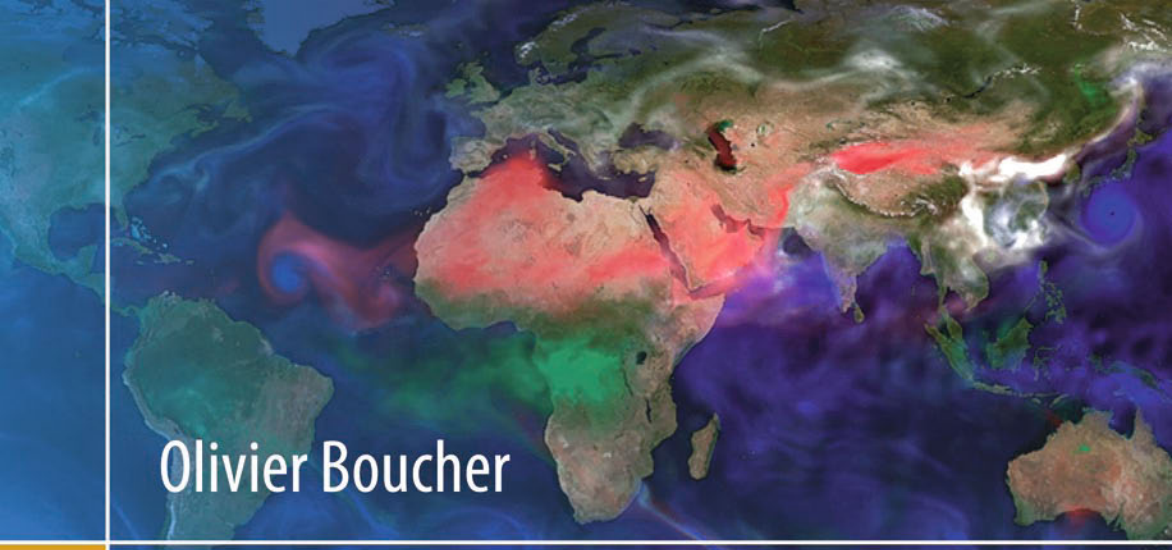


Olivier Boucher

Atmospheric Aerosols

Properties and Climate Impacts

Springer

Atmospheric Aerosols

Olivier Boucher

Atmospheric Aerosols

Properties and Climate Impacts

Springer

Olivier Boucher
Laboratoire de Météorologie Dynamique
Centre National de la Recherche Scientifique
Tour 45-55, 3e étage, Case courrier 99
Université Pierre et Marie Curie
4 place Jussieu
75252 Paris Cedex 05
France

ISBN 978-94-017-9648-4 978-94-017-9649-1 (eBook)
DOI 10.1007/978-94-017-9649-1

Library of Congress Control Number: 2015930841

Springer Dordrecht Heidelberg New York London

This book is an updated and extended translation of Aérosols atmosphériques: Propriétés et impacts climatiques.

The original work in French was published by Springer France in 2012.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

This textbook is an updated and extended version of an original work written in French language, which was also published by Springer upon the invitation of Bruno Sportisse. The textbook stems from many years of learning, research and teaching in the field of aerosol and climate sciences. My ambition has been to provide a one stop shop for students, teachers and other professionals interested in atmospheric aerosols and their role in the climate system. In many aspects, this textbook is only scratching the surface, as many of its chapters have become a field of research by themselves. Whenever, possible earlier influential papers that laid the foundations of aerosol and climate sciences are cited along with more recent publications and review articles. I hope this textbook will meet its objectives and provide useful material to the readers.

Olivier Boucher

Acknowledgments

I would like to thank all the people who contributed directly or indirectly to the accumulation of knowledge which is presented here. Some friends and colleagues will certainly recognize some of the results that appear in this textbook! In particular some of the lead authors of the "Clouds and Aerosols" chapter of the Intergovernmental Panel on Climate Change (IPCC) fifth assessment report which I have coordinated, have influenced my thinking on the topic. I would also like to thank particularly those who have prepared some figures for this book or have given me permission to reproduce their figures, a complete list of whose is provided below.

I would like to thank the "Editions Belin" for their permission to reproduce the chapter on "Radiative transfer" which was originally published in another textbook entitled "Physique et Chimie de l'Atmosphère". This book had been originally written with Michèle Vesperini who unfortunately passed away far too young and for whom I have a special thought here.

Thank you very much to Nicolas Bellouin from the University of Reading for his complete proof-reading of the French and English versions of this textbook. Nicolas has made many suggestions to improve the text and this textbook owes him a lot. Johannes Quaas from the University of Leipzig, Paola Formenti and Stéphane Alfaro from the Laboratoire interuniversitaire des systèmes atmosphériques, Fabien Waquet and Didier Tanré from the Laboratoire d'optique atmosphérique, Slimane Bekki from the Laboratoire atmosphères, milieux, observations spatiales, Frédéric Chevallier from the Laboratoire des sciences du climat et de l'environnement, Nicolas Huneeus from the Department of Geophysics of the University of Santiago, Patricia de Rosnay from the European Centre for Medium-Range Weather Forecasts, Prof. Yutaka Kondo from the University of Tokyo, Fabrice Jégou and Gwenaël Berthet from the Laboratoire de physique et de chimie de l'environnement et de l'espace, and Alan Robock from Rutgers University have also corrected bits and pieces of the book.

Special thanks to my colleagues and friends who provided some figures and text: Nicolas Huneeus (text on aerosol data assimilation), Thierry Fouchet (Figs. 5.6 and 5.7), Nicolas Bellouin (Fig. 10.5), Patricia de Rosnay (Fig. 2.4), Johannes Quaas (figure in the weekly cycle box, Figs. 9.1 and 9.2), Andy Jones (Figs. 10.3 and 10.6), Angela Benedetti (Fig. 2.5), Federica Pacifico (Fig. 4.4), Didier Tanré (Figs.

6.2 and 6.8), Paola Formenti (Figs. 2.3 and 2.7), Frank Dentener (Fig. 11.3), Rong Wang (Figs. 4.3 and 4.4), Eulalie Boucher (Fig. 2.5 top), ICARE (Figs. 6.8 and 6.10) MEEO s.r.l. (cover figure).

Nicolas Bellouin from the University of Reading has kindly provided me with data to produce Figs. 8.3 and 8.4; so has Andy Jones from the Met Office for Fig. 9.9, Philippe Dubuisson from the Laboratoire d'optique atmosphérique (Fig. 5.14), Jean-Paul Vernier from NASA (Fig. 12.7) and Glen Frick (Fig. 2.7). Christoforos Tsamalis provided data from the infrared atmospheric sounding interferometer (IASI) instrument to make the atmospheric radiance exercise more realistic. I am also grateful to the Met Office, the IPCC and many scientific journals for their permission to reproduce their copyrighted or licensed published material.

Contents

1 General Introduction 1
1.1 The Climate System 1
1.2 The Atmosphere 2
1.3 Energy Budget and Atmospheric Composition 3
1.4 The Water Cycle 4
1.5 Aerosols and Climate Change 5
1.6 Outline of this Textbook 7
References 8
Further Reading (Textbooks and Articles) 8

2 Atmospheric Aerosols 9
2.1 Definitions 9
2.2 Sources of Aerosols and Aerosol Precursors 12
2.2.1 Marine Aerosols 12
2.2.2 Desert Dust 12
2.2.3 Volcanic Aerosols 13
2.2.4 Biogenic Aerosols 14
2.2.5 Biomass Burning Aerosols 15
2.2.6 Aerosols from Fossil Fuel Combustion 15
2.3 Spatial and Temporal Aerosol Distributions 15
2.4 Aerosol–Cloud–Radiation Interactions 17
2.5 Climate Effects of Aerosols 21
References 24
Further Reading (Textbooks and Articles) 24

3 Physical, Chemical and Optical Aerosol Properties 25
3.1 Fine, Accumulation and Coarse Modes 25
3.2 Size Distribution 26
3.3 Chemical Composition 29
3.3.1 Aerosol Mixture 29
3.3.2 Inorganic Aerosols 30
3.3.3 Black Carbon Aerosols 30

3.3.4 Organic Aerosols 31
3.3.5 Geographic Distribution of Aerosol Chemical Composition 31
3.4 Refractive Index 32
3.5 Deliquescence, Efflorescence and Hysteresis 36
3.6 Definition of Aerosol Optical Properties 36
3.6.1 Absorption and Scattering Cross Sections 36
3.6.2 Phase Function 38
3.6.3 Upscatter Fractions 38
3.7 Calculation of Aerosol Optical Properties 40
3.7.1 Mie Theory 40
3.7.2 Extinction, Scattering and Absorption 41
3.7.3 Optical Depth and Ångström Coefficient 43
3.8 Optical Properties of Nonspherical Aerosols 45
3.9 Aerosols and Atmospheric Visibility 45
References 48
Further Reading (Textbooks and Articles) 49

4 Aerosol Modelling 51
4.1 Introduction 51
4.2 Emissions 52
4.2.1 Generalities 52
4.2.2 Fossil Fuels, Biofuels, and Other Anthropogenic Sources .. 53
4.2.3 Vegetation Fires 54
4.2.4 Sea Spray 55
4.2.5 Desert Dust 56
4.2.6 Dimethylsulphide 60
4.2.7 Biogenic Volatile Organic Compounds 62
4.2.8 Volcanoes 63
4.2.9 Resuspension 63
4.3 Atmospheric Processes 65
4.3.1 Nucleation 65
4.3.2 Condensation of Semi-Volatile Compounds 67
4.3.3 Coagulation 67
4.3.4 In-Cloud Aerosol Production 68
4.3.5 Wet Deposition 69
4.3.6 Dry Deposition 71
4.3.7 Sedimentation 71
4.3.8 Aerosol Transport 73
4.4 Modelling Approaches 73
4.4.1 Bulk Approach 74
4.4.2 Sectional Approach 74
4.4.3 Modal Approach 75
4.5 Example: The Sulphur Budget 77
References 80
Further Reading (Textbooks and Articles) 81

5 Interactions of Radiation with Matter and Atmospheric Radiative Transfer 83
5.1 Introduction 83
5.2 Electromagnetic Radiation 84
5.2.1 Generalities 84
5.2.2 Definitions 85
5.3 Interactions of Radiation with Matter 87
5.3.1 Matter, Energy and Spectral Lines 88
5.3.2 Intensity of Spectral Lines 92
5.3.3 Spectral Line Profiles 93
5.3.4 Processes of Interactions of Radiation with Matter 96
5.4 Modelling of the Interaction Processes 97
5.4.1 Molecular Absorption Coefficient 97
5.4.2 Scattering Phase Function 98
5.4.3 Molecular Scattering 99
5.4.4 Absorption and Scattering by Aerosols 100
5.4.5 Thermal Emission 102
5.5 Atmospheric Radiative Transfer 105
5.5.1 Equation of Radiative Transfer 105
5.5.2 Extinction Only 107
5.5.3 Scattering Medium 108
5.5.4 Plane-Parallel Atmosphere 109
5.5.5 Resolution of the Equation of Radiative Transfer 110
5.6 Absorption Bands, Energy, and Actinic Fluxes 113
5.6.1 Main Molecular Absorption Bands in the Atmosphere 113
5.6.2 Radiative Flux 115
5.6.3 Two-Flux Method 116
5.6.4 Stefan–Boltzmann Law 117
5.6.5 Radiative Budget 117
5.6.6 Actinic Fluxes 119
5.6.7 Polarization of Radiation 121
References 126
Further Reading (Textbooks and Articles) 126

6 In Situ and Remote Sensing Measurements of Aerosols 129
6.1 Introduction to Aerosol Remote Sensing 129
6.2 Passive Remote Sensing: Measurement of the Extinction 134
6.2.1 General Principles 134
6.2.2 Ground-Based Photometry 135
6.2.3 Spaceborne Occultation Measurements 136
6.2.4 Retrieval of Aerosol Size Distribution 136
6.3 Passive Remote Sensing: Measurement of the Scattering 137
6.3.1 General Principles 137
6.3.2 Ground-Based Measurement of Scattered Radiation 138
6.3.3 Spaceborne Measurements of Scattered Radiation 138

6.4 Measurement of Infrared Radiation 143
6.4.1 General Principles 143
6.4.2 Spaceborne Nadir Measurement of Infrared Radiation 144
6.4.3 Spaceborne Limb Measurement of Infrared Radiation 146
6.5 Active Remote Sensing: Lidar 146
6.5.1 General Principles 146
6.5.2 The Lidar Equation 146
6.5.3 Raman Lidar 148
6.6 In Situ Aerosol Measurements 150
6.6.1 Measurement of Aerosol Concentrations 150
6.6.2 Measurement of Aerosol Chemical Composition 152
6.6.3 Measurement of Aerosol Scattering 153
6.6.4 Measurement of Aerosol Absorption 153
6.7 Conclusions 154
References 158
Further Reading (Textbooks and Articles) 160

7 Aerosol Data Assimilation 161
7.1 Introduction 161
7.2 Basic Principles of Data Assimilation 162
7.3 Applications of Data Assimilation for Aerosols 165
References 170
Further Reading (Textbooks and Articles) 172

8 Aerosol–Radiation Interactions 173
8.1 Introduction 173
8.2 Atmospheric Radiative Effects Due to Aerosols 175
8.2.1 Simplified Equation for Scattering Aerosols 175
8.2.2 Simplified Equation for Absorbing Aerosols 178
8.2.3 Radiative Transfer Calculations 179
8.2.4 Global Estimates and Sources of Uncertainty 183
8.3 Rapid Adjustments to Aerosol–Radiation Interactions 185
8.4 Radiative Impact of Aerosols on Surface Snow and Ice 187
References 191
Further Reading (Textbooks and Articles) 192

9 Aerosol–Cloud Interactions 193
9.1 Introduction 193
9.1.1 Cloud Formation 193
9.1.2 Cloud Distribution 194
9.1.3 Aerosol–Cloud Interactions 196
9.2 Aerosol Effects on Liquid Clouds 198
9.2.1 Saturation Pressure of Water Vapour 198
9.2.2 Kelvin Effect 199
9.2.3 Raoult's Law 200

9.2.4 Köhler Theory 200
9.2.5 Extensions to the Köhler Theory 202
9.2.6 CCN and Supersaturation in the Cloud 204
9.2.7 Dynamical and Radiative Effects in Clouds 205
9.2.8 Principle of the Cloud Albedo Effect 205
9.2.9 Observations of the Cloud Albedo Effect 206
9.2.10 Adjustments in Liquid Water Clouds 208
9.2.11 Rapid Adjustments Occurring in Liquid Clouds 210
9.3 Aerosols Effects on Mixed-Phased and Ice Clouds 212
9.3.1 Elements of Microphysics of Ice Clouds 213
9.3.2 Impact of Anthropogenic Aerosols on Ice Clouds 214
9.4 Forcing Due to Aerosol–Cloud Interactions 215
9.5 Aerosols, Contrails and Aviation-Induced Cloudiness 216
9.5.1 Formation of Condensation Trails 216
9.5.2 Estimate of the Climate Impact of Contrails 218
References 221
Further Reading (Textbooks and Articles) 225

10 Climate Response to Aerosol Forcings 227
10.1 Introduction 227
10.2 Radiative Forcing, Feedbacks and Climate Response 227
10.2.1 Radiative Forcing 227
10.2.2 Climate Feedbacks 228
10.2.3 Rapid Adjustments and Effective Radiative Forcing 231
10.2.4 Climate Response and Climate Efficacy 234
10.3 Climate Response to Aerosol Forcings 235
10.3.1 Equilibrium Response 235
10.3.2 Past Emissions 237
10.3.3 Detection and Attribution of Aerosol Impacts 239
10.3.4 Future Emissions Scenarios 240
10.4 Nuclear Winter 242
References 244
Further Reading (Textbooks and Articles) 246

11 Biogeochemical Effects and Climate Feedbacks of Aerosols 247
11.1 Introduction 247
11.2 Impact of Aerosols on Terrestrial Ecosystems 248
11.2.1 Diffuse Radiation and Primary Productivity 248
11.2.2 Aerosols as a Source of Nutrients 249
11.2.3 Acidification of Precipitation 250
11.3 Impact of Aerosols on Marine Ecosystems 252
11.4 Aerosols–Atmospheric Chemistry Interactions 253
11.4.1 Interactions with Tropospheric Chemistry 253
11.4.2 Impact of Stratospheric Aerosols on the Ozone Layer and Ultraviolet Radiation 255

11.5 Climate Feedbacks Involving Marine Aerosols 255
11.5.1 Sulphate Aerosols from DMS Emissions 255
11.5.2 Marine Aerosols 257
11.5.3 Other Aerosols of Maritime Origin 258
11.6 Climate Feedbacks Involving Continental Aerosols 258
11.6.1 Secondary Organic Aerosols 259
11.6.2 Primary Aerosols of Biogenic Origin 260
11.6.3 Aerosols from Vegetation Fires 261
11.6.4 Desert Dust 262
11.7 Climate Feedbacks Involving Stratospheric Aerosols 264
References 265
Further Reading (Textbooks and Articles) 269

12 Stratospheric Aerosols 271
12.1 Introduction 271
12.2 Atmospheric Cycle of Stratospheric Aerosols 271
12.2.1 Sources of Stratospheric Aerosols 271
12.2.2 Transport in the Stratosphere 273
12.3 Physics and Chemistry of Stratospheric Aerosols 275
12.4 Volcanic Aerosol Records 277
12.4.1 Volcanic Aerosols During the 1750–2010 Period 277
12.4.2 A Few Recent Climate-Relevant Eruptions 279
12.4.3 Mega-Eruptions 281
12.5 Role of Stratospheric Aerosols on the Climate System 281
References 284
Further Reading (Textbooks and Articles) 286

13 Aerosol-Based Climate Engineering 287
13.1 Introduction 287
13.2 Stratospheric Aerosol Injection 288
13.3 Marine Cloud Brightening 290
13.4 Role of Rapid Adjustments and Feedbacks 291
13.5 A Thought on Timescales 292
References 293
Further Reading (Textbooks and Articles) 294

Appendix A Units and Physical Constants 295

Appendix B Properties of the Log-Normal Size Distribution 297

Appendix C Mie Theory 299

Appendix D Radiative Impact of Aerosols on Snow and Ice 305

Index 307

Acronyms

3D-Var	Three-dimensional variational (data assimilation)
3MI	Multi-Viewing Multi-Channel Multi-Polarisation Imaging Mission
4D-Var	Four-dimensional variational (data assimilation)
6S	Second Simulation of the Satellite Signal in the Solar Spectrum
AAOD	Aerosol absorption optical depth
AATSR	Advanced Along-Track Scanning Radiometer
ACCMIP	Atmospheric Chemistry and Climate Model Intercomparison Project
AeroCom	Aerosol Comparisons between Observations and Models
AERONET	Aerosol Robotic Network
AI	Aerosol Index
AIRS	Atmospheric Infrared Sounder
AMMA	African Monsoon Multidisciplinary Analysis
AMSU	Advanced Microwave Sounding Unit
AOD	Aerosol optical depth
AOT	Aerosol optical thickness
ATSR	Along Track Scanning Radiometer
AVHRR	Advanced Very High Resolution Radiometer
BB	Biomass burning
BC	Black carbon
BDRF	Bidirectional reflectance distribution function
BF	Bio-fuels
BrC	Brown carbon
BVOC	Biogenic volatile organic compound
CALIOP	Cloud–Aerosol Lidar with Orthogonal Polarization
CALIPSO	Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations
CCN	Cloud condensation nucleus
CDNC	Cloud droplet number concentration
CLAW	Charlson–Lovelock–Andreae–Warren
CN	Condensation nucleus
CNES	Centre National d'Études Spatiales

CNRS	Centre National de la Recherche Scientifique
CRE	Cloud radiative effect
DMS	Dimethylsulphide
DMSO	Dimethyl sulphoxide
DMSP	Dimethylsulphoniopropionate
ECMWF	European Centre for Medium-Range Weather Forecasts
EMEP	European Monitoring and Evaluation Programme
ERA	ECMWF re-analysis
ERF	Effective radiative forcing
ERFaci	Effective radiative forcing due to aerosol–cloud interactions
ERFari	Effective radiative forcing due to aerosol–radiation interactions
ESA	European Space Agency
FAT	Fixed anvil temperature
FF	Fossil fuel
FRP	Fire radiative power
GAW	Global Atmospheric Watch
GOMOS	Global Ozone Monitoring by Occultation of Stars
GPP	Gross primary productivity
HadGEM2	Hadley Centre General Environment Model 2
IASI	Infrared Atmospheric Sounding Interferometer
IASI-NG	Infrared Atmospheric Sounding Interferometer—New Generation
IFS	Integrated Forecast System
IN	Ice nucleus
INDOEX	Indian Ocean Experiment
IPCC	Intergovernmental Panel on Climate Change
IPSL	Institut Pierre Simon Laplace
KF	Kalman filter
LAI	Leaf area index
LMD	Laboratoire de météorologie dynamique
LOA	Laboratoire d'optique atmosphérique
LSCE	Laboratoire des sciences du climat et de l'environnement
LW	Longwave
MACC	Monitoring Atmospheric Composition and Climate
MAESTRO	Measurements of Aerosol Extinction in the Stratosphere and Troposphere Retrieved by Occultation
MISR	Multi-angle Imaging SpectroRadiometer
MODIS	Moderate Resolution Imaging Spectrometer
MSA	Methanesulphonic acid
NAAPS	Navy Aerosol Analysis and Prediction System
NADP	National Atmospheric Deposition Program
NASA	National Aeronautics and Space Administration
NIES	National Institute for Environmental Studies
NPF	New particle formation
NPP	Net primary productivity
NPZD	Nutrient–phytoplankton–zooplankton–detritus

NRL	Naval Research Laboratory
NWP	Numerical Weather Prediction
OA	Organic aerosol
OD	Optical depth
OM	Organic matter
PAR	Photosynthetically active radiation
PARASOL	Polarization and Anisotropy of Reflectances for Atmospheric Sciences coupled with Observations from a Lidar
PCAS	Passive Cavity Aerosol Spectrometer
PFR	Precision filter radiometer
PFT	Plant functional type
PHOTONS	PHOtométrie pour le Traitement Opérationnel de Normalisation Satellitaire
POA	Primary organic aerosol
POLDER	Polarization and Directionality of the Earth's Reflectances
PSC	Polar stratospheric cloud
QBO	Quasi-biennal oscillation
RCP	Representative Concentration Pathway
RF	Radiative forcing
RFaci	Radiative forcing due to aerosol–cloud interactions
RFari	Radiative forcing due to aerosol–radiation interactions
RH	Relative humidity
SAGE	Stratospheric Aerosol and Gas Experiment
SAM	Stratospheric Aerosol Measurement
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SOA	Secondary organic aerosol
SP2	Single Particle Soot Photometer
SRM	Solar radiation management
SSA	Single scattering albedo
SST	Sea surface temperature
STP	Standard temperature and pressure
SW	Shortwave
TOMS	Total Ozone Mapping Spectrometer
UV	Ultraviolet
VEI	Volcanic explosivity index
VOC	Volatile organic compound
WMO	World Meteorological Organization

Chapter 1
General Introduction

Abstract This chapter forms a general introduction to the climate system. It defines the different layers of the atmosphere and their properties. General concepts on the atmospheric composition, the Earth's radiative budget, the water cycles, and climate change are introduced. Finally, this chapter also offers a road map to the other chapters of the book.

Keywords Atmosphere · Climate system · Water cycle · Earth's radiative budget · Climate change

1.1 The Climate System

Climate is about the statistical properties of meteorological conditions over a given period. Such a period ought to be long enough so that these meteorological conditions are well sampled, but not too long so that natural and anthropogenic[1] climate variations can be detected. It is considered that 30 years constitute a typical period to characterize the climate system. The study of the climate system, also called climatology, therefore differs from meteorology which is about the short-term fluctuations of the atmosphere. It should be noted that climate is not just "the average of the weather", it is also about the higher moments of the distribution of meteorological variables; this is to say about the variability in the weather including statistics of extreme weather.

The climate system is usually considered to be comprised of several components, namely the atmosphere, the ocean, continental surfaces, the cryosphere (a term that designates all forms of ice on the planet), and the marine and terrestrial biosphere (all forms of living organisms on the planet). The climate results from the intrinsic properties of these components and the multiple interactions that exist between them, given the boundary conditions that are imposed to this system. What constitutes a boundary condition depends on how one defines the climate system. We consider here that the properties of the energy radiated by the Sun, the Earth's orbit, volcanic

[1] That is caused by humans.

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_1

activity, and anthropogenic emissions of pollutants are all boundary conditions that are applied to the climate system.

1.2 The Atmosphere

The atmosphere is, with the ocean, one of the most important components of the climate system; and it is the component that is of particular interest in this book. The atmosphere is a relatively thin layer of gases that is held around the Earth by gravity. Dry air is composed of 78 % of molecular nitrogen (N_2), 21 % of molecular oxygen (O_2), argon, and other trace gases. Water vapour is also present in the atmosphere in very variable conditions both in space and time. The percentage by which water vapour can vary from close to zero in dry places, up to 4 % in wet and hot places. Pressure and air density in the atmosphere decreases with altitude. One can distinguish several atmospheric layers that go from the surface to the top of the atmosphere:

- The *troposphere* extends from the surface to an altitude of about 8–15 km according to the latitude (i.e. the troposphere is deepest in the Tropics and shallowest at the poles). It is a region of the atmosphere where temperature decreases with altitude and where convection plays an important role to mix air masses in the vertical. It is also in the troposphere where the quasi-totality of clouds are encountered. The atmospheric *boundary layer* is the portion of the troposphere which is directly influenced by the Earth's surface; it has a thickness of the order of 1–2 km. The rest of the troposphere is often called the *free troposphere* and extends from the top of the boundary layer to the top of the troposphere called the *tropopause*. The gradient of temperature in the atmosphere is called the *lapse rate* and is related to the cooling of air as it expands. It varies from typically 5 $\mathrm{K\,km^{-1}}$ in wet conditions (moist or saturated adiabatic lapse rate) to almost 10 $\mathrm{K\,km^{-1}}$ in dry conditions (dry adiabatic lapse rate). The difference between the dry and moist adiabatic lapse rates is because latent heat is released when water condenses in the atmosphere, thus decreasing the vertical temperature gradient.
- The *stratosphere* extends from the tropopause to an altitude of about 50 km. The temperature increases with altitude so that this region of the atmosphere is more stratified than the troposphere. Air density continues to decrease with altitude and is much smaller than in the troposphere. The stratosphere is also where is located what is known as the *ozone layer*. The fact that air masses have to go through the tropopause and experience cold temperatures before entering the stratosphere makes it a fairly dry region.
- The *mesosphere* located above the stratosphere extends up to 80–100 km. Temperature decreases with altitude as in the troposphere.
- The *thermosphere* extends from about 80–100 km up to 600 km. The chemical composition is no longer uniform and is enriched in the lighter chemical compounds.

- The *exosphere* is the uppermost layer of the atmosphere. Air density is extremely low.

The focus of this book is about the first two layers of the atmosphere, namely the troposphere and the stratosphere, because of their relevance for both the climate system and atmospheric aerosols.

1.3 Energy Budget and Atmospheric Composition

The Earth's climate responds to external conditions applied on its boundaries (i.e. boundary conditions in the language of mathematics and physics), and first of all to the amount and distribution of solar radiation received by the planet. Incoming solar radiation (also called insolation) depends on the quantity of radiation emitted by the Sun, but also on the characteristics of the Earth's orbit around the Sun, both of which vary in time. Volcanic activity is another boundary condition which needs to be taken into account to explain some aspects of climate variability. It is convenient to consider the climate system as an engine that converts solar radiation absorbed by the surface and the atmosphere into terrestrial radiation emitted by the surface and the atmosphere, some of which escapes to space and cools the planet. Solar radiation is also called *shortwave radiation* as it mostly consists of "shorter" wavelengths (ultraviolet, visible, and near-infrared), whereas terrestrial radiation is called *longwave radiation* as it mostly consists of "longer" wavelengths (infrared). When the climate system is at equilibrium, the quantity of solar radiation absorbed by the system is equal to the quantity of terrestrial radiation that escapes the system. Figure 1.1 shows the various components and processes involved in the energy budget of the Earth. Solar and terrestrial radiation interact with the surface and the atmosphere. The chemical composition of the atmosphere is a key parameter to understand its interactions with radiation. It is important to distinguish three different forms of matter present in the atmosphere that play a role in the energy budget:

- *Gas molecules*: Molecular nitrogen (or nitrogen gas, N_2) and molecular oxygen (or dioxygen, O_2) represent together more than 99 % of the volume of the atmosphere. They scatter solar radiation but are largely transparent to solar and terrestrial radiation. Other gases in smaller atmospheric concentrations (called *trace gases*) can absorb and emit radiation. Gas molecules that can absorb and emit terrestrial radiation are also called *greenhouse gases*.
- *Hydrometeors*: They are any liquid, water or ice particles that have formed in the atmosphere. They can be in suspension in the atmosphere, in which case they form a cloud, or be falling when their sedimentation velocity exceeds the updraft. Clouds interact strongly with both solar and terrestrial radiation.
- *Aerosols*: They are small particles in suspension in the atmosphere which can exist in solid, liquid, and semi-liquid form. While aerosols may contain water, they are usually distinguished from hydrometeors that are bigger in size. Aerosols also interact with both solar and terrestrial radiation.

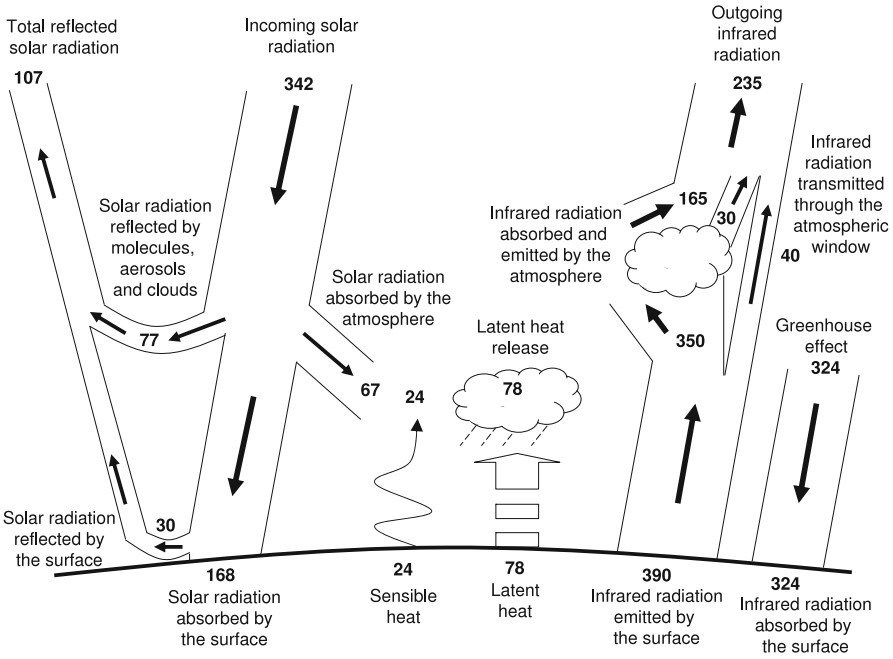


Fig. 1.1 Energy budget of the Earth's atmosphere and surface (in W m^{-2}). The fluxes are global means; they are only indicative as the energy budget of the Earth is only known to an accuracy of a few W m^{-2}. See Chap. 5 for a more in depth discussion of this figure. (Figure reproduced with the permission of the Éditions Belin)

The three different forms of matter (gases, aerosols, and hydrometeors) interact with each other and a molecule can pass from one form to another.

1.4 The Water Cycle

The water cycle is a key aspect of the climate, not only because it relates to all components of the climate system, but also because it is indispensable to life on Earth. Figure 1.2 illustrates the water cycle in the climate system and the fluxes between the different reservoirs that are the ocean, the atmosphere, continental surfaces, and the cryosphere. Schematically, water evaporates from the ocean, and to some extent from continental surfaces. It can undergo several condensation-evaporation cycles in the atmosphere before falling down to the Earth's surface in the form of precipitation. Water is present in the atmosphere in the gas, liquid, and solid phase. Water in the gas phase is called *water vapour*. Water is found in the liquid and solid phase in clouds, and precipitation can also be in the liquid phase (rain) and in the solid phase (snow and graupel). The water cycle is intrinsically linked to the energy cycle: on the one hand

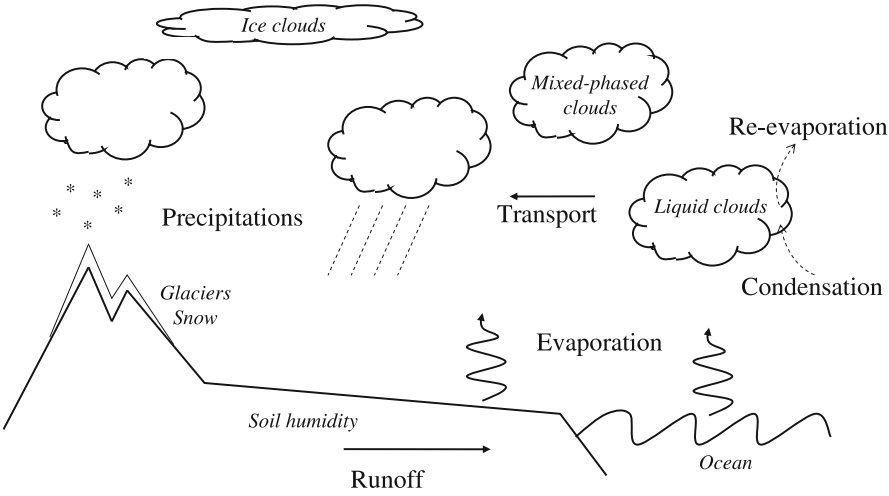


Fig. 1.2 Schematic of the water cycle in the climate system

a phase change is associated with the absorption or release of latent heat, and on the other hand water vapour and clouds interact with both solar and terrestrial radiation (Fig. 1.1). Biogeochemical cycles (carbon, nitrogen, sulphur) are also central in the past and future climate of the Earth inasmuch as the carbon, nitrogen, and sulphur species have an impact on the atmospheric chemical composition and on marine and terrestrial ecosystems.

1.5 Aerosols and Climate Change

The Earth's climate varies on numerous timescales that range from very long geological timescales and glacial–interglacial timescales typically of the last million years, to shorter timescales caused by external factors, such as volcanism, or the internal variability of the system. The latter includes a number of quasi-periodic oscillations involving the cryosphere, the ocean, and the atmosphere and their interactions.

Human activity has modified profoundly the chemical composition of the atmosphere, well beyond variations observed over the last thousands of years or those encountered during the succession of glacial and interglacial conditions of the last million years. The rate of anthropogenic changes is also very large in comparison to past changes. Among greenhouse gases of anthropogenic origin, one can mention carbon dioxide, methane, and nitrous oxide that are responsible for an additional greenhouse effect (Fig. 1.3). Observations show a warming of the planet since preindustrial times that has amplified and accelerated since the 1950s. The observed warming cannot be explained by natural factors (changes in the sun, volcanism, and natural variability) alone but requires anthropogenic factors including an enhanced

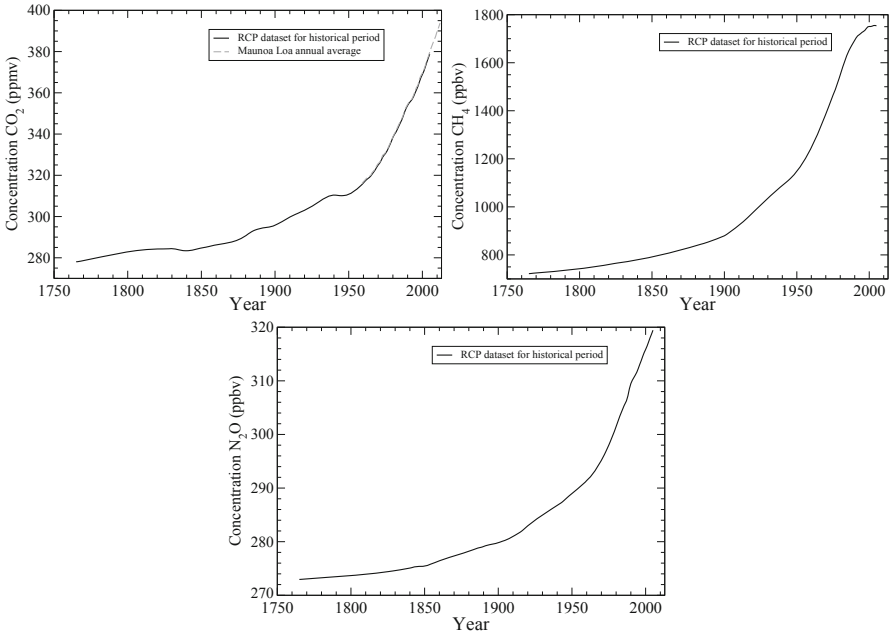


Fig. 1.3 Temporal evolution of the atmospheric concentrations of the three main anthropogenic greenhouse gases: carbon dioxide (CO_2, ppmv), methane (CH_4, ppb v), and nitrous oxide (N_2O, ppbv) for the period 1765–2005. (Data from the ACCMIP dataset for the historical period and from Mauna Loa observatory)

greenhouse effect (Fig. 1.4). This is how the Intergovernmental Panel on Climate Change (IPCC) has reached the conclusion that the observed climate change of the past century is extremely likely due in part to anthropogenic factors (Bindoff et al. 2013), confirming earlier results by Hegerl et al. (2007).

