover, 1966]. Porch et al. examined this hypothesis and showed that ship tracks are not only characterized by greater brightness but also by clear bands on the edges of the cloud tracks. They speculated, and provided some modeling evidence, that the heat and moisture fluxes from ship effluents excite a dynamic mode of instability that, in some marine stratocumulus, leads to enhanced upward and downward motions associated with the cloud circulations. However, measurements during the Monterey Area Ship Track experiment (MAST) contradicted the hypothesis by observing that heat and moisture emissions from ships are much too small to have any impact on clouds [Hobbs et al., 2000]. An alternate explanation for the formation of deeper and wetter clouds in ship tracks, and perhaps clearing to the sides of ship tracks, is that when drizzle is suppressed, cloud top radiative cooling is enhanced. The enhanced cloud top cooling destabilizes the cloud layer, which results in strong ascending motions in the clouds that transports more moisture aloft making the clouds wetter and deeper. In addition, in compensation for the enhanced upward motions, sinking motions surrounding the regions of enhanced ascent could cause clearing just outside of ship tracks. An extensive series of papers on ship tracks can be found in a special issue of the Journal of the Atmospheric Sciences (volume 57, 15 August 2000) devoted to the MAST.
Ship tracks nicely illustrate the complexity of the aerosol effects on clouds by showing that impacts on precipitation formation can hardly be examined independently of concurrent dynamical and radiative feedback processes. Note however, that ship tracks are an extreme realization of the aerosol impact on clouds because of the amount of CCN released and their narrow spatial extent that may facilitate the development of very specific dynamical modes that are not representative of extended boundary layer clouds. Their narrow extent also facilitates detection through the radiance contrast with the surrounding cloud.

6.2.3.2 Extended Cloud Systems
During ACE2 CLOUDYCOLUMN, the aerosol properties and cloud microphysical properties discussed above in Sec. 6.2.1.5 and 6.2.2.2 were measured by two instrumented aircraft flying below and inside the stratocumulus, respectively. Cloud radiative properties were measured simultaneously by a third aircraft equipped with a nadir pointing multispectral radiometer, flying above the cloud layer, and following the same track as the in situ aircraft. Figure 6-12 shows the vertical profiles of the mean volume droplet radius measured in situ and the conditional frequency distributions of the remotely sensed visible (VIS) versus Near IR (NIR) reflectances for the 26 June and 9 July ACE2 cases. The grids on the background reproduce radiative transfer simulations in adiabatic clouds for various values of CDNC and of the geometrical thickness [Schüller et al., 2003]. They corroborate the Twomey hypothesis that polluted clouds have a higher reflectance than pristine clouds at constant liquid water path and that the optical thickness scales as LWP$^{5/6}$N$^{-1/3}$ [Brenguier et al., 2000b].

While the impact of aerosols on cloud microphysics has been clearly established since the sixties, the ACE2 closure experiment, relying on collocated and independent measurements of aerosol properties, cloud microphysics and cloud radiative properties, provided compelling observational evidence of the link between these microphysical changes and radiation at the scale of a cloud system, i.e. at a much larger scale than the ship tracks. This quantitative assessment of the link between cloud microphysics and radiation is a crucial step because retrievals of cloud properties from ground or satellite remote sensing instruments that are discussed in the next sections and in more detail in Chapter 5 rely on similar measurements of cloud radiative properties.
6.2.4 Remote sensing of the effects of Aerosols on Cloud Properties

Field experiments are designed to address specific questions and test hypotheses in detail by focusing a comprehensive suite of in situ and remote sensing instruments on a selected atmospheric phenomenon. But they are limited in time and spatial coverage and only provide a few case studies. Long term observations of clouds in contrast rely on remote sensing that can be operated continuously from the ground or on board satellites. Clouds vary widely in space and time due to variation of atmospheric conditions. The aerosol effect can be only a small perturbation on the large natural variability. The wide spatial scale of satellite data can provide the large statistics necessary to separate the aerosol forcing on cloud properties from natural variability. With the advancement in remote sensing techniques, it is now possible to estimate the effects of the aerosol on clouds over much greater spatial and temporal extents. Of course, due to their larger field of view and coarser spatial resolution, satellite
measurements cannot resolve aerosol effects on as small scales as in situ measurement can (e.g., due to issues related to cloud contamination at the scale of individual clouds).

Because of the link between cloud microphysics and radiation, remote sensing techniques can be used to retrieve cloud microphysical properties and further establish statistical correlations with collocated aerosol properties. The response of remote sensing instruments depends on the technique: for active remote sensing instruments, LIDAR backscattering is sensitive to the second moment of the cloud particle size distribution while radar backscattering is rather proportional to the 6th moment. For passive remote sensing instruments, the reflected radiance in the visible range depends on the cloud optical thickness, while in the near-infrared range it is sensitive to the particle effective radius (the ratio of the third to the second moment) and microwave radiation emitted by warm clouds is proportional to the liquid water path. Consequently, the retrieval of cloud microphysical properties from satellite often involves a combination of passive and active instruments operating at different wavelengths (see Chapter 7).

6.2.4.1 Retrievals of Aerosol Properties

Long term records of aerosol concentrations were obtained from TOMS and AVHRR that showed a systematic increase in the aerosol concentration in the developing world; e.g. 10% a year increase was reported for India [Massie et al., 2004]. Global observations by TOMS and AVHRR show the impacts of volcanic eruptions and seasonal variability but no clear global trend [Torres et al., 2002; Geogdzhayev et al., 2005]. The AVHRR study also showed evidence of reduction in the aerosol emissions in the former Soviet Union. AVHRR, METEOSAT and TOMS observations were used to track emission of dust from Africa [Carlson and Prospero, 1971; Moulin, 1997; Israelevich et al., 2002] and correlate them with changes in the air circulation,

These satellite observations used mainly the aerosol optical thickness (AOT) as a measure of the aerosol column concentration. There is an issue with that approach when applied to aerosol interaction with clouds. Indeed, the CCN particles that interact with cloud microphysics represent a small and highly variable fraction of the total aerosols that determine AOT. Nakajima et al. [2001] used the product of AOT with a measure of the aerosol size – the Angstrom exponent to represent better the spatial variation of column CCN concentration. The usefulness of the aerosol index as a CCN proxy has recently been
evaluated using in-situ data by Kapustin et al. [2006], who pointed out that aerosol particle response to humidification may be a significant compounding factor.

Data from the new era satellites can be better linked to CCN and cloud processes, by reporting not only the AOT but also separating it into fine and coarse aerosols. MODIS does the separation using spectral properties [Tanré et al., 1997], MISR using the angular properties [Kahn, 2001] and POLDER using the strong polarization of the fine aerosols [Breon et al., 2002]. The ability to separate fine from coarse aerosol from space enables an estimate of the anthropogenic contribution to the global aerosol system. It is estimated that 20% of the aerosol AOT over the ocean is of anthropogenic origin [Kaufman et al., 2005a].

Advance analysis of a combination of POLDER data from the PARASOL mission, together with the wide spectral MODIS data shows the potential to retrieve not only higher accuracy optical thickness of the fine and coarse aerosols but also to estimate the size of the fine aerosol, needed for precise conversion into number concentration of particles and estimate of a CCN proxy [Gerard et al., 2005]. These methods are yet to be applied to study aerosol-cloud interaction.

6.2.4.2 Retrievals of Cloud Optical Thickness and Droplet Effective Radius

The Twomey and Cocks [1989] and Nakajima and King [1990] technique described in Chapter 5 has been extensively used to retrieve cloud optical thickness and droplet effective radius from satellite multispectral radiances and to examine the aerosol impact on clouds. Kaufman and Nakajima [1993] used AVHRR 1 km resolution data to study well-developed clouds in the Amazon region during the dry-burning season. An increase in the smoke optical thickness (measured at a wavelength of 0.55 μm) from an average background value of 0.2 to 1.0 in very hazy conditions decreased the effective cloud droplet radius from 15 to 8 μm. Though the absolute droplet size is difficult to measure from satellites, the decrease in cloud droplet size by 40% with five-fold increase in aerosol concentration (represented by the aerosol optical thickness measured in clear pixels) is a good indication of the overall effect of the aerosols on clouds. The increase in the smoke optical thickness and the corresponding decrease in the average drop size were surprisingly associated with a decrease in the cloud reflectance from 0.71 to 0.68. This may be due to absorption of sunlight by black carbon in the smoke, which reduces cloud brightness.
To detect if the aerosol affects cloud brightening due to increases in CCN in more susceptible (less reflective clouds), *Kaufman and Fraser* [1997] used a different cloud selection technique to highlight less developed clouds. The results showed that smoke increases cloud reflectance from 0.35 to 0.45, while reducing droplet size from 14 to 9 µm. However, *Kaufman and Fraser* found that the smoke effect varies systematically with the location of the measurements: in the moist tropical forest in the north of Brazil, the smoke had a maximum effect on clouds, while in the drier Cerrado in the south of Brazil the effects were minimal.

Further analysis of two years data over the Amazon region by *Feingold et al.* [2001] confirmed the latitudinal dependence of the effect of smoke on clouds, although they showed that it was unrelated to the humidity gradient. They also found a substantial difference between the two years they analyzed (1987 and 1995). Saturation of the response of cloud droplet size to aerosol occurred at an aerosol optical depth of 0.8 in 1987 and 0.4 in 1995.

*Han et al.* [1994] examined data from the International Satellite Cloud Climatology Project (ISCCP) and found “the expected systematic decreases of the droplet effective radius over land compared with over ocean and in the Northern Hemisphere compared with the Southern Hemisphere”, but not the corresponding expected cloud albedo increase. *Han et al.* [1998] extended the statistical approach to the seasonal and geographic variability of cloud albedo and droplet effective radius, in relation with the LWP variability. They concluded “that cloud albedo increases with decreasing droplet size for most clouds over continental areas and for all optically thicker clouds, but that cloud albedo decreases with decreasing droplet size for optically thinner clouds over most oceans and the tropical forest”. Observations using AVHRR satellite data over the oceans by *Nakajima et al.* [2001] however, do not show any systematic trend of liquid water path (derived from cloud optical depth and drop size) on column aerosol number concentration over the full range of column aerosol number concentrations.

*Wetzel and Stowe* [1999] used NOAA PATMOS data to evaluate the relationship between aerosol optical depth and cloud droplet size. POLDER satellite data were used by *Breon et al.* [2002] to derive aerosol concentrations and cloud droplet effective radii from 8 months of spaceborne measurements, and to explore the aerosol effect on cloud microphysics. It was found that the cloud droplet size decreased with increasing aerosol index (a measure of the
aerosol column number concentration), showing that the effect of aerosols on cloud microphysics occurs on a global scale. However, the extent of the drop-size response to changes in aerosol was much lower than expected from theory. Rosenfeld and Feingold [2003] suggested that this low response was caused by the POLDER ability to measure only extensive homogeneous layer clouds and not convectively generated clouds.

Similar cloud optical thickness and droplet effective radius retrievals were examined against sulfate burden simulated with a chemical transport and transformation model [Harsvardhan et al., 2002; Schwartz et al., 2002]. The analysis of two episodes of substantial influx of sulfate aerosol from industrial regions of Europe and North America to remote areas of the North Atlantic revealed “a decrease of the droplet effective radius concomitant with the increase in modeled sulfate burden”, while “cloud optical thickness and albedo exhibit little evident systematic trend over the episodes”. Han et al. [1998], Harsvardhan et al. [2002], and Schwartz et al. [2002] speculated that the variability of LWP was the most likely reason for this puzzling observation, but these studies were missing the independent LWP measurements required to corroborate their hypothesis.

These contradictory observations reflect the fact that cloud radiative properties are primarily determined by the cloud macrophysical properties, such as LWP. CDNC only modulates the optical thickness. Indeed, within the idealized framework of the Twomey hypothesis, the cloud optical thickness increases as $N^{1/3}LWP^{5/6}$. The two radiative parameters that are measured from satellite, optical thickness and droplet effective radius, are therefore not sufficient to distinguish between changes in the cloud macrostructure and possible impacts of aerosols on cloud microphysics and precipitation. To overcome the ambiguity due to the natural variability of LWP, two different approaches have been developed.

6.2.4.3 Retrievals of Liquid Water Path and Droplet Concentration

If the observed cloud field contains non-precipitating convective cells, optical thickness and droplet effective radius can be combined to derive information on CDNC and the cloud geometrical thickness or LWP. Brenguier et al. [2000b] developed the technique using a vertically stratified adiabatic cloud model to build up the neural network training data base necessary for the processing of the measured radiances (see Chapter 5). In this scheme, LWP is proportional to $H^2$ ($H$ is the cloud geometrical thickness). The technique was validated
[Schüller et al. 2003] against the ACE2 data set, where cloud geometrical thickness and CDNC were directly measured in situ. Figure 6-13 shows the comparison between retrieved and in situ measured values of these two parameters. The CDNC differences between pristine (blue and purple) and polluted (yellow orange and red) are precisely captured. The cloud geometrical thickness, and hence LWP are always overestimated. The origin of this bias, which is connected to the current overestimation of the droplet effective radius in most retrieval techniques, has not yet been identified. The technique has been extended to the retrieval of LWP (or cloud geometrical thickness) and CDNC from MODIS radiances (Schüller et al., 2005).

Recently Boers et al. [2005] have improved the technique by taking into account sub-adiabatic LWC values in the cloud layer and tested it against long term observations of the seasonal CCN variations recorded at Cape Grim. Over the four years, the retrieved CDNC closely follows the seasonal variations of the CCN concentration at the ground (Figure 6-14).

Using these techniques thus allows discriminating among the observed variations of the cloud albedo, the respective contributions of the LWP natural variability (which is related to cloud depth, especially under adiabatic conditions) and of CDNC, that reflects the impact of the aerosol.
Figure 6-14: Time series of MODIS-derived cloud optical depth and effective radius (first two panels), retrieved cloud droplet concentration (third panel), CCN concentration at Cape Grim (fourth panel) and retrieved cloud depth (fifth panel).

Note however that these techniques are only applicable to quasi adiabatic cloud layers, i.e. to non precipitating clouds that are not significantly affected by mixing with the environmental air. When mixing is prevalent, especially at the top of stratocumulus clouds, its impact on the droplet size distribution, hence on cloud radiative properties depend on the thermodynamic conditions at the mixing level. Either the droplet concentration is preserved while droplet sizes decrease (homogeneous mixing), or the concentration decreases while the sizes remain unchanged (heterogeneous mixing) [Latham and Reed, 1977; Baker and Latham, 1979; Baker et al., 1980, Burnet and Brenguier, 2007]. When mixing is of the inhomogeneous type
(very dry inversion), the retrieved droplet concentration may be significantly underestimated and polluted clouds can be interpreted as pristine [Chosson et al., 2007].

6.2.4.4 Retrievals of Droplet Effective Radius Vertical Profile

Rosenfeld and Lensky [1998] applied a method based on retrieval of the droplet effective radius at cloud top from satellite observations. The method assumes that all the clouds in the field of view of the satellite behave the same way, namely, that they all have the same base altitude and the same droplet growth history. This assumption was validated by in situ aircraft measurements [Freud et al., 2005] and by microphysical analyses of sequences of 3-minute rapid scans of METOSAT Second Generation satellite measurements [Lensky and Rosenfeld, 2006]. By looking at different cloud tops and their temperatures, they build profiles of effective radius and show that the effective radius increases as cloud depth increases. However, in continental clouds with higher droplet concentrations, the growth of the particles is slower than in maritime clouds. As a result, in continental clouds, drops commonly grow beyond an effective radius of 14 µm (assumed by Rosenfeld and Gutman, 1994 to represent the demarcation between clouds that may precipitate and those that do not) at much higher altitudes than in maritime clouds. In addition, the appearance of ice in the clouds appears as a sharp increase in effective radius with altitude. Rosenfeld and Lensky concluded that drops remain liquid to greater heights in continental clouds, leading to ice formation at much higher altitudes.

By analyzing the slope of the effective radius ($r_e$) versus temperature, Rosenfeld and Lensky [1998] defined five microphysical zones in convective clouds based on the temperature dependence of the $r_e$ and therefore suggest that the method can diagnose changes in clouds microphysical properties in response to different aerosol types.

6.2.4.5 Ground based remote sensing retrievals of aerosol and cloud properties

To overcome the limitations of satellite observations, surface-based remote sensing (LIDAR, radar, radiometer) data sets have been recently used to examine aerosol effects on clouds [Okamoto et al., 2002; Feingold et al., 2003a]. Although these methods can only be applied at a limited number of locations, they yield high temporal resolution data. They offer some interesting insights, and are complementary to the global satellite view. Feingold et al. [2003a] used data collected at the Atmospheric Radiation Measurement (ARM) site to allow simultaneous retrieval of aerosol and cloud properties, with the combination of a Doppler
cloud radar and a microwave radiometer to retrieve cloud drop effective radius $r_e$ profiles in non-precipitating (radar reflectivity $Z < -17$ dBZ), ice-free clouds. Simultaneously, sub-cloud aerosol extinction profiles were measured with a LIDAR to quantify the response of drop sizes to changes in aerosol properties. The microwave radiometer made it possible to sort the cloud data according to LWP, consistent with Twomey’s [1977] conceptual view of the aerosol impact on cloud microphysics. With high temporal/spatial resolution data (on the order of 20's or 100’s of meters), realizations of aerosol-cloud interactions at the large eddy scale were obtained. Moreover, by examining updrafts only (using the radar Doppler signal), the role of updraft in determining the response of $r_e$ to changes in aerosol (via changes in drop number concentration $N_d$) was examined. Analysis of data from 7 days showed that turbulence intensifies the aerosol impact on cloud microphysics.

In addition to radar/microwave radiometer retrievals of aerosol and cloud properties, surface based radiometers such as the Multi-filter Rotating Shadow-band Radiometer [MFRSR – Michalsky et al., 2001] have been used in combination with a microwave radiometer to measure an average value of $r_e$ during daylight when the solar elevation angle is sufficiently high [Min and Harrison, 1996]. Using this retrieval, Kim et al. [2003] performed analyses of the $r_e$ response to changes in aerosol at the same continental site, and instead of using extinction as a proxy for CCN, they used a surface measurement of the aerosol single scattering coefficient $\sigma_s$. Their analysis spanned much longer time periods and their data included a range of different aerosol conditions. A similar study was conducted by Garrett et al. [2004] at a location in the Arctic. The advantage of the MFRSR/microwave radiometer combination is that it derives $r_e$ from cloud optical depth and LWP and it is not sensitive to large drops as the radar is. Its drawback is that it can only be applied to clouds with extensive horizontal cover during daylight hours. It should be emphasized that none of these methods can directly address the effects of aerosols on precipitation.

### 6.2.5 Satellite Observations of Aerosol Impact on Precipitation Formation in Clouds

Since the 1960s, in-cloud measurements have clearly established that aerosols influence cloud microphysics. The overall feature is that increased CCN concentration leads to increased CDNC, although, as mentioned previously, the relationship also depends on vertical velocity at cloud base and on how CCN are distributed in size, among the Aitken, accumulation and the coarse modes.
Since that time, many other field studies have revealed that the increased CDNC and decreased droplet sizes have an impact on the formation of precipitation embryos by droplet collection. For example, Boers et al. [1998], analyzing the SOCEX data set, found that the fraction of LWC contained in drizzle drops sharply increases, when the droplet effective radius increases above 12 µm (Figure 6-15).

![Figure 6-15: Drizzle liquid water path fraction as a function of the maximum average effective radius observed in the case study profiles. Each data point is representative of individual stacks that were flown during the research flights. All winter and summer data have been included, except the 19 July 1993 (from Boers et al., [1998]).](image)

Similarly, Pawlowska and Brenguier [2003] showed (Figure 6-16) that, for the 8 case studies of the ACE-2 experiment, the averaged precipitation rate R, normalized by the cloud geometrical thickness H, scales with $H^3/N$. This feature, that may be useful for parameterizing precipitation in general circulation models, has further been corroborated with the data sets of the EPIC experiment [Comstock et al., 2005] and of the DYCOMS-II experiment [Van Zanten et al., 2005].
The effective radius profiling technique of Rosenfeld and Lensky [1998] described in Sec. 6.2.4.4 has been extensively used, in combination with TRMM remote sensing of precipitation, to examine the impact on precipitation of forest fires [Rosenfeld, 1999]; desert dust [Rosenfeld et al., 2001], sea spray [Rosenfeld et al., 2002], salt dust [Rudich et al., 2002], oil fires [Rudich et al., 2003], and the aerosol impact on the dynamics of convective clouds [Koren et al., 2005] and of storms [Fromm et al., 2006].

Overall, these satellite studies corroborate in situ measurements. Using TRMM [Rosenfeld, 1999; Rosenfeld et al., 2001, Rosenfeld et al., 2002], reported that precipitating particles were detected when the retrieved effective radius exceeded 14 µm threshold. For example, Rosenfeld [2000] using AVHRR images identified the signature of a pollution plume in a cloud layer (area 2 in Figure 6-17a). Using TRMM measurements he pointed out that this polluted area did not develop any precipitation unlike areas 1 and 3 in Fig 6-17a and b. Ayers [2005] argued against the conclusion of Rosenfeld [2000] that rainfall was suppressed in the case shown in Figure 6-17. Using radiosonde data he showed that the clouds depicted in the satellite retrievals were thin, located atop the boundary layer and were capped by a strong temperature inversion and low humidity. Therefore, they could not have produced much precipitation. This was borne out by 24h rainfall that was measured by the Australian Bureau of Meteorology, showing that rainfall was almost non-existent or minimal and isolated everywhere across the region. The debate over the interpretation of this study Ayers [2005], Rosenfeld et al., [2006a] underscore the need for further examination of this issue in a large number of cases with different methodologies.
Figure 6-17: (a) Pollution tracks (yellow) detected by AVHRR satellite imagery in clouds over South Australia. The higher reflectivity is due to reduced droplets sizes. (b) Radar shows precipitation as white patches outside pollution tracks, although these clouds have the same depth as adjacent non-polluted clouds, as shown in the vertical cross section. The gray clouds are the silhouettes obtained from the TRMM VIRS. The colors represent the radar echo intensities. No radar echoes occur within the clouds with reduced effective radius in Area 2. (c) Average effective radius of particles (probably water drops) at cloud tops in these regions (1, 2, and 3). The dashed lines represent the 15 and 85 percentiles of the distributions. From Rosenfeld [2000].

6.2.6 Correlations between aerosol and cloud macrophysical properties

In addition to its impact on droplet concentration, the aerosol also affects radiative transfer through clouds via absorption of solar light (semi-direct effect), which may accelerate cloud dissipation (cloud burning effect). Using satellite observations, Koren et al. [2004] reported
that smoke reduced daytime boundary layer cloud cover over the Amazon forest, from 38% in clean conditions to 0% in heavy smoke (optical thickness of 1.3). This response reverses the regional smoke instantaneous forcing of climate from \(-28 \text{ W m}^{-2}\), when only the direct effect is taken into consideration, to \(+8 \text{ W m}^{-2}\), when the dissipation in cloud cover is accounted for Figure 6-18.

\[\text{Figure 6-18: Reduction of cloud cover due to absorption of solar radiation by particles from biomass burning in Brazil. Photograph taken from the NASA Columbia shuttle STS-107 tragic flight on January 2003.}\]

MODIS aerosol data over the ocean [Tanre et al., 1997; Remer et al., 2002a], and cloud data from King et al. [2003] and Platnick et al. [2003] from the Terra satellite, have also been used for evaluating aerosol-cloud interactions for shallow, stratiform and trade-wind cumulus clouds and for deep convective clouds over the Atlantic Ocean. Kaufman et al. [2005b] show that each of four aerosol types covering the Atlantic Ocean during June through August (maritime aerosol from 30°S -20°S; smoke from 20°S-5°N; dust from 5°N-30°N; and, pollution aerosols from North America and Europe from 30°N-60°N) are correlated to a 20-40% cloud coverage variability of shallow stratiform and trade-wind cumulus clouds. Using co-variability analysis, they suggest that the aerosol particles might be responsible for ~40-100% of observed increases in cloud coverage. The increase in cloud cover attributed to
aerosol is consistent with Albrecht's [1989] hypothesis that enhanced CCN concentration suppresses precipitation and thus might increase cloud lifetime. This hypothesis has been criticized recently (Ackerman et al., 2004) by showing using large eddy simulations (LES) of boundary layer clouds, that increased CCN concentrations, hence increased CDNC, might reinforce entrainment at cloud top, leading to a decrease in the LWP and in cloud cover.

Analysis of convective clouds by Koren et al. [2005] show systematic effects of pollution, desert dust and biomass burning aerosols on the development and coverage of convective clouds over the Atlantic. As expected, cloud droplet sizes are decreased by ~20\% in high aerosol concentrations. This change has been accompanied by an increase in the coverage of convective and high clouds from ~30\% for low aerosol concentrations to ~60\% in hazy atmospheres. This sharp increase in cloud cover is associated with an increase in cloud top horizontal extent and altitude (the latter by as much as 1500 m), apparently through inhibition of early precipitation and the resultant strengthening of updrafts. These results are in agreement with the mechanism proposed in Williams at al. [2002] and Andreae et al. [2004] and simulated by Khain et al. [2005a] and Teller and Levin [2006] (this last reference reported only relatively small increase in cloud top, about 500 m, but a large increase in horizontal extent).

In most of the papers discussed thus far, the effects of aerosols on clouds have been confined to warm non- or slightly raining clouds, and in a few cases to convective clouds that may have contained ice. Overall, these studies are lacking information on small changes in the moisture content, enthalpy and available convective instability that could also explain the observed changes in cloud cover between different air masses. Such thermodynamical fluctuations are presently not detectable from satellites, nor are they predictable from re-analyses of the meteorological fields.

Satellite monitoring of aerosol and cloud properties thus provides evidence of correlations between aerosol optical thickness and cloud macrophysical properties (cloud fraction, cloud extend, liquid water path). However, progressing from statistical correlations to cause and effect relationships remain a challenge. Satellite measurements of cloud optical thickness, as a proxy of CCN number concentration, does not allow discriminating between aerosol layers that may interact with clouds and those that do not interact because they are located higher up. Saharan dust for example is often exported towards the North-East Atlantic Ocean in the
free troposphere, i.e. at an altitude much higher than the boundary layer where clouds develop. Such mineral dust is unlikely to interact with cloud microphysics but it may affect clouds in other ways. For example, a high dust layer may slightly decrease the incoming solar radiation, heat the atmospheric layer above the clouds and consequently increase the cloud cover of stratiform clouds [Johnson, 2004] as measured off the coast of Africa [Kaufman et al., 2005b]. Additionally, falling dust particles from above may nucleate ice particles near cloud top, thus modifying the development of precipitation via the ice phase. Quantifying the contribution of the meteorological circulation to cloud variability is also a serious obstacle because of the high sensitivity of the liquid water path to presently undetectable changes of the temperature and humidity in the boundary layer. Statistical analysis of the respective contributions of the aerosol and meteorology to cloud variability thus remains questionable because there are presently no robust meteorological predictors of shallow cloud cover.

The ACE-2 CLOUDYCOLUMN data base illustrates the paradox, with the 8 situations documented during the campaign. Figure 6-19 shows correlations between the cloud geometrical thickness \( H \), the optical thickness \( \tau \), the droplet mean volume radius \( r_v \) at cloud top, as a proxy for the effective radius, and the droplet number concentration \( N \), that characterize each case. Pristine cases are in blue, polluted cases in red, and intermediate cases in green. If \( H \) is constant, as in the Twomey hypothesis, or \( H/N \) correlations are restricted to the grey area corresponding in Figure 6-19-a to \( H \propto N \) and \( H \propto N^{-1/5} \), a negative correlation is expected between optical thickness and effective radius \( (\tau \propto r_v^{-1}, \text{when } H \text{ is constant}) \). In fact, because polluted cases are thinner than the other cases, the data set shows a \( H \propto N^{-1/2} \) correlation. In Figures 6-19-b, c, and d, the dotted line corresponds to correlations anticipated by Twomey (\( H \) constant) while the solid line illustrates the expected correlations when \( H \propto N^{-1/2} \). Figure 6-19-d shows a positive correlation between optical thickness and effective radius instead of the expected negative one.

This experiment thus demonstrates that the unexpected positive correlations observed from satellite between optical thickness and effective radius are not contradictory to the Twomey effect. They only reflect fortuitous correlations between pollution and geometrical thickness. In the ACE-2 specific case, back trajectories and meteorological analysis showed that the reduced thickness of the polluted cases was not due to their increased aerosol content, but
only to the fact that these air masses were transported over the European continent for a few days before entering the maritime area of the experiment, north of Tenerife. As a consequence, the liquid water path, hence the optical thickness, of the polluted cloud systems sampled during ACE-2 was lower than the ones of the pristine or intermediate cases. In this case, the aerosol is not the cause of the cloud LWP variability, both being determined, rather, by the mesoscale circulation.

Figure 6-19: Summary of the ACE-2 CLOUDYCOLUMN 8 case studies; a) cloud geometrical thickness $H$ as a function of droplet number concentration $N$; (b) droplet mean volume radius as a function of CDNC; (c) cloud optical thickness $\tau$ as a function of CDNC; and (d) optical thickness as a function of $r_v$.

6.3 Field Studies of The Effects Of Ice Nuclei On Clouds And Precipitation

Ice crystals may form in clouds whenever a cloud reaches altitudes where the temperature falls far below 0ºC. However, at temperatures above −40ºC, ice principally forms by what is called heterogeneous nucleation on particles that are effective as ice nuclei (IN). The role of the ice phase in the development of clouds and precipitation has been discussed in Chapter 2. Furthermore, Chapters 3 and 4 elaborate on the characteristics and the sources of IN. In spite
of decades of research in this field it has not been clarified how observed ice concentrations in clouds relate to ice nuclei in the air. This is in part because IN measurements are known to be imprecise and in part because of secondary processes of ice generation also discussed in Chapter 2,

\subsection{Observations of Ice Particle Concentrations}

Measurements show large variations in ice concentrations in different types of clouds, even at a given temperature. In some clouds, the ice concentrations increase monotonically with decrease in temperature, as would be expected from activation of ice nuclei \cite{Cooper1981, Blyth1993}. This parallel variation is a strong indication for the role of IN in determining ice crystal concentrations in the clouds observed (winter orographic and cold-based cumulus, respectively, for the two works cited), but it is still not a full proof for that to be the case. For that, direct observations of IN are needed for each cloud in which ice crystal concentrations are being measured.

\cite{Bower1996} conducted measurements of ice crystal concentrations in a number of cloud types, such as: frontal clouds over southern England and the sea areas around the British Isles, maritime convective clouds over the North Atlantic, and continental convective clouds over New Mexico and Montana in the USA. Ice concentrations were found to be several orders of magnitude higher than those that could be attributed to primary nucleation of ice nuclei at cloud-top temperatures. \cite{Bower1996} concluded, in agreement with the conclusions of \cite{Blyth1993}, that secondary ice multiplication processes, most likely the Hallett-Mossop process (see Chapter 2), must be operating in each of the cloud types examined in these studies.

Our understanding of the processes of ice formation in clouds become even less clear following the reports of \cite{Gultepe2001, Korolev2003} who showed that the average ice crystal concentration between about -10C and -30C in many stratiform clouds associated with frontal systems is almost constant (between about 2 to 5 cm$^{-3}$). The relatively constant concentrations with height that are also much higher than those predicted by the parameterizations of \cite{Fletcher1962} and \cite{Meyers1992}, suggest that
processes other than nucleation by IN must play an important role in ice crystal formation in these clouds. *Field et al.*, [2005] and *Gayet et al.*, [2002] measured similar ice crystal concentrations as a function of temperature in different geographical areas in frontal and cirrus clouds, respectively.

*Hobbs and Rangno* [1985, 1990, 1998] and *Rangno and Hobbs* [1988, 1991, 1994, 2005] discussed ice particle initiation and multiplication in maritime and continental cumulus clouds in a number of geographic locations. They suggested that the ice particles originate in two-stages. During stage 1, initial ice particles seem to originate through the freezing of larger cloud droplets in concentrations comparable to that estimated with the formula of *Meyers et al.*, [1992] derived from IN measurements. During stage 2, however, ice particle concentrations of 10's to 1000's per liter form in less than ten minutes at temperatures as high as –13ºC. Such high concentrations appear coincident with or soon after the formation of graupel particles. Supercooled raindrops or drizzle drops need not be present. In some cases, the time-scales for the formation of high ice particle concentrations are too short for the riming-splintering process to account for the observations.

It is recognized that ice crystal measurements in clouds suffer from the lack of adequate instrumentation for detecting small ice crystals (a few 10s of microns), suggesting that ice crystal concentrations in clouds might be higher than the reported values [*Korolev et al.*, [2003].

With regard to the vertical distribution of ice particles, *Blyth and Latham* [1993] reported that in the summertime convective clouds in New Mexico primary-ice nucleation was found to occur when the temperature within the cloud reached a value of between -10 and -12ºC irrespective of whether this was in the updraught or downdraught. Drops with diameters of about 0.5 mm were often observed in concentrations of about 10L⁻¹ before the formation of ice, which suggests a nucleation mechanism involving large drops. On the other hand, there are some reports from vigorous convective clouds showing that cloud drops sometimes remain unfrozen until they reach the homogeneous freezing levels [*Rosenfeld and Woodley*, 2000; *Rosenfeld et al.*, 2006b]. The authors argue that pollution in clouds would result in smaller cloud drops many of which would remain unfrozen until they are lifted to temperatures as low as -38ºC, the homogeneous freezing temperature. However, if the
pollution contains some particles that are effective IN, the formation of ice should start at higher temperatures.

The emission from steel mills in the Chicago area was used to at least partly explain the increase in rain and hail around La Porte Indiana, downwind of Chicago (in what was named the La Porte anomaly. See a review by Changnon, 1980). Furthermore, model results show [Yin et al, 2000a; Teller and Levin, 2006] that if among the pollution particles large and giant CCN are present, the large drops that are formed enhance ice formation at warmer temperatures and lower altitudes. The presence of such ice crystals in regions with high liquid water content would expedite the process of riming and the formation of graupel particles.

6.4 Aerosol Impact on Rainfall on the ground

As seen in the previous sections, aerosol impacts on cloud drop evolution are reasonably well understood while the impact of aerosols on ice crystal formation is not yet clear. This makes it difficult to assess the impact of pollution on precipitation, especially from clouds where ice processes contribute to precipitation formation. One also has to consider cloud and cloud system dynamics since they synergistically interact with microphysical changes.

Most measurement systems described in previous sections detect precipitation particles in clouds and do not provide direct measurements of precipitation rates and amounts at the surface. In this section, discussion is extended to field experiments in which rain gauge networks or remote sensing from satellites were used to assess the impact of aerosol on rainfall amounts.

6.4.1 The effect on precipitation from paper mill smoke, sugar cane and forest fires

Warner [1968] summarized the potential effects of sugarcane smoke on rainfall, by looking at multi-decadal rainfall records from stations upwind and downwind of these prolific anthropogenic aerosol sources. His first summary, which appeared in 1968, suggested the possibility of rainfall suppression by the biomass burning aerosols that affected cloud properties (downwind versus upwind properties) as shown by Warner and Twomey (1967). However, a more detailed analysis [Warner, 1971] led him to conclude that no such signal
could be found in the data. In other words, his analysis gives no support to the idea that any
association found between cane fires and rainfall at Bundaberg was due to inhibition of the
coalescence process by a reduction in average cloud droplet size. He stated, "It would be
surprising if the microphysics of a cloud played no part in determining its rainfall, but we
must await further results if this is to be adequately demonstrated".

Hobbs et al. [1970] reported on an increase of up to 30% in precipitation from warm clouds
downwind of paper mills in Washington State. Analyzing the same case through the use of a
one dimensional numerical model Hindman et al. [1977c] concluded that the emitted GCCN
from the paper mill could not by themselves account for the observed large increase in
rainfall and that the combined effects of heat, water vapor and CCN from the paper mill in
combination may be responsible for the increased precipitation.

It is valuable to note that Mather [1991] reported an increase in radar echo from clouds
affected by particles emitted from paper mills in South Africa. This observation led to a large
field experiment for rain enhancement using hygroscopic particles (see Chapter 8 for more
detail).

Through the use of MODIS and TRMM satellite data, Lin et al. [2006] analyzed the effects
of forest fires on precipitation in the dry season in the Amazon region. They report on
increases in cloud heights and in precipitation with increases in aerosol optical depth. The
increase cloud height led to enhanced growth of ice crystals, which culminated in heavier
precipitation. However, in spite of the good correlation between these variables, the authors
could not unequivocally establish causal links between aerosols and the observed changes in
cloud height or with precipitation increases. The role of enhanced convection due to the heat
from the fires and/or from heat due to absorption of solar radiation by the smoke itself could
not be ruled out.

6.4.2 Effects of Urban Pollution on Rainfall

Extensive studies were conducted to explain the anomalous behavior of the precipitation
around La Porte, downwind of Chicago. In that case, local records suggested an upward shift
in warm season rainfall, thunderstorms and hail from the late 1930s to about 1965. The
puzzling thing about this case is the fact that the anomaly appeared and then disappeared.
Changnon [1980] reviewed the observations and concluded that the microphysical effects must have played a role, but without the dynamical or meteorological pattern this effect would not have occurred.

A large field experiment was carried out around St Louis Missouri, motivated by the examination of historical data that revealed summer increases in the immediate downwind area of the city (Figure 6.20). The records show increases in: 1) rainfall (10–17%); 2) moderate rain days (11–23%); 3) heavy rainstorms (80%) 4) thunderstorms (21%); and 5) hailstorms (30%) [Changnon et al., 1971]. The results of this experiment were reported in 1974 in the Bulletin of the American Meteorological Society. In his summary of METROMEX, Braham [1974b] reported that the CCN production from the city was about $10^4 \text{ m}^{-2}\text{s}^{-1}$, much higher than the surrounding rural areas, accompanied by an increase in cloud drop concentrations and a decrease in drop size. However, the radar echoes from these clouds were usually lower than their counterparts in the rural surroundings. This seems to contradict our physical understanding of cloud growth, but it was concluded that one way to explain the observations is to assume that GCCN, that are not easily detected by the sampling methods, could be responsible for the increased precipitation.

Figure 6.20: Five year moving averages and time trend of Centerville (downwind of St. Louis) summer rainfall, 1941-1968. From Changnon et al. [1971].
Jin et al. [2005] analyzed diurnal, weekly, seasonal, and interannual variations of urban aerosols with an emphasis on summer months using 4-years of the National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) Moderate Resolution Imaging Spectroradiometer (MODIS) observations, in situ AErosol RObotic NETwork (AERONET) observations, and in situ EPA PM2.5 data for one mid-latitude city (New York) and one subtropical city (Houston). The research reveals that spatial and temporal urban aerosols optical depth vary dynamically as a result of various parallel factors, such as human activity, land cover changes, cloud-aerosol interactions, and chemical processes. Diurnal, seasonal, and interannual variations of aerosol optical depths have been examined and were found to be largely affected by weather conditions, but the optical depth often peaks during the rush hours in the morning and evening. Analysis of monthly mean aerosol optical thickness and rainfall did not show strong relationships between aerosol and rainfall in a climatological sense.

In their analysis, virtually no seasonality is observed for rainfall over Houston and New York City, suggesting that aerosols affect rainfall less than cloud microphysics. Around Houston, the Tropical Rainfall Measuring Mission (TRMM) satellite-based accumulated rainfall data show that the maxima monthly mean rainfall occurred in October 2000, May 2001, and September 2002. This is consistent with the transition between seasons in this region. In general, New York’s rainfall had less month-to-month variation than Houston, with a maximum slightly above 200 mm/month in October 2002. The effective radius for water clouds was lower in New York City than in Houston, suggesting either more aerosols in New York City than in Houston, or thinner clouds. The lack of direct relationship between rainfall and urban aerosol optical thickness implies that urban rainfall anomalies are not fully related to changes in aerosol. This observation is consistent with the earlier conclusions from METROMEX [Ackerman et al., 1978].

As can be seen from the above, in spite of many measurements, there is no conclusive evidence that aerosol pollution from urban regions indeed affects precipitation.

### 6.4.2 Effects of Air Pollution on Clouds and Rain from Orographic Clouds

Borys et al. [2000] and Borys et al. [2003] provided some evidence that pollution can delay precipitation in winter orographic clouds in the Rocky Mountains. Their analysis shows that
pollution increases the concentration of CCN and therefore cloud drops, leading to the formation of smaller cloud drops. The reduced drop size leads to less efficient riming and therefore to smaller ice crystals (Figure 6-21), smaller fall velocities, and less snowfall.

**Figure 6-21:** Light riming of ice crystals in clouds affected by pollution (left) compared to heavier riming in non-polluted clouds (right). From Borys et al. [2003].

*Givati and Rosenfeld* [2004] analyzed about 100 years of precipitation records in regions downwind of pollution sources and compared them to precipitation in regions unaffected by these sources. In their study, *Givati and Rosenfeld* [2004] documented the precipitation trends in the orographic enhancement factor, Ro, which is defined as the ratio between precipitations over the hill with respect to the upwind lowland precipitation amount. Two geographical areas were chosen for this study: California and Israel. The topography in both regions is similar, although the mountains in Israel are much lower than the Sierra Nevada. The statistical results for both locations show that downwind of pollution sources, on the upslope of mountains and mountain tops, orographic precipitation is reduced by ~20% and ~7%, respectively. It was hypothesized that this decrease is due to an increase in droplet concentrations and a decrease in droplet size. Farther downwind on the lee side of mountains, the amount of precipitation is increased by ~14%. The authors postulate that this increase is due to smaller cloud particles taking longer time to grow, allowing the winds aloft to carry them over the mountain top (see earlier study of similar effects, produced by deliberate over seeding with ice-producing particles, by *Hobbs*, 1975a and b). However, they hypothesized
that the integrated rainfall amount over the whole mountain range was reduced by the progressively increased pollution over the years. Subsequent studies show similar decreasing trends in Ro over a few western States in the US [Rosenfeld and Givati, 2006, Griffith et al, 2005] and the east slopes of the Colorado Rockies [Jirak and Cotton, 2005]. They argue that although absolute precipitation amounts and Ro are affected by fluctuations in the atmospheric circulation patterns, such as those associated with the Pacific Decadal Oscillation and the Southern Oscillation Index, they cannot explain the observed trends in Ro.

6.5 Summary

Except for a very few papers, most studies to date have dealt with the effects of anthropogenic aerosol on shallow warm non- or slightly precipitating clouds. Since the 1960s, in situ measurements provided numerous and consistent pieces of evidence that an increase in CCN concentrations results in an increase in droplet concentrations, although not in a linear fashion. The quantitative relationship between aerosol physico-chemical properties, updraft intensity at cloud base, and the resulting concentration of activated nuclei (i.e. the concentration of cloud droplets at cloud base) still needs to be improved for accurate evaluation of the aerosol impacts on clouds and climate.

In situ observations also corroborate the predictions of cloud microphysics theory, namely that clouds with more numerous, and hence smaller cloud droplets produce fewer precipitation embryos, than clouds with lower CDNC and a similar liquid water path.

Monitoring of these effects from satellite is less convincing because of the present uncertainties in remote sensing measurements of the liquid water path, which are required to stratify the observed data sets and separate the aerosol impact from the natural meteorological variability.

Beyond the LWP variability, there are additional cloud features that may significantly affect the retrieval of cloud microphysical properties from satellite. Factors of importance are spatial heterogeneity, and the variability of mixing processes in the upper part of the cloud layer, that mainly determines its radiative properties.
In spite of the limitations, satellite retrieved \( T_r \) relations have been used to gain some insight into the cloud and precipitation forming processes, and possibly to the impact of aerosols on them. Combined with spaceborne radar and passive microwave measurements, this methodology has been used in studies of the effects on precipitation of smoke from biomass burning, urban and industrial pollution, desert dust, and sea salt.

Recent remote sensing observations show that increased drop concentrations might lead to larger cloud fractions and deeper clouds. Similar results have been reported from numerical model simulations. However, there are indications from remote sensing studies that an increase in aerosol may decrease the cloud fraction of small warm cumulus clouds.

While numerous publications deal with the effects of pollution on clouds, there are few that treat the effects of pollution on the amount of precipitation on the ground. Some observations suggest that orographic precipitation may be reduced by increased aerosol pollution. The overall effects of cities on precipitation appear not to be clearly connected to aerosol pollution.

Large or GCCN in concentrations on the order of 1 per liter can significantly affect the early development of precipitation embryos in clouds. Their effect on precipitation on the ground is yet to be established.

Ice formation in clouds is not well understood. Certain effects of aerosol pollution on ice processes have been observed, but an overall link between aerosol pollution, ice nuclei, ice crystal concentrations and growth to precipitation has not been established.

To reach firmer conclusions on these effects, further observational studies and new techniques are needed on the following issues:

- For each aerosol type always characterize their CCN, IN and light absorption spectral properties (single scattering albedo).
- Measure the vertical profile of the aerosol particles in order to connect between them and the microphysical processes in the clouds.
- Evaluate the role of GCCN in the precipitation development.
• Improve the satellite retrieval techniques of the liquid water path, accounting for the cloud spatial heterogeneity and the mixing process biases.
• Improve measurements of the precipitation amount at the ground.
• Design statistical analyses that separate the effects of meteorology from the effects of aerosol pollution on precipitation on the ground.
CHAPTER 7: EFFECTS OF POLLUTION AEROSOL INCLUDING BIOMASS BURNING ON CLOUDS AND PRECIPITATION: NUMERICAL MODELING STUDIES

Lead authors: Graham Feingold, William Cotton, Ulrike Lohmann, Zev Levin

7.1 Introduction

The history of numerical modeling of the effect of aerosols on clouds dates back at least 50 years to the work of Howell [1949] and Mordy [1959] who considered the growth of a population of aerosol particles in a rising parcel of air. Models such as these addressed the effects of both aerosol and dynamical parameters (i.e., updraft velocity) on the number and size distribution of cloud droplets. To this day similar models are in wide use to examine the effects of aerosol composition and atmospheric trace gases on droplet activation (e.g., [Kulmala et al., 1993; Ghan et al., 1997; Feingold and Chuang, 2002; Nenes et al., 2002]).

From the early roots of cloud parcel models, simulating droplet growth by condensation, there have been two parallel and complementary foci (Figure 7-1). The cloud physics community has pursued the modeling of precipitation formation in a variety of modeling frameworks, ranging from parcel models to 1-D, 2-D and 3-D models of both warm and cold clouds. These studies have been able to capture the salient features of cloud droplet activation and the timescales of growth to precipitation-sized drops in a reasonable manner. Simultaneously, the aerosol and chemistry communities have developed models that have placed more emphasis on the effect of aerosol composition on cloud microphysics, as well as on the role of clouds as processors of aerosol (e.g., via aqueous-phase chemistry). Although these efforts are depicted as parallel in Figure 7-1, there has been some communication between the two communities with the result that the more recent modeling efforts include representation of the coupled aerosol-cloud-chemistry system (e.g., [Barth et al., 1992; Respondek et al., 1995; Feingold and Kreidenweis, 2002; Yin et al., 2005]). The representation of all these components is required to model the effect of aerosols on precipitation because of the myriad feedbacks that can occur in the system. For example, aerosols that are modified by cloud processes, and are then released back into the atmosphere after cloud evaporation (most clouds evaporate and do not precipitate) could affect other
clouds [e.g., Scott and Hobbs, 1967; Hegg et al., 1980, 1996; Barth et al., 1992; Hobbs, 1993; Wurzler et al., 2000; Garrett et al., 2002; Tabazadeh et al., 2004]. Therefore, not only must the effects of aerosols on clouds and precipitation be considered but also the effect of clouds and precipitation on aerosols.

Figure 7-1. Brief summary of parallel efforts in aerosol-cloud-precipitation studies over the past six decades. The left branch has focused on precipitation development, while the right branch has been concerned with aerosol composition and aqueous chemistry. Communication between these efforts over the past decade has stressed the importance of considering the aerosol-cloud-chemistry-precipitation system in a coupled manner.