Atmospheric aerosols form an integral part of the climate system and interact with the atmosphere, the cryosphere, the biosphere, and the ocean. Aerosols interact strongly with the water cycle and a number of biogeochemical cycles. Greenhouse gases are indeed not the only anthropogenic driver of climate change. It is now widely understood that anthropogenic aerosols have masked a fraction of the warming effect expected from the increase in greenhouse gases, since the beginning of preindustrial times. This book reviews the different interactions between aerosols, radiation, and the hydrological cycle. It highlights the role of natural and anthropogenic aerosols not only in the climate system, but also their role in the hydrological cycle and climate-relevant biogeochemical cycles.

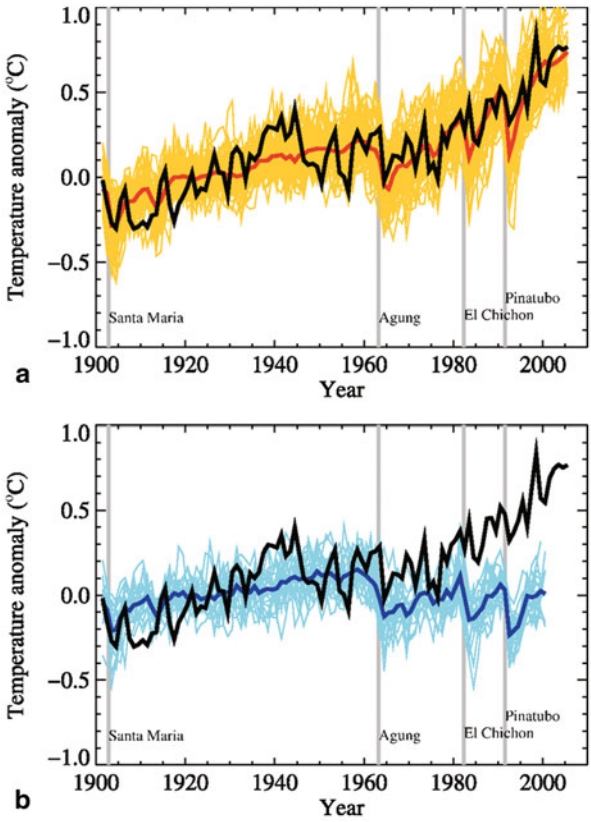


Fig. 1.4 Temporal evolution of the global-mean surface temperature as observed (*black solid line*) and as simulated by a range of climate models (*coloured solid lines* with an envelope) in response to all natural and anthropogenic factors (*top panel*) or natural factors only (*bottom panel*). (Reproduced from Fig. 9.5 in Hegerl et al. (2007). © IPCC)

1.6 Outline of this Textbook

This book is organized in the following way. Chapters 2 and 3 define atmospheric aerosols and describe their physical, chemical, and optical properties. Chapter 4 then provides the basis for aerosol modelling in global aerosol and climate models. Chapter 5 describes succinctly the interactions between aerosols and radiation, as well as the theory of atmospheric radiative transfer, which is essential to understand remote sensing techniques and the climate impacts of aerosols. Chapter 6 introduces remote sensing techniques for aerosols, as well as various techniques for in situ observations of aerosols, while Chap. 7 introduces the concept of data assimilation for aerosols. The next two chapters cover the different impacts of aerosols on the climate system, namely aerosol–radiation interactions (Chap. 8) and aerosol–cloud interactions (Chap. 9). Chapter 10 discusses the response of the climate system to the aerosol forcings, while Chap. 11 reviews how aerosols interact with the main biogeochemical cycles and the climate feedbacks that could modify the natural cycles of aerosol components in the figure. Chapter 12 deals with the specificities of stratospheric aerosols. Finally, the last chapter assesses the climate engineering schemes which

rely on solar radiation management through the injection of aerosols and aerosol precursors.

This book is enhanced by a list of references and four appendices. The list of references mixes seminal papers and a selection of more recent studies so as to give a good entry point to a more thorough literature review. The appendices include more technical material. Appendix A lists a small selection of physical and chemical data. Appendix B introduces the properties of the lognormal size distribution which is widely used to describe atmospheric aerosols. Appendix C presents briefly the Mie theory that can be used to compute the optical properties of spherical aerosols. Finally, Appendix D provides a calculation of how aerosols deposited at the surface can influence the optical properties of the snowpack.

References

Bindoff NL, Stott PA, AchutaRao KM, Allen MR, Gillett N, Gutzler D, Hansingo K, Hegerl G, Hu Y, Jain S, Mokhov II, Overland J, Perlwitz J, Sebbari R, Zhang X (2013) Detection and attribution of climate change: from global to regional. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

Hegerl GC, Zwiers FW, Braconnot P, Gillett NP, Luo Y, Marengo Orsini JA, Nicholls N, Penner JE, Stott PA (2007) Understanding and attributing climate change. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

Further Reading (Textbooks and Articles)

Chahine MT (1992) The hydrological cycle and its influence on climate. Nature 359:373–380.

Cubasch U, Wuebbles D, Chen D, Facchini MC, Frame D, Mahowald N, Winther J-G (2013) Introduction. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

Houghton J (2009) Global warming: the complete briefing, 4th edn. Cambridge University Press, Cambridge, 456 pp

Le Treut H Somerville R Cubasch U Ding Y Mauritzen C Mokssit A Peterson T Prather M (2007) Historical overview of climate change. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

McGuffie K, Henderson-Sellers A (2005) A climate modelling primer, 3rd edn. Wiley, 296 pp

Peixoto JP, Ort AH (2007) Physics of climate. Springer, Berlin and Heidelberg, 564 pp

Stephens GL, Li J, Wild M, Clayson CA, Loeb N, Kato S, L'Ecuyer T, Stackhouse PW Jr, Lebsock M, Andrews T (2012) An update on Earth's energy balance in light of the latest global observations, Nat Geosci 5:691–696

Stevens B, Bony S (2013) Water in the atmosphere. Phys Today 66:29–34

Chapter 2
Atmospheric Aerosols

Abstract This chapter provides an introduction to atmospheric aerosols by offering a definition and discussing various ways of categorizing and characterizing aerosols. The natural and anthropogenic sources of atmospheric aerosols are reviewed and discussed, leading to a description of their spatial distribution in the atmosphere. The chapter also introduces qualitatively some aerosol concepts, such as their chemical composition and size distribution, paving the way for a more complete description in the following chapters. Some examples are provided to illustrate how aerosols, despite being microscopic particles, can manifest themselves in the atmosphere. Finally, the various pathways by which aerosols impact the climate system are reviewed, introducing a revised categorization of aerosol–radiation, aerosol–cloud, and aerosol–surface interactions.

Keywords Aerosol definition · Aerosol properties · Aerosol types · Aerosol–radiation interactions · Aerosol–cloud interactions

2.1 Definitions

An aerosol is by definition a collection of solid or liquid particles in suspension in a gas. Strictly speaking the term aerosol includes both the particles and the suspending gas. In atmospheric sciences, it is usual however to use the term aerosol in its plural form to refer to the "aerosol particles" without the atmosphere as the suspending gas. We follow that convention in this book, but occasionally refer to "the aerosol" when we refer to atmospheric aerosols in a generic way. Atmospheric scientists also like to differentiate cloud particles from other types of particles in the atmosphere. For this reason we define aerosols as solid or liquid particles in suspension in the atmosphere to the exception of all hydrometeors (cloud droplets, ice crystals, raindrops, snowflakes, and graupel).

Aerosols are always present in the atmosphere but in extremely variable concentrations. This is due to the very large heterogeneity in aerosol sources and their relatively short residence time in the atmosphere (of the order of hours to weeks). The vast majority of aerosols are not visible to the naked eye because of their microscopic size but one can easily see the collective effect of aerosols in the atmosphere as soon as their concentrations are large enough. For example, a haze that reduces

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_2

Fig. 2.1 Vegetation fire in the Alps on a cloudy day. One can notice the *blueish* colour of the smoke plume which contrasts with the *white* colour of clouds. The *blueish* colour is due to the small aerosol size in the smoke plume and their larger effectiveness at scattering radiation in the shorter wavelengths of the visible spectrum. (Photograph by the author)

Fig. 2.2 Saharan dust event over the Alps. The snowpack has taken an *orange* colour due to the deposition of dust. (Photograph taken in April 2002 by the author)

the atmospheric visibility and whitens the sky is nothing else than a collection of aerosol particles that interact with solar radiation. A smoke plume is composed of microscopic particles that stem from incomplete combustion of carbonated fuels; these particles collectively darken the sky (see Fig. 2.1). Aerosols can also be visible when they get deposited in great quantity on the Earth's surface, as it is occasionally the case for Saharan dust particles (Fig. 2.2). Furthermore, one can "see" aerosol particles one by one with the help of an electron microscope as shown on Fig. 2.3. The large variety of sizes and shapes among atmospheric aerosols is already striking from these photographs.

The amount and properties of aerosols are extremely variable in space and time. This is why one is usually interested in characterizing a population of aerosols rather than individual particles. The most important characteristics of an aerosol population are the size distribution, chemical composition, and shape of the particles. It is useful to classify aerosols in different categories according to their properties. There are several possible classifications:

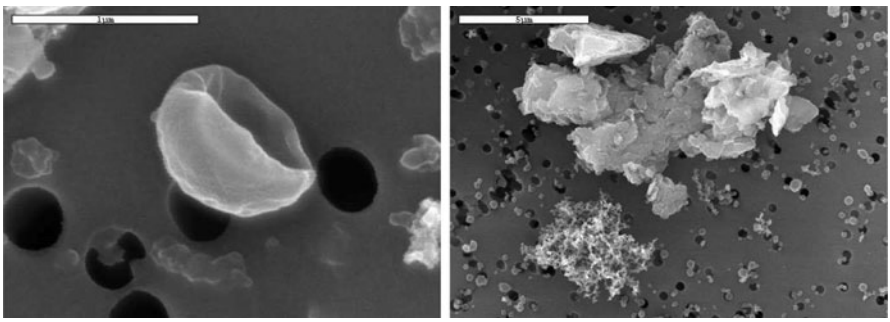


Fig. 2.3 Photographs of atmospheric aerosols collected on a filter: vegetal debris (*left*) and soot (black carbon) particle next to a mineral aggregate (*right*). The *white segments* represent 1 and 5 μm, respectively (© LISA/CNRS)

1. *Primary aerosols* have been emitted into the atmosphere as particles. This is the case of aerosols produced by the effect of the wind friction on an oceanic or terrestrial surface and aerosols produced during an incomplete combustion. *Secondary aerosols* designate those particles that have not been emitted directly in the particulate phase but come instead from the condensation of atmospheric gas-phase species. These gas-phase species, which can undergo a number of chemical transformations before they condense, are called *aerosol precursors*. The primary or secondary origin of aerosols offers a first way to categorize the atmospheric aerosol. The chemical composition of the aerosol usually provides a first idea as to whether the aerosol is primary or secondary.
2. Aerosol properties vary spatially and some of these properties can vary more or less systematically with the type of environment. One can thus speak of *urban aerosols*, *semiurban aerosols*, *continental aerosols*, *desertic aerosols*, *marine aerosols*, *volcanic aerosols* or *stratospheric aerosols*. This is an imperfect categorization in that aerosols can be transported a long way and are therefore not necessarily representative of the location where they can be found. It is for example possible to observe marine aerosols above the continents and continental aerosols above the ocean. When local effects dominate, it may be useful though to refer to such aerosol types in a broad sense.
3. Aerosols can also be classified according to their origin. One can distinguish natural from anthropogenic sources. Natural sources consist of emissions from the ocean, soils, vegetation, fires, and volcanoes. Anthropogenic sources are largely dominated by emissions from the combustion of fossil fuels (i.e. coal and oil), biofuels (plant biomass including wood, vegetable oils, animal waste), other fuels (e.g. peat), or vegetation fires caused by humans. Industrial activities, transportation, heating, or even domestic activities related to cooking in developing countries, are important sources of aerosols. Some industrial and agricultural activities can also emit primary aerosols referred to as industrial dust and arable dust, respectively.

None of these classifications can categorize fully and systematically the aerosol. The different aerosol populations mix and interact with each other in the atmosphere so that some of the terms and aerosol classes that we have just introduced are somehow misuses of language.

2.2 Sources of Aerosols and Aerosol Precursors

Let us now go through the different sources of aerosols in a more systematic way. Table 2.1 summarizes the amount of aerosols and aerosol precursors emitted by different sources.

2.2.1 Marine Aerosols

The wind friction at the ocean surface ejects fine particles of salty marine water into the atmosphere. A fraction of the water evaporates, so that the concentration of salt in the particle increases. This gives rise to sea salt particles that are more or less hydrated according to the ambient humidity. Although these particles are often called sea salt aerosols, this is yet another misuse of language because these particles may also contain biological material and other impurities. It is therefore more appropriate to refer to *sea spray aerosols*. Sea spray aerosols cover sizes that range from typically 100 nanometres (nm) to several tens of micrometres (μm). The largest particles fall back fairly quickly to the ocean surface and are therefore of lesser climatic importance.

2.2.2 Desert Dust

The wind friction on continental surfaces can detach soil particles and suspend them in the atmosphere. This is particularly the case in desertic, arid, and semiarid regions where the wind is not slowed down by the vegetation that is either completely absent or fairly sparse. The emission of soil particles to the atmosphere also requires a reduced soil humidity so that cohesive forces between soil particles are also reduced. As for sea spray aerosols, *desert dust* particles span sizes that range from typically 100 nm to tens of μm. Larger particles can also be lifted but they fall down quickly. Desert dust aerosols are also called *mineral dust* or mineral aerosols. Emissions of desert dust depend very much on environmental and meteorological conditions. They are sporadic in nature. Figure 2.4 shows a photograph of a dust storm generated by an atmospheric gravity current known as a *haboob*. Figure 6.3 shows an episode of Saharan dust aerosols advected above the Atlantic Ocean over a period of three days.

Table 2.1 Approximative emission fluxes from different types of primary aerosols and gaseous precursors of secondary aerosols. The climate importance of aerosols depends not only on the strength of their emissions, but also on their physical and chemical properties. Estimates are compiled from Penner et al. (2001), Dentener et al. (2006), Guenther et al. (1995), Jaenicke (2005), Burrows et al. (2009), Heald and Spracklen (2009). Tg = 10^{12} g = 1 million of tons. Gg = 10^{9} g = 1 thousand of tons

Aerosol type	Emission flux (per year)
Natural primary aerosols	
Desert dust	1000–3000 Tg
Sea spray	1000–6000 Tg
Biomass burning aerosols	20–35 Tg
Terrestrial primary biogenic aerosols	Order of 1000 Tg
Including bacteria	40–1800 Gg
Including spores	30 Tg
Precursors of natural secondary aerosols	
Dimethylsulphide (DMS)	20–40 Tg S
Volcanic SO_2	6–20 Tg S
Terpenes	40–400 Tg
Anthropogenic primary aerosols	
Industrial dust	40–130 Tg
Biomass burning aerosols	50–90 Tg
Black carbon (from fossil fuel)	6–10 Tg
Organic carbon (from fossil fuel)	20–30 Tg
Anthropogenic secondary aerosols	
SO_2	70–90 Tg S
Volatile organic compounds (VOCs)	100–560 Tg C
NH_3	20–50 Tg N
NO*x*	30–40 Tg N

C carbon, S sulphur, N nitrogen

2.2.3 *Volcanic Aerosols*

Volcanoes can emit fragments of pulverized rocks and minerals, usually called *volcanic ash*, during explosive eruptions. These particles have sizes typically ranging from a micrometre to millimetres. Volcanic ash can be transported over distances of a few hundreds to a few thousand kilometres but being micronic particles they tend to fall down rapidly. Hence their climate effect is limited. Volcanoes also emit sulphur-rich gases (in the form of sulphur dioxide, SO_2, and hydrogen sulphide, H_2S) that get oxidized in the atmosphere to form submicronic sulphate aerosols. If these sulphur-containing gases are emitted in the troposphere, the residence time of

Fig. 2.4 Desert dust storm in 2004 at Agoufou in Mali. These dust storms are known as haboobs in Africa but their names vary according to the country. (© Patricia de Rosnay, LEGOS/CNRS)

the subsequent aerosols will be short, a few weeks at most. However, if the eruption is powerful enough to inject the sulphur gases in the stratosphere, then the volcanic aerosols have a much longer residence time, of the order of a few months to more than a year, depending on the region and altitude of injection.

2.2.4 Biogenic Aerosols

The terrestrial biosphere is a source of *primary biogenic aerosol particles* (abbreviated PBAP). They comprise plant and insect debris, pollen (a fine powder produced by seed plants that contains the male gametes that serve for reproduction), spores (a reproduction cell or organ from many plants and fungi), bacteria and viruses. Once airborne, these particles can be transported by the wind on varying distances depending on their size. Debris are usually larger than 100 μm, pollen, spores and large bacteria are generally in the range of 1–100 μm, while small bacteria and viruses are generally smaller than 1 μm.

Seawater also can contain biological material, some of which is transferred to sea spray aerosols during the emission process. This primary organic matter is found preferentially in particles smaller than 200 nm in diameter (Leck and Bigg 2008) and its amount has been found to depend on the biological activity in ocean waters (Facchini et al. 2008).

Terrestrial and marine ecosystems are also an important source of aerosol precursors. Some species of phytoplankton produce dimethylsulphide (DMS), a gaseous compound that is oxidized in the atmosphere to form sulphur-containing aerosols. Plants and algae emit volatile organic compounds (VOCs) that are oxidized in the atmosphere and condense and contribute organic material to the atmospheric aerosol. These aerosols are referred to as *secondary biogenic aerosols*. Their sizes are typically of the order of a few tenths of a micrometre.

2.2.5 Biomass Burning Aerosols

In biology, biomass refers to material produced by living organisms, but we adjust the definition here to mean all biological (organic) material that comes from the living world and can potentially burn (i.e. vegetation, dead wood, animal dung, peat) while excluding so-called fossil fuels (coal, gas, and oil) that are formed on geological timescales. The burning of biomass generates primary aerosols that stem from the incomplete combustion of the organic matter. Biomass burning aerosols include *organic carbon*, that is associated with hydrogen and oxygen atoms, and *black carbon*, where the carbon content is very high. These aerosols are generally submicronic and are clearly visible in smoke plumes (Fig. 2.1). The sources of biomass burning aerosols are both natural and anthropogenic. The combustion of biomass also emits gaseous compounds, such as volatile organic compounds and sulphur dioxide, which are aerosol precursors.

2.2.6 Aerosols from Fossil Fuel Combustion

The combustion of coal and oil derivatives also produces black carbon and organic carbon, as well as sulphur dioxide that converts into sulphate aerosols. These are essentially submicronic particles, that are also a source of air pollution in developing and industrialized countries (Fig. 2.5). Air pollution due to particles and gas-phase pollutants, such as ozone and nitrogen oxides, is responsible for a wide range of adverse health and environmental effects. Effects on human health include increased respiratory and cardio-vascular diseases and associated mortality. Aerosols and acidic deposition are responsible for damages on historical buildings.

2.3 Spatial and Temporal Aerosol Distributions

Once they are in the atmosphere, aerosols are transported by winds but are also subject to removal pathways, which are called *atmospheric sinks*. Different aerosol types can stay for very different times in the atmosphere. It is useful to introduce the concept of atmospheric *residence time*[1], which is the average amount of time an aerosol spends in the atmosphere, and is defined as

$$\tau^* = B/S = B/R \tag{2.1}$$

[1] Residence time is a property that refers to the global system or a region (reservoir) big enough for which advective terms in and out of the considered region can be considered negligible in comparison to the source and sink terms integrated over the region. It should not be confused with the concept of *lifetime* which is defined in Chap. 4.

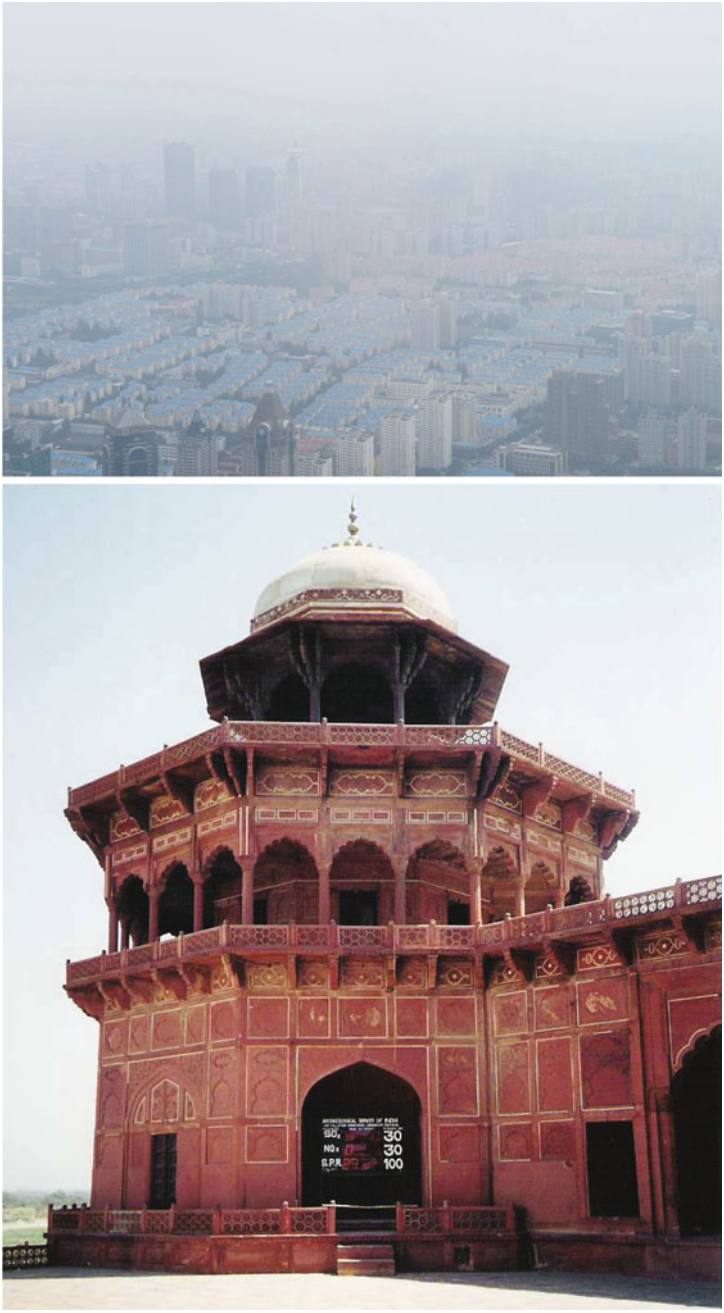


Fig. 2.5 Atmospheric pollution over the city of Shanghai in China (*top panel*, photograph by Eulalie Boucher). Air quality monitoring screen at the Taj Mahal entrance in India (*bottom panel*, photograph by the author). Atmospheric pollution has harmful effects on human health and terrestrial ecosystems. Acid deposition can also have deleterious effects on historical buildings

where B is the vertically-integrated aerosol concentration (also called the aerosol *burden* or *column*), S is the source flux and R is the removal or sink flux. The quantities are averaged over the globe and can be expressed as averages per unit surface (e.g. in units of $\mathrm{kg\,m^{-2}}$ for B and $\mathrm{kg\,m^{-2}\,s^{-1}}$ for fluxes) or as integrated quantities over the globe (e.g. in units of kg for B and $\mathrm{kg\,s^{-1}}$ for fluxes). If the source and sink terms are averaged over a long enough period, they should be balanced and the residence time can be computed from either one or the other quantity. Schematically, in the troposphere, the smallest and largest aerosols have residence times of hours to days, while the bulk of the aerosols in the intermediate size range has residence times of days to 1–2 weeks.

The spatial and temporal distribution of aerosols thus strongly depends on the distribution of sources and sinks and how these interact with transport. The most important sinks are *dry deposition* at the surface and *wet deposition* from precipitations. Figure 2.6 shows spatial aerosol distributions as predicted by an aerosol model constrained by satellite observations. One recognizes plumes of industrial aerosols over Asia and to a lesser extent over Europe and North America, plumes of biomass burning aerosols over South America and Africa during the dry season, and a plume of desert dust originating from the Saharan region. The size of the aerosol plumes reflects the strength of the sources and the time spent in the atmosphere before removal. One can note seasonal variations in the aerosol quantities (in Fig. 2.6 between the months of January and July) but also interannual variations (here between years 2004 and 2007) that are due to variations in the intensity of sources and sinks in response to variations in the meteorology and human activity.

Aerosol properties evolve during atmospheric transport. As an example, we show on Fig. 2.7 the aerosol size distribution (expressed in both aerosol number and aerosol volume concentrations) in three different locations across the Atlantic Ocean. It can be observed that the average size of the aerosols increases as the aerosols are advected off the coast. The concentration of aerosols simultaneously decreases because dry and wet depositions act as sink terms. More details on aerosol size distributions are provided in Chap. 3 and appendix B.

The concentrations and properties of aerosols also vary in the vertical. Most aerosol sources being located at the Earth's surface, one expects aerosol concentrations to decrease with height. Concentrations are generally larger in the atmospheric boundary layer than in the free troposphere. However it is not uncommon for aerosol plumes to be lifted into the free troposphere (as shown for instance in Fig. 2.8). There exists also a local maximum of concentrations in the stratosphere where aerosols are trapped by much less efficient sink terms (see Chap. 12).

2.4 Aerosol–Cloud–Radiation Interactions

Interactions between aerosols, clouds, and electromagnetic radiation are schematized in Fig. 2.9. Aerosols interact with electromagnetic radiation that propagates in the atmosphere in several ways. Solar (shortwave) radiation that interacts with aerosols is deflected in all directions but anisotropically. This process is called *scattering*. Some

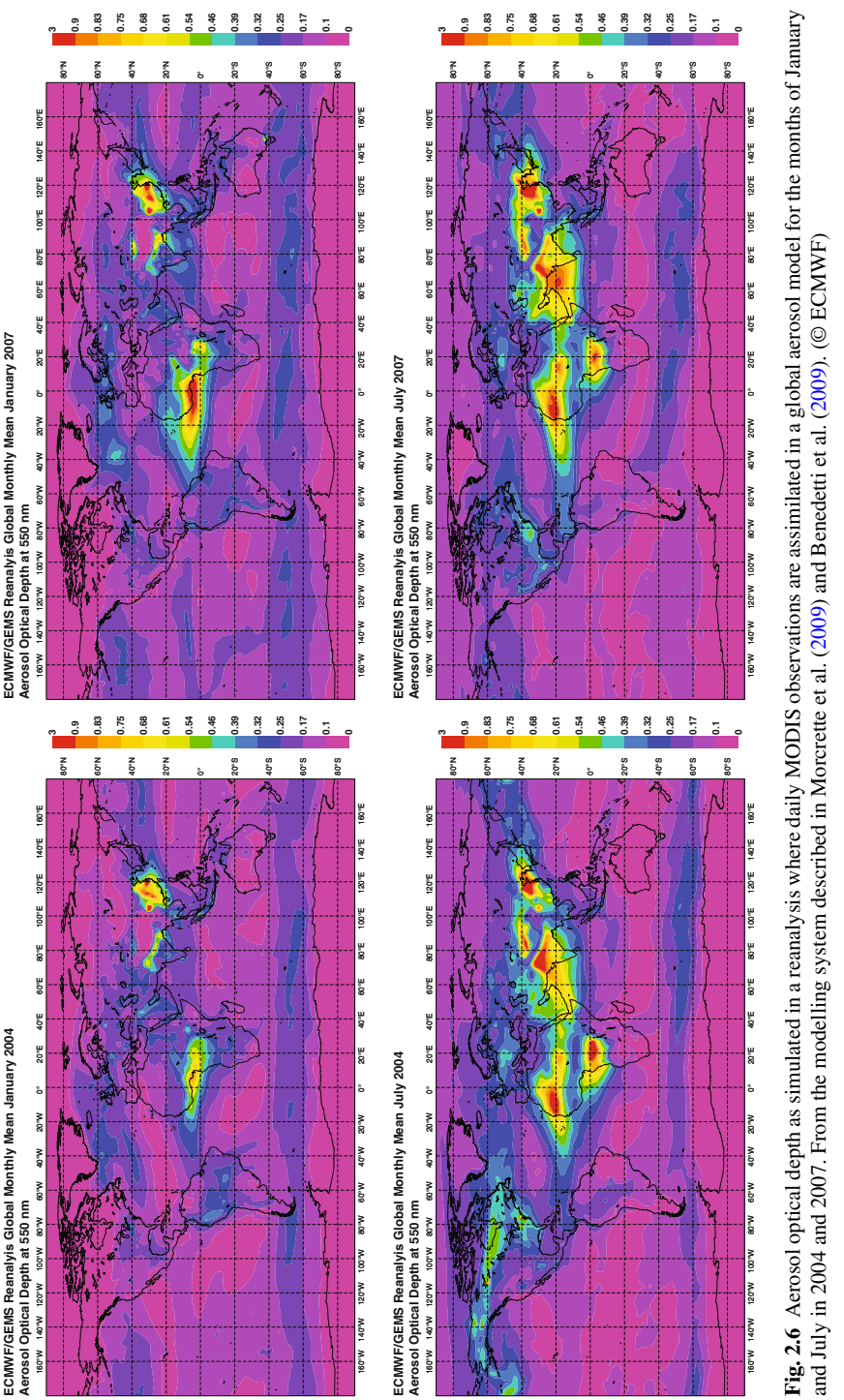


Fig. 2.6 Aerosol optical depth as simulated in a reanalysis where daily MODIS observations are assimilated in a global aerosol model for the months of January and July in 2004 and 2007. From the modelling system described in Morcrette et al. (2009) and Benedetti et al. (2009). (© ECMWF)

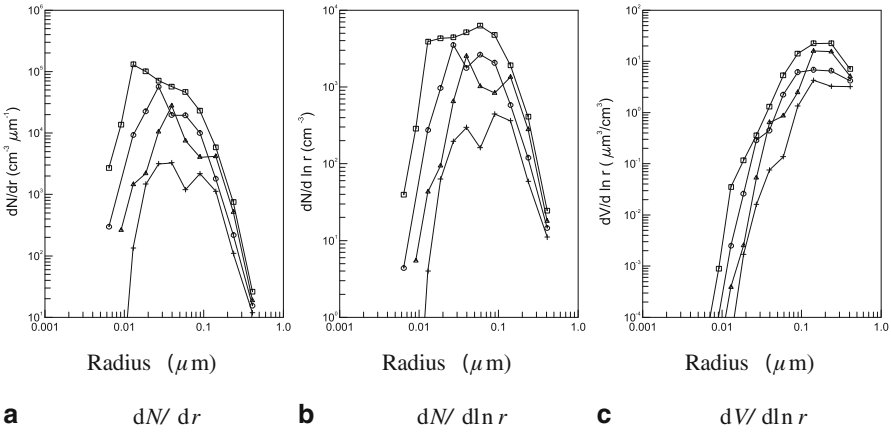


Fig. 2.7 Examples of aerosol number and volume size distributions as measured by Hoppel et al. (1990) during a field campaign across the Atlantic Ocean. These distributions illustrate the transition between a continental environment (with large concentrations of aerosols) to a maritime environment (with smaller concentrations of aerosols). The □, ◯, △, and + symbols show four different locations ranging from the East American coast to the middle of the Atlantic Ocean. (Data provided by Glen Frick. See also Fig. 3.1)

aerosols also absorb solar radiation, thereby transforming the electromagnetic energy into heat, a process called *absorption*. Aerosol also scatter and absorb terrestrial (longwave) radiation emitted by the Earth's surface and atmosphere, and can emit such radiation. These interactions are described in more detail in Chaps. 3, 5, and 8.

Aerosols influence the microphysical properties of liquid clouds, in particular through their role as *cloud condensation nuclei* upon which water vapour can condense, which controls in part the concentration and size of the population of cloud water droplets. The chemical composition of the aerosol can also determine the degree of acidity of the cloud droplets and control the rate of some of the chemical reactions that take place in the droplets. Aerosols also modify the microphysical properties of ice clouds because a small fraction of the aerosols act as *ice nuclei* that favour the formation of the ice phase. For similar reasons, aerosols influence the properties of mixed-phased clouds (i.e. those clouds that contain both liquid droplets and ice crystals). These aerosol–cloud interactions are described in more detail in Chap. 9.

Conversely clouds influence aerosol populations. Precipitating clouds contribute to remove aerosols from the atmosphere. One differentiates *in-cloud scavenging*, whereby aerosols enter water droplets within the clouds before precipitation, from *below-cloud scavenging*, whereby aerosols are captured by precipitating raindrops themselves. The same holds for ice clouds although scavenging by ice crystals and falling snowflakes is much less efficient. Clouds therefore constitutes an important sink for aerosols through wet deposition. Non-precipitating clouds will eventually evaporate and release cloud droplet residues in the atmosphere, which turn into aerosols and can serve as cloud condensation nuclei in a future cloud. Because sulphur

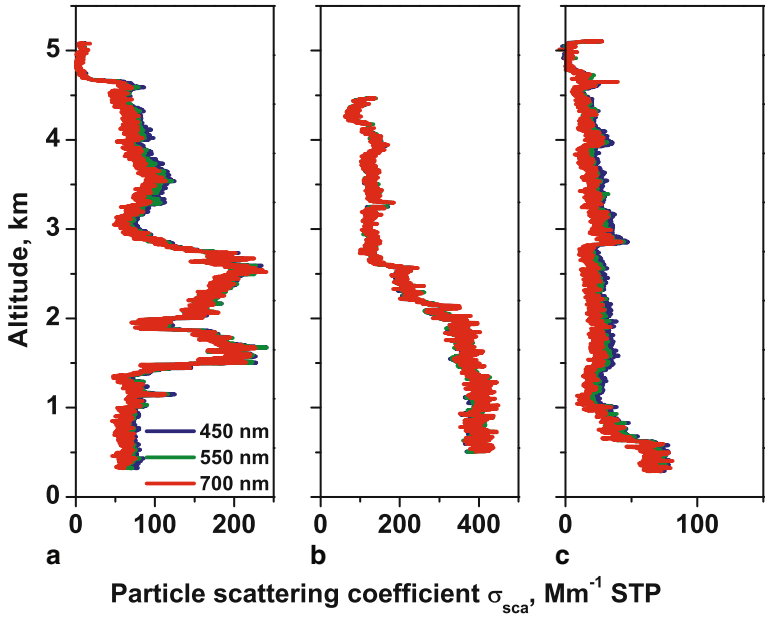


Fig. 2.8 Vertical profiles of the scattering coefficient, $\sigma^{\text{scattering}}(z)$, at 450, 550, and 700 nm (*blue*, *green*, and *red lines*, respectively) from desert dust aerosols during the AMMA field campaign in West Africa. The three profiles correspond to **a** dust aerosols transported at 14°N to the south of the intertropical discontinuity in the Saharan air layer between 1.5 and 5.5 km altitude, **b** dust aerosols transported in Saharan boundary layer north of the intertropical discontinuity, and **c** dust aerosols emitted from the Sahel region following the passage of a mesoscale convective system. The scattering coefficient is expressed in Mm^{-1} at standard temperature and pressure conditions ($1\ \text{Mm}^{-1} = 10^{-6}\ \text{m}^{-1}$). (© LISA/CNRS)

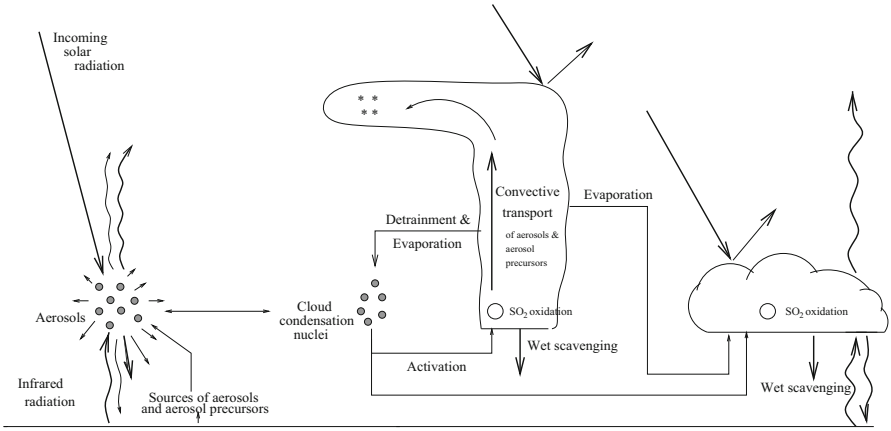


Fig. 2.9 Schematic of aerosol–cloud–radiation interactions. Aerosols interact with both solar (short-wave) and terrestrial (longwave) radiation. There are also two-way interactions with clouds as described in the text

dioxide (SO_2) is readily oxidized in cloud droplets, other interstitial aerosol material can be incorporated into cloud droplets, and coalescence of cloud droplets acts to form larger droplets, the chemical and physical properties of an aerosol population that has gone through a condensation/evaporation cycle can change substantially. The size distribution shifts to larger sizes, the aerosols get more mixed, with more soluble material in the larger size range. This cloud processing effect is particularly important in convective clouds, which are capable of activating small aerosols and release larger aerosols that can serve as cloud condensation nuclei in other clouds.