The challenge to the cloud modeler is to represent the numerous, and sometimes poorly understood cloud processes depicted in Figure 7-2 in a dynamical framework, and preferably in three dimensions (3-D). The lifetime of an individual cloud cell is on the order of an hour but, when considering precipitating cloud systems, multi-hour simulations need to be performed, with the result that computational costs quickly become prohibitive. The challenges therefore include adequate representation of (i) physics, (ii) chemistry, (iii) numerical methods for solving the equations, (iv) consideration of the most important processes and (v) consideration of temporal and spatial scales.
Below we describe some of the modeling techniques that have been used to represent precipitation, and parallel efforts at improving numerical methods, as well as the representation of the physical processes themselves. We then give examples of key results pertaining to aerosol-cloud-precipitation interactions. Finally, we discuss some of the challenges in quantifying aerosol effects on clouds and precipitation.

7.2 Numerical Methods For Modeling Cloud Microphysics

Cloud modeling requires a recognition that the aerosol cloud system comprises coupled components of dynamics, aerosol and cloud microphysics, radiation and even chemistry (Figure 7-3), and that depending on the system being modeled, neglect of some of the components may have important consequences. The correct approach to modeling the effects of aerosol on precipitation requires an appropriate balance of treatment of these components. Earlier modeling efforts tended to focus on just parts of the system (e.g., the effect of aerosol size and composition on the drop size distribution with very detailed representation of the aerosol/drop size distribution [Mordy, 1959], or the formation of precipitation in a 1-D model using simple representations of microphysics, frequently with little representation of aerosol
[Silverman and Glass, 1973]. To this day, cloud models are often used to focus on just one aspect of the system because the neglect of other components simplifies investigations and serves to elucidate specific processes. However, with the increase in computing power, and the evolution of numerical methods, the community has slowly moved towards representing both microphysics and dynamics with varying degrees of sophistication.

**The Coupled System**

![Figure 7-3](image-url)

*Figure 7-3. Schematic indicating the coupled nature of the microphysics involving clouds and precipitation. Although great gains in knowledge have been acquired through consideration of a limited number of components in this figure, an appropriate balance between these components must be considered when modeling aerosol-cloud-precipitation interactions.*

In approaching numerical cloud modeling, the importance of numerical techniques that minimize spurious particle growth must be considered hand-in-hand with the representation of the physics inherent to cloud microphysical and dynamical processes. For example, the kernel function for collision and coalescence between different types of hydrometeors is based on both theory and laboratory experiments, in which the outcome of the interactions between pairs of hydrometeors of different sizes is calculated or measured. These are then parameterized into more general functions that can be embedded in the microphysical equations. Accurate parameterization of these processes is as important as consideration of numerical methods in determining our confidence in representing microphysical processes in cloud models. In fact it can be argued that the poor knowledge of collection kernels,
particularly regarding ice-ice interactions, may limit the gains that can be realized by very
detailed microphysical schemes.

7.2.1 Bulk Microphysics

When considering cloud microphysical processes, earlier models, and many current models,
use a “bulk” representation of microphysics that represents a size distribution of
hydrometeors by one or more of its moments (e.g., mass mixing ratio, number mixing ratio,
surface area, radar reflectivity). In one approach drops and ice particles are represented by a
simple exponential function (with fixed pre-exponent) and a single-moment of the
hydrometeor spectra is predicted [Kessler, 1969]. In this case self-collection among cloud
droplets is parameterized using an autoconversion formulation, and large hydrometeors such
as raindrops are assumed to collect smaller drops and ice by continuous accretion. Moreover,
all classes of large hydrometeors are assumed to fall with a constant fall speed, usually a
water content-weighted fall speed. The autoconversion formulations have been either ad hoc
(e.g., [Kessler, 1969; Manton and Cotton, 1977; Cotton et al., 1986]) or derived from parcel
or box detailed bin-microphysics simulations (e.g., [Berry, 1967; Berry and Reinhardt, 1974;
Beheng, 1994]).

Multi-moment bulk schemes typically prescribe basis functions for the drop size
distributions, such as gamma or log-normal distributions [Clark, 1976; Clark and Hall, 1983;
Nickerson et al., 1986; Ferrier, 1994; Meyers et al., 1997; Reisner et al., 1998; Seifert et al.,
2006a,b; Milbrandt and Yau, 2005a,b], and then explicitly predict the evolution of those
basis functions (in terms of their moments) by vapor deposition/evaporation, stochastic
collection, and sedimentation. The advantage of these multimoment schemes is that they
predict number concentration, mass mixing ratio (and sometimes higher order moments) and
therefore are able to derive the broad features of the drop size distribution. In so doing they
improve the representation of growth processes and precipitation formation. Multimoment
bulk methods have also been applied to representation of aerosol growth processes (e.g.,
[McGraw, 1997]).

The two-moment scheme by Seifert and Beheng [2006a,b] predicts the evolution of mass as
well as number densities of the five hydrometeor types cloud droplets, raindrops, cloud ice,
snow and graupel. It includes novel parameterizations of autoconversion, accretion and self-
collection of water drops that have been derived by Seifert and Beheng [2001] directly from
the stochastic collection equation. In contrast to most other schemes, the new autoconversion
parameterization further considers aging of the cloud droplet size distribution with time by
relying on a dynamic similarity theory. The parameters determining the shape of the droplet
size distribution in this parameterization were chosen using droplet size distributions
simulated within the dynamical model with bin microphysics model developed at the Hebrew
University of Jerusalem [HUCM; Khain et al. 2004; Seifert et al. 2006]. This calibration
significantly increased the sensitivity of results to aerosol and allowed the authors to obtain
similar results in bulk- and bin-microphysical simulations of maritime and continental
clouds.

Another example of a multi-moment scheme implemented in a large scale model is the bin-
emulating scheme, which is implemented in RAMS (Regional Atmospheric Modeling
System, Pielke et al., 1992 and Cotton et al., 2003). The evolution of this modeling approach
can be found in Verlinde et al. [1990], Walko et al. [1995], Meyers et al. [1997], Feingold et
al. [1998], and Saleeby and Cotton [2004]. Instead of using continuous accretion
approximations, which has been common in cloud parameterizations, Feingold et al. [1998]
showed that full stochastic collection solutions for self-collection among cloud droplets and
for rain (drizzle) drop collection of cloud droplets can be obtained for realistic collection
kernels by making use of look-up tables. This approach has been extended to all hydrometeor
class interactions by collection, including the growth of graupel and hail by riming. The
philosophy of bin representation of collection has been extended to calculations of drop
sedimentation [Feingold et al., 1998]. Bin sedimentation is simulated by dividing the basis
function into discrete bins and then building look-up tables to calculate how much mass and
number in a given grid cell fall into each cell beneath a given level in a given time step.
Saleeby and Cotton [2004] refined this approach by the addition of a large cloud droplet
mode from 40-80 µm diameter, and by predicting the number concentration of cloud droplets
through explicit activation of CCN and giant-CCN. The large cloud droplet mode provides a
better depiction of self-collection of cloud droplets (or autoconversion) and permits
simulation of drizzle from fogs and marine stratocumulus clouds. The activation of CCN is
parameterized through the use of a Lagrangian parcel model that considers ambient cloud
conditions for the activation of cloud droplets from aerosol.
7.2.2 *Size-resolved Microphysics*

Size-resolved or explicit bin microphysical methods have been developed with the recognition that many cloud growth processes are highly sensitive to drop size. The first such class is the Lagrangian (or moving mass-grid) method, which represents particles at discrete sizes and allows each particle to grow by condensation on a moving mass grid. This approach eliminates numerical diffusion and allows for a smooth transition from aerosol to haze to cloud droplets without artificial distinctions between these classes. The cloud supersaturation is calculated based on source (cooling, which is related to updraft velocity) and sink (condensation) terms, enabling accurate determination of the number of activated droplets. These models typically focus on the initial growth phases from haze to droplet and include detailed representation of aerosol sizes and compositions [Mordy, 1959; Fitzgerald, 1974; Facchini et al., 1999; Feingold and Kreidenweis, 2000; Feingold and Chuang, 2002; Lohmann et al., 2004]. They frequently also consider the effect of trace gases on activation [Kulmala et al., 1993]. Aqueous production of sulfate has also been represented in studies that examine the effects of cloud processing on the aerosol size distribution [Hegg et al., 1991; Bower and Choularton, 1993; Feingold and Kreidenweis, 2000]. The Lagrangian method is used almost exclusively in kinematic cloud parcel models, where a parcel of air is moved either adiabatically or according to some known trajectory through a cloud. It is not easily adapted to the study of growth by processes such as collision-coalescence, and it is not suitable for general application in Eulerian dynamical models.

Fixed bin, or Eulerian (in size space), microphysical models have been developed to fill this need. Supersaturation calculations also follow from a balance equation for cooling (dynamical tendency) and condensation but accuracy depends on the grid resolution [Clark 1973; 1974]. Activation is not represented as accurately because the Eulerian framework does not represent the transition from dry particle to haze to droplet as in the Lagrangian framework. Condensational growth is also subject to numerical diffusion associated with the mass-grid resolution. (This also feeds back to the accuracy of the supersaturation calculation.) The bin framework is most useful for collision-coalescence calculations. Early
efforts by Telford [1955], Berry [1967], Berry and Reinhardt [1974], Kovetz and Olund [1969], and Bleck [1970] showed that great care must be taken in representing collision-coalescence to avoid numerical diffusion in the mass-transfer equations and rapid (spurious) acceleration of growth to precipitation-sized particles. More recent bin models, such as those used by Tzivion et al. [1987], Hounslow et al. [1988], and Chen and Lamb [1994], use a multi-moment representation of the cloud microphysics in each individual drop bin; this significantly reduces numerical diffusion and has the added benefit of conserving more than one moment of the size distribution. This has led to development of numerical methods that include a bin representation of aerosol in each individual hydrometeor size-bin [Bott et al., 1990; Chen and Lamb 1994; Kerkweg et al., 2003; Leroy et al., 2006]. Such methods are very accurate since they maintain knowledge of the aerosol particle upon which the drop and/or ice particles form [Leroy et al., 2006] but are too computationally intensive to be included in 3-D models. Simpler methods that track dissolved aerosol within each hydrometeor bin are more commonly used [Flossmann et al., 1985; Respondek et al., 1995; Toon et al., 1988; Feingold et al., 1996; Yin et al., 2005]. A review of aspects of bin microphysical modeling of both warm and cold cloud processes can be found in Khain et al. [2000].

Another approach to size-resolved microphysical modeling is the hybrid approach [Cooper et al. 1997; Jacobson, 1999; Pinsky and Khain, 2002] where the advantages of moving grids for condensational growth are combined with fixed grid approaches for collection processes.

While the focus here has been on numerical methods for solving microphysical interactions, it is important to consider numerical issues related to the dynamical aspects of cloud modeling. The advection equation is also susceptible to numerical diffusion (and dispersion). Modeling of the dynamical equations for hydrometeor species requires “positive-definiteness” for mass conservation. Numerical diffusion on the spatial grid has to be considered along with numerical diffusion on the mass grid; again, an appropriate level of balance is important [Clark, 1973; 1974].

7.3 Modeling Aerosol-Cloud-Precipitation Interactions

7.3.1 Warm Clouds: Aerosol Effects on Drizzle in Stratocumulus Clouds
Stratocumulus clouds have been the focus of numerous intensive field campaigns as well as detailed modeling efforts, primarily because of their importance to the Earth’s radiation budget rather than because they produce precipitation in significant quantities. Their importance derives largely from their frequency of occurrence and extensive spatial coverage, as well as their high reflective contrast with the underlying surface (particularly over oceans). Over land and ocean, stratocumulus clouds have an annually averaged coverage of ~18% and 34%, respectively [Warren et al., 1986a,b]. Stratocumulus clouds frequently produce drizzle (e.g., [Paluch and Lenschow, 1991; Stevens et al., 2003]) and the initiation of drizzle has been the focus of much modeling effort. Modeling of aerosol-cloud interactions in stratocumulus clouds received attention in the early 1990s with the development of 1-D turbulence closure models coupled to bin microphysical models [Ackerman et al., 1995; Bott et al., 1996], and modeling at the large eddy scale in 2-D and 3-D simulations [Kogan et al., 1994; Liu et al., 2000; Feingold et al., 1994; 1996; 1997; Stevens et al., 1996, 1998; Khairoutdinov and Kogan, 2000; Feingold and Kreidenweis, 2002; Jiang et al., 2002].

Ackerman et al. [1995] used a 1-D turbulence closure model coupled to an aerosol and cloud particle size-resolved model to show that progressive aerosol scavenging by droplet activation, by drop collisions and drizzle depletion of liquid water in a stratocumulus cloud can reduce the cloud liquid water path (LWP) and optical depth to a point where there is insufficient cloud-top radiative cooling to maintain the cloud. Under such conditions the cloud collapses into a fog layer. The authors suggested that coalescence scavenging and drizzle could limit the lifetime of stratocumulus clouds. They contrasted this phenomenon with the more common hypothesis that increases in aerosol concentrations suppress drizzle and help to sustain clouds with higher LWPs [Albrecht, 1989]. Bott et al. [1996] used a similar 1-D model to examine dynamical-microphysical-radiation feedbacks. Recognizing the importance of better representation of boundary layer dynamics, Feingold et al. [1994] and Kogan et al. [1994] coupled similar bin microphysical models to 2-D and 3-D Eulerian models with grid sizes on the order of 50 m–100 m. Kogan et al. [1995] demonstrated the importance of microphysical-dynamical coupling, particularly through the supersaturation field. Feingold et al. [1996] showed that in-cloud residence time is an important factor in determining precipitation development in stratocumulus. Stevens et al. [1996] pointed out that air parcels that spend significant time at cloud top, where LWC is highest, are more likely to produce drizzle drops.
Feingold et al. [1997] showed that broadening of drop spectra via collision-coalescence tends to increase cloud susceptibility (defined by Twomey [1991] as the incremental change in cloud albedo for an incremental change in cloud drop concentration at constant cloud water). Khairoutdinov and Kogan [1999] compared their large eddy simulation (LES) model including bin microphysics to measurements from field experiments and showed generally good agreement in the dynamical and microphysical fields. Driven by the large computational expense of resolving drop size distributions in LES, Feingold et al. [1998] and Khairoutdinov and Kogan [2000] studied drizzle formation in stratocumulus using both bin and bulk (2 moment) schemes and showed that bulk schemes can reproduce many of the features of drizzle formation. Liu et al. [2000] studied ship tracks using LES and demonstrated the effects of boundary layer stability on both inclusion of the stack effluent into the clouds as well as maintenance of the ship track. Decoupling of the boundary layer suppresses transport of the effluent into cloud. Because boundary layer stability has a diurnal cycle, which modifies both cloud LWC and the amount of aerosol affecting the cloud, persistence of ship tracks is not easily predicted.

The LES/bin-microphysics simulations performed by Stevens et al. [1998] illustrate the importance of drizzle to the cloudy marine boundary layer. For strongly drizzling cases where the drizzle reaches the surface, evaporation results in cooling of the surface and stabilization of the sub-cloud layer [Paluch and Lenschow, 1991]. As a result, the cloud structure changes to a cumulus-under-stratus regime, in which lower-based cumulus clouds penetrate into the stratus layer supplying it with water vapor. The cumulus-under-stratus regime has very different optical properties than a pure solid stratocumulus field. However, in Jiang et al.’s. [2002] weakly drizzling LES model simulations, the response was quite different. The small amount of drizzle that was produced evaporated just below the cloud base, and had a destabilizing effect on the cloud layer. Addition of higher concentrations of aerosol suppressed this drizzle and consequently, the destabilization. This resulted in weaker penetrating cumulus and an overall reduction in the LWC and LWP of the clouds.

Other aspects of the effect of aerosol on cloud liquid water path (LWP) were considered by Ackerman et al. [2004]. They also showed that increases in aerosol do not necessarily result in increases in LWP in stratocumulus clouds, as proposed by Albrecht [1989]. A primary factor affecting the LWP response to aerosol changes appears to be the profile of humidity
above the inversion. Only when the humidity above the inversion was high did increases in aerosol result in an increase in LWP. When dry air overlies the inversion, increases in aerosol tend to decrease LWP because of enhanced entrainment drying. Similar results were obtained by Lu and Seinfeld [2005]. In large eddy simulations of marine trade cumulus clouds, Xue and Feingold [2006] showed that although aerosol has a strong effect on precipitation, aerosol effects on cloud LWP and cloud fraction are relatively small compared to the dynamical variability of the clouds at given aerosol concentrations. In fact their simulations showed a decrease in cloud fraction with increasing aerosol, apparently due to stronger evaporation of the smaller cloud droplets, and an evaporation-entrainment feedback that reduced cloud size (see also [Wang et al., 2003]). Note that this result is for small cumulus clouds (order few 100 meters) and that the result is contrary to the Albrecht hypothesis and the satellite-measured increase in cloud fraction with increasing aerosol for larger clouds [Kaufman et al. 2005b]. The small aerosol effects on LWP in warm cumulus have also been modeled by Grabowski [2006] based on very long simulations (120 days) with a two-dimensional cloud model and bulk microphysics. The latter work stressed the importance of these extended simulations over large domains to obtain a statistically representative sample of aerosol-cloud interactions. It was shown that when viewed as a long time series, polluted clouds produced similar amounts of precipitation to clean clouds. This result is influenced by the fact that sea surface temperature was held fixed in the model, which predetermines the surface fluxes and precipitation rates. The author stressed the importance of results for ensembles of clouds vs. those for single clouds. We will return to this theme in Section 7.3.4.2.

7.3.2 Role of Giant Cloud Condensation Nuclei

An early hypothesis for droplet spectral broadening and the development of precipitation in warm clouds is that the aerosol population includes giant cloud condensation nuclei (GCCN) and/or ultragiant aerosol particles that can act as the embryos for initiating coalescence growth (Chapter 2). Johnson [1979; 1982] using a parcel model representing droplet growth on a population of aerosol particles via condensation and coalescence calculated that the addition of GCCN at concentrations on the order of 1 per liter (i.e., about 1 particle in $10^5$ or $10^6$; typical measured concentrations) can account for rapid development of precipitation-sized particles, even in colloiddally stable, continental clouds. Woodcock et al. [1971] and
Takahashi [1976] found that giant salt nuclei do not contribute substantially to warm rain initiation in maritime cumulus clouds because drizzle is active anyway under clean conditions. Feingold et al. [1999] used several models, including parcel models and 3-D large eddy simulations of marine stratocumulus clouds with detailed bin microphysics, to study the effect of GCCN on drizzle formation in stratocumulus. They found that observed GCCN concentrations of $10^{-4} \text{ cm}^{-3}$ to $10^{-2} \text{ cm}^{-3}$ were sufficient to transform a non-precipitating stratocumulus into a precipitating one when CCN concentrations were in the range 50–250 cm$^{-3}$. At lower CCN concentrations, addition of GCCN has little impact on drizzle formation, in agreement with earlier work in clean cumulus clouds. Note however, that although the relative importance of GCCN increases with increasing pollution, there is a point of diminishing returns since precipitation amounts tend to decrease steadily with increasing pollution (all else being equal).

Similar results were obtained by Yin et al. [2000a] for warm convective clouds (using a 2-D slab symmetric cloud model). Yin et al. found the intriguing result that a polluted cloud with CCN concentrations of $\sim 1700 \text{ cm}^{-3}$ and GCCN of $0.02 \text{ cm}^{-3}$ can precipitate more readily than a cleaner cloud ($\sim 1000 \text{ cm}^{-3}$ CCN) with no GCCN. This modeling result is supported by the observations of Eagan et al. [1974a,b] and Hindman et al. [1977] on the effects of effluents from paper mills on cloud micro-structures and of Rudich et al. [2002] on the effects of salt on polluted clouds near the Aral Sea. Wurzler et al. [2000] showed that cloud-processed dust particles coated with a soluble material could enhance precipitation. Rosenfeld et al. [2002] suggested that GCCN could be an effective mechanism for controlling precipitation, in agreement with earlier studies. Segal et al. [2004] used a high resolution (2000-bin) Lagrangian parcel model describing diffusional and collisional growth of aerosols and drops to investigate the comparable effects of small and large CCN on the droplet spectrum and raindrop formation. It was shown that when updrafts increase in strength above cloud base, bimodal spectra may form as a result of activation of CCN associated with the increasing supersaturation [see also Erlick et al. 2005]. As in prior studies, it was shown that the collision-coalescence process begins to play an important role even at the stage of condensational growth and GCCN play an important role at higher CCN concentrations.

In addition to influencing the warm rain process, GCCN also affect cold clouds by increasing the concentration of larger graupel particles [e.g., Yin et al., 2000a; Teller and Levin, 2006]. Yin et al. [2000a] showed that by enhancing collision-coalescence in polluted clouds (CCN ~
1700 cm$^{-3}$), the inclusion of GCCN decreased both drop and graupel concentrations; however, graupel mass was substantially increased. Precipitation was initiated earlier, rain intensities were higher, and the total accumulated precipitation as much as 4 times higher. Moderately polluted clouds (CCN $\sim$ 1000 cm$^{-3}$) exhibited weaker, but still significant enhancements in rain due to GCCN (increase on the order of 60 %).

The importance of GCCN requires consideration of the body of modeling studies of hygroscopic seeding that consider the addition of giant salt particles into clouds as a means of artificially enhancing precipitation [e.g., Hindman et al., 1977; Cooper et al., 1997; Segal et al. 2004]. The studies of Cooper et al. [1997], Yin et al. [2000b], and Segal et al. [2004] are of special interest because they elucidate the role not only of the giant seeded particles, but also of the other hygroscopic particles present in flares that have been used for artificial seeding. (See Chapter 8 section 3 for a more detailed discussion of these modeling results in the context of deliberate cloud seeding.)

The relative sensitivity of a cloud to GCCN has to be considered in the context of the non-linear response of the collection process to cloud liquid water content, the sensitivity to cloud drop concentration, and the time available for collection. Quantification of precipitation in terms of GCCN alone is therefore not feasible; no study to date has quantified the effect of GCCN in terms of all the pertinent parameters.

### 7.3.3 Cloud Processing – Cloud Effects on Aerosol

The recognition that aerosols not only affect clouds but that clouds affect aerosols (e.g., [Easter and Hobbs, 1974; Hegg et al., 1980]) provided challenges to the modeling community by requiring that CCN be treated as size-resolved prognostic species and be tracked as soluble material inside drops [Flossmann et al., 1985; Toon et al. 1988; Bott et al., 1990; Chen and Lamb, 1994; Feingold et al., 1996]. Modeling of aqueous sulfate production, and the resultant modification of the aerosol size distribution, has most often been simulated in Lagrangian parcel models (e.g., [Easter and Hobbs, 1974; Hegg and Larson, 1990; Roelofs, 1993; Bower and Choularton, 1993; Roelofs, 1992; Wurzler et al., 2000]). These models have reproduced the creation of a bimodal aerosol size distribution as observed by Hoppel et al. [1990]. The implications of this processing range from increases in light...
scattering [Hegg et al., 1996] to effects on CCN and drop number concentration [Bower and Choularton, 1993; Feingold and Kreidenweis, 2000; Yin et al., 2002]. Eulerian models have also been used to study cloud processing. Bott [2000] used a 1-D model, with a very detailed 2-D representation of the joint aerosol-drop distribution function, to study the modification of aerosol size distributions by aqueous chemistry for a range of aerosol and gas phase conditions. Feingold et al. [1996] incorporated bin microphysics for aerosols and drops in 2-D simulations of stratocumulus and studied the rate of aerosol scavenging via drop collection. Drop collection reduces the total drop number concentration, so that on subsequent evaporation the available solute mass is redistributed among fewer, and therefore larger, particles. These authors compared their results to aqueous chemistry processing and suggested that at high LWC, coalescence processing of aerosol would likely dominate aqueous chemistry processing. Similar inferences were made by Hatzianastassiou et al. [1998] and Wurzler et al. [2000]. Feingold and Kreidenweis [2002] used a 3-D LES model with size-resolved aerosol and cloud drops, as well as aqueous sulfate production. They showed that the addition of sulfate mass could increase or decrease drizzle formation in stratocumulus depending on the size distribution of the background aerosol. When mass addition occurred at sizes significantly larger than the background CCN, drizzle was increased; conversely, when mass addition occurred close to the mode of the background CCN, drizzle was reduced.

In precipitating clouds, the most direct form of cloud processing of aerosol is through wet deposition to the Earth’s surface. It has been shown that 50-90% of aerosol number concentration is removed by in-cloud nucleation (activation) processes [Flossmann et al. 1985; Flossmann and Pruppacher, 1988, Taylor 1989]. Brownian motion is effective at removing the smaller and more numerous particles. For example, Jacobson [2003] found that 42% of the number concentration of aerosol was scavenged through this process, but with little effect on aerosol mass. The reduction in aerosol mass concentration by in-cloud activation is greater than 99%.

Below cloud, impaction scavenging (washout) has a negligible effect on particle concentration since it affects the larger particles that exist in very low concentrations. The reduction in mass is also much smaller than in-cloud losses (5-10%) [Flossmann and Pruppacher 1988; Taylor 1989]. In spite of this, both of these studies suggest similar
contributions of in-cloud and below-cloud scavenging to the mass of aerosol deposited to the surface by rain. This is a result of the low precipitation efficiencies (~ 10-15%) for these simulated clouds and the fact that much of the in-cloud scavenged aerosol mass does not reach the surface as rain. Note that impaction scavenging removes both hydrophobic and hydrophilic aerosols while nucleation scavenging removes only hydrophilic aerosols.

These results have been confirmed by later studies such as those by Flossmann et al. [1998] and Yin et al. [2005]. The latter study coupled a bin microphysical model to an aqueous sulfate chemistry model. Their simulations showed that the addition of sulfate mass had an insignificant effect on cloud evolution, even for polluted conditions (10 ppbv SO$_2$ and 1 ppbv H$_2$O$_2$).

Convective clouds also act to transport boundary layer aerosol to the middle and upper troposphere [e.g., Cotton et al., 1995, Flossmann, 1998]. This can have important implications for precipitation as well as for upper tropospheric heterogeneous reactions and the life cycle of aerosols [Clarke et al., 1999]. Additionally, Teller and Levin [2006] showed that polluted clouds containing high concentrations of CCN (with about 1000 droplets cm$^{-3}$ as compared 100 cm$^{-3}$) not only lead to reduced precipitation but also leave behind as much as three times more water vapor from the evaporation of cloud drops and ice crystals at higher altitudes after precipitation stops. This could affect radiation transfer as well as various chemical reactions. See Wang [2003] for discussion on the role of convective storms in redistributing water vapor.

Cloud processing of aerosol also has important consequences on regional scales. Aerosol is processed through cloud during long-range transport, with the extent of this processing dependent on cloud contact time, cloud liquid water content, and the availability of SO$_2$ and oxidants.

7.3.4 Mixed-phase Clouds: Effects of Aerosols on Precipitation

7.3.4.1 Single Clouds

A number of early studies have shown that the speed of glaciation of a cloud is highly dependent on the presence of large supercooled raindrops or on the concentrations of CCN.
When CCN concentrations are high, warm-cloud collection processes are suppressed and supercooled raindrops are few in number, small ice crystals must first grow to several hundred micrometers in diameter before they begin collecting cloud droplets. Then the riming process proceeds slowly until the ice particles have grown to millimeter-size. Furthermore, since riming is suppressed in clouds forming in air masses with high CCN concentrations, secondary ice particle production by the riming-splintering process [Hallett and Mossop, 1974; Mossop and Hallett, 1974] is suppressed as well.

Reisin et al. [1996a,b] studied the effects of changes in CCN on precipitation in an axisymmetric cloud model with a comprehensive bin microphysical representation of water and ice hydrometeors. They showed that a simulation with relatively low CCN (100 cm\(^{-3}\) at 1% supersaturation) produced precipitation efficiently through the freezing of large droplets interacting with ice crystals. With increasing CCN concentrations and decreasing drop sizes, graupel growth was suppressed and the precipitation efficiency decreased (for CCN=900 cm\(^{-3}\), precipitation was reduced by 85%). This study also showed that for the sounding considered, increases in ice nucleus (IN) concentrations tended to produce higher precipitation amounts but the changes were very small for a three order of magnitude change in IN in a moderately polluted cloud (CCN=600 cm\(^{-3}\) at 1% supersaturation). They found a greater sensitivity to IN concentrations when CCN concentrations (1100 cm\(^{-3}\)) were high.

Khain et al. [1999] used a 2-D (slab-symmetric) model with bin microphysics to simulate aerosol effects on precipitation in an Eastern Mediterranean coastal setting. They too found that warm rain was significantly reduced by an increase in the aerosol concentrations. However, they noted that in the case of polluted clouds, the lofting and downwind transport of smaller precipitation particles could produce more rain a few tens of kilometers from the convective region than was produced by the cleaner clouds with their efficient warm rain process. The result points to the importance of both dynamics and microphysics through particle “size sorting” and spatial redistribution of ice and water phases in the spatial domain.

In a recent study, Teller and Levin [2006] using a 2-D slab symmetric model with detailed treatment of cloud microphysics also showed that increases in CCN due to pollution in Mediterranean type winter convective clouds decrease rain amounts on the ground. When the ingested air into the cloud also contained GCCN in concentrations similar to those found in
the atmosphere (a few per liter) the amount of rain increased. They showed that increasing CCN concentration from 600 to 1000 cm\(^{-3}\) resulted in a decrease in rainfall by a factor of 3, as compared to a decrease by a factor of only 1.7 when GCCN were present (Figure 7-4). Furthermore, they showed that increasing ice concentrations above those produced through the use of Meyers et al. [1992] parameterization resulted in a decrease in rainfall. This decrease was more pronounced in the clean than the polluted clouds.

**Figure 7-4**: Dependence of total precipitation on CCN concentration based on 2-D single cloud simulations including bin microphysical representation of water and ice [Teller and Levin, 2006]. The inclusion of GCCN increases precipitation under polluted conditions but the trend of decreasing precipitation with increasing CCN is robust.

Khain et al. [2001] simulated the effect of varying the concentration of CCN on ice-crystal concentrations in deep convective clouds. Clean clouds (CCN ~ 100 cm\(^{-3}\)) produced large drops and rapid freezing at ~ -15ºC, resulting in significant depletion of cloud water. In a polluted cloud, (CCN =1260 cm\(^{-3}\)), freezing was delayed until much lower temperatures (~ -35ºC) because the small drop sizes rendered mechanisms that convert water droplets to ice inefficient, until the level of homogeneous freezing was reached. These simulations compare well with in-situ measurements exhibiting high LWC at temperatures as low as -38ºC under polluted conditions [Rosenfeld and Woodley, 2000]. Precipitation amounts were about half of those in the clean case. Khain et al. suggested that the general assumption that droplet-graupel collision efficiencies are equal to droplet-droplet collision efficiencies strongly overestimates the rate of collection of small droplets by small graupel; a sensitivity test showed that cloud LWC is depleted at much higher temperatures if this assumption is made.
This result points to the importance of quantifying collision efficiencies and the importance of commensurate efforts in both modeling methodology and laboratory studies.

The importance of treating aerosol and cloud microphysics as a coupled system has been highlighted in recent papers by Fridlind et al. [2004], Yin et al. [2005], and Seifert et al. [2006]. Fridlind et al. [2004] combined measurements from Florida cumulonimbus with 3-D modeling (with bin-microphysics for aerosol and hydrometeors) to infer that most anvil ice-crystals form on mid-tropospheric aerosol and not boundary layer aerosol. Lateral entrainment incorporates these particles into the updrafts where they are exposed to supersaturations and activated. Similar conclusions were drawn by Yin et al. [2005] with a 2-D axisymmetrical model employing bin microphysics for aerosol and cloud microphysics, as well as aqueous sulfate chemistry. They showed that aerosol regeneration, which returns aerosol to the mid-troposphere upon evaporation of hydrometeors, has a significant effect on microphysical pathways; when this process was neglected, graupel concentrations were nearly doubled and surface precipitation increased by ~50%. Another example of the importance of tracking aerosol budgets can be found in Seifert et al. [2006]. Squall line simulations that excluded budgets on the aerosol evolved quite differently from those that did not. When aerosol depletion was allowed, the simulation resulted in cleaner conditions and suppression of secondary convection; when aerosol depletion was not allowed, secondary convection did occur.

### 7.3.4.2 Multicell Clouds

A number of modeling groups using very different microphysics models and dynamic configurations are suggesting complicated dynamical interactions with aerosols with significant impacts on simulated precipitation. For example, Seifert and Beheng [2006b] showed that the effect of changes in CCN on mixed phase convective clouds is quite dependent on cloud type. They found that for small convective storms, an increase in CCN decreases precipitation and the maximum updraft velocities. For multicellular storms, the increase in CCN has the opposite effect – namely, promoting secondary convection, and increasing maximum updrafts and total precipitation. Supercell storms were the least sensitive to CCN. Their study also showed that the most important pathway for feedbacks from microphysics to dynamics is via the release of latent heat of freezing. Likewise, Lynn et al. [2005a] performed mesoscale simulations with bin-resolving microphysics for Florida.
deep convective clouds and found that for shallow clouds, high concentrations of CCN delayed the formation of precipitation in initial clouds relative to clean clouds. Thus clouds in a maritime airmass precipitated sooner than clouds forming in a continental airmass but this led to weaker secondary clouds than those formed in the continental airmass. As a consequence clouds that formed in the continental airmass exhibited stronger updrafts, higher cloud tops, greater peak rainwater amounts and heavier precipitation rates.

When the environment can support deeper convective clouds, cloud modeling studies with bin-resolving microphysics by Khain et al. [2004] found that pollution-induced smaller cloud droplets reduce the production of drizzle drops. When these droplets froze, the associated latent heat release resulted in more vigorous convection. In contrast in a clean cloud, drizzle depleted the cloud liquid water so that less latent heat was released when the cloud glaciated, resulting in less vigorous convection. Thus, they found that a squall line developed under continental aerosol conditions and produced more precipitation after two hours whereas this did not happen with clean aerosol conditions. Zhang et al. [2005] came to similar conclusions in their model simulations for different three-week periods over the ARM site in Oklahoma.

Similarly, mesoscale simulations of deep convection over Florida Lynn et al [2005b] showed that higher CCN concentrations delayed the onset of precipitation but led to more intense convective storms with higher peak precipitation rates. However, the accumulated precipitation was largest for the cleaner atmosphere.

In mesoscale simulations of entrainment of Saharan dust into Florida thunderstorms with bin-emulating bulk microphysics, Van den Heever et al. [2006] found that dust not only impacts cloud microphysical processes but also the dynamical characteristics of convective storms. Dust may serve as CCN, GCCN, and IN. The effect of dust on cloud microstructure and storm dynamics in turn alters the accumulated surface precipitation and the radiative properties of anvils. These results suggest that the dynamic structure of the storms is influenced by varying dust concentrations. In particular, the updrafts are consistently stronger and more numerous when Saharan dust is present compared with a clean airmass. Like Seifert and Beheng [2006b] they found that dust results in enhanced glaciation of convective clouds, which then leads to dynamical invigoration of the clouds, larger amounts of processed water, and thereby enhanced rainfall at the ground. However, Van den Heever et
al.’s simulations suggested that rainfall is enhanced by dust ingestion during the first two hours of the formation of deep convective cells, but it is reduced on the ground later in the day. Thus the clean aerosol simulations produced the largest surface rain volume at the end of the day. This is a result of complex dynamical responses of clouds to aerosol changes associated with subcloud evaporation of rain in which low-level cold pools influence storm propagation and to scavenging of dust, so that few GCCN and IN remained late in the day.

Another study that illustrates the complexity of aerosol interactions with convective storms is van den Heever and Cotton’s [2007] examination of the impacts of urban-enhanced aerosol concentrations on convective storm development and precipitation over and downwind of St. Louis, MO. In the van den Heever and Cotton [2007] study RAMS was set up as a cloud-resolving, mesoscale model with both sophisticated land-use processes and aerosol microphysics using a bin-model emulation approach. The results indicate that urban land-use forced convergence downwind of the city, rather than the presence of greater aerosol concentrations, is the dominant control on the locations and amounts of precipitation in the vicinity of an urban complex. Once convection is initiated, urban-enhanced aerosols can exert a significant effect on the dynamics, microphysics and precipitation produced by these storms. The model results indicate, however, that the response to urban-enhanced aerosol depends on the background concentrations of aerosols; a weaker response occurs with increasing background aerosol concentrations.

It was found that when aerosol concentrations were enhanced that cloud water and rain formed more rapidly than in the clean control simulation in which only observed rural aerosol concentrations were utilized. The updrafts were also stronger initially, and the downdrafts developed more quickly. The larger amounts of supercooled liquid water available, together with the stronger updrafts, led to the generation of greater ice mixing ratios earlier in the storm development. Greater amounts of surface precipitation were also produced in this case during the first hour and a quarter to hour and a half of convective storm formation. However, the greater and more rapid production of surface precipitation generates stronger downdrafts and more intense cold pools earlier in the storm lifecycle than in the clean control simulation. This is detrimental to the updraft development and strength, the evidence of which is the earlier demise of the storm closest to the urban region following storm splitting.
In the clean control simulation, the updrafts develop later in association with the delayed hydrometeor development, but they are eventually stronger than those in the simulation in which aerosol concentrations are enhanced. The storms last longer following storm splitting, and new storm development occurs downwind of the city later on in the simulation. This results in increased amounts of accumulated surface precipitation during this time. The variations in storm dynamics in response to variations in aerosol concentrations, result in the greatest accumulated surface precipitation when aerosols are increased early in the afternoon. However as the simulation progresses, this trend reverses, and later in the afternoon, the largest accumulated precipitation occurred in the clean control case. See Figure 7-5. The extremely complex, non-linear relationships between the microphysics and dynamics therefore make it difficult to make absolute statements regarding the impacts of urban-enhanced aerosol on downwind convection and precipitation.

![Figure 7-5: Time series of the accumulated volumetric precipitation in the downwind region expressed as a percentage of the RURAL-Clean background simulation. Rural-L represents the simulation with clean background aerosols; CCN-L represents the affects of increasing CCN concentrations only over a clean background concentration; GCCN-L represents the affects of pollution serving as giant CCN only relative to a clean background; Urban-L represents the affects of pollution acting to enhance CCN and precipitation.](image-url)
It is worth pointing out at this stage that there is very little modeling work on aerosol effects on hail with the exception of Danielsen et al.’s. [1977] 1D parcel model simulations. He concluded that hail formation was sensitive to input CCN and IN concentrations. (We refer to Chapter 2 for discussion on hail.) It is possible that if pollution invigorates storm updraft velocities, as discussed above, then the size of hailstones will likely be increased since hail size correlates with the strength of storm updraft velocities.

In summary, all the above reviewed studies on aerosol impacts on convective storm precipitation illustrate that the ultimate determining factor controlling whether pollution increases or decreases precipitation is the dynamic response of the storm to enhanced or delayed precipitation formation. Moreover, the particular responses depend on the unique properties of the storm environment (i.e. static stability, wind shear, regional convergence) and that many of those complex responses occur through the formation of secondary cells and cold-pool dynamics often many hours following initial convective cell formation.

7.3.5 Orographic Clouds

The influence of aerosol pollution on orographic precipitation has been investigated in three dimensions with bin-emulating microphysics and in two-dimensions with full bin microphysics.

Using the three dimensional version of RAMS with bin-emulating microphysics, Saleeby and Cotton [2005] examined the influence of varying concentrations of CCN and GCCN on simulated wintertime orographic clouds and precipitation over the Park Range of Colorado. They found that higher CCN concentrations lead to the formation of smaller, more numerous droplets and reduced riming consistent with the observations by Borys et al. [2000, 2003] as discussed in Chapter 6. Total precipitation on the ground was therefore reduced. Moreover, with higher CCN concentrations the smaller droplets evaporate more readily when ice crystals grow at the expense of cloud droplets (the Bergeron-Findeisen processes) thus depleting cloud liquid water contents and further decreasing riming growth. [See also Teller and Levin, 2006.] Higher concentrations of GCCN resulted in enhanced surface precipitation when CCN concentrations were high. When CCN concentrations were low, however, higher
GCCN concentrations caused greater vapor competition between ice particles and droplet activation and vapor deposition growth, which led to a reduction in surface precipitation. Those simulations were performed using the Meyers et al. [1992] formula for IN activation. Observations of IN concentrations over the Park Range using a continuous flow diffusion chamber by DeMott et al. [2003b] suggest that the Meyers formula overestimates IN concentrations. Using the lower IN concentrations produced very small changes in cloud mixing ratios, but lowered ice water contents, and reduced surface precipitation. Recent observations in the Arctic basin in the fall by Prenni et al. (2007) also suggest lower IN concentrations than given by the Meyers formula. They found that when these IN observations are inserted in a mesoscale model the predictions of cloud structure are consistent with observations.

Rasmussen et al. [2002] investigated the formation of freezing drizzle in the mesoscale model MM5 coupled to a bin microphysics model fashioned after Reisin et al. [1996a]. They showed that the formation of freezing drizzle is significantly enhanced in conditions of low aerosol concentrations, high LWC (or cloud depth), long in-cloud residence times, and a dearth of ice crystals (< 0.08 L\(^{-1}\)). The latter condition was supported by both model results and observations.

Recently, Lynn et al. [2007] presented results designed to interpret the study of Givati and Rosenfeld [2004] (Chapter 6). Using the Weather and Research Forecast model (WRF) coupled to their bin microphysics model they presented a series of two-dimensional simulations of a Sierra mountain (California) case study, in which they varied CCN input, horizontal wind velocity, and relative humidity to examine the effect of these factors on the nature and distribution of condensate and precipitation relative to the mountain slope. In the control case, a five-fold increase in CCN (at 1% supersaturation) resulted in greater amounts of suspended ice condensate, and significant downwind displacement of this condensate, to the lee of the mountain. The associated decrease in surface precipitation was 27% [consistent with the analysis of precipitation trends in the Sierra Mountains by Givati and Rosenfeld 2004], although there is no observational evidence to support the assumed five-fold increase in CCN. Interestingly, the downwind shift in the surface precipitation associated with the highest topographical feature was small (~ 10 km). Sensitivity studies showed that under high relative humidity the surface precipitation amounts were a factor of about 10 higher than for the control (drier) conditions, and the differences in precipitation between clean and
polluted orographic clouds were negligible. A decrease in horizontal wind speed to 0.75 of that in the control case resulted in very significant reduction in precipitation and a stronger difference between clean and polluted conditions. This study places the aerosol effects on precipitation in the context of some important dynamical controlling factors such as horizontal wind speed and relative humidity. These factors had a stronger effect on precipitation than the (very large) five-fold increase in CCN.

### 7.3.6 Arctic Stratus Clouds

A recent modeling study suggests that Arctic stratus clouds may also be susceptible to aerosol pollution. This should not be too surprising as they have many properties similar to orographic clouds especially low liquid water contents and long lifetimes. Cloud-resolving simulations of Arctic boundary layer clouds using RAMS [Cotton et al., 2003; Saleeby and Cotton, 2004] were carried out for a particular day [Carrió et al., 2005a] and for an entire spring season [Carrió et al., 2005b]. The model was initialized either with clean sub-cloud aerosol concentrations throughout the boundary layer or with observed moderately polluted aerosol concentrations above the inversion and clean below. During the spring season simulations were performed with the model coupled to a sea-ice model. The multi-month simulations were performed using 2-3 daily soundings nudged into the cloud-resolving model to represent daily variations in the synoptic atmosphere. Mixed-phase clouds prevailed during the first two months of simulation, while predominately liquid clouds were simulated during the last month. The effects of IN entrainment in the presence of mixed-phase clouds decreased LWC while ice water paths increased. Even though IN concentrations above the inversion were much lower than estimates from mid-latitudes using the Meyers et al. [1992] formula, the clouds were essentially over-seeded (e.g. ice crystal concentrations were so high that the crystals were very small). As a result crystal fall speeds were reduced, as were precipitation rates. This resulted in longer residence times of the ice particles and increased total condensate paths. This produced enhanced downward longwave radiation. Enhanced albedo associated with enhanced CCN concentrations (hence droplet concentrations) also occurred, but net surface radiation was still higher. As a consequence, sea-ice melting rates were greater. Changes in surface fluxes of the opposite sign were simulated for short time periods when predominately liquid clouds were present, and during the last month. The effect on melting rates associated with the presence of enhanced CCN concentrations above the
boundary layer is opposite although less important than that of IN entrainment. Overall these model simulations suggest that the entrainment of a polluted layer of air overriding an inversion enhances sea-ice melting rates. Melting rates were approximately 4% higher when the air above the boundary layer was polluted than when the entire layer was composed of clean sub-cloud air.

7.3.7 Semi-direct Aerosol Effects

Heating of the air by absorbing aerosols can result in stabilization of a moist moderately stable layer and weaken convection and precipitation. This could have important climatic implications through the hydrological cycle. In regions where there is insufficient moisture or instability to support deep convection, aerosol impacts should be less. This effect of aerosols has been termed the semi-direct effect [Grassl, 1975; Hansen et al., 1997]. The reduction in cloud cover associated with this effect can alter the surface energy budget significantly. If the aerosol contains a large fraction of soot, such as the south Asian haze, then warming of the aerosol layer can desiccate stratocumulus cloud layers and alter the properties of the trade-wind cumulus layer [Ackerman et al., 2000]. The influence of black carbon dominates via its absorption of solar radiation within the atmosphere, which leads to lower surface temperatures [Ramanathan et al., 2001; Lohmann and Feichter, 2001], and reduced outgoing fluxes. General circulation model simulations by Menon et al. [2002b] suggest that black carbon emissions over China may be producing changes in the general circulation, which contribute to the observed increases in summer flooding in south China and drought in north China. Thus, in spite of the fact that anthropogenic aerosol emissions are regionally concentrated, their potential for global impacts is great. Wang [2004] studied the radiative effects of black carbon in a coupled aerosol-climate model and also showed that black carbon modifies precipitation patterns. Interestingly, the model shows an increase in low clouds and a decrease in mid level clouds as a result of black carbon.

Johnson et al. [2004] conducted large eddy simulations (with bulk microphysics) of stratocumulus clouds and imposed a range of heating rate profiles. They found that the location of the absorbing aerosol was important in determining cloud response. When the absorbing aerosol was located within the well mixed boundary layer, cloud LWP and cloud fraction were reduced. However, when the absorbing aerosol resided above the cloud layer,
LWP was increased due to a strengthening of the inversion and reduction in entrainment of dry air. Using the same model, Johnson et al. [2005] also considered the semi-direct effect for trade cumulus clouds with cloud fractions that were either high (0.1-0.5) or low (0.05-0.20). Their results indicated significant semidirect effects for the higher cloud fraction but rather insignificant effects for the low cloud fraction. Note however, that their absorbing aerosol was considered to be a separate class of aerosol from the CCN (external mixture), and fixed in space and time.

In response to Koren et al.’s [2004] observations that smoke in biomass burning regions significantly reduces cloud fraction, Feingold et al. [2005] modeled the semi-direct effect using a LES model with bin microphysics for aerosol and cloud drops. Predictive equations included aerosol advection and therefore heating rates evolved with the boundary layer and cloud fields. (Recall, previous work had imposed fixed heating profiles.) Through a series of numerical experiments, they established that in addition to the importance of the location of the smoke, the reduction in net surface radiation, and therefore land-surface latent and sensible heat fluxes associated with the smoke was likely the most important factor in suppressing cloud formation (see Figure 7-6). This reduction in surface fluxes was also a robust feature in Wang [2004]. Note that the fact that the presence of aerosol, and particularly absorbing aerosol, modifies the sensible and latent heat fluxes is of great importance since it points to possible non-monotonic responses of LWP to aerosol. As aerosol increases, it may act to increase LWP and suppress precipitation, but with increasing aerosol loading a point may be reached where the associated reduction in surface fluxes will act to suppress convection and therefore reduce LWP [Jiang and Feingold, 2006]. This result can be broadened to stress the importance of land-surface characteristics and land management for cloud development. For example, as large tracts of forest are converted to pasture in the Amazon, clouds and precipitation are expected to be modified [Fisch et al., 2004].

### 7.4 Aerosol Effects on Precipitation and The Hydro-Logical Cycle From A GCM Perspective

General circulation models (GCMs) represent the primary tool for simulation of global climate change but they also play an important role in evaluating the regional effects of pollution on modifying precipitation distribution and amounts. By integrating atmospheric,
radiative, oceanic, and land-surface processes on the global scale they provide an indication of expected changes in the coupled system, including possible consequences of coupled greenhouse gas warming and aerosol on the climate system.