Clouds scatter solar radiation much more than they absorb it, and this results in clouds reflecting a significant fraction of solar radiation back to space. This leads to a cooling of the climate system. Clouds also absorb and emit terrestrial radiation. They are therefore responsible for a greenhouse effect that leads to a warming of the climate system. One traditionally defines the *cloud radiative effect*[2] as the difference in net[3] radiative fluxes at the top of atmosphere with and without clouds:

$$\mathrm{CRE} = F^{\text{all sky}} - F^{\text{clear sky}} \approx f\,(F^{\text{cloudy sky}} - F^{\text{clear sky}}) \tag{2.2}$$

where f is the fractional cloud cover. The cloud radiative effect can be decomposed in a (negative) shortwave component and a (positive) longwave component. For low clouds, the shortwave component wins over the longwave component and the net effect of such clouds is to cool the climate system. Inversely, for high clouds that are not too thick, the longwave components wins and the net effect of such clouds is to warm the climate system. Overall the net effect between cooling and warming depends on the thickness and height of the cloud, but also on the cloud microphysical and optical properties.

Conversely, radiation is an important factor that governs the evolution of a cloud. Heating, through absorption of radiation, and cooling, through emission of terrestrial radiation, modify the vertical temperature profile in the cloud, and hence its stability.

In conclusion there is a complex set of interactions between aerosols, clouds, and radiation. It is important to have a good grasp of these processes to understand the physics and chemistry of the atmosphere in general, and the role of aerosols in the climate system in particular.

2.5 Climate Effects of Aerosols

As shown in Fig. 2.10, aerosols influence climate in a number of ways:

1. They scatter and absorb solar radiation. Backscattering of solar radiation towards space results in a reduction of incoming solar radiation at the Earth's surface, a loss of energy and a cooling of the climate system. Absorption of solar radiation

[2] The term *cloud radiative effect* is preferred to the more usual term of *cloud radiative forcing* because clouds do not exert a climate forcing as defined in Chap. 1.

[3] The term "net" refers here to energy gained by the surface–atmosphere system.

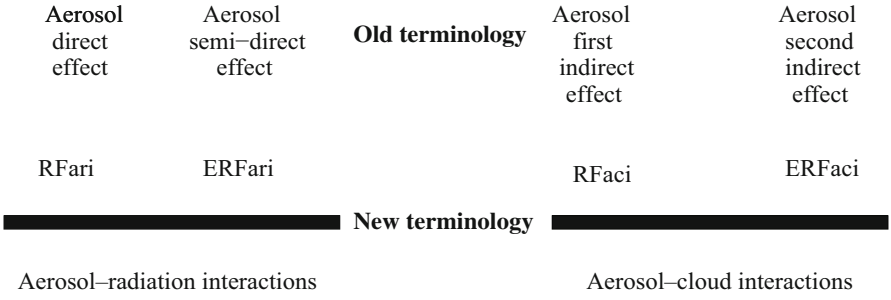


Fig. 2.10 Aerosol–radiation interactions and aerosol–cloud–radiation interactions and how they relate to the original terminology

is accompanied by a heating within the aerosol layer, but also by a reduction of incoming solar radiation at the Earth's surface. Such effects occur preferentially, but not uniquely, in clear sky conditions. These processes have traditionally been called the *aerosol direct effect* but the more precise term of *aerosol–radiation interactions* is preferred in this book.

2. The absorption of solar radiation by aerosols modifies the vertical temperature profile. This impacts the relative humidity, atmospheric stability, and therefore cloud formation. This effect has traditionally been called the *aerosol semi-direct effect*, but it can also be seen as a *rapid adjustment* of the atmospheric state that follows aerosol–radiation interactions.
3. Aerosols serve as cloud condensation nuclei in liquid water clouds, thus exerting a partial control of cloud microphysical and optical properties. An increase in the concentration of aerosols leads to an increase in the concentration of cloud condensation nuclei, and generally to an increase in the concentration of cloud droplets. For a fixed cloud liquid water content, this is accompanied by a reduction in the cloud droplet size and an increase in the cloud reflectivity. Altogether this leads to less solar energy absorbed and a cooling of the climate system. This effect has been traditionally called the aerosol *first indirect effect*, but we use here the more general term of *aerosol–cloud interactions* that includes this and related effects.
4. The modification of cloud microphysical properties is expected to have an impact on cloud evolution, in particular in terms of the ability of clouds to generate droplets that are large enough to initiate precipitation. This effect is traditionally called the aerosol *second indirect effect*, but is viewed here as a rapid adjustment (or a series of rapid adjustments) that stem from the role of aerosols as cloud condensation nuclei.
5. Aerosols can impact the properties of mixed-phased and ice clouds, in particular through their role as ice nuclei. This has been referred to in the past as yet another indirect effect of aerosols, or a *glaciation effect*. It is also part of *aerosol–cloud interactions*.
6. Absorbing aerosols may deposit onto snow and ice surfaces, thus making these surfaces less reflective (i.e. more absorbing). This contributes to warm the surface

and thus the climate system. This effect is known as the *aerosol on snow effect* but can be categorized as part of *aerosol–radiation interactions* or *aerosol–surface interactions* in our new terminology.

7. Finally, aerosols also interact with vegetation through changes in incoming solar radiation, fraction of diffuse radiation and as a source of nutrients. These are other contributions to *aerosol–surface interactions*.

As shown in Fig. 2.10, these multiple effects can be grouped into three main types of interactions: aerosol–radiation interactions (including the direct effect, the semi-direct effect and the aerosol on snow effect), aerosol–cloud interactions (including the first and second aerosol effect and the glaciation effect), and aerosol–surface interactions (effects on snow and vegetation). We follow this classification and discuss aerosol radiative effects in Chap. 8 and aerosol–cloud interactions in Chap. 9. However it is now necessary to introduce a number of concepts and tools to study and quantify the effects of aerosols on climate, which is the purpose of the next chapter.

Exercises

1. The flux of desert dust aerosols to the atmosphere is estimated to be 1000 Tg per year. Their residence time in the atmosphere is 2 days. Compute the burden of desert dust aerosols in units of Tg and mg m^{-2}. The radius of the Earth is taken equal to 6370 km.

2. The flux of SO_2 to the atmosphere is equal to 80 Tg S per year. Only 75% of the emitted SO_2 is oxidized into sulphate aerosols, the rest gets deposited to the surface by dry and wet deposition with a 4:1 ratio. Compute the wet deposition flux of SO_2 in units of mol S m^{-2} $year^{-1}$ and mg SO_2 m^{-2} $year^{-1}$?

3. An aerosol is eliminated from the atmosphere at a rate of 10 % day^{-1} due to dry scavenging and 20 % day^{-1} due to wet scavenging. Other aerosol sinks are neglected. Compute the aerosol lifetime with respect to dry scavenging, τ_{dry}, and its lifetime with respect to wet scavenging, τ_{wet}, and its residence time, τ^*.

Solutions

1. $B = 5.5$ Tg $= 11$ mg m^{-2}.

2. $F^{SO_2}_{\text{wet deposition}} = 2.5\ 10^{-4}$ mol S m^{-2} $year^{-1} = 16$ mg SO_2 m^{-2} $year^{-1}$.

3. $\tau_{\text{dry}} = C/[\partial C/\partial t]_{\text{dry}} = 1/0.1 = 10$ days.
$\tau_{\text{wet}} = C/[\partial C/\partial t]_{\text{wet}} = 1/0.2 = 5$ days.
$1/\tau^* = 1/\tau_{\text{dry}} + 1/\tau_{\text{wet}}$, i.e. $\tau^* = 3.3$ days.

References

Benedetti A, Morcrette J-J, Boucher O, Dethof A, Engelen RJ, Fischer M, Flentjes H, Huneeus N, Jones L, Kaiser JW, Kinne S, Mangold A, Razinger M, Simmons AJ, Suttie M, GEMS-AER Team (2009) Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part II: data assimilation. J Geophys Res 114:D13205. doi:10.1029/2008JD011115

Burrows SM, Butler T, Jöckel P, Tost H, Kerkweg A, Pöschl U, Lawrence MG (2009) Bacterial in the global atmosphere—part 2: modeling of emissions and transport between different ecosystems. Atmos Chem Phys 9:9281–9297

Dentener F, Kinne S, Bond T, Boucher O, Cofola J, Generoso S, Ginoux P, Gong S, Hoelzemann JJ, Ito A, Marelli L, Penner JE, Putaud J-P, Textor C, Schulz M, van der Werf GR, Wilson J (2006) Emissions of primary aerosol and precursor gases in the years 2000 and 1750, prescribed datasets for AeroCom. Atmosp Chem Phys 6:4321–4344

Facchini MC, Rinaldi M, Decesari S, Carbone C, Finessi E, Mircea M, Fuzzi S, Ceburnis D, Flanagan R, Nilsson ED, de Leeuw G, Martino M, Woeltjen J, O'Dowd CD (2008) Primary submicron marine aerosol dominated by insoluble organic colloids and aggregates. Geophys Res Lett 35:L17814. doi:10.1029/2008GL034210

Guenther A, et al (1995) A global model of natural volatile organic compound emissions. J Geophys Res 100:8873–8892

Heald CL, Spracklen DV (2009) Atmospheric budget of primary biological aerosol particles from fungal spores. Geophys Res Lett 36:L09806. doi:10.1029/2009GL037493

Hoppel WA, Fitzgerald JW, Frick GM, Larson RE, Mack EJ (1990) Aerosol size distributions and optical properties found in the marine boundary layer over the Atlantic Ocean. J Geophys Res 95:3659–3686

Jaenicke R (2005) Abundance of cellular material and proteins in the atmosphere. Science 308:73

Leck C, Bigg EK (2008) Comparison of sources and nature of the tropical aerosol with the summer high Arctic aerosol. Tellus 60B:118–126

Morcrette J-J, Boucher O, Jones L, Salmond D, Bechtold P, Beljaars A, Benedetti A, Bonet A, Kaiser JW, Razinger M, Schulz M, Serrar S, Simmons AJ, Sofiev M, Suttie M, Tompkins AM, Untch A (2009) Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Forward modelling. J Geophys Res 114:D06206. doi:10.1029/2008JD 011235

Penner JE, Andreae M, Annegarn H, Barrie L, Feichter J, Hegg D, Jayaraman A, Leaitch R, Murphy D, Nganga J, Pitari G (2001) Aerosols, their Direct and Indirect Effects, in: Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, edited by: J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden and D. Xiaosu, Cambridge University Press, UK, Chap. 5, pp 289–348

Further Reading (Textbooks and Articles)

Haywood JM, Boucher O (2000) Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: a review. Rev Geophys 38:513–543

Chapter 3
Physical, Chemical and Optical Aerosol Properties

Abstract This chapter provides the tools to describe aerosol properties qualitatively and quantitatively. It introduces a number of characteristic radii of the aerosol size distribution and different possible expressions for the size distribution. The main chemical (inorganic, organic and mineral) species composing the atmospheric aerosols are reviewed and the concepts of aerosol mixture (external and internal), hygroscopic growth, deliquescence and efflorescence are introduced. Aerosol optics is first discussed for a single spherical particle, introducing the concepts of refractive index, scattering and absorption cross sections, phase function, upscatter fraction. Optical properties for an aerosol population are also presented leading to the definitions of aerosol mass and volume extinction efficiency, extinction coefficient, optical depth and Ångström coefficient. Nonspherical particles and the link between aerosol extinction and visibility are briefly mentioned at the end of the chapter.

Keywords Size distribution · Chemical composition · Refractive index · Aerosol mixture · Scattering · Absorption · Cloud condensation nucleus · Ice nucleus · Phase function

This chapter introduces, in a simplified manner, the physical, chemical and optical properties of the aerosol. The reader looking for a more complete presentation of aerosol optics can go back and forth between this chapter and Chap. 5 that deals with radiative transfer in the atmosphere and makes the link to aerosol optics.

3.1 Fine, Accumulation and Coarse Modes

While manufactured aerosols in a laboratory can be monodisperse, atmospheric aerosols are polydisperse, i.e. they exhibit a range of sizes. Therefore, an aerosol population is first of all characterized by its size distribution. Aerosol sizes span several orders of magnitude, from just a few nanometres for new particles produced by nucleation, to tens or hundreds of micrometres for the largest particles produced by the wind friction on the land and the ocean surface. It is possible to distinguish several aerosol classes according to their size (see Whitby (1978) for early work on this). More specifically, one observe up to five modes, here defined from the smallest to the largest: the *nucleation mode* (also called the *ultrafine mode*), the

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_3

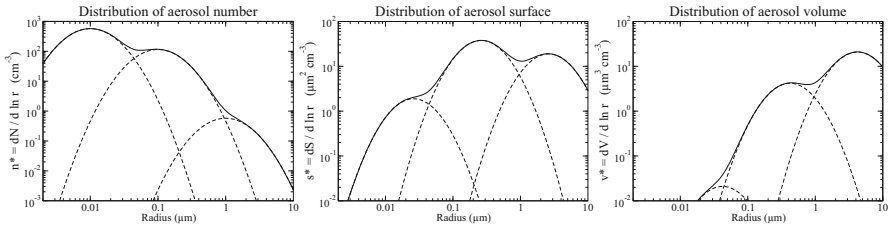


Fig. 3.1 Schematic of the three modes of the aerosol size distribution. The fine mode dominates the aerosol number distribution ($n^*(r) = \mathrm{d}N/\mathrm{d}\ln r$), the accumulation mode dominates the aerosol surface distribution ($s^*(r) = \mathrm{d}S/\mathrm{d}\ln r$) while the coarse mode dominates the aerosol volume distribution ($v^*(r) = \mathrm{d}V/\mathrm{d}\ln r$). For illustration we have superposed three log-normal distributions with mean geometric number radius r_g of 0.01, 0.1 and 1 μm, geometric standard deviation σ_g of 2 and with total concentrations of 1000, 200 and 1 cm^{-3}, respectively. The ultrafine and supercoarse modes that may also exist are not represented in this schematic

Aitken mode (named after the Scottish meteorologist and physicist John Aitken), the *accumulation mode* (so named because mass accumulates in this size range by coagulation and condensation), the *coarse mode*, and the *supercoarse mode*. Each of these modes corresponds, at least in principle, to a relative maximum of the aerosol size distribution (Fig. 3.1). The number of modes that are visible in the observations can however vary, but it is usual to see three or more modes. The nucleation and Aitken modes appear clearly in the aerosol number distribution, with particle radius in the range from a few nanometres to 0.05 μm. The accumulation mode appears generally in the aerosol surface distribution, but is sometimes also visible in the aerosol number distribution. It is comprised of particles with radii ranging between 0.05 and 0.5 μm. The coarse mode generally dominates the aerosol volume size distribution. It includes particles beyond 0.5 μm radius (or 1 μm diameter). A supercoarse mode can be found close to the source point but is generally absent from a more aged aerosol population. Sometimes, the Aitken and accumulation modes are referred together as the *fine mode* but the terms are used somewhat loosely and interchangeably.

3.2 Size Distribution

The aerosol size distribution is represented by a function $n(r)$ which is such that $n(r)\,\mathrm{d}r$ is the number of particles per unit volume whose radius lies between r and $r + \mathrm{d}r$. The total aerosol concentration can be obtained by integrating over the size distribution: $N_0 = \int n(r)\,\mathrm{d}r$. An aerosol population covers several orders of magnitude so that it is more practical to use a logarithmic scale to describe the aerosol size distribution. Letting $n^*(r)$ be such a size distribution, $n^*(r)\,\mathrm{d}\ln(r)$ is the number of particles per unit volume whose logarithm of the radius lies between $\ln r$ and $\ln r + \mathrm{d}\ln r$. The total aerosol concentration can be obtained by integration: $N_0 = \int n^*(r)\,\mathrm{d}\ln r$.

As for any distribution, one can define a number of statistical parameters. The *median* radius is such that half of the particles have a smaller radius and the other half have a larger radius than this median radius. The *mode* is the radius for which $n(r)$ presents a maximum. There could be several relative maxima in the $n(r)$ function, in which case one talks about a bimodal, trimodal or more generally a multimodal size distribution. By definition, the i^{th}-order moment of the size distribution is:

$$M_i = \int r^i \, n(r) \, \mathrm{d}r = \int r^i \, n^*(r) \, \mathrm{d} \ln r \tag{3.1}$$

The mean radius of the size distribution is given by:

$$r_{\text{mean}} = \frac{\int n(r) \, r \, \mathrm{d}r}{\int n(r) \, \mathrm{d}r} \, . \tag{3.2}$$

The mean surface radius is defined as:

$$r_{\text{surface}} = \left(\frac{\int n(r) \, r^2 \, \mathrm{d}r}{\int n(r) \, \mathrm{d}r} \right)^{\frac{1}{2}} = \left(\frac{M_2}{M_0} \right)^{\frac{1}{2}} . \tag{3.3}$$

And finally the mean volume radius can be expressed as:

$$r_{\text{volume}} = \left(\frac{\int n(r) \, r^3 \, \mathrm{d}r}{\int n(r) \, \mathrm{d}r} \right)^{\frac{1}{3}} = \left(\frac{M_3}{M_0} \right)^{\frac{1}{3}} . \tag{3.4}$$

However, these quantities are not practical to describe log-normal size distributions and the concept of mean geometric radii is preferred. The mean geometric (number) radius, noted r_g or sometimes r_{gn}, verifies the following equation:

$$\ln r_{gn} = \frac{\int \ln r \; n^*(r) \, \mathrm{d} \ln r}{\int n^*(r) \, \mathrm{d} \ln r} \, . \tag{3.5}$$

Similarly, the mean geometric volume verifies:

$$\ln r_{gv} = \frac{\int \ln r \; r^3 \, n^*(r) \, \mathrm{d} \ln r}{\int r^3 \, n^*(r) \, \mathrm{d} \ln r} \, . \tag{3.6}$$

By analogy to the standard deviation, the geometric standard deviation verifies:

$$(\ln \sigma_g)^2 = \int (\ln r)^2 \, n^*(r) \, \mathrm{d} \ln r - \left(\int \ln r \; n^*(r) \, \mathrm{d} \ln r \right)^2 . \tag{3.7}$$

Finally, the effective radius and effective variance are further quantities that can be found in the scientific literature. They are defined as:

$$r_e = \frac{\int r \, \pi \, r^2 \, n(r) \, \mathrm{d}r}{\int \pi \, r^2 \, n(r) \, \mathrm{d}r} \tag{3.8}$$

and

$$v_e = \frac{\int (r - r_e)^2 \, \pi \, r^2 \, n(r) \, \mathrm{d}r}{r_e^2 \int \pi \, r^2 \, n(r) \, \mathrm{d}r}. \tag{3.9}$$

Various size distributions have been proposed to represent aerosols:

The *Junge law* describes the aerosol size distribution by a simple dependence of $n^*(r)$ in $r^{-\alpha}$, where α is generally between 2 and 3 and r is limited to vary between two radii r_1 and r_2. If N_0 is the total aerosol concentration the Junge law can be expressed as:

$$n^*(r) = N_0 \frac{\alpha \, r_1^\alpha \, r_2^\alpha}{r_2^\alpha - r_1^\alpha} \, r^{-\alpha}. \tag{3.10}$$

This law explains reasonably well the spectral dependence of aerosol light scattering but does neither permit to extend the size distribution to small sizes ($r < 0.1$ μm), nor describe a multimodal size distribution.

A *modified Gamma law* can describe a large number of actual size distributions through a rather large number of fitting parameters. The superposition of several distributions allows to represent multimodal distributions. However the number of fitting parameters becomes large, which is inconvenient. The *Gamma law* is often preferred to the modified Gamma law because it has a more limited number of fitting parameters and because the effective radius and effective variance are readily identifiable. The Gamma law can be written under the form:

$$n(r) = \frac{N_0 \, \beta^{\alpha+1}}{\Gamma(\alpha + 1)} \, r^\alpha \, \exp(-\beta \, r) \tag{3.11}$$

where α and β are two parameters, and Γ is the mathematical Gamma function that generalizes the factorial operator to real numbers. Alternatively, one can write the Gamma law in the following way:

$$n(r) = \frac{N_0 \, (ab)^{(2b-1)/b}}{\Gamma[(1 - 2b)/b]} \, r^{(1-3b)/b} \, \exp(-r/ab) \tag{3.12}$$

where a and b are two adjustable parameters that control the mean size and the width of the size distribution. It can be shown by differentiating the function $n(r)$ that the mode of the size distribution is equal to $a(1 - 3b)$. The advantage of this formulation is that the effective radius and effective variance are precisely a and b, respectively.

Finally, the *log-normal law* is the most used distribution law because of its many advantages. It appears to be very universal and can describe observed size distributions fairly well especially those spanning many orders of magnitude. The log-normal size distribution is entirely determined by the choice of two parameters, a characteristic radius r_0 and a measure of the dispersion of the distribution σ_0:

$$n^*(r) = \frac{\mathrm{d}N}{\mathrm{d}\ln r} = N_0 \, \frac{1}{\sqrt{2\pi} \, \sigma_0} \exp\left(-\frac{1}{2} \left(\frac{\ln(r/r_0)}{\sigma_0} \right)^2 \right). \tag{3.13}$$

This distribution is convenient because r_0 happens to equal the mean geometric radius and $\sigma_g = \exp(\sigma_0)$ is the geometric standard deviation as defined above. Moreover, the lognormal law has interesting mathematical properties: the distributions of the various moments also follow lognormal laws, while both the median and the mode are equal to r_0. The effective radius, r_e, is equal to $r_0 \exp\left(\frac{5}{2}\sigma_0^2\right)$ and the effective variance, v_e, is $\exp(\sigma_0^2) - 1$. The parameters and properties of the lognormal law are described in more details in Appendix B.

3.3 Chemical Composition

The chemical composition of the aerosol is another key parameter that controls the *hygroscopicity* of the aerosol (i.e. its ability to take up water and grow in size as ambient relative humidity increases), the ease with which it can be scavenged by rain, its ability to serve as cloud condensation nucleus or ice nucleus, or its optical properties.

The previous chapter has already provided a quick overview of the chemical fractions of certain aerosol types such as mineral aerosols (inorganic), sea spray (mostly inorganic with some organic additions), biogenic aerosols (mostly organic), industrial aerosols (a mix of inorganic and organic material, with volatile and non-volatile components), and biomass burning aerosols (mostly organic material, with again volatile and nonvolatile components). We now introduce the concept of mixing before discussing further the chemical composition of atmospheric aerosols.

3.3.1 Aerosol Mixture

It is usual to distinguish and oppose *external* and *internal mixtures*. In an external mixture, particles are chemically pure and the mixture is comprised of particles of different chemical compositions. It is therefore possible to differentiate the particles according to their distinct chemical composition: e.g., black carbon, ammonium sulphate, sodium chloride, organics. In the case of an external mixture, one can define a different size distribution for each aerosol type. In an internal mixture, different chemical species are mixed within each particles. If the internal mixture is perfect, then all particles have the same chemical composition and the aerosol population can be characterized by a single size distribution.

External and internal mixtures are conceptual models: reality lies somewhere in between these two extremes. As a general rule, the aerosol chemical composition varies both as a function of size (some chemical species are found preferentially in some size ranges) and within a size class (there are different degrees of mixture for a given size). The different types of aerosol mixtures are illustrated on Fig. 3.2.

Aerosol mixing occurs through coagulation of particles of different chemical composition, both outside and inside clouds, and by condensation of semi-volatile

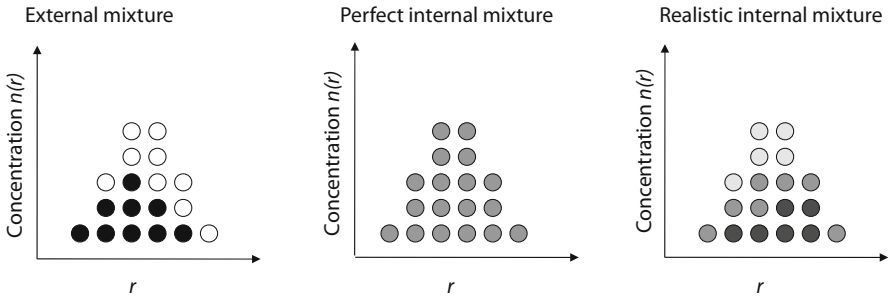


Fig. 3.2 Schematic representation of three different aerosol mixtures. The more realistic internal mixture is comprised of internally-mixed particles but with varying degrees of mixing for a given aerosol size and across the size distribution

compounds on pre-existing particles. For instance, non-soluble primary aerosols, such as black carbon and mineral aerosols, can be covered by a coating of soluble species, which change their properties such as hygroscopicity.

3.3.2 Inorganic Aerosols

The presence of inorganic material is very usual in aerosols. This is the case of sea spray composed mainly of sodium chloride, NaCl, but also of sulphate aerosols whose chemical composition depends on the degree of neutralization of the sulphuric acid. The chemical composition of sulphate therefore varies between pure sulphuric acid (H_2SO_4), ammonium bisulphate (NH_4HSO_4) and ammonium sulphate ($(NH_4)_2SO_4$). Beyond the sulphate (SO_4^{2-}) and ammonium (NH_4^+) ions, one can also find the nitrate ion (NO_3^-) so that a SO_4^{-2}–NH_4^+–NO_3^- mixture is common. The physical, chemical and optical properties of inorganic aerosols (density, water uptake, refractive index) are relatively well-known from measurements and modelling of inorganic mixture (e.g. Tang and Munkelwitz 1994; Wexler and Clegg 2002). Inorganic aerosols are generally very hygroscopic, and do not absorb light much, if at all, in the solar spectrum.

3.3.3 Black Carbon Aerosols

Black carbon is an aerosol species with unique properties. It is produced during the combustion of long carbon-chain fuels, where and when the combustion is incomplete because there is a lack of oxygen. Incomplete combustions are common, e.g., when biomass is burned, but also in stoves or even in certain parts of the most sophisticated engine. The reaction chain starts with the production of polycyclic aromatic hydrocarbons which serve as nuclei to form small spherical particles of

a few nanometres with graphite layers and a large C:H ratio (Bond et al. 2013). These spheres coagulate to form a black carbon chain. The density of black carbon is not well measured and is likely to vary with the combustion conditions. Bond and Bergstrom (2006) recommend a value of 1.7–1.9 gcm^{-3}. The size of black carbon particles vary with the source type but is typically of 100 nm. Very quickly the black carbon gets coated with organic carbon and sulphate which condense at their surface. This creates an internal mixture which modifies the microphysical and optical properties of the black carbon, that also becomes more hygroscopic.

3.3.4 Organic Aerosols

Although the presence of organic material in the aerosol has been known for a long time, it has taken decades for aerosol research to measure the concentration and composition of organics more accurately (Jacobson et al. 2000). Mass spectrometry on bulk and individual aerosol particles have permitted substantial progress in the characterization and understanding of organic aerosols. They contain a tremendous variety of chemical compounds (alcanes, alcenes, alcohols, aromatic compounds, carbonylated compounds, organic acids, amines, organosulphurated compounds, ...), for which it is illusionary to monitor and represent the full complexity. It is therefore important to seek, to simplify the way the chemical composition of these compounds is characterized. One can distinguish primary from secondary organic aerosols. Secondary organic aerosols originate from volatile organic compounds that are oxidized in the atmosphere, some of which can actually evaporate from primary organic aerosols. The chemical composition of organic aerosols originating from the combustion of fossil hydrocarbons (*hydrocarbon-like organic aerosols*) differs from that of organic aerosols originating from the combustion of biomass (*biomass burning organic aerosols*). Organic aerosols containing oxygen-rich compounds are called *oxygenated organic aerosols*. The O:C ratio in organic aerosols is actually a useful indicator of the chemical composition. This ratio is larger than 0.25 in oxygenated organic aerosols and can reach values as large as 1 for the more aged aerosols. Jimenez et al. (2009) showed that, as they age in the atmosphere, organic aerosols become more oxygenated, less volatile and more hygroscopic. There is a clear relationship between the aerosol hygroscopicity and the O:C atomic ratio (Fig. 3.3). Jimenez et al. (2009) also showed an inverse relationship between volatility of the organic aerosol and the O:C atomic ratio. Organic aerosols can also been split between semi-volatile compounds and low-volatility compounds.

3.3.5 Geographic Distribution of Aerosol Chemical Composition

The chemical composition of the aerosol can hardly be measured from space. Its knowledge therefore relies on in situ measurements whose aggregation requires an

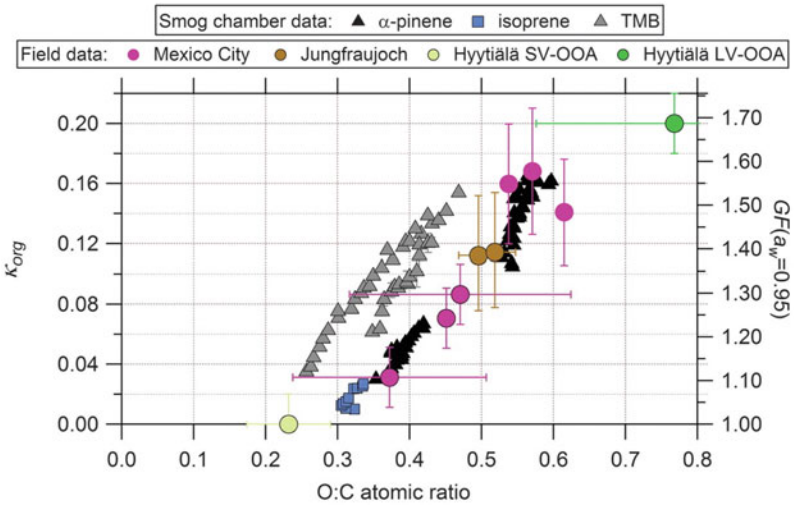


Fig. 3.3 Relationship between aerosol hygroscopicity and the O:C atomic ratio in ambient and laboratory aerosols. The hygroscopicity is characterized here by the size growth factor at 95 % relative humidity (*right axis*) or the κ factor (*left axis*). (From Jimenez et al. (2009). Reproduced with the permission of the AAAS)

important work (Putaud et al. 2004; Jimenez et al. 2009; Boucher et al. 2013). Figure 3.4 shows the chemical composition of the non-refractory fraction of the aerosols in different regions of the northern hemisphere. Figure 3.5 extends this climatology to other continents. It appears very clearly that both the inorganic and organic fractions of the aerosol are important and one or the other can dominate the mass concentration. Sulphate is an important component of the inorganic fraction of the aerosol, and is systematically accompanied by ammonium. Nitrate can be present in very variable quantities but concentrations decrease rapidly outside source regions. Mineral dust dominates over and downwind the emitting desertic regions. Sea-salt, largely composed of sodium chloride, dominates in marine environment.

3.4 Refractive Index

The refractive index is another important parameter to characterize the aerosols because, combined with information on size and shape, it can be used in principle to characterize the aerosol optical properties (see Sect. 3.7 and Appendix C).

The refractive index of a medium characterizes the way this medium, solid or liquid, interacts with electromagnetic radiation. The refractive index is by definition a complex number, denoted $m = n_r - i\, n_i$ (the minus sign results from a convention and a plus sign can also be found in the literature). The real part, n_r, is the ratio of the speed of light in the medium to that in the vacuum. It is a number greater than 1

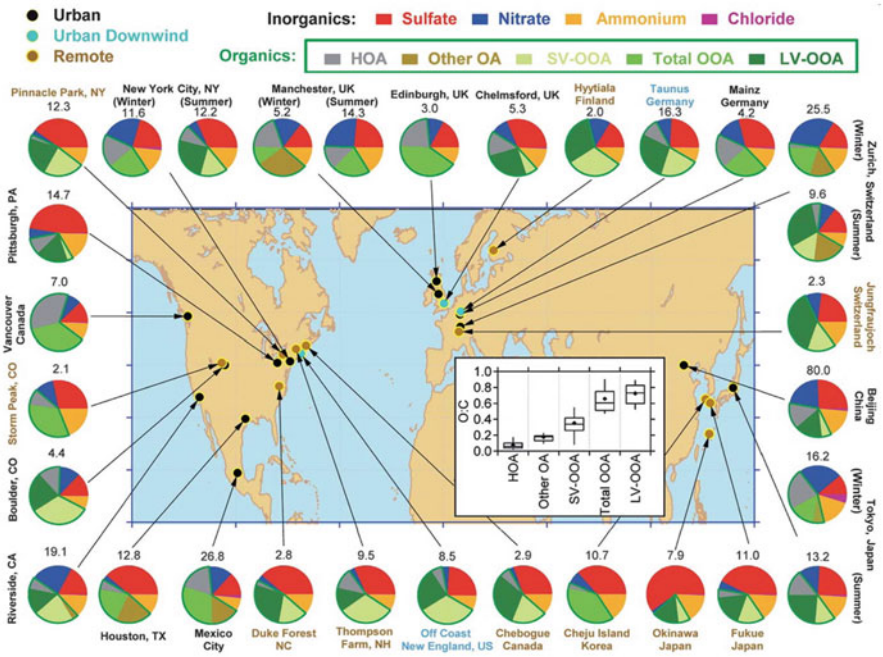


Fig. 3.4 Concentration ($\mu g m^{-3}$) and chemical composition of the non-refractory (inorganic and organic) fraction of the sub-micronic aerosols in different locations of the Northern Hemisphere. *HOA* hydrocarbon-like organic aerosols, *SV-OOA* semi-volatile oxygenated organic aerosols, *LV-OOA* low-volatility oxygenated organic aerosols. (From Jimenez et al. (2009). Reproduced with the permission of the AAAS)

(to a very few exceptions not discussed here). The imaginary part characterizes the degree of absorption of the medium. The refractive index of a medium is a function of the wavelength of the radiation under consideration.

The refractive index of the aerosol depends on its chemical composition. It can be measured reasonably accurately in the laboratory for inorganic compounds, so that it is known to have a good accuracy for inorganic salts as a function of the water content (which itself depends on the ambient relative humidity if the aerosol is in equilibrium with the gas phase). This is the case for sulphate and nitrate salts (Tang and Munkelwitz 1994; Tang 1997) and for sea-salts (Tang et al. 1997). However, pure salts are hardly found in the atmosphere. The partial molar approach allows to estimate the real part of the refractive index of a mixture of inorganic material under certain conditions (Stelson 1990). As an example, Fig. 3.6 shows the spectral variation of the refractive index of sulphate ammonium salt and pure water.

The refractive indices of organic carbon and black carbon aerosols are fairly variable and less well-known than those of inorganic aerosols. Laboratory measurements are not necessarily representative of ambient atmospheric aerosols. As a result, their refractive index is not well characterized and vary in space and time. A typical value of $1.95-0.79i$ for black carbon (soot) aerosols in the wavelengths of the visible can

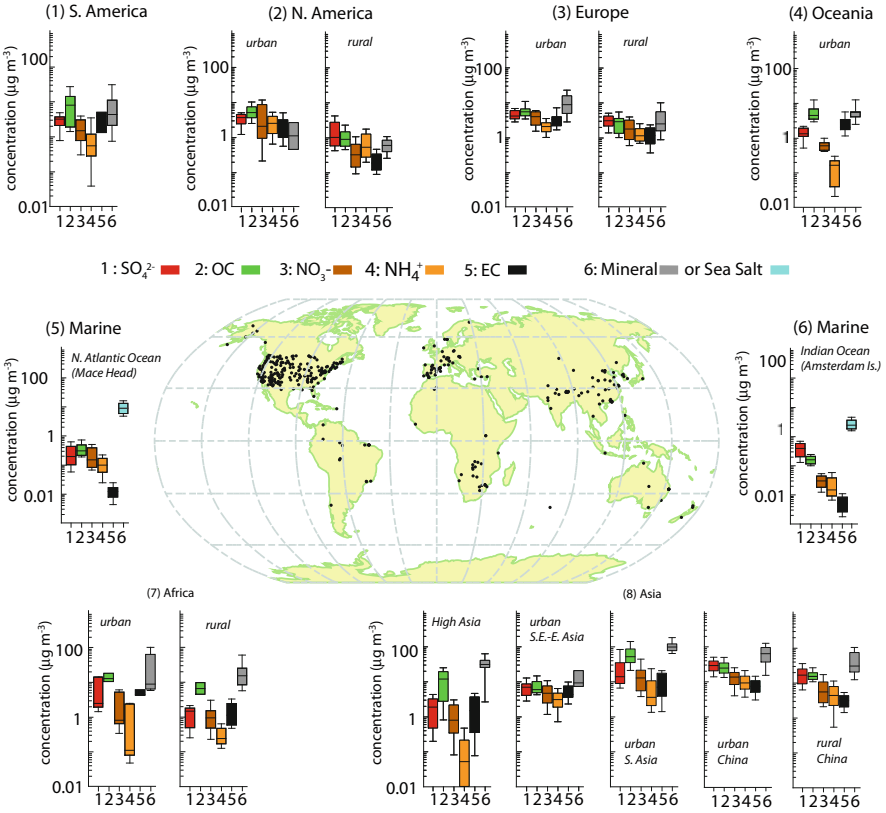


Fig. 3.5 Climatology of the mass concentrations ($\mu g m^{-3}$) of seven major chemical species of the atmospheric aerosol in different regions of the world. The locations of the stations are represented on the central map. The *bars* and *whiskers* represent for each region the 10^{th}, 25^{th}, 50^{th}, 75^{th} et 90^{th} percentiles of the distribution. (Reproduced from Boucher et al. (2013). © IPCC)

be quoted, which implies a strong absorption that characterizes this aerosol type (Bond and Bergstrom 2006). The refractive index of desert dust aerosols varies with the chemical composition, which itself varies with the source region (e.g., Wagner et al. 2012). Figure 3.6 shows the spectral variation of the refractive index as estimated by Balkanski et al. (2007).