**Figure 7-6:** Simulations of smoke aerosol effects on LWP and cloud fraction for warm cumulus clouds, where the smoke is mostly confined to the boundary layer. The control simulation ignores the absorption and heating due to the smoke (no smoke heating). The first sensitivity test includes absorption and heating and shows that LWP and cloud fraction are modified but not significantly reduced. The second sensitivity test (reduced surface flux) tests the importance of the reduction of the surface latent and sensible heat fluxes due to the presence of the absorbing aerosol and shows that this is the primary reason for the reduction in LWP and cloud fraction. Adapted from Feingold et al. [2005].

GCMs have been used to examine the influence on global climate of widespread sources of CCN. First it must be recognized that GCMs have grid spacings of 150 km to 250 km and that clouds often cover a small fraction of the grid-cell area of a GCM, and the average vertical velocities in a grid cell are very small (~0.01 m s\(^{-1}\)) whereas actual cloud-scale vertical velocities are ~1 m s\(^{-1}\) or greater. The poor representation of convection is likely a major source of the disparity in model-predicted LWP amongst GCMs (Figure 7-7). Space-based measurements of LWP are also difficult, although uncertainties in observations are smaller than the inter-model differences in Figure 7-7; nevertheless, models are not well constrained by observations. Model predictions of total precipitation show less variability
reflecting the fact that models are heavily tuned to observations [Stephens et al., 2002] and because the global annual mean amount of precipitation is constrained by the global mean evaporation.

The problem of prediction of condensate is not limited to GCMs; much of the ensuing discussion applies to mesoscale and regional scale models as well. For example, comparisons between total liquid water content (ice plus water) generated by three different cloud schemes in a forecast model show very large variability amongst themselves, and as compared to in-situ aircraft data [Guan et al., 2002]. Similar discrepancies in condensate between single column models (SCMs) were obtained by Menon et al. [2003] in their simulations of clouds observed during the ACE-2 campaign.

Cloud-scale velocities generate the peak supersaturations in clouds, which together with the number of CCN available determine the number of activated drops. It is therefore necessary to estimate cloud-scale vertical velocities to predict cloud drop concentrations in all models that do not resolve these cloud scales. Some modelers have assumed an empirical relationship between predicted sulfate mass concentrations and droplet concentrations [Martin et al., 1994; Boucher and Lohmann, 1995; Kiehl et al., 2000], which is equivalent to assuming there is only a single-value of cloud updraft velocity for all clouds in the model. Others have estimated vertical velocity based on predicted turbulent kinetic energy from boundary layer models [Lohmann et al., 1999]. This is a step in the right direction, but it does not take into account the fact that cloudy updrafts are at the tail of the probability density function (PDF) of vertical velocity. Ghan et al. [1997] and Chuang et al. [1997] assumed a normal distribution of vertical velocity with a mean value given by the GCM grid point mean. They then determined the velocity-weighted mean droplet concentration that takes into account the tails of their assumed PDF of vertical velocity. However, observed PDF's of vertical velocity in clouds in the boundary layer are multimodal and are better represented by double-Gaussian PDF's [Larson et al., 2001] with a mean that is a function of the RMS vertical velocity rather than a GCM grid point mean [Peng et al., 2005]. Moreover, the complications of precipitation or drizzle processes on cloud lifetime, cloud water contents, and cloud radiative properties discussed above cannot be simulated well in GCM cloud parameterization schemes. For example, precipitation processes are non-linear functions of total condensate water contents. As a result, the mean LWC from a GCM model grid-box is essentially meaningless for the representation of precipitation production [Stevens et al.,
A PDF approach to subgrid modeling may be the optimum approach to resolving these deficiencies. PDFs of subgrid quantities, such as vertical velocity and LWP, are determined from prescribed basis functions in which various moments of the basis functions are predicted in the models [Pincus and Klein, 2000; Golaz et al., 2002a,b; Larson et al., 2005].

Another option, though considerably more computationally expensive, is to use what has been called "super-parameterizations" in which cloud resolving models are activated at model grid points [Grabowski et al. 1999; Randall et al., 2003]. These models have the capability of predicting cloud-scale vertical velocities and LWC and thus explicitly representing precipitation processes but have yet to be applied to aerosol effects on precipitation. If the representation of aerosol and clouds in such models, or others using new and innovative techniques for representing subgrid processes can be improved, they should improve predictions of the role of aerosols in changing clouds and precipitation spatial distribution and amounts as well as climate. With these caveats in mind, we will now describe some of the results of GCM simulations of the effects of aerosol on clouds and precipitation.
7.4.1 Aerosol Impacts on Clouds; the Twomey and Cloud Lifetime Effects in GCMs

Aerosol-cloud interactions in GCMs are usually divided into the first indirect effect, (also referred to as Twomey effect [Twomey, 1977] or cloud albedo effect), and the second indirect effect, or “lifetime effect”. The albedo effect is posed for constant cloud water, which is met in GCMs by doing two radiation calculations each timestep. One radiative transfer calculation is conducted with the cloud droplet number concentration that is obtained from present-day aerosol concentrations and one with the cloud droplet number concentration that is obtained from pre-industrial aerosol concentrations keeping all other variables the same.

The cloud lifetime effect is based on Albrecht’s [1989] hypothesis that higher concentrations of cloud droplets resulting from higher concentrations of CCN will decrease droplet size, suppress drizzle, and increase liquid water contents and thereby result in clouds with greater lifetimes. As noted in 7.3.1, however, there is recent evidence from large eddy simulations...
that under certain conditions higher concentrations of cloud droplets can result in lower liquid water content clouds and reduced cloud liquid water path. [Jiang et al. 2006] analyzed hundreds of clouds generated by large eddy simulations, and showed that aerosol effects on cloud lifetime are minimal compared to the variability in cloud lifetime associated with convective variability. Single cloud models even suggest a decrease in cloud lifetime. It is also noted that there are only a few studies (cloud-scale modeling or observational) of cloud lifetime and that the spatial/temporal resolution of GCMs makes them ill-suited to calculation of cloud lifetime. The “cloud lifetime effect” in GCMs is therefore just an attempt to study the effect of autoconversion parameterizations on climate simulations.

There is considerable difference amongst models in predicted forcing due to the albedo effect, and even more significant disagreement in the lifetime effect. The global mean magnitude of the cloud albedo effect from different climate models since pre-industrial times is estimated to be between -0.5 and -1.9 W/m$^2$ from different climate models and the cloud lifetime effect to be between -0.3 and -1.4 W/m$^2$ (Lohmann and Feichter, 2005). The importance of the cloud albedo effect versus the cloud lifetime effect varies even when the models use the same aerosol fields [Penner et al., 2006]. In most simulations shown in Figure 7-8 the total aerosol effect is restricted to warm clouds except for the simulations by Jacobson [2006] and Lohmann and Diehl [2006] who also include aerosol effects on mixed-phase and ice clouds. The radiation effect at the top of the atmosphere ranges from −0.2 W m$^{-2}$ in the combined GCM+satellite simulations [Quaas et al., 2005] to −1.8 W m$^{-2}$ in the simulations by Ming et al. [2007] and Kristjansson [2002] with an average forcing of -1.1 W m$^{-2}$. The total aerosol effect is larger on the Northern Hemisphere than on the Southern Hemisphere in all models (Figure 7-8). The models, however, disagree on the dominance of the indirect effect over ocean or land and also the two combined GCM+satellite simulations provide conflicting answers (Figure 7-8). Although most model estimates also include the direct and semi-direct effects, their contribution to the TOA radiation is generally small compared with the indirect effect ranging from +0.1 to −0.5 W m$^{-2}$ due to variations of the different locations of black carbon with respect to the cloud [Lohmann and Feichter, 2005]. Other differences among the simulations include an empirical relationship between aerosol mass and cloud droplet number concentration versus a mechanistic relationship, the dependence of the indirect aerosol effect on the assumed background aerosol or cloud droplet number concentration, the competition between natural and anthropogenic aerosols as CCN, and differences in the cloud microphysics schemes, especially in the autoconversion rate.
7.4.2 Aerosol Influences on Precipitation in GCMs

Several studies have considered the response of a GCM with a mixed-layer ocean to indirect aerosol effects [Rotstayn et al., 2000; Williams et al., 2001; Rotstayn and Lohmann, 2002] or to a combination of direct and indirect aerosol effects [Feichter et al., 2004; Takemura et al., 2005; Kristjansson et al., 2005]. All of these found a substantial cooling that was strongest in the Northern Hemisphere, with a consequent southward shift of the Intertropical Convergence Zone (ITCZ) and the associated tropical rainfall belt. Rotstayn and Lohmann [2002] suggested that aerosol effects might have contributed to the Sahelian droughts of the 1970s and 1980s. The southward shift of the ITCZ was less pronounced in the Feichter et al. [2004] study than in the other studies, perhaps due to their more complex treatment of cloud droplet nucleation, which tends to give less weight to the effects of sulfate than simpler schemes do. If in turn the Northern Hemisphere is warmed, for instance due to the direct forcing by black carbon aerosols, the ITCZ was found to shift northward [Chung and Seinfeld, 2005].

Global climate model estimates of the change in global mean precipitation due to the total aerosol effects are summarized in Figure 7-9. Consistent with the conflicting results from detailed cloud system studies, the global mean precipitation decreases between 0.1 and 4.6% from pre-industrial times to the present-day. These differences are amplified over the Northern Hemisphere, ranging from +0.7% to -14.3% of their present-day values. The decreases in precipitation are larger when the atmospheric GCMs are coupled to mixed-layer ocean models (green bars), where the sea surface temperature and, hence, evaporation is allowed to vary. When aerosol effects on warm convective clouds are included in addition to their effect on warm stratiform clouds, the overall indirect aerosol effect and the change in surface precipitation can be larger or smaller than if just the aerosol effect on stratiform clouds is considered [Nober et al., 2003; Menon and Rotstayn, 2006]. Besides changes in the distribution of precipitation the frequency of extreme events may also be reduced by the presence of aerosols [Paeth and Feichter, 2006].
Figure 7-8. Global mean total anthropogenic aerosol effect (direct, semi-direct and indirect cloud albedo and lifetime effects) defined as the response in net radiation at the top-of-the-atmosphere from pre-industrial times to present-day and its contribution over the Northern Hemisphere (NH), Southern Hemisphere (SH), over oceans and over land, and the ratio over oceans/land. Red bars refer to anthropogenic sulfate [Easter et al., 2004], green bars refer to anthropogenic sulfate and black carbon [Kristjánsson, 2002*+, blue bars to anthropogenic sulfate and organic carbon [Ming et al., 2007*; Quaas et al., 2004*+, Rotstayn and Liu, 2005*], turquoise bars to anthropogenic sulfate, black, and organic carbon [Menon and Del Genio, 2005; Takemura et al., 2005; Johns et al., 2006, Storelmo et al., 2006], dark purple bars to the mean and standard deviations of anthropogenic sulfate, black, and organic carbon effects on water and ice clouds [Jacobson, 2006; Lohmann and Diehl, 2006], teal bars refer to a combination of GCM and satellite results [ECHAM+POLDER, Lohmann and Lesins, 2002; LMDZ/ECHAM+MODIS, Quaas et al., 2006] and olive bars to the mean plus standard deviation from all simulations. Vertical black lines for individual results refer to ± one standard deviation in case of multiple simulations/results.

*refers to estimates of the aerosol effect deduced from the shortwave radiative flux only.
+ refers to estimates solely from the indirect effects.

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 reduced if the supercooled cloud has more cloud droplets (see also Chapter 6). 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 because of feedbacks within the climate system. The pollution induced increase in aerosol and cloud optical thickness reduces the solar radiation at the surface and causes a cooling that favors precipitation formation via the ice phase [Lohmann, 2004].

7.4.3 Coupling of the Hydrological Cycle and Greenhouse Gas Warming

Greenhouse warming, primarily as a result of enhanced CO₂ concentrations, is amplified when the global hydrological cycle is enhanced and greater amounts of water vapor are evaporated into the air principally over the oceans but also over land. Because water vapor is a much more powerful greenhouse gas than CO₂, the increased amount of water vapor in the air, in turn, results in a strong positive feedback to CO₂ warming. Recent GCM simulations of greenhouse warming and direct and indirect aerosol effects [Liepert et al., 2004] show that in accord with the discuss in Section 7.3.7, the indirect and direct cooling effects of aerosols reduce surface latent and sensible heat transfer (Figure 7-10) and, as a consequence, act to spin-down the hydrological cycle and thereby substantially weaken greenhouse gas warming. This is important since most investigators compare top of the atmosphere radiative differences for greenhouse gas warming and aerosol direct and indirect effects separately. Since greenhouse warming causes a spin-up of the hydrological cycle, and aerosol direct and indirect cooling counteracts the spin-up, the potential influence of aerosols on global climate could be far more significant than previously thought. The simulated decrease in global mean precipitation from pre-industrial times to the present is in contrast to the observed precipitation evaneyslution in the last century and points to an overestimation of aerosol influences on precipitation in these simulations. Nevertheless, this decrease in precipitation changes sign to an increase of about 1% in simulations to 2031–2050 as compared to 1981–2000, because the increased warming due to black carbon and greenhouse gases then dominates over the sulfate cooling [Roeckner et al., 2006].
Figure 7-9. Global mean change in precipitation (present-day – pre-industrial)/present-day in % due to the total anthropogenic aerosol effect (direct, semi-direct and indirect cloud albedo and lifetime effects) from pre-industrial times to present-day and its contribution over the Northern Hemisphere (NH), Southern Hemisphere (SH) and over oceans and over land. Red bars refer to anthropogenic sulfate [Easter et al., 2004]+, blue bars to anthropogenic sulfate and organic carbon [Ming et al., 2007+; Rotstayn and Liu, 2005+], turquoise bars to anthropogenic sulfate, black, and organic carbon [Quaas et al., 2006; Menon and Del Genio, 2005; Takemura et al., 2005; Johns et al., 2006; Storelvmo et al., 2006], dark purple bars to the mean and standard deviations of anthropogenic sulfate, black, and organic carbon effects on water and ice clouds [Jacobson, 2006; Lohmann and Diehl, 2006], teal bars refer to a combination of GCM and satellite results [LMDZ/ECHAM+MODIS, Quaas et al., 2006], green bars refer to results from coupled atmosphere/mixed-layer ocean (MLO) experiments [Feichter et al. [2004] - sulfate, black, and organic carbon; Kristjansson et al. [2005] - sulfate and black carbon; Rotstayn and Lohmann, [2002]+ - sulfate only; Ming et al., personal communication – sulfate and organic carbon) and olive bars to the mean plus standard deviation from all simulations. Vertical black lines refer to ± one standard deviation.

+ refers to estimates solely from the indirect effects

We summarize by stressing that the results of these GCM simulations should not be viewed as quantitative forecasts of the effects of aerosols on patterns and amounts of regional precipitation. As noted previously, there are many uncertainties in the distribution and concentrations of aerosols and condensate. In addition, there are many simplifications in the
models that limit their ability to realistically simulate the indirect effects of aerosols. However, these model simulations demonstrate the potential effects of direct and indirect aerosol forcing on clouds and precipitation in a coupled global system.

Figure 7-10. Simulated changes in (a) evaporation (colors) and net surface solar radiation (contours) and (b) precipitation (colors) and net surface solar radiation (contours). Model results are differences of present-day minus pre-industrial climate simulations with modified anthropogenic aerosol and greenhouse gas concentrations (from Liepert et al. [2004]).

7.5 Summary

- Cloud-resolving models are able to capture the salient features of cloud droplet activation and the timescales of growth to precipitation-sized drops in a reasonable fashion (warm clouds).
- There are still unresolved questions regarding the nature and causes of droplet growth to precipitation embryos. Convective clouds entrain clear environmental air, which results in dilution of the cloud water, modification of cloud buoyancy, and introduction of aerosol particles into the cloud. Cloud models require meter-scale resolution to adequately represent
the details of entraining convective clouds; some of these details may have important consequences for formation of precipitation embryos.

- Size-resolved microphysical treatment of cloud microphysics is the preferred approach to representing aerosol effects on precipitation because it provides more flexibility for representing the tail of precipitation embryos than bulk schemes.

- Uncertainties in collection kernels, especially those involving ice particles, are a major source of uncertainty in modeling of development of precipitation. This limits the gains that can be realized by size resolved microphysical schemes.

- Aerosol effects on precipitation must be considered together with cloud and precipitation effects on aerosol. These include precipitation scavenging, aqueous chemistry, new particle formation and cloud processing. Models that include these processes produce different amounts of precipitation than those that do not. This places strong demands on models because of added computational overhead.

- Evaluation of aerosol effects on clouds and precipitation must be considered in a dynamical/meteorological context. Aerosol-cloud interactions occur in a tightly coupled system and it is important that models couple dynamics, microphysics, radiation, chemistry, and surface characteristics. Some model results suggest that even the sign of aerosol effects on precipitation can change depending on the dynamical context.

- Simulations of individual clouds show differences from a field of interacting clouds, e.g. as a result of cold pool outflow and convergence zones. Historically the cloud modeling community has focused on aerosol effects on single clouds. While these studies are still of great value there is need to pursue more studies of cloud ensembles;

- There are major challenges in modeling precipitation at the appropriate range of spatial/temporal scales, primarily due to limitations in computing power.

- Mesoscale, regional scale and global scale models do not always resolve the appropriate spatial and temporal scales of the physical processes represented, and their use in predicting the effects of aerosol on clouds and precipitation should be regarded with caution. The scaling-up of bulk parameterizations developed for cloud-scale models to regional and global scale models does not have solid scientific basis.

- Because of these uncertainties and limitation in simulating aerosol effects on precipitation, climate model estimates of aerosol-induced changes in precipitation since pre-industrial vary tremendously.
7.6 Recommendations for Future Modeling Studies

To place our understanding of aerosol-precipitation interactions on a firmer footing, we recommend the following for future modeling efforts.

1. Improve model physics

There is a strong need for an improvement in the representation of cloud microphysical processes in models. This includes both the representation of the physics of hydrometeor behavior, hydrometeor interactions (e.g., laboratory studies of collision and coalescence efficiencies for various combinations of hydrometeor types), the parameterization of these processes so that they can be applied in numerical algorithms, and numerical methods. Advances in representation of hydrometeor collection kernels have generally lagged the development of new bin and bulk microphysical methods.

Specific recommendations include:

- Simulations of aerosol-cloud interactions should be done at the appropriate scale so that the development of cloud supersaturations, droplet activation, and growth processes are resolved. This will yield greater confidence in the models’ ability to represent actual cloud responses to variable aerosol inputs. Because of the coarse grid spacing of mesoscale, regional and global models, clouds and cloud properties essentially have to be parameterized as sub-grid-scale processes. Parameterizations at the full range of scales should be pursued.

- Aerosol effects on clouds must be considered within the appropriate dynamical and meteorological context. It is likely that the response of clouds to aerosol will vary depending on input thermodynamic sounding, surface fluxes, subsidence, large-scale advective tendencies, etc. Simulations of these interactions should be extended to larger scales to enable dynamical interactions between neighboring clouds.

- Land surface characteristics determine the surface latent and sensible heat fluxes and drive boundary layer and cloud development. Current and future changes in vegetation and soil moisture are therefore an important consideration. Moreover, because the presence of atmospheric aerosol (particularly absorbing aerosol) modifies the surface radiation budget, it also changes the surface fluxes, and these changes are likely to differ from one land-surface regime to another. Consideration of the coupled aerosol-cloud-land-surface system is therefore of great importance and may have consequences at the range of spatial scales.
• Neither GCMs nor regional-scale models resolve topography sufficiently well to depict cloud-scale vertical motions responsible for the activation of aerosol and the formation of precipitation. Solving this problem goes beyond traditional thinking for cloud parameterization schemes.

• Model evaluation should be performed through a series of case study intercomparisons of modeled and observed aerosol-cloud-precipitation interactions. These should occur at a large range of spatial and temporal scales.

2. Need for model input data

• To model aerosol-cloud-precipitation interactions input data on aerosol that have the potential to act as CCN, GCCN, and IN are needed. Ideally these data should include not only surface measurements but also vertical profiles. Long-term data sets should form the backbone, with intensive field campaigns being supplementary. Where possible, CCN proxies such as accumulation mode aerosol concentration together with composition information should be considered.

• Ice nucleation deserves specific mention. New methods of measuring ice nuclei and analyzing their composition need to be developed. Only a few of the known modes of ice nucleation are represented in current field-deployable devices. The representation of ice formation in cloud models is still rudimentary and relies on a very limited number of measurements and parameterizations.

• Analyze satellite data to establish constraints on aerosol effects on clouds and precipitation. This may require combination of passive radiometry to characterize aerosol and cloud drop sizes, and radar to characterize precipitation. The launch of CLOUDSAT and CALIPSO in the spring of 2006 as part of the A-Train (see Chapter 7) promises to provide important data on aerosol, clouds, and precipitation.

3. Assessment, Attribution of Cause-and-Effect

• Models are particularly useful for providing quantified estimates of parameters such as precipitation amounts for a given simulation. However, the skill level of models in so doing should be quantified statistically. Models should be used to isolate the role of aerosols from the influence of the meteorology and to quantify the relative importance of each process.

• Models should be used to develop metrics to quantify aerosol effects on precipitation efficiency (Chapter 2).
• Models should consider the role of clouds as aerosol scavengers and the scavenging efficiency of clouds and rain. Cloud scavenging, both through coalescence scavenging (7.3.1) and precipitation washout are important aerosol sinks. The links between precipitation scavenging of aerosol and precipitation efficiency should also be explored further. Consideration of these processes should be included in clouds of all scales.

• Identify cause-and-effect. Models are very effective tools for isolating cause-and-effect in a physical system. It is important to correctly ascribe the reasons for agreement or disagreement with observations. As models become more complex, it becomes increasingly difficult to isolate causal relations. Effort should be put into applying existing and new methodologies for quantifying the sensitivity of a given parameter to a change in an input parameter or change in a combination of inputs.
8.1 Introduction

Deliberate cloud seeding, with the goal of increasing precipitation by the injection of specific types of particles into clouds, has been pursued for over 50 years. Efforts to understand the processes involved have led to a significant body of knowledge about clouds and about the effects of the seeding aerosol. A number of projects focused on the statistical evaluation of whether a seeding effect can be distinguished in the presence of considerable natural variability. Both the knowledge gained from these experiments, and the awareness of the limitations in that understanding, are relevant to the general question of aerosol effects on precipitation. Definite proof from the seeding projects for an induced increase in precipitation as a result of the addition of seeding material to the clouds would represent a powerful demonstration of at least one type of dominant aerosol-precipitation link in the clouds involved. Therefore, in this chapter we review the fundamental concepts of cloud seeding and overview the parallels and contrasts between evaluations of deliberate and inadvertent modification of precipitation by aerosols. It is not our intent to provide a comprehensive assessment of the current status of cloud seeding research. We direct the reader to more comprehensive weather modification assessments in NRC (2003), Cotton and Pielke [2007], Silverman [2001, 2003], and Garstang et al. [2005].

Deliberate cloud seeding experiments can be divided into two broad categories: glaciogenic seeding and hygroscopic seeding. Glaciogenic seeding occurs when ice-producing materials (e.g., dry ice [solid CO₂], silver iodide, liquid propane, etc.) are injected into a supercooled cloud for the purpose of stimulating precipitation by the ice particle mechanism (see Section 2.2.2). The underlying hypothesis for glaciogenic seeding is that there is commonly a deficiency of natural ice nuclei and therefore insufficient ice particles for the cloud to produce precipitation as efficiently as it would in the absence of seeding.

The second category of artificial seeding experiments is referred to as hygroscopic seeding. In the past this type of seeding was usually used for rain enhancement from warm clouds [see
Cotton, 1982 for a review of early hygroscopic seeding research]. However, more recently this type of seeding has been applied to mixed phase clouds as well. The goal of this type of seeding is to increase the concentration of collector drops that can grow efficiently into raindrops by collecting smaller droplets and by enhancing the formation of frozen raindrops and graupel particles. This is done by injecting into a cloud (generally at cloud base) large or giant hygroscopic particles (e.g., salt powders) that can grow rapidly by the condensation of water vapor to produce collector drops (see Section 2.2.1).

8.2 Static Glaciogenic Cloud Seeding

Static cloud seeding refers to the use of glaciogenic materials to modify the microstructures of supercooled clouds and precipitation. Many hundreds of such experiments have been carried over the past 50 years or so. Some are operational cloud seeding experiments (many of which are still being carried out around the world) which rarely provide sufficient information to decide whether or not they modified either clouds or precipitation. Others are well designed scientific experiments that provide extensive measurements and modeling studies that permit an assessment of whether artificial seeding modified cloud structures and, if the seeding was randomized, the effects of the seeding on precipitation. While there still is some debate of what constitutes firm “proof” [see NRC, 2003; Garstang et al., 2005] that seeding affects precipitation, generally it is required that both strong physical evidence of appropriate modifications to cloud structures and highly significant statistical evidence be obtained.

8.2.1 Glaciogenic Seeding of Cumulus Clouds

The static seeding concept has been applied to supercooled cumulus clouds and tested in a variety of regions. Two landmark experiments (Israeli I and Israeli II), carried out in Israel, were described in the peer-reviewed literature. The experiments were carried out by researchers at the Hebrew University of Jerusalem (HUJ), hereafter the experimenters. These two experiments were the foundation for the general view that under appropriate conditions, cloud seeding increases precipitation [e.g., N.R.C, 1973, Sax et al., 1975; Tukey et al., 1978a,b; Simpson, 1979; Dennis, 1980; Mason, 1980, 1982; Kerr, 1982; Silverman, 1986; Braham, 1986; Cotton, 1986a,b; Cotton and Pielke, 1992, 1997; Young, 1993].
Nonetheless, reanalysis of those experiments by Rangno and Hobbs [1993] suggested that the appearance of seeding-caused increases in rainfall in the Israel I experiment was due to “lucky draws” or a Type I statistical error. Furthermore, they argued that during Israel II naturally heavy rainfall over a wide region encompassing the north target area gave the appearance that seeding caused increases in rainfall over the north target area. At the same time, lower natural rainfall in the region encompassing the south target area gave the appearance that seeding decreased rainfall over that target area. But this speculation could not explain the positive effect when the north target area was evaluated against the north upwind control area. Details of this controversy can be found in the March 1997 issue of the Journal of Applied Meteorology [Rosenfeld, 1997; Rangno and Hobbs, 1997a; Dennis and Orville, 1997; Rangno and Hobbs, 1997b; Woodley, 1997; Rangno and Hobbs, 1997c; Ben-Zvi, 1997; Rangno and Hobbs, 1997d]. Some of these responses clarified issues; others have left a number of questions unanswered.

It is interesting to note that in the Israeli experiments the effects of artificial seeding with silver iodide appeared to be an increase in the duration of precipitation, with little if any effect on the intensity of precipitation [Gagin, 1986; Gagin and Gabriel, 1987], a finding compatible with the “static” seeding hypothesis.

Givati and Rosenfeld [2005] wrote, "that cloud seeding with silver iodide enhances precipitation especially where the orographic enhancement factor (see Chapter 6) was the largest. Likewise, the pollution effects reduced precipitation by the greatest amount at the same regions”. They suggested that this is because the shallow and short-living orographic clouds are particularly susceptible to such impacts. This suggests that attempts to alter winter precipitation should be concentrated on orographic clouds. Or interpreted in terms of inadvertent modification of clouds; winter orographic clouds may be the most susceptible to precipitation modification by pollution.

This also suggests that the conceptual model on which the Israeli cloud seeding experiments was based is not exactly as postulated. The seeding was originally aimed at the convective clouds that formed over the narrow coastal plain, with the intent of nucleating ice crystals and forming graupel earlier in the cloud lifecycle [Gagin and Neumann, 1974] thus leading to increased rainfall in the catchment basin of the Sea of Galilee to the east of the Galilee Mountains. However, the report of Givati and Rosenfeld [2005] concluded that “cloud
seeding did not enhance the convective precipitation, but rather increased the orographic precipitation on the upwind side of the Mountains, probably by the Bergeron-Findeisen process.

The lack of enhancement of the convective clouds in Israel might be explained by their tendency to mature and dissipate inland during the winter storms. Seeding of mature convective clouds cannot affect them much. The lack of enhancement is also consistent with the microphysically maritime nature of the convective clouds. This appears to be caused mainly due to the natural hygroscopic seeding by sea spray or mineral dust particles coated with soluble material [Levin et al., 1996; 2005] in the winter storms that enhance the warm precipitation [Rosenfeld et al., 2001] as well as promoting the formation of ice hydrometeors that is followed by ice multiplication [Hallett and Mossop, 1974]. These suggestions are supported by the results of glaciogenic cloud seeding in Tasmania, which targeted a hilly area by seeding along an upwind coastline. The seeding in Tasmania was shown to enhance precipitation from the stratiform orographic clouds, but not from the convective clouds [Ryan and King, 1997]. This contradicts the microphysical conclusions of Rangno and Hobbs [1993], who asserted that cloud seeding as done in Israel could not have possibly caused the statistically documented rain enhancement from the convective clouds there.

Recently, Givati and Rosenfeld [2005] carried out a study in which the effects of pollution on rainfall suppression in orographic clouds were separated from the effects of cloud seeding in Israel. They concluded that the two effects have the opposite influence on rainfall, demonstrating the sensitivity of clouds to anthropogenic aerosols of different kinds. By analyzing the rainfall amounts in northern Israel during the last 53 years during days in which no seeding was carried out, they observed a decreasing trend of the orographic factor R₀ (discussed in Chapter 6) with time from the beginning of the study. They associated this decrease with the increase in aerosol pollution. The same trend but shifted upward by 12%-14% was observed for days in which seeding was carried out. Thus, it appears that the opposing effects of air pollution and seeding appear to have nearly canceled each other.

Another noteworthy experiment was carried out in the high plains of the U.S. [High Plains Experiment (HIPLEX-1) Smith et al., 1984]. Analysis of this experiment revealed the important result that after just five minutes, there was no statistically significant difference in the precipitation between seeded and non-seeded clouds, [Mielke et al., 1984]. Cooper and
Lawson [1984] found that while high ice crystals concentrations were produced in the clouds by seeding, the cloud droplet region where the crystals formed evaporated too quickly for the incipient artificially produced ice crystals to grow to appreciable sizes. Instead they formed low density, unrimed aggregates having the water equivalent of only drizzle drops, which were too small to reach the ground before evaporating. Schemenauer and Tsonis [1985] affirmed the findings of Cooper and Lawson in a reanalysis of the HIPLEX data emphasizing their own earlier findings [Isaac et al., 1982] that cloud lifetimes were too short in the HIPLEX domain for seeding to have been effective in the clouds targeted for seeding (i.e., those with tops warmer than −12°C). Although the experiment failed to demonstrate statistically all the hypothesized steps, the problems could be traced to the physical short lifetimes of the clouds [Cooper and Lawson, 1984; Schemenauer and Tsonis, 1985]. This in itself is a significant result that shows the ability of physical measurements and studies to provide an understanding of the underlying processes in each experiment. The results suggested that a more limited window of opportunity exists for precipitation enhancement than was thought previously. Cotton and Pielke [1995] summarized this window of opportunity as being limited to: Clouds that are relatively cold-based and continental; Clouds with top temperatures in the range −10 °C to −25 °C, and a timescale confined to the availability of significant supercooled water before depletion by entrainment and natural precipitation processes.

Today, this window would even be viewed too large, since many cold based continental clouds with tops >−25°C have copious ice particle concentrations [e.g., Hobbs and Rangno, 1985; Rangno and Hobbs, 1988]. The HIPLEX results also indicated that small clouds make little contribution to rainfall.

This begs the question, should we expect a similar “window of effectiveness” for inadvertent IN pollution?

### 8.2.2 Seeding Winter Orographic Clouds

The static mode of cloud seeding has also been applied to orographic clouds. Precipitation enhancement of orographic clouds by cloud seeding has several advantages over cumulus clouds. The clouds are persistent features that produce precipitation even in the absence of
large-scale meteorological disturbances. Much of the precipitation is spatially confined to high mountainous regions thus making it easier to set up dense ground based seeding and observational networks. Moreover, orographic clouds are less susceptible to a ‘time window’ as they are steady clouds so they offer a greater opportunity for successful precipitation enhancement than cumulus clouds. A time window of a different type does exist for orographic clouds, which are related to the time it takes a parcel of air to condense to form supercooled liquid water and ascend to the mountain crest. If winds are weak, then there may be sufficient time for natural precipitation processes to occur efficiently. Stronger winds may not allow efficient natural precipitation processes but seeding may speed up precipitation formation. Stronger winds may not provide enough time for seeded ice crystals to grow to precipitation before being blown over the mountain crest and evaporating in the sinking subsaturated air to the lee of the mountain. A time window related to the ambient winds, however, is much easier to assess in a field setting than the time window in cumulus clouds.

The landmark randomized cloud seeding experiments at Climax, near Fremont Pass, Colorado (referred to as Climax I and Climax II), Colorado, reported by Grant and Mielke [1967] and Mielke et al. [1970; 1971] suggested increases in precipitation of 50% and more on favorable days [e.g. Grant and Mielke, 1967; Mielke et al., 1970, 1971], and the results were widely viewed as demonstrating the efficacy of cloud seeding [e.g., NRC, 1973; Sax et al., 1975; Tukey et al., 1978a,b; American Meteorological Society, 1984], even by those most skeptical of cloud seeding claims [e.g., Mason, 1980; 1982]. Nonetheless, Hobbs and Rangno [1979], Rangno and Hobbs [1987; 1993] question both the randomization techniques and the quality of data collected during those experiments and conclude that the Climax II experiment failed to confirm that precipitation can be increased by cloud seeding in the Colorado Rockies. Even so, in their reanalysis, Rangno and Hobbs [1993] did show that precipitation increased by about 10% in the combined Climax I and II experiments. This should be compared, however, to the original analyses by Grant and Mielke [1967], Grant and Kahan [1974], Grant and Elliott [1974], Mielke et al. [1971], Mielke et al. [1976] and Mielke et al. [1981] that indicated greater than 100% increase in precipitation on seeded days for Climax I and 24% for Climax II.

Two other randomized orographic cloud seeding experiments, the Lake Almanor Experiment [Mooney and Lunn, 1969] and the Bridger Range Experiment (BRE) as reported by Super and Heimbach [1983] and Super [1986] suggested positive results. However, these particular
experiments used high elevation AgI generators, which increase the chance that the AgI plumes get into the supercooled clouds. Moreover, both experiments provided physical measurements that support the statistical results \[Super \text{ and Heimbach, 1983; 1988}\].

There have been a few attempts to use mesoscale models to evaluate cloud seeding programs. \textit{Cotton et al.} \[2006\] applied the Colorado State University Regional Atmospheric Modeling System (RAMS) to the simulation of operational cloud seeding in the central Colorado Mountains in the 2003-2004 winter season. The model included explicit representation of surface generator production of AgI at the locations, burn rates, and times supplied by the seeding operator. Moreover, the model explicitly represented the transport and diffusion of the seeding material, its activation, growth of ice crystals and snow, and precipitation to the surface. Detailed evaluation of model forecast orographic precipitation was performed for 30 selected operational seeding days. It was shown that the model could be a useful forecasting aid in support of the seeding operations. But, the model over-predicted natural precipitation, particularly on moist southwest flow days. The model also exhibited virtually no enhancement in precipitation due to glaciogenic seeding. There are a number of possible causes for the lack of response to seeding, such as over prediction of natural precipitation, which prevented the effects of seeding from being seen. In addition, the background CCN and IN concentrations are unknown, therefore lower CCN concentrations than occurred would make the clouds more efficient in precipitation production, thus reduces seeding effectiveness.

Finally, \textit{Ryan and King} \[1997\] reviewed over 14 cloud seeding experiments covering much of southeastern, western, and central Australia as well as the island of Tasmania. They concluded that static seeding over the plains of Australia is not effective. They argue that for orographic stratiform clouds, there is strong statistical evidence that cloud seeding increased rainfall, perhaps by as much as 30% over Tasmania when cloud top temperatures are between -10 and -12 C in southwesterly airflow. The evidence that cloud seeding had similar effects in orographic clouds over the mainland of southeastern Australia is much weaker. Note that the Tasmanian experiment had both strong statistical and physical measurement components and thus meets or at least comes close to meeting the \textit{NRC} \[2003\] criteria for scientific “proof.” Benefit/cost analysis of the Tasmanian experiments suggests that seeding has a gain of about 13/1. This is viewed as a real gain to hydrologic energy production.
A complication revealed in the analysis of some of the Australian seeding experiments is that precipitation increases were inferred one to three weeks following seeding in several seeding projects [e.g., Bigg and Turton, 1988]. Bigg and Turton [1988] and Bigg [1988; 1990; 1995] suggested that silver iodide seeding can trigger biogenic production of additional ice nuclei. The latter research suggests that fields sprayed with silver iodide release secondary ice nuclei particles at intervals of up to ten days.

In summary, the 'static' mode of cloud seeding has been shown to cause the expected alterations in cloud microstructure including increased concentrations of ice crystals, reductions of supercooled liquid water content, and more rapid production of precipitation elements in both cumuli [Isaac et al., 1982; Cooper and Lawson, 1984] and orographic clouds [Reynolds, 1988; Super and Boe, 1988; Super et al., 1988; Super and Heimbach, 1988; Reynolds and Dennis, 1986]. The documentation of increases in precipitation on the ground due to static seeding of cumuli, however, has been far more elusive with the Israeli experiment [Gagin and Neumann, 1981] providing the strongest evidence that static seeding of cold-based, continental cumuli can cause significant increases of precipitation on the ground. The evidence that orographic clouds can cause significant increases in snowpack is far more compelling, particularly in the more continental and cold-based orographic clouds [Mielke et al., 1981; Super and Heimbach, 1988].

Perhaps, however, the most challenging obstacle to evaluating cloud seeding experiments to enhance precipitation, is the inherent natural variability of precipitation in space and time, and the inability to increase precipitation amounts to better than ~10%. This last obstacle puts great demands on the measuring accuracy and the duration of the experiments. Shouldn’t we expect similar obstacles in evaluating inadvertent effects of IN pollution on precipitation?

8.3 Dynamic Glaciogenic Seeding

So far we have considered only static seeding, in which the principle thrust is to modify the microstructures of clouds generally for the purpose of enhancing precipitation. There is, however, another glaciogenic seeding hypothesis in which the cloud-scale dynamics of a cloud is enhanced by stimulating buoyancy and upward motions of air. This is referred to as dynamic cloud seeding. In principal, this can be done by glaciating convective clouds so that
large quantities of latent heat are released by the freezing of copious liquid water, to invigorate updrafts in the cloud. This can be particularly effective if prior to seeding, the tops of the clouds are restricted by a shallow stable layer produced by a temperature inversion. In this case, the sudden release of a large quantity of latent heat might provide enough buoyancy to push the top of the cloud through the stable layer and into a region where the air is naturally unstable. The cloud might then rise too much greater heights than it would have done naturally.

In a series of randomized experiments carried out in Florida in 1968 and 1970–1973 (called the Florida Area Cumulus Experiment or FACE), it was found that precipitation (measured by radar) from isolated cumulus clouds ~5 km in diameter, which were artificially seeded to induce explosive growth, was about twice that from the unseeded control clouds [e.g., Simpson and Woodley, 1975; Woodley et al., 1982]. The seeded clouds rained more than the control clouds because they were bigger and lasted longer, rather than their rainfall rates being significantly greater.

In FACE II, the attempt was made to confirm and replicate the results of FACE I by going the additional step of specifying the manner in which clouds would respond to seeding based on what was understood to have been the response in FACE I. While there were several suggestions of seeding effects on some clouds and some days [e.g., Woodley et al., 1983], the overall experiment officially failed to confirm the results of FACE I [Flueck et al., 1981; Nickerson, 1979, 1981].

In recent years the dynamic seeding strategy has been applied in Thailand and West Texas. Results from exploratory dynamic seeding experiments over west Texas have been reported by Rosenfeld and Woodley [1989; 1993]. Analysis of the seeding of 183 convective cells suggests that seeding increased the maximum height of the clouds by 7%, the areas of the cells by 43%, the durations by 36%, and the rain volumes of the cells by 130%. The results are encouraging but such small increases in vertical development of the clouds are hardly consistent with earlier exploratory seeding experiments.

As a result of their experience in Texas, Rosenfeld and Woodley [1993] proposed an altered conceptual model of dynamic seeding in which explosive vertical development of seeded clouds is not emphasized. As pointed out by Silverman [2001], however, application of the revised hypothesis in Thailand [Woodley et al., 2003a, 2003b] indicated rainfall
enhancement, but the results did not reach statistical significance. Moreover the enhanced downdraft presumably produced by it did not appear to be delayed [Woodley et al., 1999b].

In summary, the concept of dynamic seeding is a physically plausible hypothesis that offers the opportunity to increase rainfall by much larger amounts than simply enhancing the precipitation efficiency of a cloud. It is a much more complex hypothesis, however, requiring greater quantitative understanding of the behavior of cumulus clouds and their interaction with each other, with the boundary layer, and with larger-scale weather systems. As discussed in Chapter 7, similar complex dynamical responses to pollution aerosols is found in modeling studies, thus requiring greater quantitative understanding of the behavior of cumulus clouds and their interaction with each other, with the boundary layer, and with larger-scale weather systems as well.

8.3 Hygroscopic Cloud Seeding

Hygroscopic seeding was mainly used in the past in warm clouds without any ice in them. However, more recently, this type of seeding methods have been tried in mixed-phase clouds. The aim in seeding warm clouds is to enhance drop growth by coalescence and thus improving the efficiency of rainfall formation. On the other hand, seeding mixed-phase clouds seems to affect both drop growth and ice formation, probably through the efficient formation of graupel particles. Appropriately sized salt particles, water droplets from sprays of either water or saline solution [Bowen, 1952; Biswas and Dennis, 1971; Cotton, 1982; Murty et al., 2000; Silverman and Sukarnjanasat, 2000], and hygroscopic flares [Mather et al., 1997; WMO, 2000] have been used for seeding. Essentially hygroscopic seeding is similar in concept to inadvertently infecting clouds with higher than normal GCCN and recent enthusiasm for the concept was motivated by Mather’s [1991] study showing the effects of paper mill effluent on precipitation. Statistical results, observations and modeling results have provided some evidence that under certain conditions and with optimal seed drop size spectra, precipitation may be enhanced [Farley and Chen, 1975; Rokicki and Young, 1978; Young, 1996, Reisin et al., 1996c; Yin et al., 2000a,b]. Hygroscopic flare particle seeding experiments have provided statistical support for rainfall increases (Figure 8-1) due to seeding based on single cloud analyses [Mather et al., 1997; Bigg, 1997; WMO, 2000; Silverman, 2003]. Model simulations suggest that the increase in rainfall amounts stems from the increase in graupel numbers and masses, which are generated by the
increased concentrations of large drops [Yin et al., 2000a,b]. Such increases could generate more rain but it is not clear how these procedures can affect the clouds for such a length of time as some of the measurements suggest [e.g. Silverman, 2003].

The principle of enhancing the coalescence process via hygroscopic seeding is dependent on three important parameters: the chemistry (hygroscopicity), size and concentration of the particles (CCN) produced from the flares or large particle salt seeding. In addition, the effectiveness of seeding will depend on the natural background particles and their characteristics with regard to the same three parameters. The principle of flare seeding is to have the flares produce effective CCN (usually salts such as sodium chloride, potassium chloride, or calcium chloride) particles in larger sizes (large or giant nuclei) than occur in the natural environment.

![South African vs. Mexico Results](image)

**Figure 8-1.** Quartile values of Radar Derived Rain Mass (Precipitation Flux integrated over time) versus time after “decision time” for seeded and non-seeded cases for the South African (red lines) and Mexican (blue lines) experiments. Rain Mass is the Precipitation Flux integrated over 5-minute intervals. 1st quartile is the value of Rain Mass that is larger than the value for 25% of the storms; 2nd quartile value exceeds the value for 50% of the other storms; and 3rd quartile value exceeds the value for 75% of the storms.

The majority of the hygroscopic cloud seeding flares, including the flares used in South Africa, are based on the formula described by Hindman [1978] that was developed to initiate fog cover for military vessels. The flares used in the South African experiment [Mather et al.,
provide large CCN (>0.3 µm diameter) to a growing cloud, influencing the initial condensation process and allowing fewer CCN to activate to cloud droplets as shown in the modeling study by Cooper et al. [1997]. If the CCN that are introduced into the cloud from the flare are larger in size than the natural CCN, the introduced CCN will activate preferentially over the natural CCN and change the character of the cloud drop size distribution to favor the collision-coalescence process and the formation of rain. The larger artificial CCN inhibit the smaller natural CCN from nucleating, resulting in a broader droplet spectrum at cloud base. The fewer cloud droplets grow to larger sizes, and are often able to start growing by collision-coalescence with other cloud droplets within ~15 minutes, initiating the rain process earlier within a typical cumulus cloud lifetime of ~30 minutes.

Model simulations [Reisin et al., 1996c; Yin et al., 2000a, b; Caro et al., 2002; Segal et al., 2004] suggest that in addition to the size of the hygroscopic particles the time of seeding is a crucial parameter. Seeding too early or too late can lead to reduced rain amounts or to be ineffective at all, respectively. The models suggest that the window of opportunity for effective hygroscopic seeding is small, although it is longer than that of glaciogenic seeding.

In both South Africa [Mather et al., 1997] and Mexico [WMO, 2000], hygroscopic flares have been applied to mixed-phase convective cloud systems in limited physical and statistical experiments. Aircraft microphysical measurements were made to verify some of the processes involved. Radar-measured 30 dBZ volumes produced by the convective complexes were tracked by automated software and various storm and track properties were calculated. These two sets of experiments produced remarkably similar results in terms of the difference in radar-estimated rainfall between the seeded and non-seeded groups. The South African data have been re-evaluated independently by Bigg [1997] and Silverman [2000]; both concluded that there is statistically significant evidence of an increase in radar-estimated rainfall from seeded convective cloud systems.

Mather et al. [1997], Bigg [1997], and Silverman [2000] all allude to apparent dynamic effects of seeding clouds, manifest in the seeded cloud systems being longer-lived. It was speculated that the relation between precipitation loading and evaporation, and the characteristics of the downdraft that is generated, and between the downdraft and the storm organization, evolution, and lifetime, determines the dynamic effect of seeding on rainfall.
Another factor not mentioned is the possible consequences of altered raindrop size distributions. If seeding shifts the raindrop size distribution to smaller raindrops, then greater sub-cloud evaporation would ensue, which would alter cold-pool dynamical effects. If seeding shifts the raindrop spectrum to larger drops the opposite response would be expected [Yin et al., 2001].

The individual storms selected for the experiment, almost without exception, extended well above the freezing level. For the exploratory analyses on the South African data, marked differences were found in storm properties above 6 km [Mather et al., 1997]. The 6 km level generally corresponds to the −5° to −10°C level, which points to probable ice-phase processes being part of the apparent seeding effect. Some indication of how the microphysical changes of broadening the droplet spectrum can be brought about by hygroscopic flare particles, as well as supporting measurements, are given by Cooper et al. [1997], Reisin et al. [1996c] and Yin et al. [2000b]. Although the modeled and observed mixed-phase cloud responses to hygroscopic seeding are not fully understood, they are consistent with earlier modeling and observational studies, which indicated:

- High concentrations of large ice particles or graupel particles are found in maritime clouds (in which collision-coalescence is active) that extend above the freezing level [Koenig, 1963; Cotton, 1972; Koenig and Murray, 1976; Scott and Hobbs, 1977], compared to the relatively low concentrations of large ice particles (graupel) in continental clouds (in which coalescence is not active);
- Freezing temperatures increase with an increase in droplet size, due to the higher probability that larger droplets will contain (or come in contact with) ice nuclei and riming is thereby enhanced [Johnson, 1987]; and,
- Secondary ice particle formation by ice splinter formation during riming is enhanced in the presence of relatively large cloud droplets [Hallett and Mossop, 1974].

It appears that continental convective storms are remarkably sensitive to changes in the CCN ingested at cloud base. For example, both the South African and Mexican experiments with hygroscopic flares show very strong signals in terms of increased storm lifetime in seeded storms, increases in reflectivity aloft, and increases in storm densities. Thus, these hygroscopic flare seeding experiments suggest that it is possible, under appropriate
conditions, to produce large differences in cloud properties by injection of hygroscopic particles into cloud bases.