For internally-mixed aerosols, one usually defines an equivalent refractive index that depends on the properties and concentrations of the different components. These physical models, that can be more or less empirical, are known under the umbrella term of *effective medium theory* (Chýlek et al. 1988). The most simple of these theories consists in estimating the equivalent refractive index of the aerosol mixture, m_e, as the volume-weighted refractive indices of the different aerosol components. In the case of aerosols made of a scattering matrix with refractive index m_1 and small inclusions with refractive index m_2, the equivalent refractive index can be

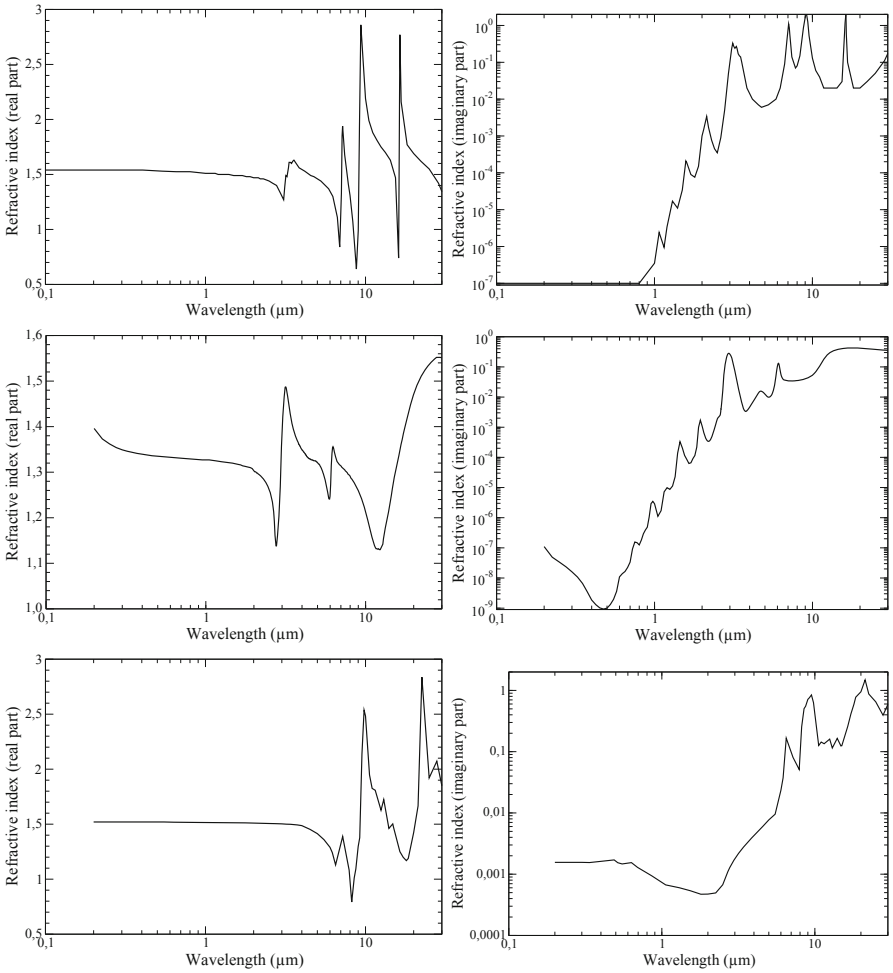


Fig. 3.6 Real (*left column*) and imaginary (*right column*) parts of the complex refractive index of ammonium sulphate (Toon et al. 2012; *top* panels), pure water (Hale and Querry 1973; *middle* panels) and Saharan dust particles (Balkanski et al. 2007; *bottom* panels). The wavelength and imaginary part of the refractive index are shown on logarithmic scales

approximated by:

$$m_e = v_1\, m_1 + v_2\, m_2 \tag{3.14}$$

where v_1 and v_2 are the volume fractions of the two components ($v_1 + v_2 = 1$). This approximation is not very accurate and other approximations are usually preferred, such as the Bruggeman approximation or the Maxwell–Garnett equation. The latter determines the equivalent refractive index for a matrix with small inclusions through

the following relationship:

$$m_e^2 = m_1^2 \frac{m_2^2 + 2m_1^2 + 2v_2(m_2^2 - m_1^2)}{m_2^2 + 2m_1^2 - v_2(m_2^2 - m_1^2)} \tag{3.15}$$

and is known as being more accurate. It should be stressed, however, that an equivalent refractive index remain an approximation and cannot give an account of all optical properties of the aerosol.

3.5 Deliquescence, Efflorescence and Hysteresis

Hygroscopicity is a key aerosol property in relation to their climate effect. It measures their ability to attract and hold water molecules in the condensed phase, and determines the variations of aerosol size, physical and optical properties with relative humidity. Aerosol hygroscopicity is directly related to aerosol solubility and it can be useful to distinguish insoluble and soluble aerosols even though in reality there is a spectrum of solubility.

Aerosol growth with relative humidity is due to the transfer of water molecules from the gas phase to the particulate phase. It is neither a linear, nor a continuous function of relative humidity. Rather an aerosol particle can experience a sudden size increase at the *deliquescence* point, where it goes from the solid to the liquid phase. The size variations experienced by such an aerosol particle are different for increasing and decreasing relative humidities, thus creating a *hysteresis* effect. For decreasing relative humidities, crystallization does not occur at the deliquescence point. The aerosol particle stays in a metastable state, until the relative humidity reaches a critical value called *efflorescence* or crystallization humidity. These phenomena are illustrated on Fig. 3.7, they are relatively well understood and are observed for pure salt aerosols (Tang 1997) but also for ambient atmospheric aerosols (Rood et al. 1989). The presence of insoluble material in an otherwise soluble aerosol particle can modify its deliquescence and efflorescence points.

3.6 Definition of Aerosol Optical Properties

3.6.1 Absorption and Scattering Cross Sections

Aerosols interact with electromagnetic radiation through the processes of scattering and absorption. The scattering and absorption cross sections, s^{sca} and s^{abs}, represent the "surface" of particle that interacts with the radiation in a plane perpendicular to the direction of propagation. They measure the degree of interaction of the particle with the radiation.

The scattering and absorption factors, Q^{sca} and Q^{abs} (unitless), are defined as the ratio of the scattering and absorption cross sections and the geometric cross section

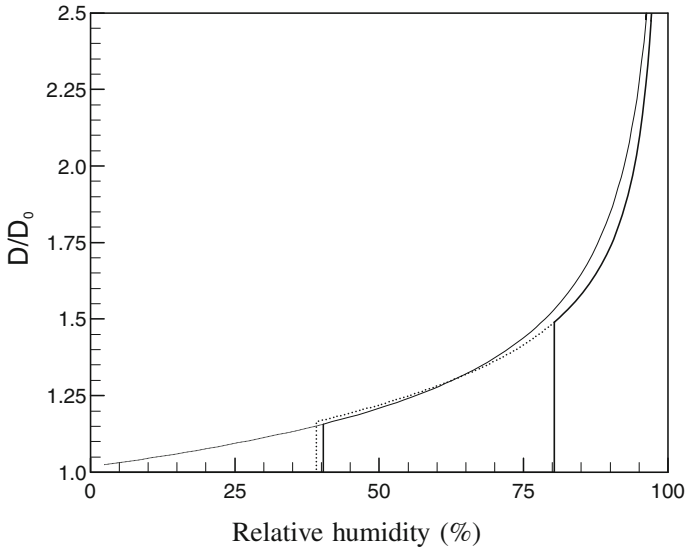


Fig. 3.7 Growth factor as a function of relative humidity (%) for sulphate ammonium (*bold line*) and bisulphate ammonium (*thin line*) particles. The hysteresis is shown with solid lines for increasing relative humidity and dotted lines for decreasing relative humidity. (Drawn from data published by Tang and Munkelwitz (1994))

s_g of the particle ($s_g = \pi\, r^2$ for a spherical particle with radius r):

$$Q^{\text{sca}} = \frac{s^{\text{sca}}}{s_g} \text{ and } Q^{\text{abs}} = \frac{s^{\text{abs}}}{s_g}. \tag{3.16}$$

Because extinction results from the combination of scattering and absorption, the extinction factor is equal to $Q^{\text{ext}} = Q^{\text{sca}} + Q^{\text{abs}}$. One further defines the single scattering albedo, ϖ_o, as the ratio between scattering and extinction: $\varpi_o = Q^{\text{sca}}/Q^{\text{ext}}$.

The scattering and absorption cross sections can be normalised by the particle volume v (or particle mass m) to obtain the *scattering and absorption volume* (or *mass*) *efficiencies*[1]:

$$\begin{cases} \alpha_v^{\text{sca}} = s^{\text{sca}}\, v \\ \alpha_v^{\text{abs}} = s^{\text{abs}}\, v \\ \alpha_m^{\text{sca}} = s^{\text{sca}}\, m = \alpha_v^{\text{sca}}\, \varrho \\ \alpha_m^{\text{abs}} = s^{\text{abs}}\, m = \alpha_v^{\text{abs}}\, \varrho \end{cases} \tag{3.17}$$

[1] These quantities are sometimes called volume (or mass) scattering and absorption cross sections.

where ϱ is the particle density. The volume scattering, absorption and extinction efficiencies are often reported in m^2cm^{-3} while the mass scattering, absorption and extinction efficiencies are usually measured in m^2g^{-1}.

3.6.2 Phase Function

The scattering and absorption factors are not sufficient to characterize the interaction of radiation with particles. One needs to know how the radiation is scattered. The *phase function* describes the angular distribution of the scattered radiation. For a spherical particle, it only depends on the angle between the directions of propagation of the incident and scattered beams. This angle is called the *scattering angle*, usually noted Θ, and ranges between 0 and π. In this case the phase function $P(\Theta)$ is normalised to a value of 2:

$$\int_0^{\pi} P(\Theta) \sin \Theta \, \mathrm{d}\Theta = 2. \tag{3.18}$$

It is sometimes more convenient to define the phase function P as a function of the cosine of the scattering angle, $\mu = \cos(\Theta)$, in which case the previous equation becomes:

$$\int_{-1}^{1} P(\mu) \, \mathrm{d}\mu = 2. \tag{3.19}$$

The reader is redirected to Chap. 5 for a more general definition of the phase function that also applies to nonspherical particles.

The asymmetry parameter is defined as the first-order moment of the phase function:

$$g = \int_{-1}^{1} P(\mu) \, \mu \, \mathrm{d}\mu. \tag{3.20}$$

It measures the degree of symmetry of the phase function between the forward and backward hemispheres. If g is equal to 1, all the scattered radiation is scattered in the forward direction; if g is equal to -1, all the scattered radiation is scattered in the backward radiation; if g is equal to 0, there is as much scattering in the forward than in the backward hemispheres. It should be noted that the asymmetry parameter only reflects a fraction of the information contained in the phase function. For an optically thin medium, an accurate calculation of diffuse radiation requires to take into account higher-order moments of the phase function.

3.6.3 Upscatter Fractions

Upscatter fractions give a more intuitive picture of the forward–backward asymmetry of the aerosol phase function (see Fig. 3.8). Considering downward incident (solar)

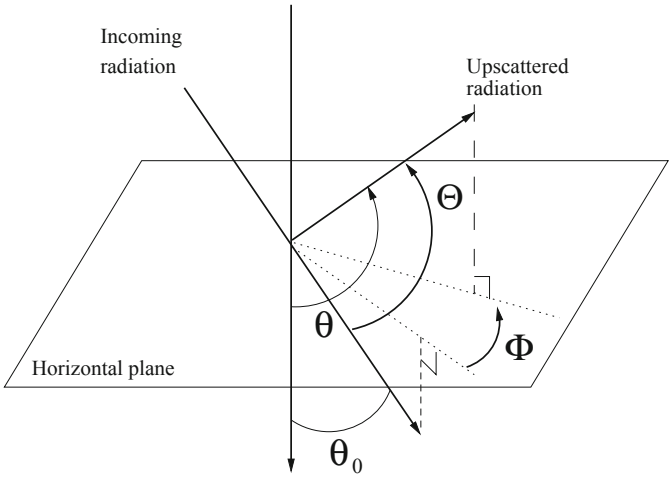


Fig. 3.8 Schematic diagram defining the upscatter fraction for an atmospheric aerosol. The zenith angle is measured relative to the downward vertical axis. The upscatter fraction corresponds to scattered radiation with a zenith angle between 90 and 180°

radiation with a zenith angle θ_0 measured relative to the downward vertical axis, one defines a monodirectional upscatter fraction as the fraction of scattered radiation that is scattered upwards, which can be expressed mathematically as:

$$\beta(\mu_0) = \frac{1}{2}\int_{-1}^{0}\frac{1}{2\pi}\int_{0}^{2\pi} P\left(\mu\mu_0 + (1-\mu^2)^{\frac{1}{2}}(1-\mu_0^2)^{\frac{1}{2}}\cos\Phi\right)\,\mathrm{d}\Phi\,\mathrm{d}\mu \quad (3.21)$$

where $\mu_0 = \cos\theta_0$, $\mu = \cos\theta$ and Φ is the azimuthal angle. Wiscombe and Grams (1976) have shown that Eq. 3.21 can be rewritten as:

$$\begin{aligned}\beta(\mu_0) = \quad & \frac{1}{2\pi}\int_{\frac{\pi}{2}-\theta_0}^{\frac{\pi}{2}+\theta_0} \arccos\left(\mathrm{cotan}\theta_0\ \mathrm{cotan}\theta\right)\ P(\cos\theta)\ \sin\theta\ \mathrm{d}\theta\ + \\ & \frac{1}{2}\int_{\frac{\pi}{2}+\theta_0}^{\pi} P(\cos\theta)\ \sin\theta\ \mathrm{d}\theta.\end{aligned} \quad (3.22)$$

The following notation

$$b = \beta(1) = \frac{1}{2}\int_{-1}^{0} P(\mu)\,\mathrm{d}\mu \quad (3.23)$$

is often used for the upscatter fraction for vertical downward incident radiation (i.e. corresponding to a sun in zenith[2] position). This parameter b is also called

[2] The zenith refers to an imaginary point directly "above" a given location.

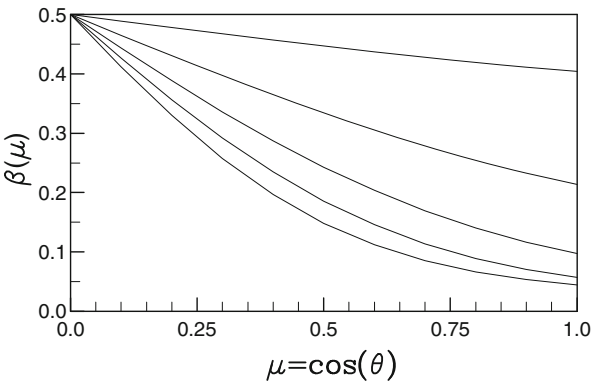


Fig. 3.9 Upscatter fraction, $\beta(\mu)$, as a function of the cosine of the zenith angle, μ, for a range of log-normal size distributions for homogeneous spheres with a refractive index $m = 1.33$, mean geometric volume diameters (from *top* to *bottom* on the graph) of 0.1, 0.2, 0.3, 0.4 and 0.5 μm, and a geometric standard deviation $\sigma_g = 1.4$

the *backscatter fraction* as it measures the fraction of scattered radiation that is backscattered. Finally the *isotropic upscatter fraction* (i.e. the upscatter fraction for isotropic downward radiation with an average μ value of 0.5) is:

$$\overline{\beta} = \int_0^1 \beta(\mu)\,\mathrm{d}\mu. \tag{3.24}$$

Wiscombe and Grams (1976) have shown that $\overline{\beta}$ can be expressed as a single integral of the phase function:

$$\overline{\beta} = \frac{1}{2\pi}\int_{-1}^{1} \arccos(\mu)\; P(\mu)\,\mathrm{d}\mu. \tag{3.25}$$

The functions $\beta(\mu)$ are represented on Fig. 3.9 for typical submicronic size distributions of spherical particles. It can be seen that $\beta(\mu)$ is larger for smaller particles and increases when θ increases from 0 to 90°.

3.7 Calculation of Aerosol Optical Properties

3.7.1 Mie Theory

Computing optical properties of an aerosol particle is relatively easy for a sphere because the well-known Mie theory applies (van de Hulst 1982; Bohren and Huffman 1998). One can also treat analytically the case of concentric homogeneous spheres (i.e the core-shell model, see Toon and Ackerman 1981), chains of adjacent spheres (Fuller 1994, 1995a), spheres containing arbitrarily located spherical spherules (Fuller 1995b) and some particular types of nonspherical particles such as spheroids and ellipsoids (Mishchenko and Travis 1994). More complex geometries require solving Maxwell's equations numerically, which is computationally expensive for a population of aerosols.

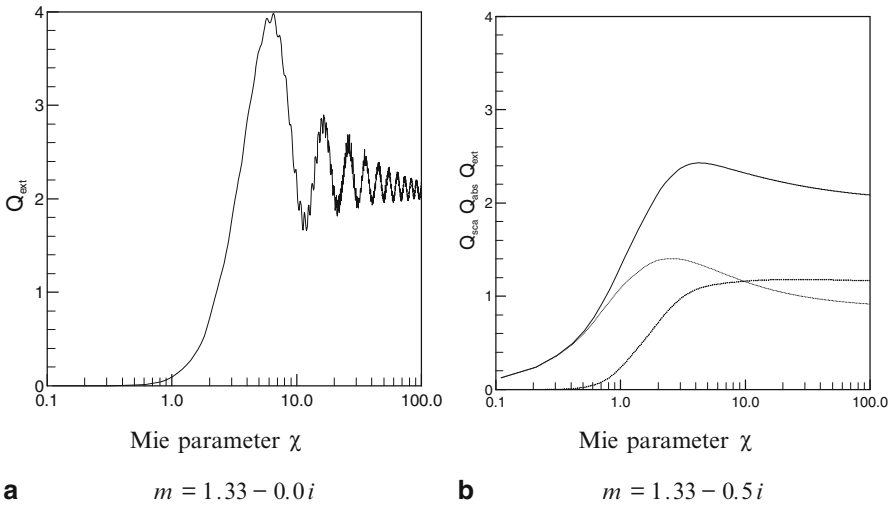


Fig. 3.10 Extinction, scattering and absorption factors as a function of the Mie size parameter, $x = 2\pi r/\lambda$, for homogeneous spheres. **a** The refractive index is real, $m = 1.33$, so there is no absorption and the extinction factor is equal to the scattering factor. In the Rayleigh regime (x small and $x(m-1)$ small), one obtains $Q^{\text{sca}} \propto x^4$. **b** The refractive index is complex, $m = 1.33 - 0.5i$, which implies a strong absorption. The extinction coefficient is represented by a *solid line*, while the scattering and absorption factors appear as *dotted lines*. The absorption factor dominates for small x while the scattering factor dominates for large x. In the Rayleigh regime, the extinction coefficient is due to absorption and $Q^{\text{ext}} \propto x$

Mie theory[3] has been formulated by Gustav Mie in 1908. It describes the scattering of electromagnetic radiation by spherical particles and fills the gap between the theory for Rayleigh scattering that applies to particles much smaller than the wavelength of light, and geometrical optics that applies to particles much larger than the wavelength. The reader is referred to the textbooks of van de Hulst (1982) and Bohren and Huffman (1998) for a full exposé of the theory. An algorithm of the Mie solution is presented in Appendix C. Mie theory thus describes the interaction of spherical particles of arbitrary sizes with electromagnetic radiation. An important limitation of Mie theory is that it only applies to homogeneous spheres or stratified spheres. According to Mie theory, optical properties only depend on a dimensionless size parameter $x = 2\pi r/\lambda$ and the refractive index m of the particle. For a fixed refractive index, the dependence of aerosol optical properties with size is inverse to the dependence to the wavelength.

3.7.2 Extinction, Scattering and Absorption

The extinction factor is represented as a function of the size parameter for a non-absorbing spherical particle in Fig. 3.10a. It can be noted that Q^{ext} is proportional to

[3] Mie theory is also referred to as Lorenz–Mie or Lorenz–Mie–Debye theory from the names of other physicists who developed the solution independently.

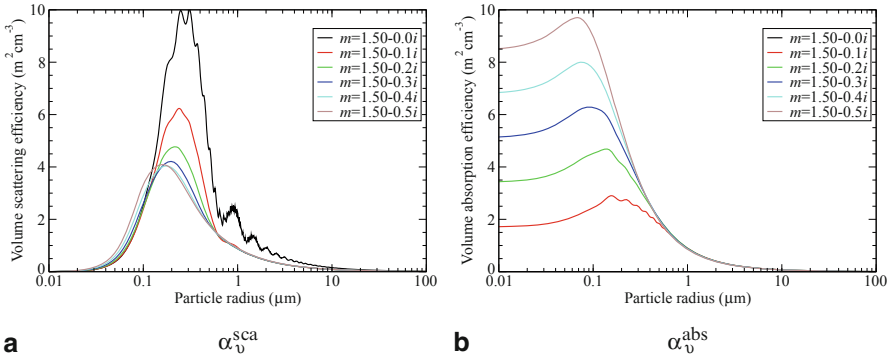


Fig. 3.11 **a** Volume scattering efficiency, α_v^{sca}, and **b** volume absorption efficiency, α_v^{abs}, in units of m^2cm^{-3}, as a function of particle radius. The volume scattering and absorption efficiencies are for homogeneous spheres with varying complex refractive indices at a wavelength of 550 nm

x^4 for small size parameters and tends towards 2, the value predicted by geometrical optics, when x tends towards $+\infty$. Resonance phenomena can be observed with Q^{ext} which can reach values as large as 4 for some size parameters. If the particle is absorbing, these resonance phenomena disappear. The extinction factor is proportional to x, reaches a maximum and tends towards 2 for larger size parameters (Fig. 3.10b).

Figure 3.11a shows the volume scattering efficiency for homogeneous spheres with varying degree of absorption. At a wavelength of 550 nm, the volume scattering efficiency presents a peak for particle radius between 0.1 and 0.3 μm. The maximum depends on the real and imaginary parts of the refractive index and is more pronounced when the particle is non absorbing. It is noteworthy that for the wavelengths of the visible spectrum, the particle radius where the volume scattering efficiency is maximum corresponds to aerosol sizes typical of the accumulation mode (Fig. 3.12).

The volume absorption efficacy also presents a maximum, but it is much less pronounced than for scattering (Fig. 3.11b). Contrary to what is observed for scattering, the absorption efficiency tends towards a nonzero value when the particle size tends towards 0. This means that, for a constant mass or volume, the aerosol size distribution does not matter so much as long as aerosol particles are not too large. These calculations are only illustrative but are consistent with observations. For instance, measurements indicate a mass absorption efficiency of the order of 7.5 m^2g^{-1} at 550 nm for fresh black carbon aerosols. As the black carbon ages and becomes coated with soluble material, the aerosol becomes more hygroscopic and internally-mixed, and its absorption efficiency per unit mass black carbon can reach 10–15 m^2g^{-1}.

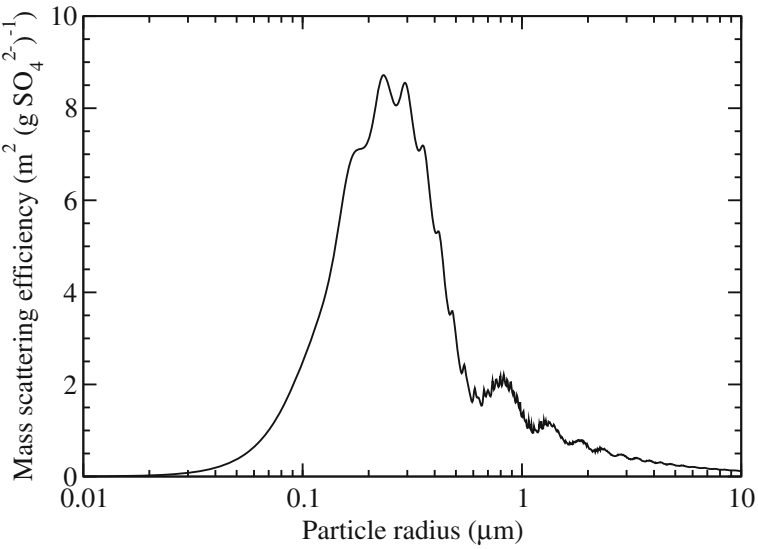


Fig. 3.12 Mass scattering efficiency (m^2 (g SO_4^{2-})$^{-1}$) as a function of the particle radius, r, in μm for dry ammonium sulphate particles with a real refractive index, m=1.53, and a wavelength, $\lambda = 0.53$ μm. Note that the scattering efficiency is expressed here per mass of sulphate rather than per mass of ammonium sulphate

3.7.3 Optical Depth and Ångström Coefficient

For a population of aerosol particles optical properties need to be integrated over the size distribution $n(r)$. As an example of an intensive aerosol property, we provide the expression for the mass scattering and absorption efficiencies (unit m^{2}g^{-1}):

$$\begin{cases} \alpha_m^{\text{sca}} = \dfrac{\displaystyle\int_0^\infty \pi\, r^2\, Q^{\text{sca}}(r)\, n(r)\ \mathrm{d}r}{\displaystyle\int_0^\infty \frac{4}{3}\pi\, r^3\, \varrho\, n(r)\ \mathrm{d}r} \\ \alpha_m^{\text{abs}} = \dfrac{\displaystyle\int_0^\infty \pi\, r^2\, Q^{\text{abs}}(r)\, n(r)\ \mathrm{d}r}{\displaystyle\int_0^\infty \frac{4}{3}\pi\, r^3\, \varrho\, n(r)\ \mathrm{d}r} \end{cases} \tag{3.26}$$

Without normalization by the total volume or mass of particles, the scattering and absorption cross sections are integrated to obtain the scattering and absorption

coefficients (unit m^{-1}) following:

$$\begin{cases} \sigma^{\text{sca}} = \displaystyle\int_0^{\infty} \pi\ r^2\ Q^{\text{sca}}(r)\, n(r)\, \mathrm{d}r \\ \\ \sigma^{\text{abs}} = \displaystyle\int_0^{\infty} \pi\ r^2\ Q^{\text{abs}}(r)\, n(r)\, \mathrm{d}r. \end{cases} \tag{3.27}$$

Moreover the extinction coefficient verifies:

$$\sigma^{\text{ext}} = \sigma^{\text{sca}} + \sigma^{\text{abs}}. \tag{3.28}$$

The aerosol scattering, absorption and extinction coefficients are local properties of the atmosphere (i.e. they vary with the location and altitude). The integral of the extinction coefficient over the vertical is called the aerosol optical thickness (AOT) or *aerosol optical depth* (AOD)[4]:

$$\tau = \int_{\text{surface}}^{\text{top of atmosphere}} \sigma^{\text{ext}}(z)\, \mathrm{d}z. \tag{3.29}$$

Aerosol optical properties depend on the wavelength of the radiation, and so does the aerosol optical depth. The variations with wavelength can be substantial, which is why it is important to always specify the wavelength associated to a given aerosol optical depth. The *Ångström coefficient*, also called Ångström parameter or Ångström exponent, measures the spectral dependence of the aerosol optical depth (or extinction coefficient). It is calculated in the following way:

$$\alpha = -\frac{\ln(\tau_1/\tau_2)}{\ln(\lambda_1/\lambda_2)} \tag{3.30}$$

where τ_1 and τ_2 are the aerosol optical depths at wavelengths λ_1 and λ_2. The Ångström coefficient varies with the choice of wavelengths, λ_1 and λ_2, however this variation is usually relatively small, at least in the visible spectrum. In practice the aerosol optical depth at a wavelength λ can be estimated from the knowledge of the aerosol optical depth, τ_r, at a reference wavelength λ_r and the Ångström coefficient using the equation:

$$\tau = \tau_r \left(\frac{\lambda}{\lambda_r}\right)^{-\alpha}. \tag{3.31}$$

The Ångström coefficient contains information on the size of the scatterer. Scattering by molecules results in an Ångström coefficient of 4. Particles of the accumulation

[4] Strictly speaking there is a distinction between the two quantities. The AOD represents a vertical coordinate, while the AOT refers to the integral from the surface to the top of atmosphere. However in practice the term AOD is often used to mean AOT and we follow this practice in this textbook.

mode have an Ångström coefficient of about 2, while values close to 0 or even slightly negative are typical of coarse mode aerosols. There is not, however, a direct correspondence between aerosol size and Ångström coefficient because the aerosol refractive index (and its variation with the wavelength) also plays a role in determining the spectral dependence of the extinction.

One can also define an Ångström coefficient for absorption, β, that verifies

$$\tau^{\text{abs}} = \tau_r^{\text{abs}} \left(\frac{\lambda}{\lambda_r} \right)^{-\beta}. \tag{3.32}$$

where τ^{abs} and τ_r^{abs} are the aerosol absorption optical depths (AAOD) at wavelengths λ and λ_r. The absorption Ångström coefficient varies with the aerosol type and can be a way to discriminate between dust, black carbon and organic carbon aerosol (Chung et al. 2012).

3.8 Optical Properties of Nonspherical Aerosols

Mie theory only applies to spherical particles. Nonspherical particles whose size is comparable to the wavelength can present scattering properties that are very different to those from spherical particles. In particular, the phase function presents neither a backscattering peak at 180°, nor a pronounced minimum around 120°. Several theories have been developed that allow an exact calculation of the optical properties for some types of nonspherical particles of homogeneous compositions. Measurements are also helpful to confirm the findings from aerosol scattering models. Figure 3.13 illustrates the discrepancy between observations and Mie theory applied to nonspherical desert dust aerosols.

3.9 Aerosols and Atmospheric Visibility

Atmospheric visibility is defined as the distance at which a black object can be discerned when observed against a white background. It is usually reported as a horizontal distance. Visibility can be approximated from the total (molecules and aerosols) extinction coefficient through the Koschmieder equation:

$$x = \frac{3.912}{\sigma_{\text{ext}}} \tag{3.33}$$

where x is the visibility in m and σ_{ext} is the extinction coefficient, averaged over the wavelengths of the visible spectrum, in units of m^{-1}. In practical terms, σ_{ext} can be taken at 520 or 550 nm. At sea level, in the absence of aerosols, the extinction coefficient is 13.2×10^{-6} m^{-1}, which corresponds to a visibility of approximately 300 km. The visibility increases with altitude, but decreases with the presence of aerosols or in foggy and cloudy conditions. It can be as low as 1 km (or even below) in the presence of a very thick aerosol plume.

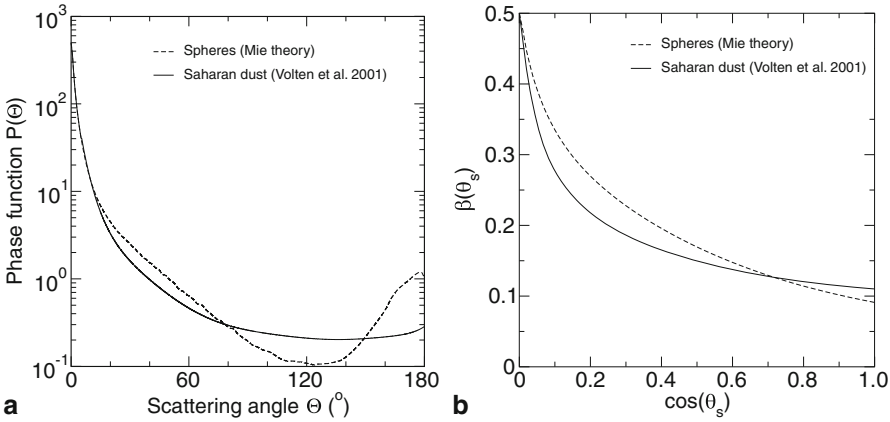


Fig. 3.13 **a** Phase function for desert dust aerosols as a function of the scattering angle. **b** Upscatter fraction as a function of the zenith angle according to the measurements of Volten et al. (2001) in comparison with results from Mie theory for the same refractive index and an equivalent size distribution. (From Bellouin et al. (2004). © Royal Meteorological Society)

Exercises

1. Integrate Eq. 3.10 to confirm that the total aerosol concentration is N_0.

2. The Gamma distribution can be written as

$$n(r) = \frac{N_0\,(ab)^{(2b-1)/b}}{\Gamma[(1-2b)/b]}\, r^{(1-3b)/b}\, \exp(-r/ab)$$

where a and b are two adjustable parameters. Demonstrate that the mode of the size distribution is equal to $a(1-3b)$, the effective radius is a and the effective variance is b.

3. An aerosol population follows a log-normal size distribution with a mean geometric radius $r_0 = 0.15$ μm and a geometric standard deviation $\sigma_g = 2$. Compute the mean radius, the median radius, the mean surface radius, the mean volume radius and the number, surface and volume geometric radii. Rank these characteristic radii from the smallest to the largest. It is recommended to make use of Appendix B.

4. The atmosphere contains a 1-km thick layer of monodisperse aerosols with radius $r = 0.175$ μm, concentration 50 μgm^{-3}, density 1.7 gcm^{-3}, refractive index 1.50–0.1i. Estimate the scattering, absorption and extinction optical depths at 550 nm using Fig. 3.11.

5. An aerosol layer has an optical depth of 0.5 at the 550 nm wavelength and an Ångström coefficient of 1.8. Compute its optical depth at 440 and 870 nm.

6. One considers the log-normal distribution $n^*(r)$:

$$n^*(r) = \frac{N}{\sqrt{2\pi}\,\sigma_0}\,\exp\left[-\frac{1}{2}\left(\frac{\ln r - \ln r_0}{\sigma_0}\right)^2\right]$$

where N is the total concentration in aerosols, r_0 is the modal aerosol and σ_0 a measure of the dispersion of the distribution.
a. Show that the distribution of the variable $z = r^n$ also follows as log-normal law. What is its mode?
b. By definition, the n^{th} order moment of the distribution is equal to

$$M_n = {r_n}^n = \int r^n\, n^*(r)\, \mathrm{d}\ln r \;/\; \int n^*(r)\, \mathrm{d}\ln r$$

Express M_n as a function of r_0, σ_0 and n and deduce that the equivalent radius for the n^{th} order moment is:

$$r_n = {M_n}^{1/n} = r_0\,\exp\left(\frac{n\,\sigma_0^2}{2}\right).$$

Solutions

3. $r_0 = r_{\text{median}} = r_{g0} = 0.15 < r_{\text{mean}} = 0.19 < r_{\text{surface}} = 0.24 < r_{\text{volume}} = 0.31 < r_{gs} = 0.39 < r_{gv} = 0.63\ \mu\text{m}$.

4. $x \approx 2$, $\alpha_{s-V} \approx 5.7\ \text{m}^2\text{cm}^{-3}$, $\alpha_{a-V} \approx 2.8\ \text{m}^2\text{cm}^{-3}$, $\tau_s \approx 0.17$, $\tau_a \approx 0.08$ and $\tau = \tau_s + \tau_a \approx 0.25$.

5. $\tau_{440} = 0.75$ and $\tau_{870} = 0.22$.

6a. The reader is referred to Appendix B. The mode of the power n of the radius is $r_0\,\exp(n\,\sigma_0^2)$.

6b. One needs to apply the substitution $X = \dfrac{\ln(r) - \ln(r_0)}{\sigma_0}$ followed by $Y = X - n\,\sigma_0$.

6c. The same substitutions should be done as in 6b.

References

Balkanski Y, Schulz M, Claquin T, Guibert S (2007) Reevaluation of mineral aerosol radiative forcings suggests a better agreement with satellite and AERONET data. Atmos Chem Phys 7:81–95

Bellouin N, Boucher O, Vesperini M, Tanré D (2004) Estimating the aerosol direct radiative perturbation: impact of the ocean surface representation and aerosol non-sphericity. Q J Royal Meteorol Soc 130:2217–2232

Bond TC, Bergstrom RW (2006) Light absorption by carbonaceous particles: an investigative review. Aerosol Sci Technol 40:27–67

Bond TC, Doherty SJ, Fahey DW, Forster PM, Berntsen T, DeAngelo BJ, Flanner MG, Ghan SJ, Kärcher B, Koch D, Kinne S, Kondo Y, Quinn PK, Sarofim MC, Schultz MG, Schulz M, Venkataraman C, Zhang H, Zhang S, Bellouin N, Guttikunda SK, Hopke PK, Jacobson MZ, Kaiser JW, Klimont Z, Lohmann U, Schwarz JP, Shindell D, Storelvmo T, Warren SG, Zender CS (2013) Bounding the role of black carbon in the climate system: a scientific assessment. J Geophys Res 118:5380–5552. doi:10.1002/jgrd.50171

Chung CE, Ramanathan V, Decremer D (2012) Observationally constrained estimates of carbonaceous aerosol radiative forcing. Proc Natl Acad Sci U S A 109:11624–11629

Chýlek P, Srivastava V, Pinnick RG, Wang RT (1988) Scattering of electromagnetic waves by composite spherical particles: experiment and effective medium approximation. Appl Opt 27:2396–2404

Fuller KA (1994) Scattering and absorption cross sections of compounded spheres I. Theory for external aggregation. J Opt Soc Am 11:3251–3260

Fuller KA (1995a) Scattering and absorption cross sections of compounded spheres II. Calculations for external aggregation. J Opt Soc Am 12:881–892

Fuller KA (1995b) Scattering and absorption cross sections of compounded spheres III. Spheres containing arbitrarily located spherical inhomogeneities. J Opt Soc Am 12:893–904

Hale GM, Querry MR (1973) Optical constants of water in the 200-nm to 200-micron wavelength region. Appl Opt 12:555–563

Jacobson MC, Hansson H-C, Noone KJ, Charlson RJ (2000) Organic atmospheric aerosols: review and state of the science. Rev Geophys 38:267–294

Jimenez JL et al (2009) Evolution of organic aerosols in the atmosphere. Science 326:1525–1528

Mishchenko MI, Travis LD (1994) Light scattering by polydispersions of randomly oriented spheroids with sizes comparable to wavelength of observation. Appl Opt 33:7206–7225

Putaud J-P, Raes F, van Dingenen R, Brüggemann E, Facchini M-C, Decesari S, Fuzzi S, Gehrig R, Hüglin C, Laj P, Lorbeer G, Maenhaut W, Mihalopoulos N, Müller K, Querol X, Rodriguez S, Schneider J, Spindler G, ten Brink H, Tørseth K, Wiedensohler A (2004) A European aerosol phenomenology–2: chemical characteristics of particulate matter at kerbside, urban, rural and background sites in Europe. Atmos Environ 38:2579–2595

Rood MJ, Shaw MA, Larson TV, Covert DS (1989) Ubiquitous nature of ambient metastable aerosol. Nature 337:537–539

Stelson AW (1990) Urban aerosol refractive index prediction by partial molar refraction approach. Environ Sci Technol 24:1676–1679

Tang IN (1997) Thermodynamic and optical properties of mixed-salt aerosols of atmospheric importance. J Geophys Res 102:1883–1893

Tang IN, Munkelwitz HR (1994) Water activities, densities, and refractive indices of aqueous sulfates and sodium nitrate droplets of atmospheric importance. J Geophys Res 99:18801–18808

Tang IN, Tridico AC, Fung KH (1997) Thermodynamic and optical properties of sea salt aerosols. J Geophys Res 102:23269–23275

Toon OB, Pollack JB, Khare BN (1976) The optical constants of several atmospheric aerosol species: ammonium sulfate, aluminium oxide, and sodium chloride. J Geophys Res 81:5733–5748

Volten H, Muñoz O, Rol E, de Haan J, Vassen W, Hovenier J, Muinonen K, Nousiainen T (2001) Scattering matrices of mineral aerosol particles at 441.6 nm and 632.8 nm. J Geophys Res 106:17375–17401

Wagner R, Ajtai T, Kandler K, Lieke K, Linke C, Müller T, Schnaiter M, Vragel M (2012) Complex refractive indices of Saharan dust samples at visible and near UV wavelengths: a laboratory study. Atmos Chem Phys 12:2491–2512

Wexler AS, Clegg SL (2002) Atmospheric aerosol models for systems including the ions H^+, NH^{4+}, Na^+, $SO4^{2-}$, NO^{-3}, Cl^-, Br^-, and H_2O. J Geophys Res 107(D14). doi:10.1029/2001JD000451

Whitby KT (1978) The physical characteristics of sulfur aerosols. Atmos Environ 12:135–159

Wiscombe WJ, Grams GW (1976) The backscattered fraction in two-stream approximations. J Atmos Sci 33:2440–2451

Further Reading (Textbooks and Articles)

Bohren CF, Huffman DR (1998) Absorption and scattering of light by small particles. Wiley Science Paperback Series, New York, 544 pp

Boucher O, Randall D, Artaxo P, Bretherton C, Feingold G, Forster P, Kerminen V-M, Kondo Y, Liao H, Lohmann U, Rasch P, Satheesh SK, Sherwood S, Stevens B, Zhang XY (2013) Clouds and aerosols. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, pp 571–657

Després VR, Huffman JA, Burrows SM, Hoose C, Safatov AS, Buryak G, Frohlich-Nowoisky J, Elbert W, Andreae MO, Pöschl U, Jaenicke R (2012) Primary biological aerosol particles in the atmosphere: a review. Tellus 64B:15598. doi:10.3402/tellusb.v64i0.15598

Penner JE, Andreae M, Annegarn H, Barrie L, Feichter J, Hegg D, Jayaraman A, Leaitch R, Murphy D, Nganga J, Pitari G (2001) Aerosols, their direct and indirect effects. In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Xiaosu D (eds) Climate change 2001: the scientific basis, contribution of working group I to the third assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, UK, pp 289–348 (Chapter 5)

Prospero JM, Charlson RJ, Mohnen V, Jaenicke R, Delany AC, Moyers J, Zoller W, Rahn K (1983) The atmospheric aerosol system: an overview. Rev Geophys 21:1607–1629. doi:10.1029/RG021i007p01607

van de Hulst HC (1982) Light scattering by small particles. Dover Publications, New York

Chapter 4
Aerosol Modelling

Abstract This chapter introduces some concepts of aerosol modelling, going through the different terms of the continuity equation. The procedure for producing emissions inventories of aerosols and aerosol precursors from fossil fuel and biomass burning is briefly discussed. The basis for parametrizing the emissions of other aerosol types and aerosol precursors, such as sea spray and desert dust, dimethysulphide, and volatile organic compounds is also presented. The most relevant atmospheric processes are then reviewed starting from nucleation, new particle formation, and condensation of semi-volatile compounds, and continuing with coagulation, production in the liquid phase, dry deposition, wet deposition, and sedimentation. For each of these processes, a conceptual model is presented. The various approaches to aerosol modelling (bulk, sectional, and modal) are then presented with their advantages and disadvantages. Finally, an example is provided with the global sulphur cycle.

Keywords Emissions · Sinks · Scavenging · Coagulation · Sulphur cycle · Modelling

4.1 Introduction

Modelling aerosols, at any spatial and temporal resolution, requires to represent the sources, transformation processes, transport and sinks of aerosols and their precursors. The relative importance of the different processes depend on the scale considered. Some processes which are important to consider at the fine scale, may be less important at the large-scale, and vice versa. The modelling strategy, therefore, needs to be adapted to the objectives of the model. This chapter presents some fundamentals of aerosol modelling but only considers processes that are important for the large scale as our aim is primarily to understand the climate effects of aerosols.

The fate of a tracer in the atmosphere is governed by the conservation equation which applies to the aerosol concentration (in number or in mass) per unit volume, here denoted c:

$$\frac{\partial c}{\partial t} + \text{div}\,(c\,\mathbf{V}) = \text{div}\,(K_m\,\boldsymbol{\nabla} c) + S - P \tag{4.1}$$

where div represents the divergence operator, $\boldsymbol{\nabla}$ is the gradient operator, $\mathbf{V}$ is the wind field, K_m the diffusion coefficient, S the source term (emission and chemical production), and P is the sink term (deposition and chemical sinks). There are as

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_4

many conservation equations as there are variables that one wishes to predict. It is possible to define a *lifetime* for each sink process i composing P as the ratio between the concentration c and P^i. It corresponds at first order to the time it would take to exhaust the concentration c with the sink process P^i and can be computed locally. A sink process can be efficient in one location but not in another one.

One can then write a system of n partial derivative equations that can be solved numerically. For a variable whose lifetime is short as compared to the timestep one wishes to resolve, a stationarity hypothesis can be made, $\partial c/\partial t = 0$, which also neglects the transport term. In that case, Eq. 4.1 simplifies to a balance between the source and sink terms.

If one adopts a description in terms of mixing ratio $C = c/\varrho$ where ϱ is the density of dry air, the conservation equation can be written in the form:

$$\frac{\partial C}{\partial t} + \mathbf{V} \cdot \nabla C = \frac{1}{\varrho} \operatorname{div} (K_m \nabla \varrho \, C) + \frac{S - P}{\varrho}. \tag{4.2}$$

This equation can be obtained from Eq. 4.1, the conservation equation for dry air (also called the continuity equation) in its Eulerian form:

$$\frac{\partial \varrho}{\partial t} + \operatorname{div} (\varrho \mathbf{V}) = 0 \tag{4.3}$$

and the fact that $\operatorname{div} (a \, \mathbf{x}) = a \operatorname{div} \mathbf{x} + \nabla a \cdot \mathbf{x}$.

There are as many Eqs. 4.1 or 4.2 as there are variables to predict. If one variable is enough to predict the concentration of a trace gas in the atmosphere, predicting the time evolution of an aerosol population requires a number of variables that increases with the amount of complexity (size distribution, chemical composition, degree of mixing) that one wants to represent. The choice of variables and the number of variables are the result of a compromise between the complexity that one wants to resolve, the spatial resolution that one wants to achieve on the simulated domain, the number and length of the simulations that are to be performed, and the available computing power. A large number of processes play a role in the evolution of an aerosol population and need to be taken into account in the source and sink terms of each of the relevant variables. We provide here a quick description of these processes and their physical representation.

4.2 Emissions

4.2.1 Generalities

Modelling aerosols starts with a good characterization of the source term. Unfortunately, it often happens that this term is not known to a sufficient accuracy. Depending on the aerosol type, we rely on emission inventories—that tabulate the emissions as a function of the location, time, and other relevant parameters—or

physical parametrizations–that seek to express the source terms as a function of meteorological and other relevant variables. We provide here a general view of the source estimation for different aerosol types. The global annual fluxes of aerosols and aerosol precursors are provided in Table 2.1 of Chap. 2.

4.2.2 Fossil Fuels, Biofuels, and Other Anthropogenic Sources

The anthropogenic sources of aerosols that stem from the combustion of fossil fuels and biofuels can be estimated from inventories of the usage of fossil fuels and biofuels by economic sector and fuel type (index i below) and the *a priori* knowledge of emission factors for each emission type. The emission type (in $\mathrm{kgm^{-2}s^{-1}}$) for a given species can be written as:

$$E = \sum_i A_i \, \mathrm{EF}_i \, (1 - \alpha_i) \tag{4.4}$$

where A_i is the fuel consumption for activity i (kg of fuel per unit time and unit surface), EF_i the emission factor for activity i (in kg of species emitted per kg of fuel burned), and α_i is the efficiency of the mitigation technology for emission reduction (α_i equals 0 in the absence of mitigation technology). What is at stake here is to get a dataset that is as accurate as possible and as disaggregated as possible in space, time, and activity. In practice, one can work with annual statistics and correct those for seasonal, weekly, and diurnal variations. Similarly, one tries to get good statistics on the A_i quantities at a spatial resolution commensurate or higher than the resolution at which one wants to model the aerosol. The emission factors, EF_i, are compiled from measurements performed at the source point for things like gasoline cars, diesel cars, diesel light truck, diesel heavy truck, off-road vehicles, cement factories, coal-fired power plant, oil-fired power plant, steel plant, domestic heating, etc. For mobile sources, these emission factors should represent an average over a fleet of vehicles and include the contribution of so-called super-emitters, old vehicles that may contribute disproportionally to some emissions. Emission factors for different aerosol precursors and aerosol species vary across the industrial sectors; and so does the aerosol size distribution at the source point depending on the fuel and combustion type. It can be interesting to factor in processes that occur in the exhaust plume itself (such as condensation of some volatile species and/or coagulation) into the emission factors themselves given that such processes are generally not resolved in regional and global aerosol models.

The above equation can be generalized to include all activity types and not only those related to fossil fuel and biofuel combustion. In this case, A_i is the "quantity" of activity per unit time and unit surface and EF_i is the emission factor per unit activity. This applies for instance for dust aerosols generated from road traffic and the building sector, agricultural sources of aerosols, and aerosol precursors. In this latter case, inventories may disaggregate emissions by crop type and agricultural practice. Again, the aerosol size distribution at the source point depends on the source type.

4.2.3 Vegetation Fires

The sources of aerosols and aerosol precursors that come from vegetation fires can be estimated from burned surface areas, vegetation types, and emission factors. The emission flux of a particular species (in $\mathrm{kgm^{-2}s^{-1}}$) can be written as:

$$E = \sum_i A_i \; B_i \; C_i \; \mathrm{EF}_i \tag{4.5}$$

where A_i is the fraction surface area for ecosystem i that burns at a given time ($\mathrm{m^2m^{-2}}$), B_i is the quantity of biomass that is susceptible to burn ($\mathrm{kgm^{-2}}$), C_i is the rate of combustion of the biomass ($\mathrm{s^{-1}}$), and EF_i is the emission factor ($\mathrm{kg\ kg^{-1}}$) for species i. As a general rule, A_i can be obtained from satellite monitoring while the other parameters come from in situ measurements during field experiments. C_i is the inverse of the timescale needed for ecosystem i to burn. In practice, the two terms C_i and EF_i may be estimated together as a product. They represent an average of emission factors during the phase of *flaming* (early stage of the fire with more complete combustion) and *smoldering* (later stage of the fire with more incomplete combustion).

Because of their high temperature, fires emit infrared radiation that can be measured by those Earth's observation satellites that have channels in the near-infrared atmospheric window near 3.7 μm (Wooster et al. 2005). Such satellite measurements can be used to retrieve the radiative power of the fire throughout the electromagnetic spectrum, which itself is a good indicator of the energy released by the fire. The emission flux can then be estimated as:

$$E = m \; \mathrm{FRP} \; \mathrm{F}_{\mathrm{CO}_2} \; \mathrm{EF}/\mathrm{EF}_{\mathrm{CO}_2} \tag{4.6}$$

where FRP is the fire radiative power (in $\mathrm{Wm^{-2}}$) as measured by satellite or as interpolated in time between two satellite measurements, m is a multiplicative factor greater than 1 to take into account the smaller fires which are "missed" by the satellite instrument, $\mathrm{F}_{\mathrm{CO}_2}$ is the emission factor for $\mathrm{CO_2}$ per unit of energy radiated seen by the satellite ($\mathrm{kgJ^{-1}}$) and $\mathrm{EF}/\mathrm{EF}_{\mathrm{CO}_2}$ is the ratio between the emission factors for the species under consideration and for $\mathrm{CO_2}$. Like for the previous technique, the various emission factors are measured in situ during field experiments and can differ for different ecosystems. The advantage of this method is that what the satellite sees is more directly related to a rate of combustion and therefore an emission rate than the burned surface area, but it nevertheless requires some calibration as only a small fraction of the energy released during a fire is radiated to space.

In both cases, it is important to resolve the seasonal cycle of emissions which dominate during the dry season. Fires, at least the smaller ones, also have a diurnal cycle, which can be important to represent in models because this diurnal cycle interacts with the diurnal cycle in meteorological conditions (convection, boundary layer mixing) that govern the dispersion of the aerosol plume. It is also important to model the vertical distribution of the emission flux from vegetation fires. Fires generate energy that can trigger convection and lift the aerosol plume quickly in the

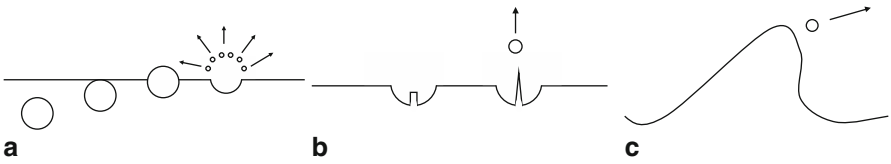


Fig. 4.1 Schematic description of the three different source mechanisms for sea spray aerosols: **a** ejection of droplets from the rupture of the water film located at the top of a bursting bubble on the ocean surface (film droplets), **b** ejection of droplets from the breaking of the water jet that follows the bursting of the bubble (jet droplets), and **c** production of droplets from the wind friction on the wave crests (spume droplets). (Adapted from Lewis and Schwartz (2004))

vicinity of the fire. This can be represented in an *ad hoc* way or with a plume rise model that depends on environmental conditions.

4.2.4 Sea Spray

The ocean is a source of liquid particles composed of sea water, which once in the atmosphere shrink as they evaporate and get enriched in sea salt and other substances that are present in the sea water. These particles are called sea salt or more generally sea spray as they may also contain organic material. There are two main mechanisms responsible for the emission of sea spray in the atmosphere: the bursting of air bubbles in the ocean and the wind tearing off wave crests (see Fig. 4.1). When the wind is strong enough, the friction of the wind on the ocean surface moves the water of the surface more rapidly than the water of the subsurface, which causes the wave to break. This incorporates air bubbles in the ocean, that can subsequently rise to the surface, and break up. When there are many air bubbles, the ocean is covered with foam also called whitecaps (Monahan et al. 1983); however air bubbles can also be present in much smaller concentrations. Biological processes may also form air bubbles in the ocean, but this source is thought to be less important than that from the wind. The bursting of the air bubbles at the ocean surface gives rise to the production of seawater particles by two distinct mechanisms. Droplets get formed during the bursting of the water film around the bubble itself. This ejects particles of relatively small size in all directions. These are called *film droplets*. Some droplets are also ejected vertically when the interior of the bursted bubble fills up with water and a vertical jet is formed. These are called *jet droplets*. This process only ejects a few droplets per air bubble, but there is a large number of air bubbles whose size is optimal to cause the emission of jet droplets. Finally, droplets can be teared off the crests of breaking waves. These droplets are larger than those coming from the other two mechanisms and are called *spume droplets*.

The windspeed is the primary factor that affects the production rate of sea spray droplets. Existing parametrizations rely on a small number of field measurements. They relate in some empirical way the surface wind speed (typically taken at 10 m

height and denoted U^{10}) to the rate of emissions of sea spray particles with radius r. The flux of emitted particles (in particles $\mathrm{m}^{-2}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$) is often approximated by an equation of the type:

$$\frac{\mathrm{d}F}{\mathrm{d}r} = f(U^{10}, r). \tag{4.7}$$

It is important to specify the altitude, z, of the wind speed that is considered because wind speed varies rapidly close to the surface, following the expression:

$$U(z) = \frac{U^*}{\kappa} \ln\left(\frac{z}{z_0}\right) \tag{4.8}$$

where U^* is the wind friction (which expresses the surface turbulent flux in the boundary layer under the form of an equivalent wind speed), z_0 is the surface roughness length (i.e. the altitude where $U(z)$ is equal to zero, generally of the order of one tenth of the size of the obstacles), and κ is the von Kármán constant ($\kappa \approx 0.35$–0.40). The reader is redirected to the book of Lewis and Schwartz (2004) for a review of existing parametrizations. Another important factor is atmospheric stability that governs vertical motions in the boundary layer and thus the capacity of droplets emitted at the surface to be entrained in the vertical and stay in suspension in the atmosphere. Rain or the presence of a film of organic substances at the surface of the ocean also play a role in the emission process of sea spray particles. The sea state is not just determined by the local wind speed, and would seem to be an important parameter to consider when designing more physically-based parametrizations of sea spray emissions. However, no parametrizations exist yet that use the information on the sea state available from one of the existing wave models.

4.2.5 Desert Dust

Deserts, arid, and other semi-arid regions host an important source of mineral aerosols, often called desert dust. The emission mechanism of such aerosols is depicted on Fig. 4.2. The wind friction on bare soil (i.e. soil without any vegetation) induces a horizontal flux of large (sand-sized) particles which are in fact aggregates containing smaller (clay or silt-sized) particles and can break when they fall back to the surface. The movements of the coarse soil particles are known as *creeping* and *saltation*, respectively, and their breaking as *sandblasting*. The wind can then generate a vertical flux of particles that are sufficiently small to be entrained in the boundary layer. Emissions of desert dust depend on characteristics of the parent soil, soil humidity, the amount of vegetation, and the surface wind speed. They are inhibited in the presence of snow (which can fall in some high-altitude deserts) or a too large soil humidity.

The study by Marticorena and Bergametti (1995) provides a good starting point to understand the physics of desert dust emissions. The motion of the soil particle is

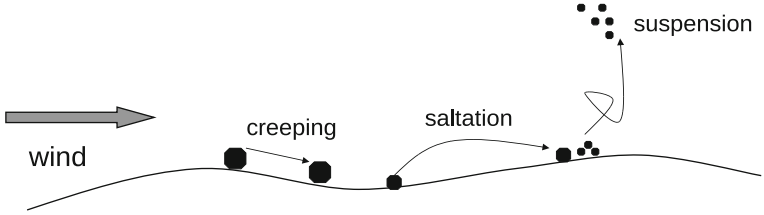


Fig. 4.2 Schematic description of the source mechanism for desert dust particles: creeping of the bigger particles which are rolling along the ground, saltation of the particles small enough to be lifted by the wind before falling, bouncing against and breaking up some of the soil aggregates into smaller particles, and suspension of the smallest particles

controlled by the sum of the forces that are exerted upon it, namely the friction force of the wind, the gravity force, and cohesion forces with neighboring particles. There is a threshold wind speed beyond which a particle starts moving. This wind speed threshold depends on the size of the particles being considered, and goes through a minimum for typical sizes of about 50 μm because cohesion forces increase strongly when particle size decreases, whereas the gravity force increases strongly with particle size. Only the smallest particles (< 60 μm) can be lifted up and transported by wind gusts, whereas larger particles (> 60–2000 μm) are seldom lifted higher than 1 m before falling down again. These larger particles, as well as those that are too big to be lifted but can roll or *creep*, contribute to the horizontal flux. One distinguishes the process of saltation, whereby the larger particles bombard the soil aggregates and generate smaller particles, from the process of suspension itself, whereby the smaller particles actually become suspended in the atmosphere.

Laboratory and field measurements have shown that the horizontal flux depends on an exponent of the friction wind speed beyond a threshold friction wind speed, denoted here U_t^*. The presence of non-erodible elements can modify the emission mechanism. These influence the threshold wind speed in two different ways: first by limiting the soil fraction that can be eroded and contribute to the emission of dust particles, and second by slowing down the wind at the surface, thus diminishing the momentum that is available to erode the erodible fraction of the soil. This effect can be described physically by partitioning the wind friction between erodible and non-erodible elements or empirically by adjusting the threshold friction wind speed according to the region, its elevation and/or soil type. Marticorena and Bergametti (1995) showed that the horizontal flux can be expressed in the following way:

$$G = c \frac{\varrho_a}{g} U^{*3} \left(1 + \frac{U_t^*}{U^*}\right) \left(1 - \frac{U_t^{*2}}{U^{*2}}\right) \tag{4.9}$$

where c is a constant, ϱ_a is the aerosol density, g is the gravity constant, and U^* is the friction wind speed. In practice, the horizontal flux is a function of the particle size and Eq. 4.9 must be integrated analytically or numerically over the size distribution of the parent soil particles. The emission flux of desert dust particles, i.e. the vertical flux or F, is composed of the finest particles and represent the mass of particles that

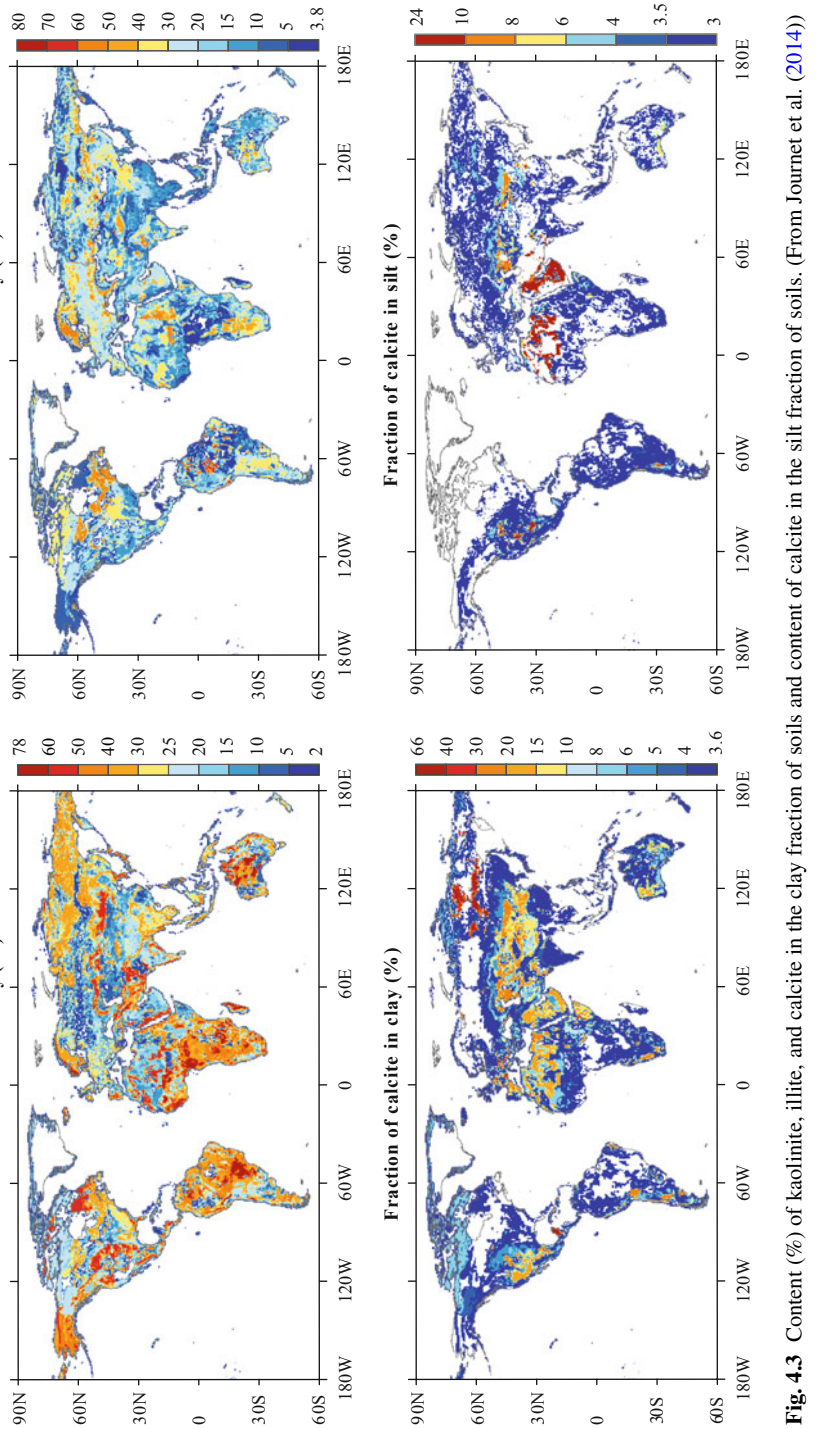


Fig. 4.3 Content (%) of kaolinite, illite, and calcite in the clay fraction of soils and content of calcite in the silt fraction of soils. (From Journet et al. (2014))

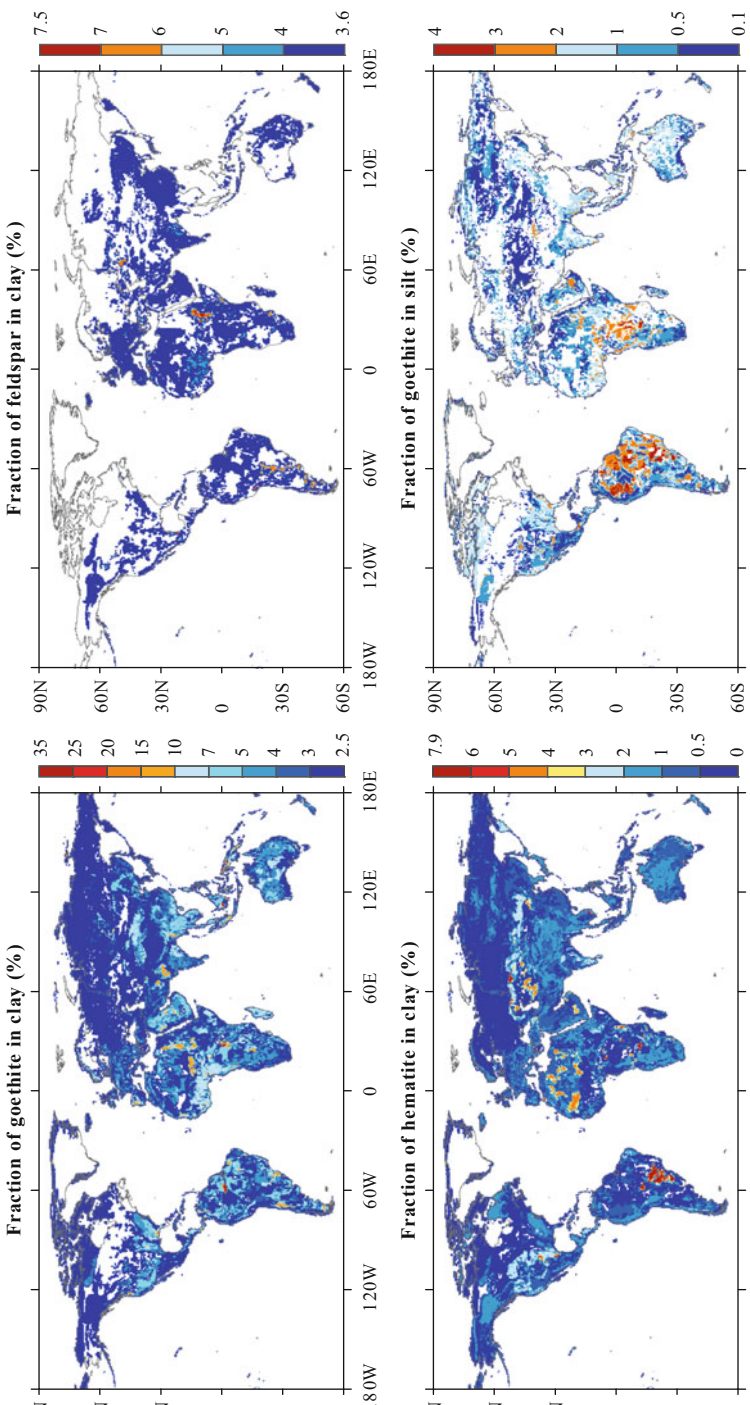


Fig. 4.4 Content (%) of goethite, feldspar, and hematite in the clay fraction of soils and content of goethite in the silt fraction of soils. (From Journet et al. (2014))

flow upward per unit time and unit surface through a horizontal surface. This surface should be high enough so that aerosols that flow through it can be considered as being emitted in the boundary layer. In first approximation, one usually considers that the vertical flux is proportional to the horizontal flux integrated over the size distribution. The vertical flux has a size distribution that depends on the size distribution of the parent soil particles and on the wind speed (Alfaro and Gomes 2001) but that is different from that of the horizontal flux because of the size-selectivity of both the saltation and sandblasting processes described earlier. The ratio between the vertical and horizontal fluxes, F/G, depends on the mineralogical composition of the soil in general, and the clay content in particular. Gillette (1979) showed that the F/G ratio tends to increase with the clay content up to contents of 20 %. However, soils containing large amounts of clay can develop a crust that inhibits saltation and thus prevents soil erosion.

More comprehensive dynamical model exist that aim to represent the saltation process. Increasing computer time makes it possible to account for the motions and collisions of sand particles. The non-spherical shape of the particles can also play a role (Wang et al. 2014).

We have seen already that the texture and granulometric composition of the parent soil impact on the dust uplift (Kok et al. 2011a, b). It is usual to distinguish the soil particles according to their sizes, illustrated by the relative importances of the clay, silt and sand fractions. The soil mineralogical composition, for which Journet et al. (2014) gathered a unique dataset, has an impact on the chemical and microphysical properties of the emitted mineral dust. Figure 4.3 shows the contents of some minerals (two silicates namely kaolinite and illite as well as calcite of chemical composition $CaCO_3$) in the clay fraction, and the content of calcite in the silt fraction. This latter species is one of the main alcaline material contained in soils that can contribute to decrease the acidity of the aerosol and rain. Figure 4.4 further shows the contents of two iron oxides (goethite and hematite) in the clay fraction and the content of goethite in the silt fraction. These species are important for at least two reasons: first they absorb solar radiation and their concentration has an impact on the single scattering albedo of mineral dust; second they constitute a source of iron for the ocean through wet scavenging. The fraction of feldspar, reported by Atkinson et al. (2013) as an important contributor to ice nucleation, is also shown.

4.2.6 Dimethylsulphide

Dimethylsulphide (DMS) is an organosulphurated compound of molecular formula CH_3SCH_3. It is the oxidation product of dimethylsulphidepropionate (DMSP) which comes from marine phytoplankton. The DMS produced in the ocean can be emitted to the atmosphere or destroyed in the ocean by the ultraviolet radiation that penetrates the surface layer. If the DMS concentration in oceanic surface waters is known, either from a climatology or from a marine biological model, the flux of DMS to the atmosphere can be estimated using one of the parametrizations of air–sea fluxes

available in literature. In a general manner, the flux of oceanic DMS towards the atmosphere is expressed by:

$$F_{\mathrm{DMS}} = K(T)\ \left(c_{\mathrm{DMS,ocean}} - k^{H}_{\mathrm{DMS}}\ p_{\mathrm{DMS,atm}}\right) \approx K(T)\ c_{\mathrm{DMS,ocean}} \tag{4.10}$$

where $K(T)$ is the velocity of the DMS transfer through the air–sea interface (in m s^{-1}), $c_{\mathrm{DMS,ocean}}$ is the concentration of DMS in the surface ocean (in mol m^{-3}), $p_{\mathrm{DMS,atm}}$ is the atmospheric pressure of the DMS (in Pa), and k^{H}_{DMS} is the Henry constant for DMS (expressed here in mol m^{-3} Pa^{-1}). The DMS solubility being weak, the DMS flux depends mostly on the transfer velocity and the DMS concentration in the ocean surface layer. The transfer velocity can be parametrized as a function of the surface wind speed (for example U^{10}) and the characteristics of the air–sea interface (see e.g. Liss and Merlivat 1986; Kettle and Andreae 2000).

There exist many in situ measurements of the concentration of oceanic DMS, however their spatial coverage is insufficient and a global scale climatology of oceanic DMS, therefore, has to rely on somewhat crude assumptions. It is possible to estimate the chlorophyll content of sea water by satellite imagery and relate this quantity to the concentration of phytoplankton and the primary marine productivity, but it is difficult to convert these quantities in DMS concentration because different species of phytoplankton produce DMS in very different quantities. Coccolithophores[1], for example, produce large quantities of DMS in comparison to other species. In general, non-diatoms, to which coccolithophores also belong, produce more DMS than diatoms. Figure 4.5 shows the seasonal variation in the climatological DMS sea surface concentration (Lana et al. 2011).

Modelling of marine biology for the study of climate remains embryonic. The large majority of current models rely on the NPZD approach that simulates the concentration in nutrients (N), phytoplankton (P), zooplankton (Z), and detritus (D). Some models start to incorporate a representation of the competition occurring between species, either between diatoms and non-diatoms, or in a more general manner between a predefined number of families of species that share specific biological functions.

The concentration of DMS in the ocean can then be modelled either in an empirical way from the concentration of phytoplankton and the mixed layer depth, or in a prognostic way by resolving the equilibrium between sources (from phytoplankton) and sinks (destruction by the ultraviolet radiation and emission to the atmosphere).

We also refer to Chap. 11 to learn more about the response of DMS to climate change.

[1] Coccolithophores are unicellular algae that protect their cell under a layer of calcite plates.