The calculations of Reisin et al. [1996a,b, c], Cooper et al. [1997], Yin et al. [2000a,b], Caro et al. [2002] and Segal et al. [2004] show that for clouds with maritime cloud droplet spectra, hygroscopic seeding should have no effect, since the collision-coalescence acts very efficiently in such clouds. Relatively high concentrations of large natural CCN prevent the seeded particles from dominating the growth. Model calculations show that for clouds with natural CCN concentrations < 500 cm$^{-3}$ should not respond favorably to hygroscopic seeding.

It has been postulated that the initial spreading of the seeding effect through a cloud occurs through the formation of drizzle drops. The modeling study by Cooper et al. [1997] indicated that the formation of drizzle is highly dependent on the particle size spectra produced by the flares and that the concentrations of drizzle drops produced can vary by several orders of magnitude depending on the particle size spectra. Cooper et al. found that the optimum particle size for seeding with hygroscopic flares in order to obtain higher concentrations of drizzle size drops is ~1 µm diameter. Reisin et al. [1996c], Yin et al. [2000a, b] and Segal et al. [2004] showed that particles larger than 1 µm radius are the most effective for seeding. All these modeling studies also indicate that background CCN (size and concentrations) properties are crucial for determining the effectiveness of the seeded particles because the seeded nuclei compete with the background aerosols for the available water vapor.

The NRC [2003] in their assessment of weather modification research concluded that the South African and Mexican experiments have demonstrated responses in clouds to treatment in accordance with understanding of the chain of physical reactions leading to precipitation. But since the analyzed statistical results are for radar-defined floating targets, they still do not prove that rainfall can be increased by hygroscopic seeding on the ground for specific watersheds. Moreover, since seeding may alter the size-spectrum of raindrops, which alters the radar return, uncertainties, exist in the evaluation of actual rain amounts for seeded versus not-seeded floating targets. Finally, since the main response to seeding found in the South African, Mexican, and Thailand experiments is delayed in time for as much as 1 and 6h,
respectively, following the cessation of seeding, we lack a clear understanding of the actual processes that can lead to such a physical response.

Thus while the results of the hygroscopic seeding experiments are quite promising they still do not constitute a “proof” that hygroscopic seeding can enhance rainfall on the ground over an extended area. The areas affected by cloud seeding remain an open question. In after-the-fact analyses several rain enhancement projects have reported evidence for physical effects outside the area or timing originally designated as the target, or beyond the time interval when seeding effects were anticipated. For example, in recent large particle hygroscopic and glaciogenic seeding trials involving warm-base convective clouds in Thailand and Texas, increases in rain were reported three to 12 hours after seeding was conducted, well beyond the time at which direct effects of seeding were expected and possibly outside the target area. In Project Whitetop the seeding appears to have decreased rain in the area immediately downwind of the seeding release line. This was followed by apparent rainfall increases well downwind in space and time. Does this mean that the scientists misjudged where seeding materials were actually reaching receptive cloud conditions or does it mean that the primary effects of seeding were followed by secondary effects well beyond the original target? Such secondary effects could occur, for instance, if seeding materials become entrained in a downdraft and then are carried outward into the updraft of other clouds. Or are seeding produced cold pools beneath the clouds colder than unseeded clouds thereby forming new cells on the boundaries of the cold pools? In the case of the hygroscopic seeding experiments the postulated dynamic effects due to microphysical and dynamical interactions in the cloud and sub-cloud region and with the environment could result in longer-lived or progeny clouds. Another related uncertainty in seeding convective systems is whether a positive effect on some individual clouds (or cloud complexes) will aggregate to result in increased area rainfall. As more is learned about the global water balance and as new tools enable the cloud scientist to better understand clouds and their response to seeding, the question of extended area affects likely will become better defined and understood.
8.4 Hail Suppression

In addition to studies using cloud seeding principles to enhance rainfall, cloud seeding with aerosols have been used to reduce hail damage. The main concepts for suppressing hail by cloud seeding are as follows:

*The Glaciation Concept*

The aim of hail suppression by glaciation is to introduce so many ice crystals via seeding that the ice crystals consume all the available supercooled liquid water as they grow by vapor deposition and riming of cloud droplets. To be effective this technique requires the insertion of very large amounts of seeding materials in the storm updrafts. Modeling studies [Weickmann, 1964; Dennis and Musil, 1973; English, 1973; Young, 1977] have suggested that unless very large amounts of seeding material are used, the strongest updrafts remain all liquid and hail growth is not substantially affected. Therefore, the glaciation concept is generally thought not to be a feasible approach to hail suppression. The glaciation concept is also not popular because many scientists think that it may result in a reduction in rainfall along with hail. Since most hail-prone areas are semi-arid, the loss of rainfall can have a greater adverse impact on agriculture than economic gains from hail suppression.

*The Embryo Competition Concept*

The competing embryo concept, first introduced by Iribarne and dePena [1962], involves the introduction of modest concentrations of hailstone embryos (on the order of 10 per cubic meter) in the regions of major hailstone growth. The idea is that millimeter-sized ice particles will then compete 'beneficially' for the available supercooled water and result in numerous small hailstones or graupel particles rather than a few large, damaging hailstones. Because it is not economically feasible to introduce hailstone embryos directly in the cloud, one must use a seeding strategy, which utilizes the storm's natural hailstone embryo manufacturing process. For example, the Soviet concept of hail suppression [Bibilashvili et al., 1974] can be considered an embryo competition strategy. In their case the hypothesized hailstone embryos are frozen supercooled raindrops. By dispersing seeding material into a region containing
supercooled raindrops, the raindrops readily freeze and immediately become millimeter-sized hailstone embryos. The numerous hailstone embryos then beneficially compete for the available supercooled water resulting in the formation of numerous small hailstones, many of which would melt before reaching the ground.

Now consider a cloud in which supercooled raindrops are not present. In such clouds millimeter-sized ice particles first must form by vapor deposition until ice crystals of the size of a few hundred micrometers (0.1 mm) form. This takes a significant amount of time, on the order of five to ten minutes. The larger vapor-grown ice crystals can then settle through a population of cloud droplets and grow rapidly by riming those droplets to form graupel particles. The larger ice crystals can also collide with each other to form aggregates. Both the graupel particles and aggregates can serve as hailstone embryos since they have significant fall velocities and cross-sectional areas to enable them to grow rapidly by accreting supercooled cloud droplets to form hailstones.

As a result of the significant amount of time for hailstone embryos to form, seeding intense updrafts, such as exist in supercell storms and the mature cell of severe multicell storms with weak echo regions, is unlikely to have any significant effect on hail growth. The ice crystals formed from seeding would probably be swept aloft into the anvil before becoming large enough to serve as embryos of hailstones. In the case of multicell storms, the recommended approach is to seed in the flanking towering cumulus clouds where updrafts are weaker and transient. If the cell is a daughter cell or a cell that eventually becomes a mature cell, it may be laden with numerous artificially produced hailstone embryos. Likewise, if the flanking cell is in the right location to serve as a feeder cell, then the natural and artificially produced hailstone embryos will be entrained into the mature cell to beneficially compete for the supercooled liquid water and reduce the size of hailstones.

Early Rainout and/or Trajectory Lowering Liquid Water Depletion

The idea behind early rainout is to initiate ice-phase precipitation lower into the feeder or daughter cell clouds where temperatures are in the range -5 to -15C. If the prematurely initiated precipitation settles below an otherwise rain-free base, it could fall into the inflow into the storm and impede the flow of moisture into the storm, which, in turn, would reduce supercooled liquid water contents deeper in the storm. In addition, initiation of ice lower in
the smaller turrets has the potential of reducing supercooled liquid water available for hail growth in the larger turrets, where updrafts are stronger and more conducive to the growth of larger hailstones.

Essentially the same seeding strategy can be used to achieve early rainout and trajectory lowering since if the precipitation initiated lower in the cloud does not precipitate from the rain-free cloud base, the number of hail embryos is increased causing beneficial competition.

Another proposed approach to hail reduction is to seed the base of clouds with hygroscopic materials and thereby initiate a warm rain process in the lower levels of the cloud. The concept behind this trajectory-lowering technique is that the precipitation settling out of the lower part of the cloud will deplete the liquid water in the cloud and therefore limit hailstone growth. It is based on the hygroscopic seeding strategies discussed above in which salt particles or some other hygroscopic material is introduced into the base of flanking clouds, thereby initiating a more vigorous warm rain process in the lower levels of the cloud. Some early cloud modeling studies [Young, 1977] suggest that this may be a feasible approach in regions such as the High Plains of the United States or Canada where cloud base temperatures are cold and cloud droplet concentrations are large.

It has also been proposed to use a combination of salt seeding and ice phase seeding to both deplete supercooled liquid water and to promote beneficial embryo competition [Dennis and Musil, 1973]. Only limited exploratory field studies have been performed to examine the feasibility of this approach.

These concepts are illustrated in Figure 8-2. Developing flanking line cells with weaker updrafts are shown on the left of the figure and the mature cell with strong updrafts on the right. One can also interpret the figure as being a time-height cross section of a storm with zero time on the left and the time of the dissipating cells on the far right. We emphasize here that it is the weaker updraft regions of developing cumulus congestus, rather than the main cumulonimbus cell that is the preferred region for seeding.
In general, operational hail suppression operations follow the embryo competition concept. While operational hail suppression programs are carried out in many countries in Europe, Canada, China, the U.S. and elsewhere, scientific field confirmation of the concepts have been unsuccessful [Atlas, 1977; Federer et al., 1986].

Only Long-term statistical analyses of non-randomized, operational programs have provided more convincing evidence suggested that seeding could significantly reduce hail frequency [Mesinger and Mesinger, 1992; Smith et al., 1997; Eklund et al., 1999; Rudolph et al., 1994; Dessens, 1998]. An advantage of evaluating an operational program is that often one can work with long-period records such as 40 years in Yugoslavia whereas randomized research programs typically cannot get funding for more than five years or so. The disadvantage is that one cannot totally eliminate concerns about natural variability in the climate [see comments in NRC, 2003].

Figure 8-2. Hail suppression concepts. [From World Meteorological Organization, 1996.]
Overall the scientific basis of hail suppression remains not fully resolved. Similarly there is no sound basis on which to base an assessment of the potential impact of pollution hailstorms. Nonetheless owing to the persistence of pollution aerosols and potential for pollution to serve as CCN, GCCN, and IN, pollution may have a more demonstratable effect on hailstorms than deliberate hail suppression.

8.5 Implications of Cloud Seeding Research to Investigations of Aerosol Pollution Effects on Precipitation

In view of the findings summarized in the foregoing sections, it is clear that aerosol from pollution and biomass burning will have certain effects on clouds and on the conditions that govern precipitation development (ice concentrations, drop sizes, etc.). It is less clear whether precipitation on the ground can be significantly altered. However, in drawing parallels between deliberate and inadvertent seeding of clouds some important differences must be kept in mind. In some respects, the situation with inadvertent seeding is more complex, but there are also simplifying factors.

Some of the complicating factors in evaluating the inadvertent effects of aerosol pollution on clouds are as follows:

- At any specific location in the atmosphere the particles that may affect cloud structures and precipitation derive from both natural and anthropogenic sources (biomass burning, industrial, cities, agricultural practices and others). In most cases as opposed to cloud seeding experiments, it is difficult, if not impossible, to determine and differentiate the type of particles (Chapters 2, 3, 6, 7).
- Particles, with different physical and chemical properties, are emitted by most pollutant sources (cf., the variety of emissions from a biomass burning, city, industrial, etc). These various particles may affect clouds in different ways, some of which could be self-canceling (See Chapters 2, 6 and 7).
- In addition to any cloud-active particles emitted directly by an anthropogenic source, various gases (e.g., SO$_2$, NO$_x$) emitted by the same source may be converted into cloud-active particles downwind (e.g., into sulfates and nitrates). One consequence of such gas-to-particle conversion is that an anthropogenic source may affect cloud structures and
precipitation well downwind of the source (see Chapters 2 and 3). This is particularly true once the anthropogenic source material is processed through clouds downwind of the source.

- In general, it is more difficult to carry out controlled randomized experiments using anthropogenic emissions as a source of particles. Alternate methods of evaluation are needed.

On the other hand, there are some advantages in utilizing anthropogenic sources to investigate the modification of clouds and precipitation by particles.

- In some cases concentrations of cloud-active particles emitted by certain industries (e.g., cloud condensation nuclei from paper mills) have been measured. Moreover, these particles are generally emitted continuously, day and night. Consequently, cloud structures and precipitation downwind of such sources may differ significantly from regions upwind of the source.
- In those cases where a new industrial site is established in what was a previously pristine location, the structures of clouds and precipitation in the vicinity of the site can be compared before and after the establishment of the industry. Also, in the case of an isolated industry, it should be possible to characterize rather well the nature of the emissions. However, a complicating factor is the potential for climate variations during the period prior to and after the construction of the new industrial site.
- Variations in pollution output diurnally and for weekdays versus weekends may provide a means of assessing cause and affect

In summary, the history of evaluating the effects of cloud seeding on precipitation has taught us a number of lessons that are useful for evaluating inadvertent modification of clouds and precipitation by aerosols.

The scientific community has established a set of criteria for determining that there is “proof” that seeding has enhanced precipitation. For firm “proof” [see NRC, 2003; Garstang et al., 2005] that seeding affects precipitation, both strong physical evidence of appropriate modifications to cloud structures and highly significant statistical evidence is required. Likewise, for firm “proof” that human production of aerosol in the atmosphere is altering precipitation, both strong physical evidence of appropriate modifications to cloud structures and significant statistical evidence is required. Unfortunately, as noted in Cotton and Pielke (2007) there is a human tendency to accept the results of assessments of inadvertent
modification of clouds and precipitation even though those studies do not meet the standards of “proof” while at the same time requiring that cloud seeding evaluations meet such “proof” criteria.

Another lesson from evaluating cloud seeding experiments is that natural variability of clouds and precipitation can be quite large and thus can inhibit conclusive evaluation of even the best designed statistical experiments. The same can be said for evaluating the effects of inadvertent introduction of aerosols into the climate system. If the signal is not strong, then to evaluate if human activity has produced some observed effect (cause and effect), one requires much longer time records than is available for most if not all data sets. Thus we often do not even have a measure of the natural variability of the climate effecting clouds and precipitation.
CHAPTER 9: SUMMARY

Aerosols impact global climate in a number of ways. First they directly affect the earth’s radiation budget by absorbing and reflecting solar radiation and to a lesser extent altering the profile of IR absorption in the atmosphere. Second, by serving as cloud condensation nuclei (CCN) and ice nuclei (IN) they determine cloud microphysics, the formation of precipitating particles and cloud radiative properties particularly cloud albedo and emission. As a consequence, these properties influence the local radiation budget, atmospheric temperatures, land-surface and ocean temperatures. Aerosols can therefore affect regional cloud properties and may affect precipitation amounts. If these effects are large enough, they could adversely affect water availability, and will impact global climate by changing land-surface radiative properties and surface energy budgets, which can alter the general circulation of the atmosphere, and climate.

This report summarizes our present knowledge of the effects of pollution on clouds and precipitation from the regional to the global scales. Although a great deal of progress has been made in the past few years about our understanding of the effects of pollution on cloud processes, the effects on precipitation on the level of a single storm, on a regional and on a global scales are not yet understood. Below we summarize some of the effects that have been observed and some of the gaps in our knowledge and the difficulties in assessing the effects of pollution on rainfall.

Although our ability to forecast precipitation based on climate models is still limited, we can already conclude that some changes in precipitation would take place. Global climate model estimates of the change in global mean precipitation due to the total aerosol effects vary between 0 and -4.5%. These differences are amplified over land, ranging from -1.5% to -8.5%. Recent coupled ocean mixed layer GCM simulations show that the cooling due to the direct and indirect effects by aerosols reduce surface latent and sensible heat transfer and, as a consequence, act to reduce the intensity of the hydrological cycle. Thus the potential influence of aerosols on climate could be far more significant than previously thought. Estimates by GCMs of the consequences of combined greenhouse gas warming and aerosol cooling in the future depend on uncertain estimates of future pollution emissions and greenhouse gas releases into the atmosphere, as well as the uncertainties associated with the parameterizations of cloud-aerosol interactions in the models.
Specific points

- The mere presence of aerosols in the atmosphere modifies the radiation balance by warming or cooling the atmosphere and thus affecting convection and precipitation.
- Ice plays an important role in precipitation development in many mid-to high latitude clouds. The role of ice in more temperate climates is probably less dominant and precipitation development by the warm cloud processes (condensation and collection) often dominates.
- The role of aerosols in clouds containing both water and ice is complex because precipitation can grow in a number of ways: by the collision-coalescence process, by ice growth by deposition, by the riming of ice particles, or by all of the above.
- There are recent satellites measurements showing positive correlations between cloud cover (cloud fraction) and the loading of scattering aerosol (aerosols from pollution and from natural sources, e.g. dust) but the reasons for this have yet to be established. Similar results from model simulations have been obtained.
- There are indications from both observations and modeling studies that a predominantly absorbing fine particle aerosol can reduce cloud fraction.

The effects of CCN on clouds (Chapters 2, 6 and 7)

- Cloud droplets are formed by nucleation on cloud condensation nuclei from natural or polluted sources.
- The CCN are a subset of the total aerosols present in the atmosphere. Most of the pollution aerosols above a certain size threshold are effective CCN (e.g. sulfate, nitrates and certain organics).
- Many particles that are not very efficient CCN become effective CCN due to various chemical reactions (gas and aqueous phase) that take place during their transport in the atmosphere, especially when they move through clouds.
- Field measurements of CCN and GCCN concentrations in dusty atmospheres exhibit higher concentrations than in corresponding dust-free environments. Moreover, dust particles coated with soluble material (sulfate and sea salt) have been measured as ice residuals further indicating their role in cloud and ice formation in the atmosphere.
• The number of cloud drops formed depends on these aerosol properties and the strength of the updraft. For a given aerosol, stronger updrafts produce higher supersaturations leading to a larger fraction of the aerosols being activated, thus higher supersaturation leads to the activation of more CCN, when they are available.

• Field measurements and modeling studies show that increasing CCN concentrations (from pollution or from natural sources) increases cloud droplet concentrations.

• At given liquid water content the increase in droplet concentrations results in smaller cloud drops.

• Very small concentrations (~ 1/litre) of large or giant CCN (GCCN) can serve as precipitation embryos and significantly affect the growth of precipitation.

**The role of ice nuclei in cloud processes**

• While it is well established that ice particles can have a major role in the formation of precipitation globally, our understanding of the details of the processes are critically short of what would be required for valid representation of these processes in cloud and climate models.

• Ice formation in the atmosphere can occur through nucleation on ice nuclei (IN) ("primary"), freezing of liquid droplets (homogeneous freezing) or by secondary processes (e.g. multiplication processes).

• Ice nuclei are a very small sub-set of atmospheric aerosols and are generally smaller than the CCN concentrations by up to five to six orders of magnitude. The concentration of activated nuclei increases with decreasing temperature and with supersaturation with respect to ice.

• Many measurements suggest that the most effective IN in the atmosphere originate from mineral dust. Some particles from industrial sources and biogenic materials have also been shown to be effective IN.

• There have been many measurements showing that the number of ice crystals in clouds exceeds the concentrations of IN by several orders of magnitude under certain conditions. To explain the discrepancy a number of ice multiplication mechanisms have been proposed.

• Recent observations suggest that in midlatitude frontal clouds the average concentrations of ice crystals seem to be independent of temperature over a large range of temperatures (-5 to -30C).
The effects of aerosol pollution on precipitation

- Obtaining quantitative estimation of the effects of aerosol pollution on precipitation necessitates the design of new methods and new experiments that could separate the aerosol effects from meteorological effects.
- There are only a few observational studies of aerosol effects on precipitation and they lack quantitative estimates of the aerosol and precipitation changes.
- Measurements in the Amazon region show that clouds affected by biomass burning have higher concentrations of cloud drops, narrower drop size spectra and slower development of precipitation embryos.
- When clouds form directly above the concentrated intense plumes of the fires, they developed strong updrafts, taller clouds with small drops, thus delaying precipitation formation.
- Precipitation downwind from paper mills has been shown to increase due to the emission of effective GCCN. Motivated by these studies, GCCN have been used to artificially seed clouds for rain enhancement (see chapter 8).
- Precipitation downwind from sugar cane fires in Australia was found to decrease, however, the role of meteorology in this reduction could not be ruled out.
- Precipitation around urban regions (such as St Louis, New York and Houston) has been shown to increase in some cases, but the connection to increased aerosol pollution could not be established. A more convincing connection was found between changes in precipitation and urban-land use and other meteorological factors.
- Modeling studies of single mixed phase clouds show both decreases and increases in precipitation. Extensive modeling of single mixed phase clouds show that precipitation from clean clouds is most susceptible to aerosol pollution. In these cases tripling the CCN concentrations decreases precipitation amounts by about the same factor. Polluted clouds, on the other hand, are less susceptible to increases in CCN and beyond about 1500 CCN per cm$^3$ (active at 1% supersaturation) very little change in precipitation is shown.
- Both observational and modeling studies suggest that orographic clouds are the most susceptible to modification in precipitation by pollution owing to the modest liquid water contents in them, the relatively short time the drops and ice crystals spend in the clouds and the large areal coverage. Some observational studies suggest that aerosol pollution may
substantially decrease orographic precipitation. Although these studies needs to be validated by more experiments, they have major ramifications in watersheds which depend upon orographic precipitation as the major water resource,

- Recent statistical analysis of the ratio ($R_0$) of orographic rainfall downwind to upwind of polluted sources shows an inverse trend between pollution and rainfall on the upslope of the mountains. Namely, increases in pollution leads to a decrease in $R_0$. Smaller decreases were reported at the mountain top and a small increase on the downwind side of mountains.
- A few giant CCN per liter of air enhance precipitation development. These effects are more pronounced in polluted clouds where the GCCN grow rapidly to precipitation size particles, both large drops and graupel. Model simulations suggest that the increase in precipitation due to GCCN pollution only moderates the larger decrease in rainfall due to small pollution aerosols.
- Models show that since pollution alters droplet distributions, this can influence glaciation and precipitation development, but only limited observational studies have quantified this process at this point.
- Models also show that the changes in cloud and precipitation development due to pollution have effects on the dynamics of clouds. Enhanced downdrafts create cold pools below clouds, which in some cases can transform individual clouds into long-lived squall lines with enhanced precipitation and in other cases lead to weakening of storms and reduced precipitation.
- Some models of cloud fields and some cloud seeding experiments suggest that the response of the clouds to aerosol changes may take place over periods of several hours and multiple cloud lifecycles, the reasons for these lasting effects are not clearly understood.

**Gaps in our knowledge**

- Ice formation in clouds is still not fully understood.
- The impact of pollution on ice formation has not been clearly demonstrated.
- Although models suggest that GCCN enhance precipitation formation, there are relatively few observations that confirm the enhanced effects on precipitation on the ground.
- Evaluation of the effects of pollution or aerosols from natural sources on precipitation has to rely on large amounts of data. Although case studies of individual storms can shed some light on the role of various processes, they cannot determine the long-term climatological
effects of pollution aerosol on precipitation. At the same time, long term measurements also suffer from difficulties stemming from climatic fluctuations that take place during the experiment. Researchers in weather modification have for years faced similar problems in their evaluation of cloud seeding experiments.

- A method to assess the effect of aerosols on precipitation can include direct measurements of the correlation between precipitation efficiency and aerosol concentration, whether from the surface, aircraft or satellites. Precipitation efficiency involves not only the measurement of surface precipitation but also horizontal water vapor flux and vertical air motion within clouds. Even though measurements and model simulations show the strong pollution interference with cloud microphysics, the magnitude or direction of the effect on precipitation and its variations with meteorological conditions have not been clearly identified.

- The expected changes in precipitation due to enhanced pollution or biomass burning could be smaller than the errors of the measurements. Therefore, in order to increase signal to noise ratio, the measurements of all variables must be accurate enough to be able to resolve the expected changes. Unfortunately, many of the instruments listed in Chapter 5 fall short of this requirement.