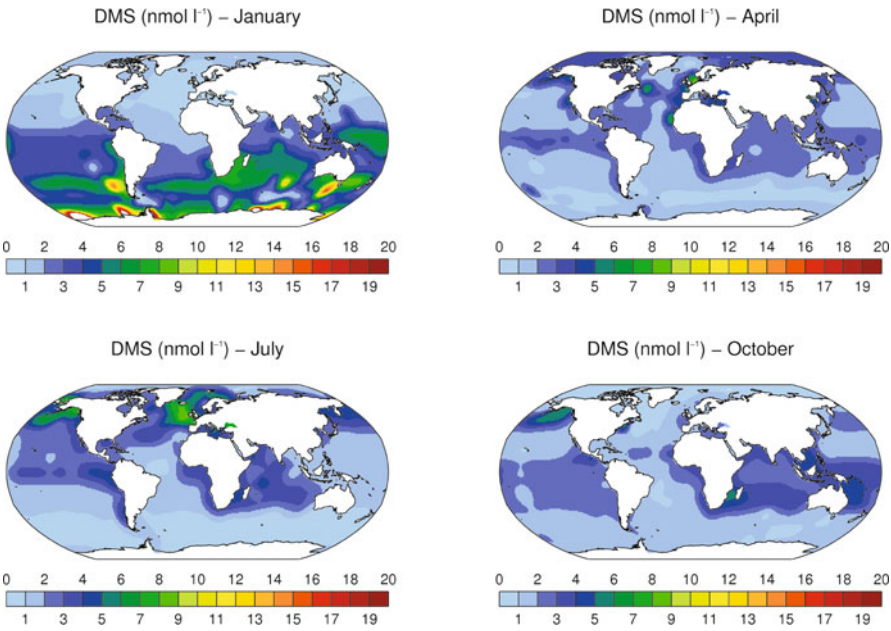


Fig. 4.5 Climatology of DMS sea surface concentration (nM or nmol l^{-1}) for January, April, July, and October. (Data from Lana et al. (2011))

4.2.7 Biogenic Volatile Organic Compounds

Vegetation is another source of volatile organic compounds (VOCs) which are emitted in very variable quantities according to plant types and environmental conditions. The role of these emissions for the plant physiology is not completely elucidated but they could be linked to protection mechanisms against external aggressions. The main compound that is emitted is isoprene, but only a small fraction of the emissions are converted to semi-volatile compounds that contribute to the formation of secondary organic aerosols. Compounds that lead to the formation of secondary organic matter in a more systematic way include monoterpenes and sesquiterpenes. Emissions of VOCs have been measured for a number of ecosystems under a number of environmental conditions so that parametrizations have been developed; such parametrizations remain empirical, however. In a general way, the flux of a given VOC at the leaf scale can be expressed by the following equation:

$$F_{\mathrm{VOC}} = k\,(p_{\mathrm{leaf}} - p_{\mathrm{air}})/R \tag{4.11}$$

where k is the diffusion coefficient, p_{leaf} and p_{air} are the partial vapour pressures of the VOC in the leaf and in the atmosphere, and R is a resistance term to the transfer. The different terms to this equation are difficult to estimate, which limits its applicability at the large scale. One often prefers a semi-empirical parametrization

whereby the flux of VOC is expressed as a function of macroscopic parameters:

$$F_{\mathrm{VOC}} = \sum_{\mathrm{PFT}} f_{\mathrm{PFT}}\ F_{\mathrm{PFT}}(T, \mathrm{PAR}, \mathrm{CO_2}) \tag{4.12}$$

where PFT designates the various plant functional types (which are thought to behave similarly in terms of their VOC emissions), f_{PFT} the areal fraction covered by a particular plant type, F_{PFT} is the VOC flux for this plant type and depends on the temperature T, photosynthetically active radiation (PAR) and the CO_2 concentration (Fig. 4.6). Arneth et al. (2007) have shown that an enhanced CO_2 concentration is able to reduce or inhibit the VOC emissions. This effect is particularly important in the context of a warmer future climate (see Chap. 11). Eventually, the VOC emission flux per unit surface of a given plant type can be expressed as:

$$F_{\mathrm{PFT}} = F_{\mathrm{leaf,PFT,standard}}\ \frac{\mathrm{NPP}}{\mathrm{NPP_{standard}}}\ f_T\ f_{\mathrm{CO2}}\ \mathrm{LAI_{PFT}} \tag{4.13}$$

where $F_{\mathrm{leaf,\ PFT,\ standard}}$ is the flux per unit surface of leaves in standard conditions, $\mathrm{NPP/NPP_{standard}}$ is the ratio between the (observed or simulated) primary productivity and the primary productivity in standard conditions, f_T is a corrective factor to account for temperature variations, $f_{\mathrm{CO_2}}$ a corrective factor to account for the dependence in the CO_2 atmospheric concentration, and LAI is the leaf area index (i.e. the total surface of leaves per unit surface at the ground). Such parametrizations remain largely empirical.

4.2.8 Volcanoes

Volcanoes emit aerosols in the form of ash and aerosol precursors in the form of SO_2, H_2S, and a few other minor species. Volcanic ash is of little climatic significance is not discussed further. Emissions of SO_2 are substantial however and have an impact on both the background tropospheric sulphur budget and the stratospheric aerosol. Emissions are separated into emissions from quiescent and explosive volcanoes. A number of volcanoes are now equipped with surface monitoring devices so that SO_2 emissions can be estimated in near-real-time. Observations of volcanic SO_2 emissions from satellite is also increasingly used for large enough eruptions that take SO_2 concentrations above the detectability thresholds of current satellite instruments. Emission inventories are built by aggregating information from the various volcanoes.

4.2.9 Resuspension

In some locations, such as in urban areas, especially near industrial plants and busy roads, and close to strong source regions of dust, it may be necessary to account

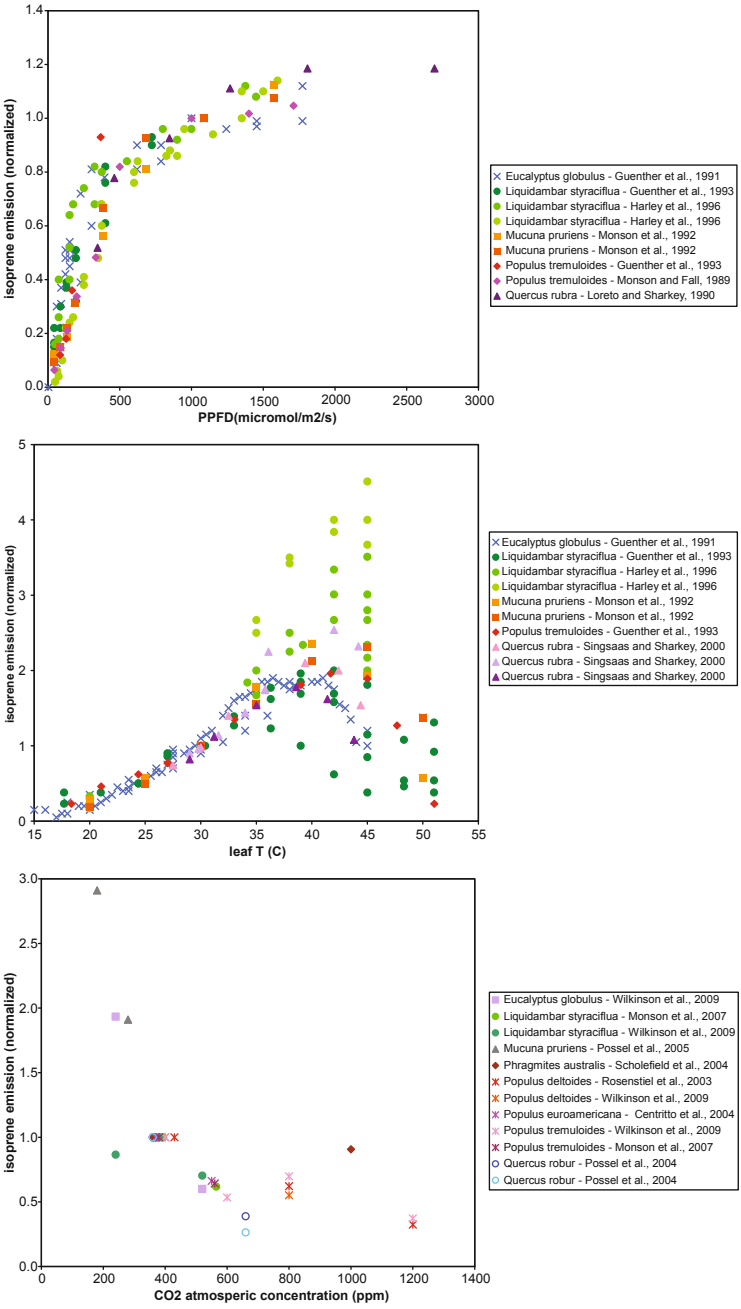


Fig. 4.6 Isoprene emissions by different types of vegetation as a function of solar radiation at the surface (in μmol m^{-2} s^{-1}), leaf temperature (in K), and the atmospheric concentration in CO_2 (ppmv). From Pacifico et al. (2009). (© British Crown Copyright 2009, Met Office)

for the resuspension of industrial and desert dust particles previously deposited at the surface. Resuspension can be treated as an additional source of aerosols that is triggered under suitable wind conditions and soil dryness.

4.3 Atmospheric Processes

Once emitted, aerosols and aerosol precursors undergo a number of atmospheric processes which determine their residence time in the atmosphere. This section reviews the various important aerosol processes.

4.3.1 Nucleation

Nucleation is the transformation of matter from one phase to another phase through the formation of nuclei. For atmospheric aerosols, nucleation refers to the transformation of gas-phase molecules into a cluster of molecules called an aerosol embryo or an aerosol nucleus. This process is sometimes called *homogeneous nucleation* to distinguish it from the process of *heterogeneous nucleation* where the phase change occurs on a pre-existing surface. However, we follow the mainstream aerosol terminology and use the term "nucleation" to designate homogeneous nucleation and the term "condensation" to refer to heterogeneous nucleation. One can further distinguish homomolecular nucleation, where only one gas-phase species is involved (e.g. sulphuric acid), from heteromolecular nucleation which implies several gas-phase species (e.g. sulphuric acid and water vapour). Heteromolecular nucleation can be binary (i.e. only two species are involved) or ternary (i.e. three different species are involved).

Nucleation is a complex process that is not completely understood even though new in situ measurements of very fine (nanometre-sized) particles now allow to better apprehend conditions favoring the formation of new particles. The nucleation process is nevertheless very important in that it influences the dynamics of the aerosol size distribution and the concentration of cloud condensation nuclei, and hence aerosol–cloud interactions.

Theoretically, nucleation requires a supersaturation of the gas phase species involved (i.e. the vapour pressure of the gas must exceed the saturation pressure that is usually taken relative to a flat surface). However not all clusters (or embryos) formed by nucleation become aerosol particles. For this to happen, it is necessary that the cluster size exceeds a critical radius beyond which there are more new molecules condensing on the cluster than molecules evaporating away from it. Strictly speaking, one should distinguish the process of *aerosol nucleation* from that of *new particle formation* because a cluster formed by nucleation can disappear before it gets a chance to lead to a new stable particle. However, we will not make this distinction in the rest of this chapter.

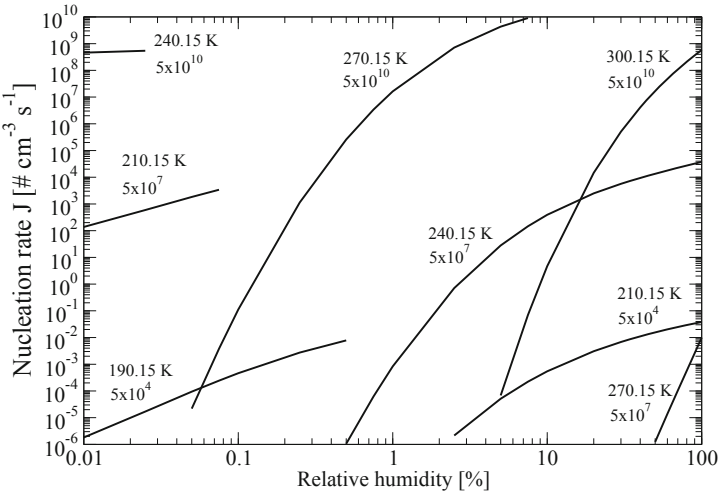


Fig. 4.7 Nucleation rate (#cm^{-3} s^{-1}) as a function of the relative humidity (%) for different values of the temperature (in K) and atmospheric concentration in gas-phase sulphuric acid (molecules cm^{-3}). (Adapted from Vehkamäki et al. (2002).

The nucleation process can be described from thermodynamic laws. The higher the supersaturation, the smaller the critical radius. Nucleation occurs when the total surface concentration of pre-existing particles available for condensation is small. However observations show that nucleation is more frequent than early theories would predict. In particular it was initially thought that nucleation would only occur in the stratosphere and in the free troposphere in very clean conditions. It is now established that nucleation events can occur in the continental boundary layer, and in particular above forested areas.

Parameterizations of aerosol nucleation are still evolving. They are generally based on theoretical models that are evaluated or adjusted based on nucleation rates observed in controlled conditions in the laboratory or ambient conditions in the real atmosphere. One can cite for instance the parametrization of Vehkamäki et al. (2002) for the nucleation of sulphuric acid which is expressed as a function of the gas-phase concentration of sulphuric acid, [H_2SO_4], temperature, T, and relative humidity, RH:

$$J(\mathrm{cm}^{-3}\,\mathrm{s}^{-1}) = f(\mathrm{H_2SO_4}, \mathrm{RH}, T). \tag{4.14}$$

The nucleation rate J is represented on Fig. 4.7. It increases very strongly with the concentration of sulphuric acid, and strongly with relative humidity which favours the apparition of clusters sufficiently large to be stable. A number of recent experiments have shown the important role of ammonia (NH_3), certain volatile or semi-volatile organic compounds and ions that can increase nucleation rates by several orders of magnitude (Kirkby et al. 2011). Ammonia and volatile organic compounds are found preferentially close to the surface where they are emitted, while ion concentrations may be larger in the upper troposphere and stratosphere where the ionization rate due to cosmic rays is larger.

4.3.2 Condensation of Semi-Volatile Compounds

Semi-volatile compounds, preferentially, condense onto the surface of pre-existing aerosol particles, and this represents a dominant process in terms of how much mass is transferred from the gas phase to the particulate phase. Among semi-volatile compounds, one can cite sulphuric acid (H_2SO_4), for which the saturation vapour pressure is low, methanesulphonic acid (MSA), and a number of organic compounds such as monoterpenes.

The rate of condensation of these semi-volatile species onto a given aerosol particle is a function of the vapour pressure of the gas phase species, its saturation vapour pressure, the curvature radius of the particle, and the affinity between the gas and the particle (for instance the hydrophilic character of a particle for a gas that is soluble in water). The condensation of a gas phase species is easier onto a plane surface than onto a curved (convex) surface. It is, therefore, easier onto a large aerosol particle than onto a smaller aerosol particle. This is expressed by the *Kelvin effect* which gives the saturation vapour pressure onto a curved surface:

$$P_{\mathrm{sat},r} = P_{\mathrm{sat}} \exp\left(\frac{2\,\sigma\,M}{\varrho\,R\,T\,r}\right) \tag{4.15}$$

where r is the particle radius (in m), σ is the surface tension (in J m^{-2}), M is the "molecular mass" of the aerosol (in kg mol^{-1}), ϱ is the density of the solution (in kg m^{-3}), $R = 8.314$ J mol^{-1} K^{-1} is the ideal gas constant, and T is the temperature. The concentration of the gas at equilibrium above the curved surface is therefore:

$$c_{\mathrm{sat},r} = \frac{P_{\mathrm{sat},r}\,M}{R\,T} \tag{4.16}$$

where M is the molecular mass of the considered gas (kg mol^{-1}). $P_{\mathrm{sat},r}$ and $c_{\mathrm{sat},r}$ are larger than P_{sat} and c_{sat} and tend towards these values when r tends towards infinity.

The condensation rate (in kg s^{-1}) can then be expressed in the following way:

$$J = 4\pi\,r\,D\,f(Kn,a)\,(c - c_{\mathrm{sat},r}) \tag{4.17}$$

where c is the atmospheric concentration of the gas-phase species (in kg m^{-3}), D is the diffusion coefficient of the gas-phase species (in m^2 s^{-1}), f is a corrective factor that depends on the Knudsen number Kn (which itself determines the flow regime of the gas around the particle) and the accommodation coefficient a.

4.3.3 Coagulation

Coagulation is a process by which aerosol particles agglomerate together to form larger particles. Brownian motion of the aerosol particles, resulting from their collisions with fast-moving air molecules, is the main cause of coagulation in the

atmosphere. This is a particularly important process when concentrations are large, as it is the case for small particles in polluted areas, and more so in aerosol plumes themselves. In the troposphere, coagulation is a process that contributes to increase the degree of mixing of the aerosol. Coagulation is also important for accumulation-mode aerosols in the stratosphere, because their lifetime is longer than in the troposphere. This process allows aerosol particles to grow to sizes where sedimentation becomes significant (see below).

The equation governing the number concentration $c(m,t)$ of particles of mass m at time t due to coagulation is given by:

$$\begin{aligned}\frac{\partial\, c(m,t)}{\partial t} = \quad & 1_{m\geq 2m_0}\ \tfrac{1}{2}\int_{m_0}^{m-m_0} K(m', m-m')\, c(m',t)\, c(m-m',t)\, \mathrm{d}m' \\ & -c(m,t)\ \int_{m_0}^{\infty} K(m,m')\, c(m',t)\, \mathrm{d}m'\end{aligned} \tag{4.18}$$

where m_0 is the mass of the smallest aerosol (typically the mass of aerosols coming from nucleation), $1_{m\geq 2m_0}$ is a function that is equal to 1 when $m \geq 2m_0$ and 0 otherwise, $K(m_1, m_2)$ is the coagulation kernel that describes coagulation between two aerosols of masses m_1 and m_2. The first term in the right-hand side of the equation describes the source of particles of mass m coming from coagulation of particles of masses m' and $m - m'$, while the second term describes the sink of particles of mass m by coagulation with all other particles. The coagulation kernel is expressed in unit of volume per unit time ($\mathrm{m}^3\ \mathrm{s}^{-1}$). The reader is redirected to Sportisse (2010) for a more detailed description of the coagulation process.

4.3.4 In-Cloud Aerosol Production

Soluble species can dissolve in liquid cloud water, undergo chemical reactions and be released in the particulate phase (i.e. within aerosol particles) when the cloud droplets evaporate. This is in particular the case for sulphur dioxide (SO_2), a significant part of which is oxidized in the aqueous phase within cloud droplets to form sulphate. To quantify, this effect requires to model not only the mass transfer between the gas phase and the aqueous phase, but also the chemical reactions that take place in the aqueous phase and the subsequent release of sulphate to non-cloudy air. Some of these chemical reactions occurring in the aqueous phase depend on the pH (i.e. a measure of the acidity) of the cloud water[2].

The equilibrium between the gas phase and the aqueous phase for a soluble gas follows Henry's law. It is expressed as:

$$c^X_{\mathrm{aq}} = k^X_H\ p^X_{\mathrm{atm}} \tag{4.19}$$

[2] $\mathrm{pH} = -\log_{10}[\mathrm{H}^+]$ where $[\mathrm{H}^+]$ is the concentration in H^+ (or equally $\mathrm{H_3O^+}$) cations in water in unit mol l^{-1}.

where c_{aq}^{X} is the concentration of species X in the aqueous phase (in mol m^{-3}), p_{atm}^{X} is the atmospheric concentration of species X in the gas phase (in Pa), and k_{H}^{X} is the Henry constant for species X (in mol m^{-3} Pa^{-1}). It should be noted that if species X dissociates in the aqueous phase, it is indeed the concentration of X –and not the total concentration in X and its dissociation products– that enters Eq. 4.19. Practically, one often considers that the mass transfer between the gas and aqueous phase is sufficiently fast so that equilibrium is reached at any time but this is not necessarily the case for all species. This assumption may also not be justified for falling hydrometeors such as raindrops.

Other processes contribute to the evolution of the properties of aerosols that are incorporated into clouds. Interstitial aerosols (i.e. aerosols not incorporated in a cloud droplet) can coagulate with liquid cloud droplets, thereby increasing the mass of soluble and insoluble material in the cloud phase. Cloud droplets also collide with each other and coagulate. When a cloud droplet evaporates, it releases the cloud condensation nucleus it nucleated upon and all the soluble and insoluble material it has collected during its lifetime in the cloud. This process is particularly important for non-precipitating (stratiform) clouds, but also, albeit to a smaller extent, for precipitating clouds and convective clouds that detrain a fraction of their liquid water in the environment outside the cloud. The net effect is to shift the aerosol size distribution towards larger particle sizes, and to increase the degree of mixing of the aerosol, and therefore to increase the concentration in cloud condensation nuclei. Figure 2.6, discussed earlier, illustrates the impact of these processes on the aerosol size distribution in a polluted continental air mass that is advected over the Atlantic Ocean.

4.3.5 Wet Deposition

The main sink of aerosols in the atmosphere, at least for soluble species, is *wet deposition*, also called *wet scavenging*. It includes both scavenging that occurs within the clouds during the formation of precipitation (i.e. in-cloud scavenging), and scavenging that occurs under the cloud because of the downward flux of precipitating water (i.e. below-cloud scavenging).

When a liquid water cloud forms in the atmosphere, a fraction of the aerosol is incorporated in the aqueous phase, either as a cloud condensation nuclei (water vapour condenses upon a hydrated aerosol particle and becomes a cloud droplet), or through impaction (an interstitial aerosol particle is incorporated into the aqueous phase during the collision with a cloud droplet). If cloud droplets grow to a size where their sedimentation velocity is large enough, the cloud starts to produce drizzle or precipitation, and aerosol matter that is incorporated in the falling droplets is scavenged. If the falling droplets (or drops) reach the surface, then the aerosol mass is removed from the atmosphere.

During their fall, raindrops sweep a volume of air and collide with a fraction of the aerosols that are present in this volume. This process, called below-cloud scavenging, is less effective than in-cloud scavenging to remove aerosols from the atmosphere, but it can be important under certain conditions. One also needs to consider the reevaporation of rain, which releases aerosol particles in the atmosphere. Reevaporation is important in situation of drizzle where raindrops are relatively small (typically 100–200 nm radius) and can easily reevaporate in the drier air below the cloud. This leads to a redistribution of the aerosol concentrations on the vertical.

In first approximation, one can parametrize in-cloud scavenging from the vertical flux of precipitation:

$$F_{\text{in-cloud scavenging}} = W\ c = \beta\ f\ x\ c \tag{4.20}$$

where W is the scavenging rate in the cloud (in s^{-1}), c is the atmospheric concentration in aerosol, β is the conversion rate of cloud (liquid) water in rainwater (also called autoconversion, in $\text{kg kg}^{-1}\ \text{s}^{-1}$), f is the cloud volume fraction, and x is the fraction of aerosols in the aqueous phase. The parameter x depends on the aerosol solubility. It is equal to 0 for aerosols that are completely interstitial and 1 for aerosols that are soluble and large enough to serve as cloud condensation nuclei. The conversion rate β can be calculated from the cloud parameters:

$$\beta = -\frac{\text{div}_z\, \mathbf{P_r}}{\varrho_{\text{air}}\ f\ q_l} \tag{4.21}$$

where div_z is the divergence operator on the vertical, $\mathbf{P_r}$ is the vertical profile of the precipitation field ($\text{kg m}^{-2}\ \text{s}^{-1}$ and positive downwards), ϱ_{air} is the air density (kg m^{-3}), and q_l is the cloud (liquid) water content (kg kg^{-1}).

Below-cloud scavenging also depends on the precipitation flux. It can be parametrized by integrating the volume of air swept by a population of precipitating raindrops:

$$F_{\text{below-cloud scavenging}} = K\ c = \frac{3\ P_r\ \alpha}{4\ R_r\ \varrho_{\text{water}}}\ c \tag{4.22}$$

where K is the below-cloud scavenging rate (in s^{-1}), R_r is the radius of the raindrops, ϱ_{water} is the water density, and α is the efficacy with which aerosol particles are collected by raindrops. This efficacy is relatively small because aerosol particles tend to follow the air flow that is deviated around a raindrop. It depends on the relative sizes of the aerosol particles and falling hydrometeors. Pruppacher and Klett (1997) of 0.001 for raindrops and 0.01 for snowflakes.

Aerosol release upon reevaporation of the rain also needs to be parametrized. The complete evaporation of a raindrop is required before an aerosol is released. This means that the smallest raindrops will release their aerosol first while the largest raindrops may only shrink and reach the surface before being completely evaporated. Aerosol release can, therefore, be parametrized as a non-linear function of the reevaporation rate.

Ice clouds and solid hydrometeors (graupel and snowflakes) also play a role in the aerosol atmospheric cycle. For these clouds in-cloud scavenging represents a less efficient sink for aerosols however. Only a small fraction of the aerosols can serve as ice nuclei upon which ice crystals can form. Aerosols can also collide and stick more efficiently to ice crystals and falling hydrometeors.

4.3.6 Dry Deposition

If wet deposition is an efficient sink for aerosols, it is of course conditional on the presence of precipitating clouds whose spatial and temporal distribution is very heterogeneous. Some regions experience very little precipitation while other regions exhibit very strong seasonal variations in precipitation. In the absence of precipitation, the direct deposition of aerosols and aerosol precursors onto the Earth's surface, called dry deposition, becomes important. One can distinguish deposition at the surface induced by the turbulent flux of aerosols from that induced by sedimentation. The former dominates for smaller aerosols, while the latter dominates for larger aerosols. The deposition flux can be represented as the product of the deposition velocity and the aerosol surface concentration:

$$F_{\text{dry deposition}} = v_{\text{dry deposition}}\, c_{\text{surface}}. \tag{4.23}$$

Figure 4.8 shows the deposition velocity as a function of aerosol size for three different aerosol densities. There is minimum in the deposition velocity for particles of intermediate sizes, with radius between 0.1 and 1 μm. This minimum corresponds to the accumulation mode which dominates the concentration of aerosols in the absence of nucleation for smaller aerosols or other production mechanism for larger aerosols.

More complex models exist to represent the dry deposition of chemical species (including aerosols) at the Earth's surface. Some models rely on an analogy with electrical engineering, the surface (ground and vegetation) being represented as a combination of resistances that act in series or in parallel (Wesely and Hicks 2000).

Aerosols can also get deposited at the surface in the presence of fog or mist. In this case, aerosol particles are deposited together with the small water droplets that constitute the fog.

Finally, we have already seen that resuspension of aerosol particles deposited at the surface could play a role in polluted and/or dry regions.

4.3.7 Sedimentation

Sedimentation or gravitational settling of the aerosols is an important process, not only because it creates a sink of aerosols at the surface, but also because it redistributes

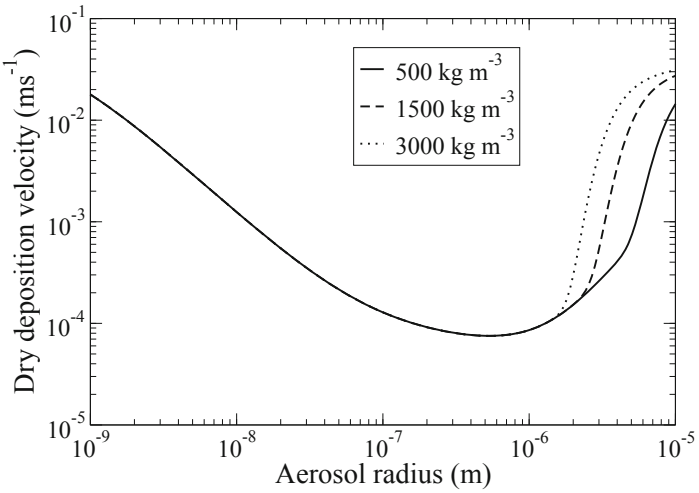


Fig. 4.8 Dry deposition velocity (m s^{-1}) as a function of particle size for three different values of aerosol density. (Reproduced from a computer code provided by Kerkweg et al. (2006))

aerosols on the vertical through a (relatively slow) downward motion. Ignoring other processes, sedimentation acts in the following way on the continuity equation for aerosols:

$$\frac{\partial\, \varrho\, C}{\partial t} + \mathrm{div}_z(\varrho\, C\, \mathbf{V}_{\mathrm{sed}}) = 0 \tag{4.24}$$

where C is the aerosol mixing ratio and $\mathbf{V}_{\mathrm{sed}}$ the sedimentation velocity.

Sedimentation is a slow process for accumulation mode particles. In the troposphere, large-scale transport, convective transport, and wet scavenging are the dominant processes. However in the stratosphere, where vertical transport is slow because of a larger stratification and aerosol residence time is long, sedimentation is an important process to transport aerosols into the troposphere.

Sedimentation is also important for aerosol particles of the coarse mode in the troposphere. In particular, desert dust particles can be transported to elevated heights during or after the emission process, and sedimentation contributes to bring them down closer to the surface where they can be deposited. Likewise, sedimentation plays a significant role for the coarse and supercoarse modes of sea spray aerosols.

The sedimentation velocity controls the sedimentation process. An aerosol particle is subject to the gravity force and to a drag (air resistance) force. The drag force acts on the particle in the opposite direction to its velocity, and its magnitude is proportional to the size of the particle. During the fall of a particle, the gravity and drag forces balance each other and one can estimate the particle terminal velocity:

$$V_s = \frac{(\varrho_p - \varrho_a)\, D_p^2\, g\, C_c}{18\, \mu}. \tag{4.25}$$

This relationship is known as the Stokes' law and the terminal velocity as the Stokes velocity. It involves ϱ_p the particle density, ϱ_a the air density, D_p the particle aerodynamic diameter, g the gravitational constant, μ the air viscosity, and C_c a correction factor. This correction factor is expressed as:

$$C_c = 1 + \frac{2\,\lambda}{D_p}\left(1.257 + 0.4\exp\left(\frac{-0.55\,D_p}{\lambda}\right)\right) \tag{4.26}$$

where λ is the mean free path of particles in the atmosphere.

4.3.8 Aerosol Transport

Aerosols are transported by the atmospheric flow, and this needs to be accounted for to explain the observed aerosol distribution. One can differentiate several transport mechanisms. First of all aerosols are advected by large-scale winds. There exist several numerical advection schemes and a discussion of their advantages and disadvantages is out the scope of this book. One usually favours an advection scheme that is mass-conserving (i.e. it does not create or lose mass) and positive (i.e. it does not create negative values), without being too diffusive. The choice of an advection scheme has to balance desired properties, accuracy, and computational cost.

One also needs to account for mixing in the boundary layer that has a direct influence on how surface emissions and deposition influence the atmospheric concentrations. Finally if convection is not resolved by the model, but is represented by a subgrid scale parametrization instead, then it is important to consider convective transport of aerosols and aerosol precursors. Once entrained into a convective cloud, only a fraction of soluble species can actually "escape" through detrainment, as the rest of the species is scavenged by convective rain.

4.4 Modelling Approaches

We have shown that a number of processes have to be considered and numerically resolved to describe the evolution of an aerosol population. Several approaches are possible depending on the modelling objectives, which we review here. One may want to model only the aerosol mass concentration and assume a constant size distribution. This approach, called the *bulk approach*, is simple and computationally cheap. To first approximation, it is justified by the fact that the accumulation-mode size distribution is relatively stable because coagulation and sedimentation are sinks that operate on both sides of the size distribution. However, the approach does not allow to simulate observed variations in aerosol size distribution, and it is becoming obsolete. If instead the objective is to model the aerosol size distribution, two approaches are possible: a *sectional approach* and a *modal approach*. In the former approach, one discretizes the aerosol size distribution in size bins and the evolution

of the size distribution is represented explicitly as the model tracks the aerosol concentration in each size bin. In the latter case, one assumes that there is a number of predefined aerosol modes whose properties (e.g. chemical composition and average size distribution) can vary within specific bounds.

4.4.1 Bulk Approach

First generation aerosol models did not seek to represent explicitly the aerosol size distribution and mixing state. In that case, only mass concentrations of the different aerosol species are prognostic model variables, for which the conservation equation is solved:

$$\frac{\partial c_i}{\partial t} + \mathrm{div}(c_i\ \mathbf{V}) = \mathrm{div}(K_m \nabla c_i) + S_i - P_i \tag{4.27}$$

where S_i and P_i represent, respectively, the sum of the sources and sinks for chemical species i.

The model assumes a typical size distribution when parametrizing processes that depend on particle size, such as sedimentation, scavenging, or the interaction of aerosols with radiation. Concentrations of cloud condensation nuclei are related to the mass concentrations of soluble aerosols using an empirical function. Coagulation between aerosol particles and condensation of soluble or semivolatile gaseous species onto the surface of pre-existing aerosols are processes that cannot be explicitly represented in such simple models. They have to be highly parametrized. For instance, the conversion of insoluble aerosols to soluble aerosols, which is particularly important for biomass burning aerosols and black carbon, can be represented through a simple aging process. This can be done by considering two aerosol classes, a (fresh) hydrophobic class and an (aged) hydrophilic class, and moving mass from one class to the other using a fixed timescale, λ, as shown in the equation below for black carbon:

$$\frac{\mathrm{d}[\mathrm{BC}_{\mathrm{hydrophilic}}]}{\mathrm{d}t} = -\frac{\mathrm{d}[\mathrm{BC}_{\mathrm{hydrophobic}}]}{\mathrm{d}t} = \frac{[\mathrm{BC}_{\mathrm{hydrophobic}}]}{\lambda}. \tag{4.28}$$

The bulk approach allows to simulate the different aerosol types and their large-scale spatial and temporal variations, but is not sufficient to characterize accurately aerosol properties. It has some capability to simulate accumulation-mode particles but is fairly limited for coarse-mode aerosols because aerosol size has a strong impact on deposition processes which cannot be parametrized accurately.

4.4.2 Sectional Approach

Some processes being sensitive to aerosol size, a more detailed approach consists in discretizing the aerosol size distribution into n classes of radius r_j and mass m_j

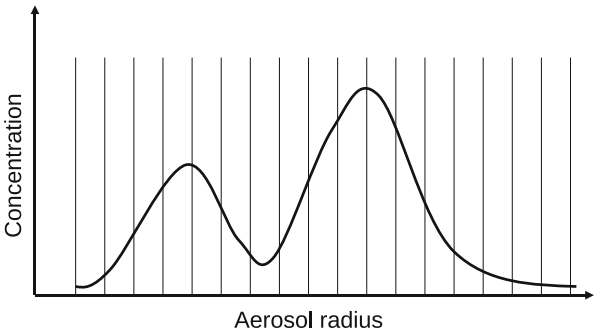


Fig. 4.9 Schematic representation of the sectional approach used for modelling atmospheric aerosols. The size distribution is discretized and the aerosol concentration in each section is a prognostic variable of the model

($j = 1, n$) (see Fig. 4.9). The number of size classes can vary from just a few to several hundreds depending on the required accuracy. The time evolution of the concentration of aerosols in size class j follows the conservation equation:

$$\frac{\partial c^j}{\partial t} + \mathrm{div}(c^j \, \mathbf{V}) = S^j - P^j \tag{4.29}$$

where c^j is the number concentration of aerosol in size class j and S^j et P^j are the source and sink terms for that size class.

Coagulation acts both as a sink of smaller particles and a source of larger particles. If the size discretization is such that $m_j + m_k = m_{j+k}$, the coagulation term can simply be written as:

$$\frac{\partial c^j}{\partial t} = \frac{1}{2} \sum_{k=1}^{k=j-1} K_{k,j-k} \, c^k \, c^{j-k} - c^j \sum_{k=1}^{k=n} K_{j,k} \, c^k \tag{4.30}$$

where $K_{j,k}$ is the coagulation kernel for aerosol with masses m_j and m_k.

The sectional approach predicts the number (or the mass) of particles in different size classes. Different chemical species can be handled through different variables but the approach does not lend itself easily to predicting the degree of mixing within each size class, unless the number of variables for each size class is increased dramatically.

4.4.3 Modal Approach

The modal approach results from a compromise that allows to represent the evolution of both the aerosol size distribution and the degree of mixing at an affordable computational cost but using a number of assumptions. It is assumed that atmospheric aerosols can be represented as the superposition of n predefined modes (Fig. 4.10). The number of modes is adjustable, depending on how many degrees of freedom one wants to introduce, but several authors consider that one can reasonably well represent the climate effect of atmospheric aerosols at the global scale with seven

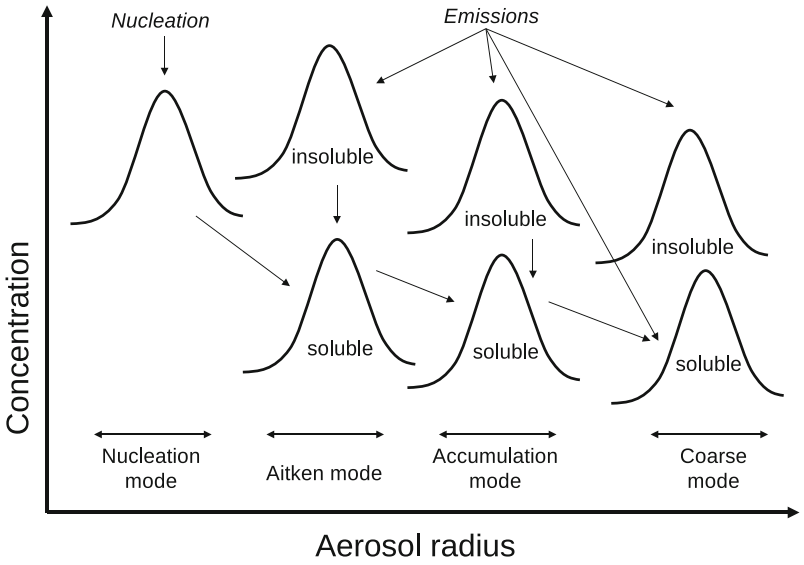


Fig. 4.10 Schematic representation of the modal approach used for modelling atmospheric aerosols. The aerosol number and mass concentrations for each mode are prognostic variables of the model. The *arrows* indicate how various processes transfer aerosol mass and number from one mode to another

modes (Stier et al. 2005). These modes are the nucleation mode, two Aitken modes, two accumulation modes, and two coarse modes. A supercoarse mode can be added if required.