**The limited accuracy of the measurements:**

- Radar measurements of precipitation can be off by up to 100%.
- Retrievals of the cloud droplet size (effective radius) from satellites are biased by up to 30% for convective clouds, due to departure of the clouds from the plane parallel model. However we can still measure the relative change of cloud properties with change in aerosol concentrations. Moreover, retrievals of effective radius are biased heavily by the presence of ice in a pixel. Therefore, effective radii are not necessarily representative of droplet sizes that affect precipitation formation.
- Aerosol column optical thicknesses measured from satellites are not yet accurate enough (±20% over the oceans, less accurate over vegetated land or over the deserts).
- Ground based remote sensing of aerosol can have a precision of ±0.02 in units of optical depth.
- Translating the optical thickness into the column aerosol concentration and into CCN is difficult due to variation in the aerosol size distribution and chemical composition, water
uptake and the vertical distribution of the aerosols. These are critical parameters needed for interpreting aerosols ingestion and effects on clouds.

- In situ measurements from aircraft have limitations in correctly sampling all aerosol sizes. In addition, such measurements only represent small volumes that may not always correctly represent the processes that affect clouds. Ground based measurements may not always represent the properties of the aerosols and the clouds above.

- Cloud properties have a diurnal cycle not well represented by polar satellite observations. Geostationary satellite measure the diurnal cycle but with limited ability to derive the cloud droplet size.

- Once a statistical connection is established between the aerosol concentration, the cloud properties, the large scale water vapor flux, and precipitation, it is still difficult to assess whether the changes in the precipitation patterns are a results of the aerosols impact, whether it is changing of the precipitation patterns that affect the aerosol properties or that both phenomena are concomitant, reflecting properties of the air mass.

- Most informative measurements so far on the effect of aerosols on precipitation do not include simultaneous quantitative measurements of both aerosols and precipitation.

- Because of the coarse grid spacing of GCMs, clouds and cloud properties essentially have to be parameterized as sub-grid-scale processes. The challenge is to parameterize aerosol activation, the influence of aerosols on precipitation processes, the precipitation processes and their feedbacks on cloud and storm dynamics, and cloud processing of aerosols all in sub-grid-scale parameterizations.
CHAPTER 10: RECOMMENDATIONS

- It is recommended that a series of international projects targeted toward unraveling the complex interactions among aerosols, clouds, and precipitation be implemented. A series of international workshops and field studies are needed to address the impacts on clouds and precipitation of aerosols from a range of sources including biomass burning, dust, and industrial pollution within different regional weather regimes in the tropics, middle latitudes, and the Polar Regions. All of these studies need to have an adequate scientific component, but should also deal with specific situations occurring in the developing world by addressing the impact of pollution/precipitation on economic, sociological, environmental and health issues.

- It is recommended that WMO/IUGG take the lead in such projects together with other UN and International Organizations. Some of those projects could be sponsored and financially supported by the countries involved. For example, the effects of an evolving industrial economy such as China on precipitation should be studied. Similarly, a study of the effects of biomass burning and dust in some of the African regions would be highly valuable for our understanding of these complex issues.

- It is also proposed to consider a project in a country where the relationship between pollution and precipitation can be explored with the biggest arsenal of the newest scientific instrumentation. The WMO/IUGG can play a key coordination role in making sure the following recommendations are implemented.

a) Better characterization of aerosols

- Emission inventories
  
- To be useful as a reference for studies of pollution effects on cloud and precipitation, size-resolved measurements of the numbers of particles emitted is required for the various primary aerosol sources, especially sea salt, mineral dust, pyrogenic aerosol and primary biogenic aerosol.

- The abundance and rate of production of submicron sea salt particles is still under dispute. Since these particles can play a significant role in cloud microphysics and precipitation in remote marine regions, observational constraints to their source parameterizations are urgently needed.
• Due to the complexity of carbonaceous aerosol composition and emission processes, a critical concern is to define an adequate classification scheme for these aerosols based on observable characteristics.
• There is a need to build a reliable inventory of emission parameterization of carbonaceous aerosols (BC, PPOC and SOA) that includes size distributions and number fluxes.
• In building an inventory of emission the injection height should be included since changes in injection height due to shifts in agricultural or technical practices will have important consequences on the lifetime and fate of particles.

❖ Chemical processes, physical properties and instrumentation

• Accurate knowledge of the rates of the processes leading from SO$_2$ to sulfate aerosols is needed for clear and cloudy atmospheres.
• The ability of different types of particles (e.g. mineral dust particles, biomass smoke, biogenic aerosols, carbonaceous aerosols) to act as CCN, GCCN, and IN as a function of aerosol size, origin, and air mass history needs to be determined.
• In order to assess the human impact on cloud physics, we need to know the aerosol, CCN and IN distributions in the pre-anthropogenic atmosphere. Novel methods of inferring or backing-out pre-industrial cloud nucleating aerosol distributions are needed. The sources, characteristics, and fluxes of natural aerosol types must be investigated by carefully conducted field and laboratory studies.
• New instruments should be developed and measurements should be carried out to determine CCN, GCCN and IN concentrations as a function of particle size, composition and supersaturation. Uncertainties are greatest for IN because only a few of the known modes of ice nucleation are represented in current field-deployable devices. These measurements should be conducted in the context of closure studies.
• A program for validation and intercomparison of CCN, GCCN, and IN measurements should be carried out. This should be in the form of workshops in which instruments are compared using the same air samples.
• Global coordination of observational networks is needed for more complete coverage of global aerosols. These could combine in situ and ground-based remote sensing methods (e.g. AERONET).
Continuity of the observations and suitable locations of the observation sites (depending on the scale of phenomena) is very important. Aerosol type classification methods using depolarization ratio, wavelength dependence, and the lidar ratio (if available) are useful for the comparison.

More accurate assessment from satellites of the aerosol column concentration, vertical and size distribution, refractive index and single scattering albedo is needed.

It is important to extend the measurements to small (100 nm particle diameter) particles in the retrievals, since many of these could be efficient CCN.

b) The effects on clouds and precipitation:

Vertical profiles of cloud microphysical processes, size, composition and phase of hydrometeors should be carried out by profiling from aircraft and from satellites and ground measurements. This is especially important in the tropics in areas affected by biomass burning and highly populated areas. Such data could be used to validate model simulations and to shed more light on the main processes leading to precipitation under clear and polluted conditions.

More accurate measurements of cloud droplet size, liquid water path and total condensate through remote sensing should be developed.

Most optical spectrometers today cannot accurately distinguish small droplets from ice crystals. In light of recent measurements showing high concentrations of very small (<50 \(\mu\)m) ice crystals it is recommended that new in situ instruments for measuring very small (5-30 micrometer diameter) ice crystals be developed.

Experiments should be designed to better understand the role of ice in precipitation development and the connection between IN and ice crystals in clouds.

Multi-year measurements from space of precipitation patterns along with retrievals of cloud nucleating aerosols are needed to assess both regional and global impacts of aerosol pollution on precipitation. These space-born studies should be supplemented with intensive field campaigns in which coordinated detailed observations of clouds and aerosols are obtained.

Because both observational and modeling studies suggest that orographic clouds are the most susceptible to precipitation modifications due to pollution aerosols, and because the ramifications on water resources are large, it is recommended that focused, coordinated observational and modeling campaigns be implemented to study aerosol precipitation
interactions for selected watersheds where pollution effects are likely to be large. Examples include the 2006 Suppression of Precipitation (SUPRECIP) study in California, the planned Convective and Orographically-induced Precipitation Study (COPS) experiment in Germany in 2007 and the Colorado River Basin in the western U.S. It is recommended that similar studies be conducted in regions such as China and India, where the effects of mega cities on precipitation could be evaluated. These studies should be basin wide and include ground-based, airborne, and satellite measurements of aerosol properties, cloud microphysical measurements, surface and radar (particularly cloud radar) measurements of cloud structures and precipitation, basin-wide hydrological measurements of liquid precipitation, snowpack and runoff, etc.

- Because of the indications of strong influences of urban land-use and aerosol pollution on precipitation and lightning, it is recommended that a coordinated modeling and observational campaign be established in a number of large metropolitan regions where convection is prevalent. Crucial to site selection is the importance of finding locations where the local physiography (i.e. topography, land-water interfaces) is relatively simple and where generally widespread aerosol pollution is minimal. Measurements should include the list given above plus greater attention to surface measurements of fluxes and land-use properties.

- Because cloud-resolving models suggest a significant dynamic response of cumulus clouds to pollution aerosols which then modifies simulated precipitation, particularly through secondary dynamic responses of clouds via cold pools and gravity waves, it is recommended that a coordinated observational and modeling campaign be organized to investigate the response of cloud systems to varying amounts and characteristics of pollution. These studies should include ground-based, airborne, and satellite measurements of aerosol properties, cloud microphysics measurements, surface mesonets, ground-based multiple Doppler radars, and airborne Doppler radars. These campaigns should be done in regions of significant biomass burning or urban aerosol emissions, and in other locations where the meteorology varies such that different cloud responses can be expected.

- Modeling studies that extend beyond single-cloud responses but consider the cloud ensemble response to variations in cloud nucleating aerosols are needed. These types of simulations should be combined with a number of runs using random variations in initial conditions, in order to determine the statistical variations of the final product, namely the precipitation on the ground.
• Models are particularly useful for providing quantified estimates of parameters such as precipitation amounts for a given simulation. However, the skill level of models in so doing should be quantified statistically.

• Models should be used to isolate the role of aerosols from the influence of the meteorology and to give a quantitative answer as to the relative importance of each process.

• Models are very effective tools for isolating cause-and-effect in a physical system. As models become more complex, it becomes increasingly difficult to isolate causal relations. Effort should be put into new methodologies for quantifying the sensitivity and response of a parameter such as precipitation to changes in input parameters such as pollution or changes in a combination of inputs. Possibly the method of factor separation should be more frequently used to determine the relative importance of each parameter.

• Because of the coarse grid spacing of GCMs, clouds and cloud properties essentially have to be parameterized as sub-grid-scale processes. The challenge is to parameterize aerosol activation, the influence of aerosols on cloud microphysics and precipitation processes, the precipitation processes and their feedbacks on cloud and storm dynamics, and cloud processing of aerosols all in sub-grid-scale parameterizations. Clearly there is a need to develop new and novel methods of parameterizing aerosol interactions with clouds in large-scale models. It is urged that aerosol-cloud-precipitation specialists in observations and modeling collaborate with GCM developers to refine the representation (parameterization) of aerosol-cloud-precipitation processes in GCMs. Such a strategy of combining observations, cloud resolving models and climate models has been put forward in the GEWEX Cloud System Study context.

c) Immediate action items:

• It is recommended that aerosol monitoring stations be equipped with instruments for measuring CCN activation spectra in order to be able to establish empirical relationships between measured aerosol properties and their potential effects on clouds.

• A Workshop should be held on developing improved instrumentation for measuring IN, small ice particles, and precipitation. Because of the complexities involved, it might be necessary to develop an overall strategy to address these issues.

• A Workshop should be held to determine a strategy for a focused, coordinated observational and modeling campaign that could address the effects that aerosols have on
orographic clouds. This might include additions to existing planned measurement campaigns or a new campaign(s).

- It is recommended that aerosol-cloud-precipitation cloud-resolving-model intercomparison studies be implemented to serve as a stimulus for model refinement research and evaluations of model performance. These should be done for a variety of regimes including warm and cold-season orographic cloud models, shallow convection and deep convection, and stratiform cloud systems. These studies could be linked to some of the campaigns suggested above or to other research campaigns where implementation of an aerosol component could be accomplished easily and economically.

- There is a strong need to assemble data sets or climatology that can be used for the assessment of climate simulations with GCMs including aerosol-cloud-precipitation parameterizations. Some of the parameters that need evaluation include seasonal and annual precipitation amounts, regional precipitation climatology, global and regional aerosol distributions, and TOA radiation budgets. This could be done in the context of GCM intercomparison studies much like that recommended for cloud resolving models.

- It is recommended that existing and new statistical methods be applied to current and future data sets to distinguish aerosol effects on precipitation from meteorological influences. Numerical models should play an integral role in this process where they can serve as “predictors” or “covariate” in the statistical evaluations.

- Finally it is recommended that the WMO creates a web site to serve as central index to archives of public domain measurements of aerosol, cloud, and precipitation parameters. Individual agencies/countries are to be encouraged to maintain long term web-accessible archives of research (including short-term field experiments) and operational measurements. Archives are encouraged to include key measured instrument parameters as well as retrieved parameters.
APPENDIX: LIST OF ABBREVIATIONS USED IN THE REPORT

ABC - Atmospheric Brown Cloud
ACE - Aerosol Characterization Experiments
ACE-2 CLOUDYCOLUMN - CLOUDYCOLUMN is one of ACE 2 projects, which took place in June-July 1997, between Portugal and the Canary Islands
AD-Net - Asian Dust Network
AEROCE – Atmospheric Ocean Chemistry Experiment
AEROCOM – Aerosol model intercomparison
AEROSIBNET – Siberian system for aerosol research - [http://sibrad.iao.ru/]
AGCM- atmospheric general circulation model
AMS-American Meteorological Society
AOD - Aerosol optical depth
APS - Aerosol Mass Spectrometer
ARM – Atmospheric Radiation Measurement [www.arm.gov/]
ASTEX - Atlantic Stratocumulus Transition Experiment
ATSR – Along Track Scanning Radiometer
AVHRR - Advanced Very High Resolution Radiometer
BASE-
BC- Black Carbon
BRE - Bridger Range Experiment
BSRN - Baseline Surface Radiation Network/World Radiation Monitoring Center
CALIPSO - Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation
CAPS - Cloud Aerosol Precipitation Spectrometer
CASTnet -Clean Air Status and Trends Network
CCM3- Community Climate Model
CCN-Cloud Condensation Nuclei
CDNC - Cloud droplet number concentration
CDP - Cloud Droplet Probe
CFDC - Continuous-flow diffusion chamber
CFDE (I and III)- Canadian Atlantic Storms Program (I and III)
CFS - Continuous flow spectrometers
CGMS - Coordination Group for Meteorological Satellites
CIS-LiNet - Commonwealth of Independent States Lidar Network
Climax – Cloud seeding experiment for rain enhancement near Climax Colorado
CLIVAR – Climate Variability and Predictability
CMAP - NOAA Climate Predication Center’s Merged Analysis of Precipitation
CMDL - Climate Monitoring & Diagnostics Laboratory.
CN-Condensation Nuclei
COSAM- Comparison of Large Scale Sulfate Aerosol Models
CPC - Condensation particle counter
CVI - Counterflow Virtual Impactor
DISORT – Discrete Ordinate Radiative Transfer Model
DMPS - Differential mobility particle sizer
DWD - German Weather Service
DYCOMS-II experiment
EARLINET - European Aerosol Research Lidar Network
EC- Elemental Carbon
ECHAM5- A general circulation model of the atmosphere, developed at the Max Planck Institute, Hamburg.

EF- emission factors
EI - Electron impact ionization
EMEP- European Monitoring and Evaluation Program
EPIC - Eastern Pacific Investigation of Climate Experiment
EUROTRAC - Project on the Transport and Chemical Transformation of Environmentally Relevant Trace Constituents in the Troposphere over Europe

FACE - Florida Area Cumulus Experiment
FAO- Food and Agriculture Organization of the United Nations
FAR - False Alarm Rate
FIRE- First ISCCP Regional Experiment
FPD - Flame photometric detectors
GAW - Global Atmospheric Watch
GAW-PFR – Global Atmospheric Watch Precision Filter Radiometer Network
GCCN- Giant CCN
GCM- Global Climate Model
GEIA- Global emission Inventory
GLAS - Geoscience Laser Altimeter System
GOES - Geostationary Operational Environmental Satellites
GPCP - Global Precipitation Climatology Project
GPI - GOES Precipitation Index algorithm
HIPEX - High Plains Experiment
HSRL - High-spectral resolution lidar
HTDMA - Humidified Tandem Differential Mobility Analyzers
HUJ - The Hebrew University of Jerusalem
HULIS- Humic-like substances
IC - Ion chromatography
ICESat - Ice, Cloud, and land Elevation Satellite
ICP-MS - Inductively coupled plasma-mass spectrometry
IDDI - Infrared Difference Dust Index
IFA - intensive flux array
IMAGES (Intermediate Model of Global Evolution of Species) - SCAPE2
IMPROVE- Interagency Monitoring of Protected Visual Environments
INAA - Instrumental Neutron Activation Analysis
INDOEX – Indian Ocean Experiment
IN-Ice Nuclei
IPCC-TAR- Intergovernmental Panel for Climate Change
IPWG - International Precipitation Working Group
ISCCP - International Satellite Cloud Climatology Project
ITCZ- Inter Tropical Convergence Zone
KNMI - Koninklijk Nederlands Meteorologisch Instituut - http://www.knmi.nl
LAMMA - Laser Microprobe Mass Analysis
LBA-SM OCC - Large Scale Biosphere-Atmosphere Experiment in Amazonia-Smoke, Aerosols, Clouds, Rainfall and Climate
LCF- Longwave cloud forcing
LCL- Lifting Condensation Level
LITE- Lidar In-Space Technology Experiment
LMDzT INCA – A model, coupling the general circulation model LMDzT of the Laboratoire de Météorologie Dynamique to the INCA (Interaction with chemistry and aerosols) module of CEA-CNRS,
LWC- Liquid Water Content
LWP – Liquid Water Path
MAAP - Multi Angle Absorption Photometer
MAST - Monterey Area Ship Track experiment
METEOSAT – Meteorological Satellite
METROMEX - An Investigation of Inadvertent Weather Modification downwind of St Louis.
MFO - Low-grade marine fuel oil
MFRSR - Multi-filter Rotating Shadow-band Radiometer
MISR - Multi-angle Imaging SpectroRadiometer
MMD- Mean Mass Diameter
MODIS - MODerate resolution Imaging Spectroradiometer
MOUDI - Micro Orifice Uniform Deposit Impactor
MPLNET - Micro-Pulse Lidar Network of NASA
MSG - Meteosat Second Generation
NAPS - National Air Pollution Surveillance
ND- Not Determined
NDF - U.S. Navy distillate fuel
NIES - National Institute for Environmental Studies
NOAA PATMOS - A project to derive atmospheric and surface climate records from AVHRR data.
NPOESS - National Polar-orbiting Operational Environmental Satellite System
NRC – National Research Council
NWP - Numerical Weather Prediction model
OPC - optical particle counter
PARASOL - Polarization and Anisotropy of Reflectances for Atmospheric Sciences coupled with Observations from a Lidar
PBAP- Primary biogenic aerosol particles
PBL- Planetary boundary layer
PCASP - Passive Cavity Aerosol Spectrometer Probe
PCTE - Polycarbonate Track Etched
PDF - Probability density function
PFR - Precision Filter Radiometers
PHOTONS - Sunphotometer network operated by the University of Lille - www-loa.univ-lille1.fr/photons
PILS - Particle-into-liquid sampler
PMOD/WRC - Physikalisch-Meteorologisches Observatorium Davos.
PMS FSSP – Particle Measuring System- Forward scattering spectrometer probe.
POC- Particulate organic carbon
POD - Probability of Detection
POLDER - Polarization and Directionality of the Earth Reflectances
POM- Particulate organic matter
PPOC- Primary particulate organic carbon
PPOM- Primary particulate organic matter
PSA – Particle surface area
PSAP - Particle Soot Photometer
QMS - Quadrupole mass spectrometry
RAMS - Colorado State University Regional Atmospheric Modeling System
RCFM- Reconstructed fine mass
REALM - Regional East Atmospheric Lidar Mesonet
RH- Relative Humidity
SCAR - Sulfates, Clouds and Radiation Experiments (A in America and B in Brazil)
SDCC - Static thermal gradient diffusion cloud chamber
SFU - Stacked Filter Units
SKYNET - Atmospheric radiation and weather observation network in Japan -
   http://atmos.cr.chiba-u.ac.jp/
SMPS - Scanning mobility particle spectrometer
SOA- Secondary organic aerosols
SOCEX - Southern Ocean Cloud Experiment
SPOT- Satellite Pour l'Observation de la Terre
SRES- IPCC Special Report on Emissions Scenarios
SURFRAD - Surface Radiation Network
SVOC- Semi-volatile organic Carbon
TEOM - Tapered Element Oscillating Microbalance
TIM - TRMM Microwave Imager
TM-3- 3-D Transport Model
TOGA COARE - Tropical Ocean Global Atmosphere Coupled Ocean Atmosphere Response
   Experiment
TOMS-Total Ozone Measuring Spectrometer
TRACE-P - Transport and Chemical Evolution over the Pacific
TRMM - Tropical Rainfall Measuring Mission
USDA - United States Department of Agriculture
VIIRS - Visible/Infrared Imager/Radiometer Suite
VMD - Volume medium diameter
VOC- volatile organic compounds
WINSIC- water insoluble inorganic carbon
WINSOC- water soluble organic compounds
WORCC - World Optical Depth Research and Calibration Center
WRC - World Radiation Center
WSOC- water soluble organic compounds
XRF - X-Ray Fluorescence
ZSR- Zdanovskii-Stokes-Robinson
REFERENCES


Boers, R., G. P. Ayers, and J. L. Gras, Coherence between seasonal cycles in satellite observed cloud optical depth ad boundary layer CCN concentration at a mid-latitude Southern Hemispheric site, Tellus, **46B**, 123 – 131, 1994.


Campuzano-Jost, P., C. D. Clark, H. Maring, D. S. Covert, S. Howell, V. Kapustin, K. A. Clarke, E. S. Saltzman, and A. J. Hynes, Near-real-time measurement of sea-salt aerosol during the SEAS campaign: Comparison of emission-based sodium


Chiapello, I., et ., Understanding the long-term variability of African dust transport across the Atlantic as recorded in both Barbados surface concentrations and large-scale


GCOS-107, Systematic Observation requirements for satellite-based products for climate, Global Climate Observing Systems, WMO/TD No. 1338, 17-26, 2006


Gérard, B., J.-L. Deuzé, M. Herman, Y. J. Kaufman, P. Lallart, C. Oudard, L. A. Remer, B. Roger, B. Six, and D. Tanré, Comparisons between POLDER 2 and MODIS/Terra


Haywood, J., P. Francis, O. Dubovik, M. Glew, and B. Holben, Comparison of aerosol size distributions, radiative properties, and optical depths determined by aircraft


Khain, A., A. Pokrovsky, M. Pinsky, A. Seifert, and V. Phillips, Simulation of effects of atmospheric aerosols on deep turbulent convective clouds using a spectral


Market, P., S. Allen, R. Scofield, R. Kuligowski, and A Gruber, Precipitation efficiency of

Marple V.A., K.L. Rubow and S.M. Behm, "A Microorifice Uniform Deposit Impactor

Marshak, A., Y. Knyazikhin, K. Evans, and W. Wiscombe, The “RED versus NIR” plane

Laboratory simulations and parameterization of the primary marine aerosol

Martin, G.M., D.W. Johnson, and A. Spice, The measurement and parameterization of

Martin, G.M., D.W. Johnson, and A. Spice, The measurement and parameterization of

Cloud screening for remote sensing of aerosol over oceans using spatial variability,

Martonchik, J. V., D. J. Diner, K. A. Crean, and M. A. Bull, Regional aerosol retrieval

Martonchik, J.V., D.J. Diner, R. Kahn, M.M. Verstraete, B. Pinty, H.R. Gordon, and T.P.
Ackerman, Techniques for the Retrieval of aerosol properties over land ocean using

Martonchik, J.V., D.J. Diner, R.A. Kahn, B.J. Gaitley, and B.N. Holben, Comparison of


Mason, B.J. and J. Maybank, The fragmentation and electrification of freezing water drops.


Rudolf, B., H. Hauschild, W. Rueth, and U. Schneider, Terrestrial Precipitation Analysis: Operational Method and Required Density of Point Measurements, in *Global*


Stith, J.L., L.F. Radke, and P.V. Hobbs, Particle emissions and the production of ozone and nitrogen oxides from the burning of forest slash, Atmos. Environ., 15, 73–82, 1981.


Wegener, A., Thermodynamik der Atmosphäre, J. A. Barth, Leipzig, 1911.

Wehrli, C., Calibration of filter radiometers for determination of atmospheric optical depth, Metrologia, 37, 419-422, 2000.


Zhao, T. L., S. L. Gong, X. Y. Zhang, and I. G. McKendry, Modelled size-segregated wet and dry deposition budgets of soil dust aerosol during ACE-Asia, 2001:

481