For each of these modes, an insoluble and a soluble modes are considered. Each mode is represented by several variables, namely the aerosol total number concentration for that mode and the mass concentrations for each species present in that mode. The shape and width of the aerosol size distribution are fixed for each mode on the basis of observations or more sophisticated models. The various processes contribute to shift the size distribution of each mode within some permitted interval, by altering the ratio between number and mass concentrations. Aerosol aging is modelled through a transfer of aerosol number and mass from the insoluble to the soluble modes. Aerosol coagulation contributes to decrease the aerosol number concentration, but not the mass concentrations, and transfers aerosol mass from the smaller to the larger modes.

The modal approach has a number of advantages, as discussed above. It does not, however, provide any information on the size variations in the degree of mixing within a mode.

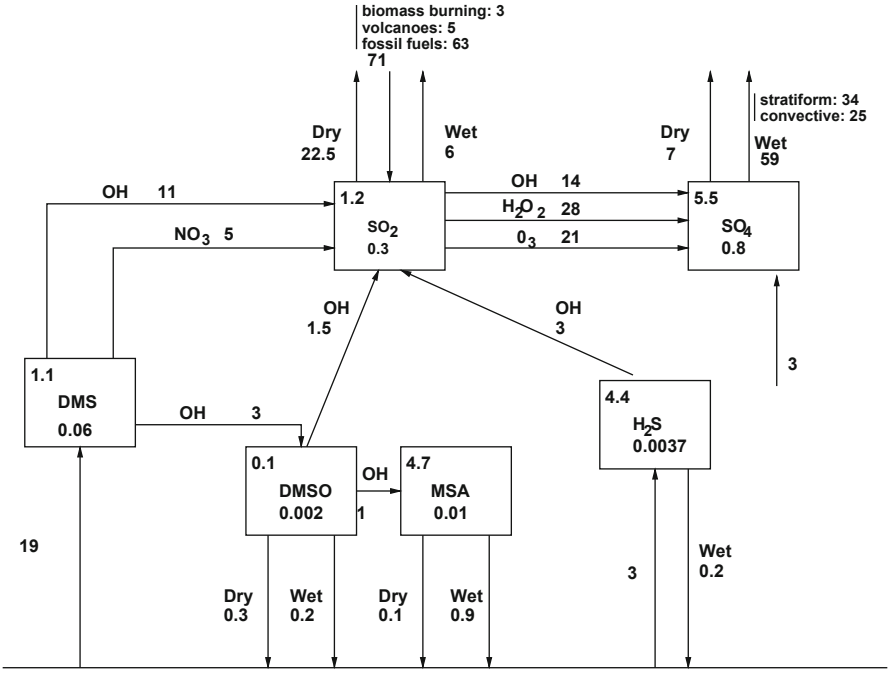


Fig. 4.11 Budget of the species involved in the sulphur atmospheric cycle. The fluxes (sources and sinks) are given in Tg S yr^{-1}, atmospheric reservoirs in Tg S, and residence time in days. (From Boucher et al. (2002))

4.5 Example: The Sulphur Budget

The reader is redirected to the scientific literature to learn about the improving capability of global models to simulate the global aerosol system (e.g. Mann et al. 2010). The particular case of sulphate aerosols is discussed below and in Fig. 4.11 using a fairly simple atmospheric model of the sulphur cycle. The main sources of sulphur (excluding sea salt sulphate) include dimethylsulphide (DMS) and sulphur dioxide (SO_2). The sources of DMS are exclusively natural while there are natural and anthropogenic sources of SO_2. It is also necessary to include a small source of hydrogen sulphide (H_2S) from biogenic and industrial origins. A fraction of the emitted SO_2 by industrial processes is oxidized rapidly at the source point (i.e. shortly after combustion takes place and before the plume gets diluted in the atmosphere), which can be parametrized as a small direct source of sulphate of anthropogenic origin. Sulphur follows an oxidation chain that goes from DMS to DMSO and SO_2, then sulphate (SO_4^{2-}) and methanesulphonic acid (MSA), the latter two species being essentially found in the particulate phase. The main oxidants playing a role in the oxidation of sulphur species are the hydroxyl radical (OH), the nitrate radical (NO_3), ozone (O_3) and hydrogen peroxide (H_2O_2). All sulphur species are subject to dry deposition at

the surface and soluble species (SO_2, H_2S, DMSO, sulphates, and MSA) are also subject to wet deposition.

The mass budget presented on Fig. 4.11 is particularly useful to understand the atmospheric dynamics of sulphur species. It is a prerequisite for an aerosol model to conserve mass and achieve a balanced mass budget. It can be verified that for each sulphur species the sink terms balance the source terms provided that fluxes are averaged on a long enough period so that a steady state has been reached (in a global mean sense). One can observe on Fig. 4.11 that DMS is completely oxidized, and that the vast majority of the DMS is eventually converted to SO_2. A fraction of the SO_2 is deposited at the surface but the majority is oxidized in sulphate aerosols. Only a small fraction of the oxidized SO_2 comes from gas-phase oxidation. In the aqueous phase, oxidation by H_2O_2 is more important than oxidation by O_3 that only takes over in situations where the H_2O_2 concentration is low. This is the case in wintertime over midlatitudes or in very polluted regions where H_2O_2 is depleted by SO_2 oxidation and is not replenished fast enough. Sulphate being a very soluble aerosol, wet deposition is the main sink term and dominates over dry deposition. Atmospheric reservoirs (given in unit of Tg S) show that sulphate and SO_2 dominate the atmospheric sulphur burden. The ratio between the average burden of a species and the sum of all the source terms (or all sink terms) is by definition the atmospheric residence time of that species. Residence times are approximately 1 day for DMS and SO_2, 4 days for H_2S, and 5 days for species in the particulate phase such as sulphate and MSA. The atmospheric lifetime of the species depends, of course, on where they are in the atmosphere.

Exercises

1. Derive Eq. 4.2 from Eq. 4.1.

2. A homogeneous spherical particle with radius r and density $\rho_p = 1500$ kg m^{-3} falls in the atmosphere. The air has a density $\rho_a = 1.3$ kg m^{-3} and a viscosity coefficient $\mu = 1.8 \times 10^{-5}$ Pa s. The Stokes' law for the friction force is given by $F_{\text{Stokes}} = 6\,\pi\,\mu\,r\,V$ where V is the particle velocity. What other forces are exerted upon the particle? Express the particle terminal fall velocity, V_p, as a function of the other variables. Compute the terminal fall velocities (in cm s^{-1}) for particles with radii 0.01, 0.1, 1, and 10 μm.

3a. An aerosol layer with mixing ratio c is scavenged below a cloud by rain. Let R be the radius of the raindrops and N_0 their concentration. The falling velocity of the raindrops is a function of their radius and follows the equation:

$$V(R) = k \left(\frac{\varrho_0}{\varrho_{\text{air}}}\right)^{\frac{1}{2}} R^{\frac{1}{2}}.$$

Estimate the volume of air swept by a raindrop per unit time and derive the scavenging rate of the aerosol. We assume that the raindrop size distribution is monodisperse (i.e. all raindrops have the same radius R and a parameter α is introduced to represent the collection efficiency of the aerosol particles by raindrops. Express the result as a function of c, R, ϱ_0, α, the air density ϱ_{air}, the water density ϱ_{water}, and the rainwater mixing ratio q_r (unit kg kg^{-1}).

3b. We assume that raindrops follow a Marshall–Palmer size distribution:

$$N(R) = N_0 \ \exp(-R/R_m).$$

Express the scavenging rate of the aerosol as a function of c, R_m, q_r, ϱ_0, ϱ_{air}, and ϱ_{water}.

It is reminded that

$$\int_0^\infty t^{z-1} \exp(-t)\, \mathrm{d}t = \Gamma(z)$$

where Γ represents the Gamma function that extends the notion of factorial to real numbers (i.e. $\Gamma(n) = n!$).

4. Consider an air mass with atmospheric concentration of SO_2 of $[SO_2] = 50$ ppbv. The temperature is 298 K. A cloud forms in this air mass with a cloud liquid water content $q_l = 1$ g m^{-3}, which leads to the partial dissolution of the SO_2 followed by a partial dissociation to HSO_3^- and SO_3^{2-}. Compute the new atmospheric concentration of SO_2 in the absence of source and sink terms for S_{IV} (= sum of SO_2, HSO_3^-, and SO_3^{2-}), and assuming equilibrium is reached between the gas and aqueous phase. It is further assumed that the new pH of cloud water after dissociation of the dissolved SO_2 is 5. Henry's constant for SO_2 at 298 K is $K_H^{SO_2} = 1.4$ mol m^{-3} Pa^{-1}, and the dissociation constants at 298 K are $K_1 = 1.3\ 10^{-2}$ mol l^{-1} for SO_2 to HSO_3^-, and $K_2 = 6.5\times10^{-8}$ mol l^{-1} for HSO_3^- to SO_3^{2-}. Take special care with units!

Solutions

2. An aerosol particle is affected by the gravity force, $P = \frac{4}{3}\,\pi\, r^3\, \rho_p\, g$, going downwards, the air buoyancy force going upwards, $A = \frac{4}{3}\,\pi\, r^3\, \rho_a\, g$, and the friction force. In the stationary regime, when the particle reaches its terminal fall velocity, there is $P = A + F_{\text{Stokes}}$, which leads to $V_p = \dfrac{2\, g\, (\rho_p - \rho_a)\, r^2}{9\, \mu}$. The terminal fall velocities can be computed as 1.8×10^{-6}, 1.8×10^{-4} 1.8×10^{-2}, and 1.8 cm s^{-1} for the four radii.

3a. $F_{\text{scavenging}} = \dfrac{3\,\alpha\,k\,\varrho_0^{\frac{1}{2}}\,\varrho_{\text{air}}^{\frac{1}{2}}\,c\,q_r}{4\,\varrho_{\text{water}}\,R^{\frac{1}{2}}}$.

3b. $F_{\text{scavenging}} = \dfrac{\Gamma(7/2)\,\alpha\,k\,\varrho_0^{\frac{1}{2}}\,\varrho_{\text{air}}^{\frac{1}{2}}\,c\,q_r}{8\,\varrho_{\text{water}}\,R_m^{\frac{1}{2}}}$.

4. The exercise is solved by calculating the concentration in H^+ cations ($h = 10^{-5}$ mol l^{-1}) and the total content in S_{IV} in the cloud water as a function of the aqueous-phase SO_2 concentration and the dissociation constants. The mass conservation equation should then be used for the total (gas and aqueous phase) S_{IV} in a common unit (for instance ppbv), which gives

$$[SO_2]' = \frac{[SO_2]}{1 + K_H^{SO_2}\left(1 + \dfrac{K_1}{h} + \dfrac{K_1\,K_2}{h^2}\right)\dfrac{q_l\,R\,T}{\varrho_{\text{water}}}} \simeq 9 \text{ ppbv}.$$

References

Alfaro SC, Gomes L (2001) Modeling mineral aerosol production by wind erosion: emission intensities and aerosol distributions in source areas. J Geophys Res 106:18075–18084

Arneth A, Miller PA, Scholze M, Hickler T, Schurges G, Smith B, Prentice IC (2007) CO_2 inhibition of global terrestrial isoprene emissions: potential implications for atmospheric chemistry. Geophys Res Lett 34:L18813. doi:10.1029/2007GL030615

Atkinson JD, Murray BJ, Woodhouse MT, Whale TF, Baustian KJ, Carslaw KS, Dobbie S, O'Sullivan D, Malkin TS (2013) The importance of feldspar for ice nucleation by mineral dust in mixed-phase clouds. Nature 498:355–358

Boucher O, Pham M, Venkataraman C (2002) Simulation of the atmospheric sulfur cycle in the Laboratoire de Météorologie Dynamique General Circulation Model. Model description, model evaluation, and global and European budgets, Note Technique de l'IPSL n°23

Gillette DA (1979) Environmental factors affecting dust emission by wind erosion. In: Morales C (ed) Saharan dust. Wiley, New York, pp 71–94

Journet E, Balkanski Y, Harrison SP (2014) A new data set of soil mineralogy for dust-cycle modeling. Atmos Chem Phys 14:3801–3816

Kerkweg A, Buchholz J, Ganzeveld L, Pozzer A, Tost H, Jöckel P (2006) Technical note: an implementation of the dry removal processes DRY DEPosition and SEDImentation in the Modular Earth Submodel System (MESSy). Atmos Chem Phys 6:4617–4632

Kettle AJ, Andreae MO (2000) Flux of dimethylsulfide from the oceans: a comparison of updated data sets and flux models. J Geophys Res 105:26973–26808

Kirkby J et al (2011) Role of sulphuric acid, ammonia and galactic cosmic rays in atmospheric aerosol nucleation. Nature 476:429–433

Kok JF (2011a) A scaling theory for the size distribution of emitted dust aerosols suggests climate models underestimate the size of the global dust cycle. Proc Natl Acad Sci U S A 108:1016–1021

Kok JF (2011b) Does the size distribution of mineral dust aerosols depend on the wind speed at emission? Atmos Chem Phys 11:10149–10156

Lana A, Bell TG, Simó R, Vallina SM, Ballabrera-Poy J, Kettle AJ, Dachs J, Bopp L, Saltzman ES, Stefels J, Johnson JE, Liss PS (2011) An updated climatology of surface dimethlysulfide concentrations and emission fluxes in the global ocean. Glob Biogeochem Cycles 25:GB1004. doi:10.1029/2010GB003850

Lewis ER, Schwartz SE (2004) Sea salt aerosol production: mechanisms, methods, measurements and models—a critical review. AGU Geophysical Monograph 152, Washington, 413 pp

Liss PS, Merlivat L (1986) In: Buat-Ménard P (ed) Air-sea gas exchange rates: introduction and synthesis. The role of air-sea exchange in geochemical cycling. D. Reidel Publishing Company, pp 113–127

Mann GW, Carslaw KS, Spracklen DV, Ridley DA, Manktelow PT, Chipperfield MP, Pickering SJ, Johnson CE (2010) Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model. Geosci Model Dev 3:519–551

Marticorena B, Bergametti G (1995) Modeling the atmospheric dust cycle: 1 design of a soil-derived dust emission scheme. J Geophys Res 100:16415–16430

Monahan EC, Fairall CW, Davidson KL, Boyle PJ (1983) Observed inter-relations between 10 m winds, ocean whitecaps and marine aerosols. Q J Royal Meteorol Soc 109:379–392

Pacifico F, Harrison SP, Jones CD, Sitch S (2009) Isoprene emissions and climate. Atmos Environ 43:6121–6135

Pruppacher HR, Klett JD (1997) Microphysics of clouds and precipitation. Kluwer, Dordrecht, 976 pp (Second revised and enlarged edition)

Stier P, Feichter J, Kinne S, Kloster S, Vignati E, Wilson J, Ganzeveld L, Tegen I, Werner M, Balkanski Y, Schulz M, Boucher O, Minikin A, Petzold A (2005) The aerosol-climate model ECHAM5-HAM. Atmos Chem Phys 5:1125–1156

Vehkamäki H, Kulmala M, Napari I, Lehtinen KEJ, Timmreck C, Noppel M, Laaksonen A (2002) An improved parametrization for sulfuric acid-water nucleation rates for tropospheric and stratospheric conditions. J Geophys Res 107:4622. doi:10.1029/2002JD002184

Wang Z, Ren S, Huang N (2014) Saltation of non-spherical sand particles. PLOS One 9(8):e105208. doi:10.1371/journal.pone.0105208

Wesely ML, Hicks BB (2000) A review of the current status of knowledge on dry deposition. Atmos Environ 34:2261–2282

Wooster MJ, Roberts G, Perry GLW, Kaufman YJ (2005) Retrieval of biomass combustion rates and totals from fire radiative power observations: FRP derivation and calibration relationships between biomass consumption and fire radiative energy release. J Geophys Res 110:D24311. doi:10.1029/2005JD006318

Further Reading (Textbooks and Articles)

de Leeuw G Andreas EL Anguelova MD Fairall CW Lewis ER O'Dowd C Schulz M Schwartz SE (2011) Production flux of sea spray aerosol. Rev Geophys 49:RG2001. doi:10.1029/2010RG000349

Seinfeld J, Pandis S (2006) Atmospheric chemistry and physics: from air pollution to climate change. Wiley, New York, 1232 pp

Sportisse B (2010) Air pollution modelling and simulation, 2nd edn. Springer, Berlin, 608 pp

Chapter 5
Interactions of Radiation with Matter and Atmospheric Radiative Transfer

Abstract This chapter provides a short yet comprehensive treatment of the interactions of electromagnetic radiation with matter as well as atmospheric radiative transfer. A brief introduction to electromagnetic radiation is followed by a discussion of molecular energy levels leading to an explanation of the molecular absorption and scattering processes. The chapter then discusses the various related physical quantities (such as particle cross section, extinction coefficient, and phase function). Physical laws relevant to the emission of radiation (Kirchhoff's and Planck's laws) are then presented. This provides all the ingredients to describe atmospheric radiative transfer in both the shortwave and longwave part of the electromagnetic spectrum. Some resolution techniques of the radiative transfer equation are presented under the assumption of single and multiple scattering. This leads naturally to discussing atmospheric windows, the atmospheric radiative budget and actinic fluxes. Finally, a short presentation of the polarization of light and its impact on scattering ends the chapter.

Keywords Radiative transfer · Scattering · Absorption · Extinction · Phase function · Optical depth · Polarization

5.1 Introduction

Electromagnetic radiation plays a key role in the regulation of the Earth's climate. We are mostly interested here in the ultraviolet, visible, and infrared domains because this is where the bulk of the energy emitted by the Sun or the Earth lies. Radiation represents the initial source of energy for many of the physical mechanisms at play in the atmosphere. We have in particular already shown in Chap. 1 how climate and climate change can be framed in terms of the radiative balance of the Earth.

Atmospheric radiation interacts with matter, both at the Earth's surface and in the atmosphere. Gases, aerosols, and cloud particles contribute to these interactions and are crucial in the Earth's radiative budget. Moreover, interactions between radiation and matter in a broad sense form the basis for remote sensing techniques for Earth's observations. The rationale for introducing atmospheric radiative processes and atmospheric radiative transfer is therefore twofold. The physical principles for aerosol remote sensing are further explained in Chap. 6.

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_5

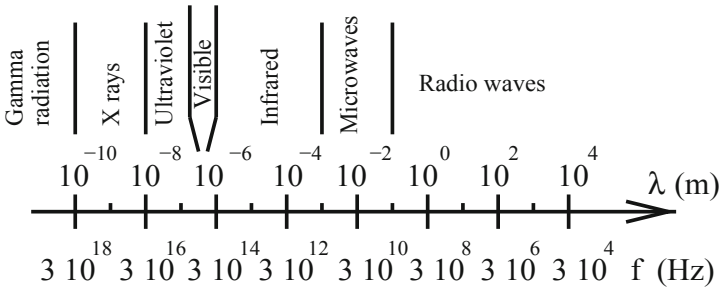


Fig. 5.1 Domains of the electromagnetic spectrum in wavelength, λ (in m) and in frequency, f (in Hz)

Some of the developments in this chapter (e.g. on molecular absorption) go beyond what is necessary to master to study and understand atmospheric aerosols. However, they form a useful addition to this chapter which is relevant to atmospheric sciences at large.

5.2 Electromagnetic Radiation

5.2.1 Generalities

Electromagnetic radiation is an important vector of energy in the atmosphere (and in the upper layer of the ocean). It is composed of oscillating waves of the electric and magnetic fields. In a neutral gaseous medium like the atmosphere, the electric and magnetic fields are perpendicular to each other and are also perpendicular to the direction of propagation of the waves, that "travel" in the vacuum at a constant speed known as the speed of light. The electromagnetic spectrum covers the domain of gamma rays, X-rays, ultraviolet, visible, infrared radiation, and microwaves and radio waves (Fig. 5.1). The human eye is only sensitive to visible radiation (from violet to red through blue, green, and yellow wavelengths). Since nature does things well, this corresponds to the emission peak of radiation from the Sun! We are mostly interested here in ultraviolet, visible, and infrared radiation because they dominate the flows of radiative energy in the atmosphere. Microwave radiation, although negligible for the atmospheric energetics, is mentioned because of its role in remote sensing of the Earth's atmosphere and surface.

Electromagnetic radiation in the ultraviolet, visible, and infrared domains is usually characterized by its wavelength, λ, expressed in nm or μm. In the field of spectroscopy, the wavenumber

$$\nu = 1/\lambda \tag{5.1}$$

is often preferred. It is usually expressed in cm^{-1} and is proportional to the energy of the wave. The following equations are convenient in that it converts wavelengths

to wavenumbers (and vice versa) in their usual units:

$$\begin{aligned} \nu\ [\mathrm{cm}^{-1}] &= 10^4/\lambda\ [\mu\mathrm{m}] \\ \lambda\ [\mu\mathrm{m}] &= 10^4/\nu\ [\mathrm{cm}^{-1}] \end{aligned} \tag{5.2}$$

so that wavelengths of 0.5, 1, 10, and 50 μm correspond to wavenumbers of 20,000, 10,000, 1000, and 200 cm^{-1}, respectively. Finally, let us note that microwave radiation is often characterized by its frequency, $f = c/\lambda = c\ \nu$, generally expressed in GHz (where c is the speed of light equal to $2.9979 \times 10^8\ \mathrm{ms}^{-1}$ in the vacuum).

By means of the wave–particle duality, electromagnetic radiation can also be approached through the notion of photons, elementary particles with zero mass that carry radiative energy. We will use both the wave and particle properties of light in the rest of chapter according to how easily one or the other concept can be used to provide a given explanation.

5.2.2 Definitions

5.2.2.1 Spherical Geometry

The position of a point P in the three-dimensional space can be specified within the spherical coordinate system by its coordinates (r, θ, φ) (refer to Fig. 5.2b). The θ and ϕ angles determine the (OP) direction and are called the polar and azimuthal angles, respectively, while r is the radial distance. As we consider the atmosphere, we refer to the polar angle as the zenith angle. It can be replaced by the coordinate $\mu = \cos\theta$ that varies between -1 and 1. Different conventions exist for θ. We prefer whenever possible to measure θ between 0 and 90° for upwelling radiation (corresponding to $\mu > 0$) and between 90 and 180° for downwelling radiation (corresponding to $\mu < 0$).

5.2.2.2 Solid Angle

A solid angle, ω, is defined by a point O and a spherical surface centred in O (Fig. 5.2a). It is by definition equal to the ratio between the surface and the square of the radius of the sphere supporting that surface, so that $\omega = S/r^2$. The solid angle is expressed in steradian (symbol sr). If the surface is the complete sphere, that is if the solid angle encompasses all directions, then it takes a value of 4π. With the help of Fig. 5.2b, it can easily be shown that the elementary solid angle corresponding to small variations $\mathrm{d}\theta$ and $\mathrm{d}\varphi$ around the zenith and azimuthal angles θ and φ is equal to:

$$\mathrm{d}\omega = \frac{\mathrm{d}S}{r^2} = \sin\theta \mathrm{d}\theta \mathrm{d}\varphi. \tag{5.3}$$

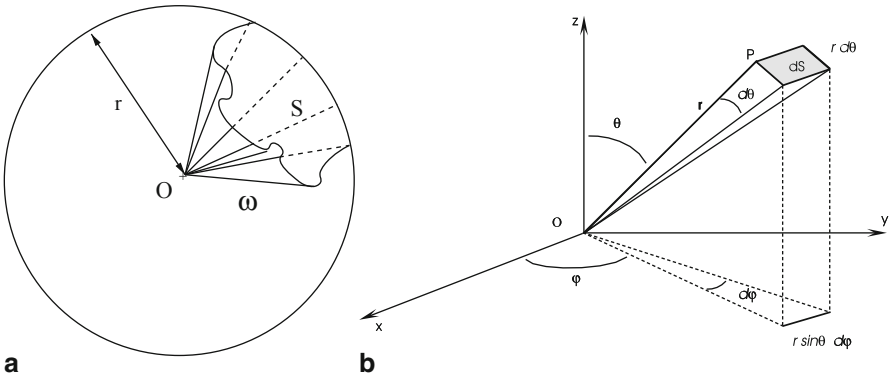


Fig. 5.2 Definition of **a** a solid angle and **b** an elementary solid angle in spherical coordinates (θ, φ)

This expression is useful as we will often consider the radiative energy contained in a light beam confined to an elementary solid angle.

5.2.2.3 Radiative Flux

Because electromagnetic radiation propagates through the atmosphere, it is quite natural to quantify it by looking at how much goes through a given surface. More specifically a radiative flux, noted F and expressed in Wm^{-2}, represents the amount of energy radiated per unit time through a unit surface.

Radiance

Electromagnetic radiation can be described in every point M of the three-dimensional space by the knowledge of the directional radiance field $L(\theta, \varphi)$. The flux of energy per unit time going through an elementary surface $\mathrm{d}S$ perpendicular to the direction of propagation and coming from an elementary solid angle $\mathrm{d}\omega$ around the point M is therefore $\mathrm{d}F = L(\theta, \varphi)\ \mathrm{d}\omega\ \mathrm{d}S$. The base quantity to describe electromagnetic radiation is in fact the *monochromatic radiance*, which is noted $L_\lambda(\theta, \varphi)$ for a given wavelength λ. The monochromatic radiance can fully describes the directionality and spectral variations of the electromagnetic radiation. It can be interpreted as the quantity of energy radiated per units time, surface, wavelength, and solid angle, expressed in $\mathrm{Wm}^{-2}\mathrm{sr}^{-1}\mu\mathrm{m}^{-1}$, and can be written as

$$L_\lambda = \frac{\mathrm{d}E_\lambda}{\mathrm{d}t\ \mathrm{d}S\ \mathrm{d}\lambda\ \mathrm{d}\omega} = \frac{\mathrm{d}E_\lambda}{\mathrm{d}t \cos i\ \mathrm{d}A\ \mathrm{d}\lambda\ \mathrm{d}\omega}, \tag{5.4}$$

where $\mathrm{d}E_\lambda$ is the radiant energy with wavelength between λ and $\lambda + \mathrm{d}\lambda$ coming from directions limited to an element of solid angle $\mathrm{d}\omega$ that goes through a surface

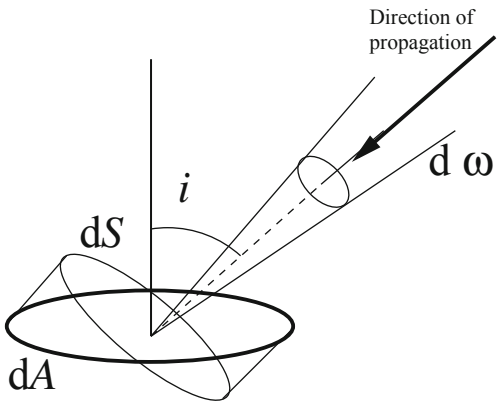


Fig. 5.3 A radiance is the quantity of radiant energy per unit time coming from a solid angle element and going through a surface element dS perpendicular to the direction of propagation. See text for further explanation

element dS perpendicular to the direction of propagation during a time interval dt (Fig. 5.3). Equivalently, dE_λ is the energy radiated through a surface element dA whose normal makes an angle i with the direction associated with the solid angle (i.e. $\mathrm{d}S = \cos i \; \mathrm{d}A$). The monochromatic radiance can also be expressed per unit wavenumber (in unit $\mathrm{Wm^{-2}sr^{-1}(cm^{-1})^{-1}}$ or equivalently $\mathrm{Wm^{-2}sr^{-1}cm}$). In this case, and with both dν and dλ being positive by definition, the two quantities are such that:

$$L_\nu \mathrm{d}\nu = L_\lambda \mathrm{d}\lambda$$

so that

$$L_\nu = \lambda^2 L_\lambda = \nu^{-2} L_\lambda.$$

5.3 Interactions of Radiation with Matter

Emission of radiation originates from desexcitation of excited atoms of molecules. Thermal collisions is the most prominent excitation mechanism, but various other mechanisms, such as chemical reactions, absorption of radiation, and electron precipitation, are also source of excitation, and hence desexcitation. Light that reaches our eye does not generally come directly from a source (such as the Sun or a lamp) but indirectly after it was scattered in different directions. Scattering is a fundamental physical process that takes place at all wavelengths but in a very variable manner. The sky appears blue because solar radiation at short wavelengths corresponding to the blue colour is scattered more efficiently than radiation from other wavelengths composing visible light[1]. In the atmosphere, solar radiation is scattered by molecules, aerosols and cloud particles. Single scattering is when radiation is (assumed to be)

[1] The sky does not appear violet because there is somewhat less violet radiation in the lower atmosphere but also because of the way colour receptors in the eye respond to different wavelengths.

scattered at most once in the atmosphere, while multiple scattering is when radiation can be scattered more than once.

Scattering is often accompanied with the physical process of absorption. Radiative energy that is absorbed is transformed into some other energy form, and is therefore no longer available as radiant energy at its original wavelength. The grass is green because solar radiation of wavelengths corresponding to the red and blue colour has been preferentially absorbed while radiation in the green wavelengths has been preferentially scattered.

5.3.1 Matter, Energy and Spectral Lines

5.3.1.1 Energy Levels and Transitions

Leaving aside the nuclear binding energy, the internal energy of atoms and molecules is comprised of kinetic energy (also called translation energy) due to their thermal motion (E_c), electronic energy (E_e), and the rotational and vibrational energies (E_r and E_v). For an isolated molecule, the internal energy is simply the sum of these components:

$$E = E_c + E_e + E_r + E_v. \tag{5.5}$$

The last three forms of energy are *quantized*, that is to say the energy can only take certain discrete values (energy levels) that are dependent on the characteristics of the molecules. The emission or the absorption of radiation corresponds to a transition between two different levels, with the wavelength of the radiation corresponding to the difference in the two energy levels. Each atom and molecule, therefore, has a set number of possible transitions that determine a spectrum of absorption lines that characterizes it.

Some important properties of atomic and molecular structure may be exemplified using a picture of an atom that is called the Bohr model. This model is clearly incomplete and insufficient to account for all observations but it is fairly simple and useful to understand the main features of electronic transitions and can be easily extrapolated to account for molecular spectra as we describe them in this chapter.

Absorption and emission arise from the interaction between matter and the electromagnetic field. Neutral molecules are electrically balanced, but the motions of atoms and electrons in their orbits create a dipolar or quadrupolar structure. The electric dipole or quadrupole couples the two energy levels and enables an energy transition which is associated with the absorption or emission of a photon.

If $E(J'')$ represents the energy of the upper level of the transition and $E(J')$ that of the lower level, the central frequency, f_o, and the associated wavenumber, ν_o, of the transition are given by

$$h\ f_o = h\ c\ \nu_o = E(J'') - E(J'), \tag{5.6}$$

where h is the Planck constant ($h = 6.62607 \times 10^{-34}$ Js). In spectroscopy, the h factor is often omitted so that the wavenumber of a photon can be identified as the difference in energy between two levels.

Electronic Energy Levels

Electronic transitions require relatively large energies that correspond to the wavelengths of ultraviolet and visible radiation. In practice, during an electronic transition, there are also vibrational and rotational transitions (see below). The wavelength corresponding to an electronic transition is, therefore, followed by a succession of lines whose associated wavelengths are smaller than that of the electronic transition. If the energy brought by a photon is large enough to dissociate the molecule, the remaining energy can be converted into thermal energy, or the resulting radicals can be in excited states.

Rotational Energy Levels

The rotational energy of a rigid linear molecule can be written under the form

$$E_{\text{rot}} = h\,c\,F(J) = h\,c\,B\,J\,(J+1), \tag{5.7}$$

where B is a constant that is inversely proportional to the moment of inertia of the molecule and J is a quantum number that only takes integer values ($J = 0, 1, 2, ...$). For the sake of simplification, the quantity $B\,J\,(J+1)$ is often assimilated to the rotational energy level itself, which is then expressed in unit cm^{-1}.

The rotational energy of a nonlinear molecule with an axis of symmetry can be written in the form:

$$E_{\text{rot}} = h\,c\,(B\,J\,(J+1) + (A-B)\,K^2), \tag{5.8}$$

where K is a second quantum number ($K \leq J$). Molecules with more complex structures are obviously more difficult to treat. For molecules without an axis of symmetry, it is not possible to provide a general analytical formula that describes the rotational energy levels. There is always however a rotational quantum number J involved that takes the values 0, 1, 2, ... associated with a characteristic energy B.

Vibrational Energy Levels

The number of vibrational modes depend on the type of molecule: 1 for a linear diatomic molecule, 3 for a nonlinear triatomic molecule (such as H_2O), 4 for a linear triatomic molecule, but only 3 including a double one if the molecule presents a symmetry (such as CO_2). Vibrational modes for triatomic molecules are illustrated

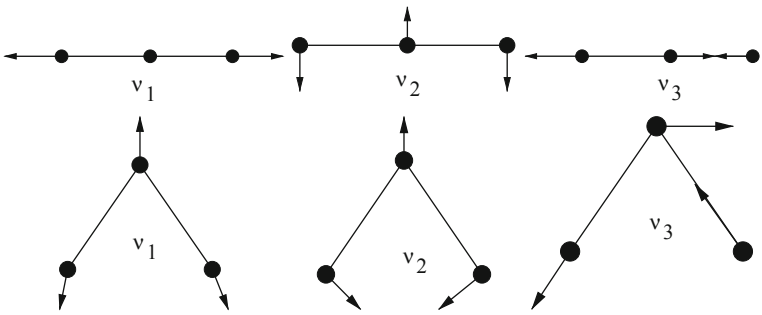


Fig. 5.4 Modes of vibration of linear and nonlinear triatomic molecules

in Fig. 5.4 and are noted ν_i [2]; the associated quantum number is noted v_i. Under the assumption of a harmonic oscillator, the vibrational energy can be written as

$$E_{\text{vib}} = hc \sum_i \left(v_i + \frac{1}{2} \right) \nu_i, \tag{5.9}$$

where v_i and ν_i are the quantum number and the wavenumber of vibrational mode i.

5.3.1.2 Selection Rules

We neglect in the following the kinetic and electronic energies of the molecules, but focus on the vibrational and rotational energies, which are important in the infrared part of the electromagnetic spectrum. As indicated earlier, absorption lines of a molecule result from transitions between two energy levels. However, at first order, only transitions that modify the electric dipole momentum of the molecule actually give rise to the emission or absorption of a photon. The transitions that are permitted follow so-called selection rules, while other transitions are forbidden. Generally speaking, permitted transitions have to obey the following selection rules:

$$\Delta v_i = \pm 1 \text{forasingle } i, \ \Delta J = 0 \text{or} \pm 1 \text{ and } \Delta K = 0. \tag{5.10}$$

Vibrational–rotational transitions are explicited in Fig. 5.5. The set of transitions for which the lower and upper vibrational energy levels have quantum numbers $v_i = 0$ and $v_i = 1$ are part of the ν_i band (called fundamental). Within a band, to a rotational quantum number J for the lower energy level can correspond a quantum number $J + 1$, $J - 1$ or J for the upper level. In the first case, the absorption lines have wavenumbers $\nu_i + 2B$, $\nu_i + 4B$, ... and form the R branch. In the second case, the absorption lines have wavenumbers $\nu_i - 2B$, $\nu_i - 4B$, ... and form the P branch.

[2] The same notation is used traditionally to indicate the vibration mode and its associated wavenumber.

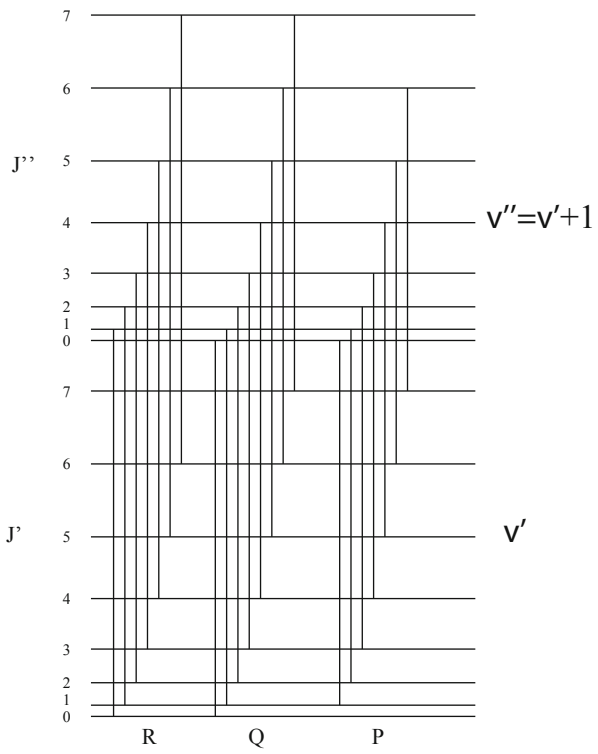


Fig. 5.5 Vibration–rotation transitions (branches *P*, *Q*, and *R*)

Finally in the third case, all the transitions should have the same wavenumber ν_i. This is not the case, however, because of a small dependence of the vibrational energy with the rotational quantum number J. This is clearly visible in Fig. 5.6 that represents the ν_5 band of the $^{12}C^{13}CH_2$ molecule. The vibrational transitions corresponding to no change in the rotational quantum number form the Q branch.

The set of transitions from the lower vibrational energy level $v_i = 1$ to the upper energy level $v_i = 2$ is called $2\nu_i - \nu_i$. A small anharmonicity in the molecular vibration leads to a deviation from the vibrational energy given by Eq. 5.9: the wavenumber of the $2\nu_i - \nu_i$ band is slightly less than that of the ν_i band. For example, in the case of the CH_4 molecule, the wavenumber of the ν_3 band is 3019 cm^{-1}, while that of the $2\nu_3 - \nu_3$ band is 2933 cm^{-1}. Finally, if the transition takes place between a lower vibrational level $v_i = 0, v_j = 1$ and an upper level $v_i = 1, v_j = 1$, the band is named $\nu_i + \nu_j - \nu_j$.

There also exist transitions between the lower and upper rotational levels J and $J + 1$, in which case all the vibrational quantum numbers remain equal to zero ($v_i = 0,\ i = 1, 2, ...$). These correspond to purely rotational absorption lines of the molecule.

The above-mentioned selection rules suffer a certain number of exceptions, that are generally related to symmetries of the molecule. Nonharmonicity can make it

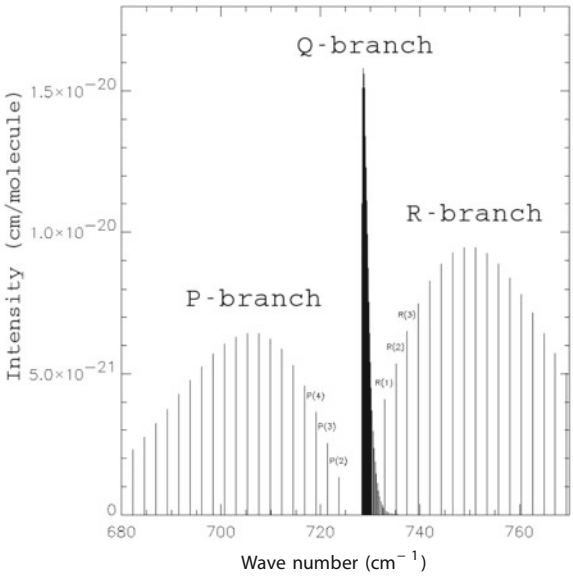


Fig. 5.6 Intensity of the absorption lines of the ν_5 band of the $^{12}C^{13}CH_2$ molecule at 296 K. (Adapted from Fouchet (2000))

possible to jump a band ($\Delta v_i > 1$) or to combine two transitions ($\Delta v_i = 1$ for two different v_i). It is also important to account for the presence of isotopologues, among which the absorption lines vary slightly.

5.3.2 *Intensity of Spectral Lines*

The intensity of a spectral absorption line is proportional to the intensity of the transition and to the number of molecules that are in the initial energy level E'. The rest of this discussion assumes local thermodynamic equilibrium (LTE) conditions. LTE is verified when atmospheric pressure is large enough so that collisions between molecules redistribute the energy quickly among the different types of molecules. This is the case in the Earth's atmosphere up to altitudes of 50–60 km depending on the molecules. The problem of estimating the energy levels of the molecules gets much more complex in non-LTE conditions (see López-Puertas and Taylor 2001). How different energy levels of a set of molecules are populated depend on the temperature, therefore the intensity of the absorption lines also depends on temperature. According to the Maxwell–Boltzmann distribution law, which expresses how energy levels are distributed in a gas at thermal equilibrium, the number of molecules N_n at the energy level E_n and degeneracy factor g_n is proportional to $g_n\ e^{-E_n/k_BT}$ where k_B is the Boltzmann constant and T is the temperature, so that:

$$N_n = \frac{g_n\ e^{-E_n/k_BT}}{\sum_i\ g_i\ e^{-E_i/k_BT}}\ . \tag{5.11}$$

The denominator $Q = \sum_i g_i \, e^{-E_i/k_B T}$ is known as the partition function. For atmospheric temperatures, only the terms relating to the rotational and vibrational energy levels are accounted for, while high electronic energy levels are neglected. The rotational and vibrational energy levels are considered to be independent to each other at first approximation. The partition function can then be written as the product of the vibrational partition function Q_v and the rotational partition function Q_r.

The vibrational partition function can be written as

$$\begin{aligned} Q_v &= \sum_{v_1 v_2 \ldots} e^{-(v_1 v_1 + v_2 v_2 + \cdots) hc/k_B T} \\ &= \sum_{v_1} e^{-v_1 v_1 hc/k_B T} \times \sum_{v_2} e^{-v_2 v_2 hc/k_B T} \times \cdots . \end{aligned} \tag{5.12}$$

By summing over all v_i and accounting for the degeneracy degree d_i of the vibration v_i, the equation becomes:

$$Q_v = (1 - e^{-v_1 hc/k_B T})^{-d_1} \, (1 - e^{-v_2 hc/k_B T})^{-d_2} \cdots . \tag{5.13}$$

It is impossible to provide a general equation for the rotational partition function unless the temperature is sufficiently large so that the inequality $k_B \, T \gg h \, c \, B$ is verified. In that case, the partition function Q_r is proportional to T for linear molecules and to $T^{3/2}$ for other molecules. In all other cases, the rotational partition function has to be computed numerically, which is outside the scope of this chapter. Molecular spectroscopic databases provide polynomial functions of the temperature that approximate the partition function.

As indicated above, the intensity of an absorption line of wavenumber v_o depends on the temperature, as it is proportional to the fraction of molecules that are in lower energy level $E(J')$ of the transition, and is proportional to $e^{-h \, c \, E/k_B \, T}/Q(T)$. The intensity of the absorption line $S(T)$ needs to corrected accordingly:

$$S_1(T) = S(T_o) \frac{Q_v(T_o)}{Q_v(T)} \left(\frac{T_o}{T} \right)^{\alpha} \exp \left[-\frac{h \, c \, E}{k_B} \left(\frac{1}{T} - \frac{1}{T_o} \right) \right], \tag{5.14}$$

where α is an exponent that depends on the linearity of the molecule and T_o is a reference temperature that depends on the spectroscopic database that is considered. A second correction needs to be applied in order to account for the induced emission. The intensity of the absorption line at temperature T then becomes:

$$S(T) = S_1(T) \, \frac{1 - e^{-\frac{h \, c \, v_o}{k_B \, T}}}{1 - e^{-\frac{h \, c \, v_o}{k_B \, T_o}}} \cdot \tag{5.15}$$

5.3.3 Spectral Line Profiles

According to Eq. 5.6, a spectral line is a Dirac function at wavenumber v_o. Several processes cause a broadening of the line.

Natural Linewidth

According to Heisenberg's uncertainty principle, if an excited state has a limited lifetime of Δt, the energy E of that state presents an uncertainty $\Delta E = h/2\pi\ \Delta t$, and the wavenumber corresponding to the transition has an uncertainty $\Delta \nu_o = 1/(2\pi c)\ \Delta t$. The associated broadening of the line (all wavenumbers in the interval $\nu_o - \Delta\nu_o$ and $\nu_o + \Delta\nu_o$ correspond to the same transition) is actually very small, of the order of 10^{-12} cm^{-1} for the CO_2 band at 15 μm for instance.

Doppler Broadening

Second, if an emitting or absorbing molecule has a speed with a component u along the direction of propagation of the radiation, there is a change in the wavenumber implied by the Doppler effect: $\Delta\nu = \nu_o\, u/c$. In this case, the shape of the absorption line follows the function:

$$g_D(\nu - \nu_o) = \frac{1}{\alpha_D\sqrt{\pi}}\ \exp\left(-\frac{(\nu - \nu_o)^2}{\alpha_D^2}\right), \tag{5.16}$$

where

$$\alpha_D = \frac{\nu_o}{c}\sqrt{\frac{2\,k_B\,T}{m}}, \tag{5.17}$$

m being the mass of the molecule. The half-width of the line at half of the maximum height is $\alpha_D\sqrt{\ln(2)}$. For atomic oxygen at 557.7 nm, $\alpha_D \simeq 3.3 \times 10^{-2}$ cm^{-1} at 300 K. For the rotational transition of H_2O around 200 cm^{-1}, $\alpha_D \simeq 3.5 \times 10^{-4}$ cm^{-1}.

Pressure Broadening

Third, the wavenumber associated with a transition depends on the initial and final energy levels following $h\ c\ \nu_o = E(J'') - E(J')$. When two molecules collide with each other—or if their electric dipoles interact with each other—the energy levels are slightly modified and it follows that ν_o varies. This cause of broadening of the absorption line is the dominant one in the troposphere.

For pressure broadening, the absorption line follows the Lorentz shape:

$$g_L(\nu - \nu_o) = \frac{\alpha_L}{\pi[(\nu - \nu_o)^2 + \alpha_L^2]}, \tag{5.18}$$

where α_L is the half-width at half of the maximum height. The α_L parameter is proportional to the number of collisions per unit time and varies according to the ambient pressure and temperature, and can be approximated as

$$\alpha_L = \alpha_{L,0}\,\frac{p}{p_o}\,\left(\frac{T_o}{T}\right)^n, \tag{5.19}$$

where n is an index that varies between 0.3 and 1 according to the molecule.

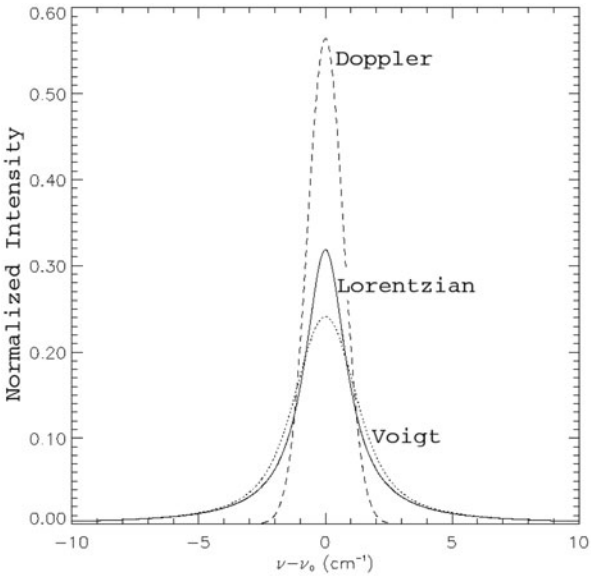


Fig. 5.7 Comparison of the Lorentzian (*solid line*), Doppler (*dashed line*), and Voigt (*dotted line*) profiles for $\alpha_L = \alpha_D = 1\ \mathrm{cm}^{-1}$. (Adapted from Fouchet (2000))

Voigt Profile

When the pressure and Doppler broadenings are of similar magnitude, the two processes not being correlated, the shape of the absorption line can be estimated by a convolution of the two shapes, which leads to:

$$g_V(\nu - \nu_o) = \int_0^\infty g_D(\nu - \nu')\, g_L(\nu' - \nu_o)\, \mathrm{d}\nu'. \tag{5.20}$$

The pressure broadening is dominant in the troposphere while the two processes take similar values at an altitude of around 33 km for the CO_2 molecule, 31 km for H_2O and 30 km for O_3.

A Voigt profile consists of a central Doppler part and Lorentz wings (Fig. 5.7). If the line is saturated, all radiation with a wavenumber close to that of the centre of the line is absorbed, and the shape of the profile close to the centre does not matter so much. Finally, it should be noted that far away from the centre of the line, the Lorentz profile generally overestimates absorption. These sub-Lorentzian far wings are generally modelled through a composite profile, that is Lorentzian up to a distance $\nu - \nu_o$ to the centre of the line ranging from 20 to 30 cm^{-1}, and a more accurate description of the wings further away.

The characteristics of the intensity and shape of the lines for different gases are available in spectroscopic databases that are regularly updated.

5.3.4 Processes of Interactions of Radiation with Matter

5.3.4.1 Gaseous Medium

For a gaseous medium, the interactions between matter and radiation can be described at the molecular scale. When a photon interacts with a molecule, two processes can occur: scattering and absorption.

1. *Scattering* usually occurs when the interaction takes place without any of the various forms of molecular energy being modified. This is the case if the energy levels of the molecule are quantified and if the energy of the incident photon is relatively far away from all the differences in the (quantified) energy levels of the molecule. The wavelength of the radiation, therefore, remains unchanged.
2. *Absorption* occurs when molecular collisions take place before re-emission takes place, that is during the lifetime of the excited molecular state. The radiative energy is, therefore, converted to kinetic (or thermal) energy. The process of *emission* is the opposite one and converts kinetic energy into radiative energy.

Molecular motion being random, the absorption and emission processes are isotropic. However, the scattering process is anisotropic and the interaction can be more or less efficient.

In practice, the interaction of a photon with a molecule is also considered as scattering when the absorption of a quantum induces a transition towards an excited state of the molecule that is very short lived. In this case, the absorbed quantum is re-emitted immediately at the same frequency and the conversion to translational energy is negligible. This is referred to as coherent scattering. If the return to the initial state of the molecule occurs through a series of transitions, then the process is referred to as incoherent scattering.

What we have described so far is known as *elastic scattering*, whereby the energies of the molecule and incident photon, and thus the wavelength of the incident photon, are conserved. Only the direction of the incident photon is changed. Occasionally, a scattered photon is accompanied by a transition in the energy level of the molecule it interacts with. This process is called *Raman scattering* or *Raman effect*. In this case, the molecule absorbs or loses energy and the photon will be less (Stokes Raman scattering) or more (anti-Stokes Raman scattering) energetic. As the upper energy level is less populated than the lower energy level, a Raman-scattered photon generally has a lower energy, and thus a longer wavelength, than the incident photon.

5.3.4.2 Dense Medium

For a dense medium (solid or liquid matter), it is not possible to describe the processes at the molecular scale, because of the many interactions between the different molecules. This is the case for aerosols in suspension in the air, and droplets and ice crystals in clouds. The medium has to be described at the macroscopic scale considering a distribution of particles, and taking into account their shape (spherical

or not), their size distribution, and their refractive index. These parameters allow us to model the scattering and absorbing properties of the medium.

Scattering by particles that are much smaller than the wavelength can be understood from the Rayleigh theory, while scattering by homogeneous spherical particles can be understood from the Mie theory, sometimes also referred to as the Lorenz–Mie theory. Other methods are required to characterize scattering by nonspherical and/or inhomogeneous particles whose size is of a similar size (or larger) than the wavelength considered. Modelling of the absorption and scattering properties of particles was briefly introduced in Chap. 3, it is described in more detail later in this chapter and also in Appendix C.

5.4 Modelling of the Interaction Processes

Describing the flows of radiative energy requires to model the above-mentioned processes of radiation–matter interaction. A number of physical quantities need to be introduced.

5.4.1 Molecular Absorption Coefficient

Molecules represent an important cause of absorption of electromagnetic radiation in the atmosphere[3]. The absorption lines, that correspond to transitions between energy levels, and the broadening effects are such that the molecular absorption can present very rapid spectral variations. In the theoretical case of an isolated absorption line, the absorption cross section (noted $k^{\text{abs}}(\nu)$ in unit cm^2 molecule^{-1}) can be written from quantities explicited above:

$$k^{\text{abs}}(\nu) = S\ g(\nu - \nu_o), \tag{5.21}$$

where S represents the intensity of the absorption line (in cm molecule^{-1}), ν_o is the wavenumber associated with the transition ($h\,c\,\nu_o = E(J'') - E(J')$) and $g(\nu - \nu_o)$ represents the normalized shape of the line such that

$$\int_{-\infty}^{+\infty} g(\nu - \nu_o)\mathrm{d}\nu = 1.$$

The *molecular absorption coefficient* expresses the absorption cross-section per unit volume (or mass):

$$\sigma^{\text{abs}}(\nu) = k^{\text{abs}}(\nu)\ U^{\text{abs}}, \tag{5.22}$$

[3] Atoms also contributes to absorption in the upper atmosphere but are omitted from the discussion here although the principles are the same.

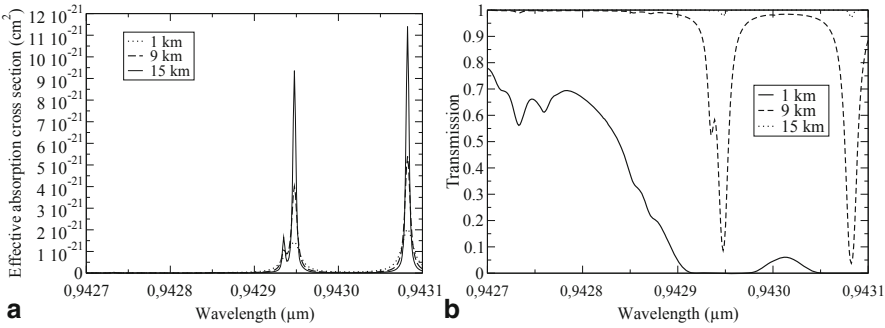


Fig. 5.8 Absorption cross section (cm^2 $molecule^{-1}$) and atmospheric transmittance through a 1-km-thick layer of water vapour at different altitudes. The data are plotted for a high spectral resolution

where U^{abs} is the gas concentration (in molecules cm^{-3} or molecules kg^{-1}). The typical unit for σ^{abs} is, therefore, cm^2 cm^{-3} (or in the case of molecular absorption coefficient per unit mass cm^2 kg^{-1}). This coefficient (expressed per unit volume) can be used to display the linear variation of absorption along the optical path (known as Bouguer's law). When multiplied by a distance, the molecular absorption coefficient becomes an absorption optical depth.

Because of the temperature and pressure dependence of the intensity and shape of the absorption line, $k^{\text{abs}}(\nu)$ varies as a function of altitude. Figure 5.8a shows the absorption cross section of the water vapour molecule in three different layers of the atmosphere: close to the surface and at altitudes of 9 and 15 km. It can be seen that the lines are more intense and sharper at altitude. Figure 5.8b will be discussed later.

5.4.2 *Scattering Phase Function*

The directional characteristics of the scattering process is described by means of the *scattering phase function* that gives the probabilities for an incident photon that is scattered to be scattered in a given direction. The incident and scattered directions, $s_{\mathbf{i}}$ and $s_{\mathbf{d}}$, are referenced by their zenith and azimuthal angles θ_i, φ_i and θ_s, φ_s. The phase function is usually normalized to 4π:

$$\int_{\omega} \mathrm{p}(\omega)\mathrm{d}\omega = \int_0^{2\pi} \int_0^{\pi} \mathrm{p}(\theta_s, \varphi_s) \sin\theta_s \mathrm{d}\theta_s \mathrm{d}\varphi_s = 4\pi, \tag{5.23}$$

with the phase function depending here only on the scattering direction relative to the incident direction. When the scattering is due to spherical particles, which is the case for hydrated particles and cloud droplets, or when the particles are randomly oriented in space, the phase function only depends on the scattering angle: $\Theta = \widehat{s_{\mathbf{i}}, s_{\mathbf{s}}}$, and is noted $P(\Theta) = \int_0^{2\pi} \mathrm{p}(\theta, \varphi)\, \mathrm{d}\varphi$, so that $\int_0^{\pi} P(\Theta) \sin\Theta \,\mathrm{d}\Theta = \int_{-1}^{1} P(\mu)\,\mathrm{d}\mu = 2$ where $\mu = \cos(\Theta)$. The definition of the scattering angle in the general case is

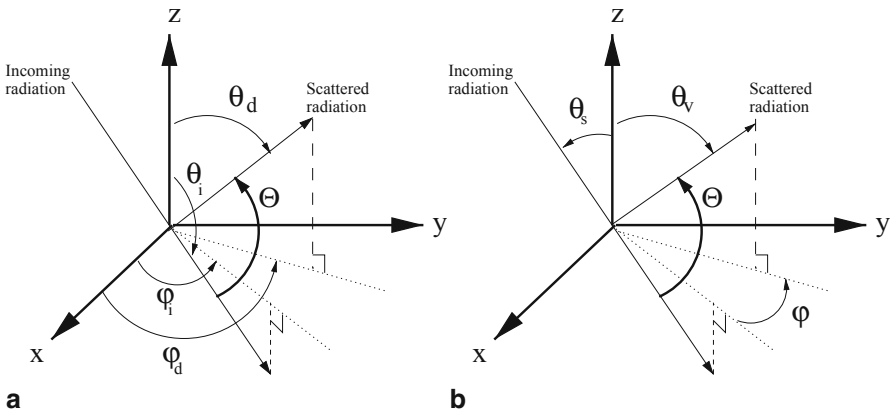


Fig. 5.9 Geometry of the scattering angle, Θ. **a** If the cosine of the zenith angle is counted negatively and positively for downward and upward radiation, respectively, the scattering angle can be calculated from the solar azimuthal and zenith angles of the (incident) downward and (scattered) upward radiation using the equation $\cos\Theta = \cos\theta_i \cos\theta_d + \sin\theta_i \sin\theta_d \cos(\varphi_d - \varphi_i)$ and notations from the left panel. **b** In remote sensing, both cosines of the zenith angle are sometimes counted positively. In this case, with the notation from the right panel, the equation is $\cos\Theta = -\cos\theta_s \cos\theta_v - \sin\theta_s \sin\theta_v \cos\varphi$

provided in Fig. 5.9. One can define a number of of parameters that inform us on the shape and properties of the phase function: backscatter fraction, upscatter fraction, Legendre moments, etc. The most usual parameter being considered is the asymmetry parameter (or asymmetry factor), that is given by the integral $g = \int_{-1}^{1} P(\mu)\,\mu\,\mathrm{d}\mu$ and is a measure of the degree of forward–backward symmetry of the phase function.

5.4.3 Molecular Scattering

Molecular scattering arises from the interaction between electromagnetic radiation and the electrical dipole of the molecules, according to a theory that was developed by Rayleigh in 1871. The phase function for the scattering by molecules is almost independent of the wavelength. It can be written as

$$P(\Theta) = \frac{3}{4}(1 + \cos^2\Theta). \tag{5.24}$$

The phase function being symmetrical between the forward and backward directions, its asymmetry parameter is 0, which is characteristics of very small particles in comparison to the wavelength.

The scattering cross section (unit m^2) of an air molecule is

$$s_\lambda^{\mathrm{sca}}(\lambda) = \frac{24\,\pi^3}{N^2\,\lambda^4}\frac{(m^2-1)^2}{(m^2+1)^2}, \tag{5.25}$$

where m is the index of refraction of the air and N is the density of molecules (in m^{-3}). Typically the scattering optical depth for the whole atmosphere can be approximated by

$$\tau^{\mathrm{sca}}(\lambda) \simeq 0.01\lambda^{-4}, \tag{5.26}$$

where λ is given in μm and the surface pressure taken to be $p_o = 1013$ hPa. For a surface pressure p then

$$\tau^{\mathrm{sca}}(\lambda, p) = \frac{p}{p_o}\tau^{\mathrm{sca}}(\lambda, p_o). \tag{5.27}$$

In fact, the dependence in λ is slightly stronger due to a small variation of the air refractive index with the wavelength.

5.4.4 Absorption and Scattering by Aerosols

For a single aerosol particle, it is usual to introduce the following quantities:

- The scattering and absorption factors, Q^{sca} and Q^{abs}, are the ratios of the scattering and absorption cross sections, s^{sca} and s^{abs}, to the geometric cross section s_g of the particle:

$$Q^{\mathrm{sca}} = \frac{s^{\mathrm{sca}}}{s_g} \quad \text{and} \quad Q^{\mathrm{abs}} = \frac{s^{\mathrm{abs}}}{s_g} \cdot \tag{5.28}$$

  The extinction being the sum of scattering and absorption, the extinction factor is $Q^{\mathrm{ext}} = Q^{\mathrm{abs}} + Q^{\mathrm{sca}}$.
- The single scattering albedo, ϖ_o, is the ratio between scattering and extinction, $\varpi_o = Q^{\mathrm{sca}}/Q^{\mathrm{ext}}$.
- The asymmetry parameter, already defined in Chap. 3, and the backscatter and upscatter fractions also describe the shape of the phase function.

The Mie theory was formulated by Gustav Mie in 1908 and bridges the gap between the Rayleigh theory, that applies to particles with a size much smaller than the wavelength, and geometrical optics, that applies to particles much larger than the wavelength. The Mie theory is based on wave optics to predict rigorously the interaction of light with spheres of arbitrary sizes. Schematically, the Mie theory relies on a decomposition of the electromagnetic waves in spherical harmonics inside and outside the particle, and the boundary conditions at the surface of the particle provide an equation for the scattered wave. Generally one is only interested in the solution in the far field, that is well away from the particle. According to the Mie theory, the scattering and absorption properties of a homogeneous spherical particle depend only on its complex index of refraction $m = n_r - in_i$, which in general depends on the wavelength, and on its size parameter (or Mie parameter) defined as

$$x = \frac{2\,\pi\,r}{\lambda}, \tag{5.29}$$

where r is the radius of the particle and λ is the wavelength under consideration. The extinction, scattering and absorption factors can be obtained by computing series that involve some sums and products of a_n and b_n coefficients that can be calculated recursively (van de Hulst 1982; Bohren and Huffman 1998). The computation of the phase function also involves other series of coefficients π_n and τ_n that depend on the scattering angle.

For a population of particles with size distribution $n(r)$ ($n(r)\,\mathrm{d}r$ being the number of particles per unit volume with radius in the range r to $r + \mathrm{d}r$), the scattering and absorption coefficients (in units of m^{-1}) and the average phase function $P(\mu)$ are computed in the following way:

$$\begin{cases} \sigma^{\mathrm{sca}} = \int_0^\infty \pi r^2 \; Q^{\mathrm{sca}}(r) n(r) \mathrm{d}r \\ \\ \sigma^{\mathrm{abs}} = \int_0^\infty \pi r^2 \; Q^{\mathrm{abs}}(r) n(r) \mathrm{d}r \\ \\ P(\mu) = \frac{1}{\sigma^{\mathrm{sca}}} \int_0^\infty \pi r^2 \; Q^{\mathrm{sca}}(r) P(r, \mu) \; n(r) \; \mathrm{d}r. \end{cases} \tag{5.30}$$

It is important to note that any average on the phase function, the asymmetry parameter, the backscatter and upscatter fractions has to be weighted by $Q^{\mathrm{sca}}(r)$ when integrated over the size distribution.

Figure 3.9 in Chap. 3 represents the extinction factor for a spherical particle as a function of the size parameter. It can be observed that $Q^{\mathrm{ext}}(x)$ tends towards 2, the value predicted by geometrical optics, when x tends towards infinity. The effect of absorption is to decrease the maximum reached by Q^{ext} and to smooth out the dependence of Q^{ext} on x.

It is then possible to define the mass (or volume) scattering and absorption efficiencies as the scattering and absorption coefficients per unit mass (or volume) of aerosol particles, noted α^{sca} and α^{abs}, with unit $\mathrm{m}^2\ \mathrm{g}^{-1}$ ($\mathrm{m}^2\ \mathrm{cm}^{-3}$). These quantities are sometimes referred to as mass (or volume) scattering and absorption cross sections. The volume scattering efficiency exhibits a peak and decreases rapidly around that peak. The value of the maximum depends on the real and imaginary parts of the refractive index (Fig. 3.10 in Chap. 3). It should be noted here that the peak in α^{sca} occurs precisely for Mie parameters corresponding to the aerosol accumulation mode for wavelengths of visible light. This is a coincidence that bears important consequences for aerosol–radiation interactions. The volume absorption efficiency also exhibits a maximum, but in contrast to what is observed for scattering, it tends to a nonzero value that is not too different from the maximum when x tends towards 0.

The phase function computed from the Mie theory for aerosol particles of varying size is reported in Fig. 5.10. It can be noted that the phase function becomes more and more asymmetrical as the particle size increases, with a forward peak that is more and more pronounced. The presence of a small peak in the backscatter direction ($\Theta = 180°$ or $\mu = -1$) should also be noted.

For cloud droplets, geometric optics can be used to approximate scattering and absorption at wavelengths that are small compared to the cloud droplet size. This

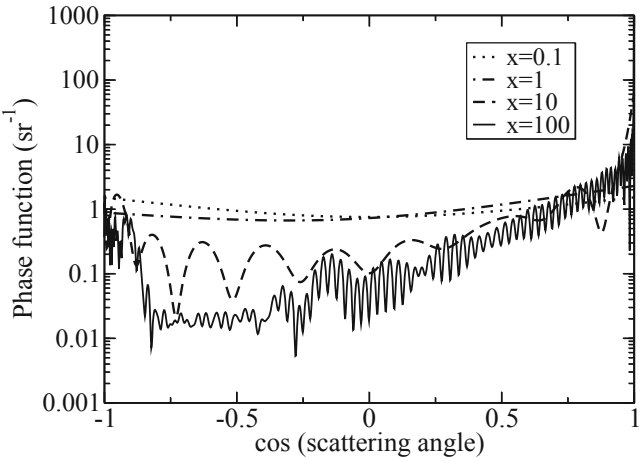


Fig. 5.10 Phase function, $P(\mu)$, for homogeneous spheres with increasing Mie size parameter ($x = 0.1$, 1, 10 and 100) and a refractive index $m = 1.45 - 0.0i$

circumvents the use of the Mie theory that becomes expensive at large particle size. For nonspherical particles and for wavelengths that are small compared to the size of the particles, such as ice crystal in the visible spectrum, it is possible to use the ray tracing method. In this case, this requires to take into account diffraction caused by the ice crystals and interferences between rays that are scattered in similar directions. For nonspherical aerosols, for which the wavelength is of the same order of magnitude as the particle size, more complicated methods, such as the T-matrix method, are needed (see Mishchenko et al. (1996) and further reading at the end of this chapter).

5.4.5 Thermal Emission

5.4.5.1 Black Body and Kirchhoff's Law

Let us consider a cavity that is impermeable to radiation and maintained at a constant temperature. The cavity is either empty or filled with a homogeneous isotropic medium; it contains a body whose state only depends on temperature. The system gets to an equilibrium upon which the cavity and the body it contains have the same temperature. The radiative field in the cavity only depends on the temperature, it is isotropic and does not depend on the shape of the inside walls of the cavity. We characterize this radiation by its spectral radiance B_λ.

Let us now consider a surface element dS of the body in the cavity. This surface element receives during a time period dt and from a beam dω around a direction that has an angle θ with the normal to the surface element dS an energy equal to

$\mathrm{d}W_\lambda^i = B_\lambda \, \mathrm{d}\lambda \, \mathrm{d}S \cos\theta \, \mathrm{d}\omega \, \mathrm{d}t$. It absorbs a fraction $\mathrm{d}W_\lambda^a = \alpha_\lambda \, \mathrm{d}W_\lambda^i$, where α_λ (≤ 1) is the absorptivity, and reflects $\mathrm{d}W_\lambda^r = (1 - \alpha_\lambda) \, \mathrm{d}W_\lambda$. Denoting ${L_\lambda}^{em}$ the emitted radiance, the body emits an amount of energy $\mathrm{d}W_\lambda^e = {L_\lambda}^{em} \, \mathrm{d}\lambda \, \mathrm{d}S \, \cos\theta \, \mathrm{d}\omega \, \mathrm{d}t$ through the surface element $\mathrm{d}S$. The second principle of the thermodynamics applied to the total energy $W = \int_{\Delta\lambda} \int_S \int_\Omega \int_{\Delta t} \mathrm{d}W$ implies conservation of energy. As it is valid for any integration domain it leads to the following equality:

$$ {L_\lambda}^{em} = \alpha_\lambda \, B_\lambda . $$

The body at a given temperature and under the assumption of local thermodynamical equilibrium emits the same amount of radiation whether it is inside or outside the cavity. It follows up a number of properties:

- If $\alpha_\lambda = 1$ for all wavelengths, then the body is said to be a *black body*.
- The radiance, ${L_\lambda}^{em}$, of a body is less or equal than that of the black body, B_λ, at the same temperature.
- The ratio ${L_\lambda}^{em}/B_\lambda$, called emissivity of the body and noted ε_λ, is equal to its absorptivity, α_λ.

This last property ($\varepsilon_\lambda = \alpha_\lambda$) is called *Kirchhoff's law*. It is important to understand the subtlety of this physical law. Kirchhoff's law stipulates that a body can only emit radiation at wavelengths which it can absorb. It does not mean however that, in the atmosphere, energy emitted is necessary equal to the energy absorbed at the same wavelength. The atmosphere also emits in a spectral domain that is function of its temperature according to Planck's law described below.

5.4.5.2 Planck's Law

We have seen that emission of radiation by molecules occurs in the form of a spectrum of lines. In the case of an opaque medium, emission is best described by Planck's law. Noting that emission of radiation by matter—which he approximated by harmonic oscillators—could only be quantified, Max Planck provided a physical explanation for the temperature dependence of the emission law of a black body that had been found empirically. This law can be expressed in terms of wavenumber ν or wavelength λ, noting that $B_\nu(T) \, \mathrm{d}\nu = B_\lambda(T) \, \mathrm{d}\lambda$:

$$ B_\nu(T) = \frac{2 \, h \, c^2 \, \nu^3}{\exp\left(\dfrac{h \, c \, \nu}{k_B \, T}\right) - 1} \quad \text{and} \quad B_\lambda(T) = \frac{2 \, h \, c^2}{\lambda^5 \left[\exp\left(\dfrac{h \, c}{k_B \, T \, \lambda}\right) - 1\right]}, \tag{5.31} $$

where T is the temperature of the black body and the values of physical constants involved are $k_B = 1.38065 \times 10^{-23}$ JK^{-1} (Boltzmann's constant) and $h = 6.62607 \; 10^{-34}$ Js (Planck's constant). B_ν and B_λ are generally expressed in $\mathrm{Wm}^{-2} \, \mathrm{sr}^{-1} \, (\mathrm{cm}^{-1})^{-1}$ and $\mathrm{Wm}^{-2} \, \mathrm{sr}^{-1} \, \mu\mathrm{m}^{-1}$, respectively. Figure 5.11 shows a few B_λ functions for various representative temperatures of the Earth–Sun system. It should be noted that the different $B_\nu(T)$ or $B_\lambda(T)$ curves do not cross each other.

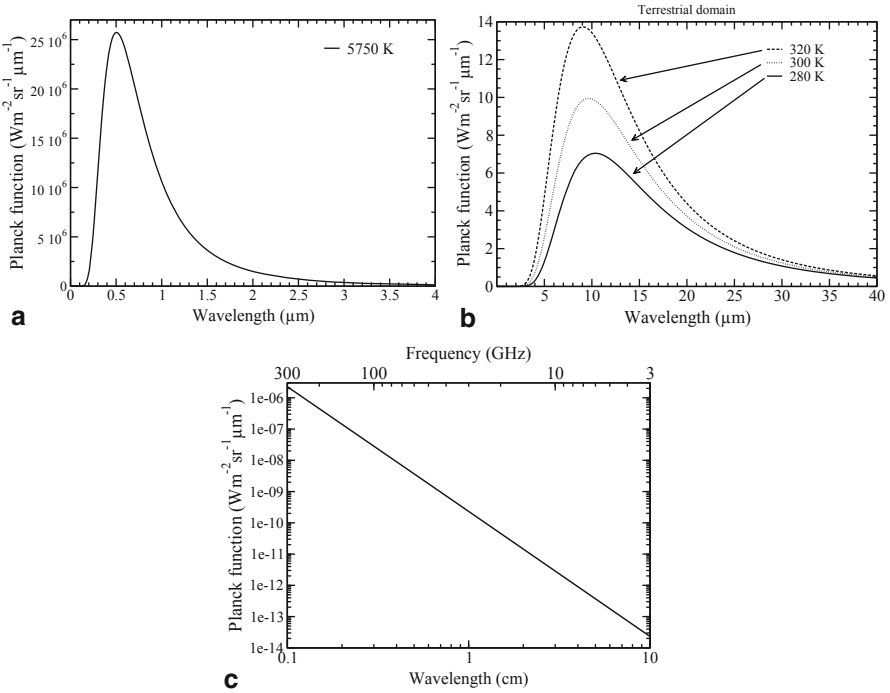


Fig. 5.11 Planck's law for various temperatures of a black body. **a** The *top left* panel shows the emission of radiation in the solar spectrum for temperature of 5750 K. **b** The *top right* panel shows emission of radiation in the thermal spectrum for temperatures of 280, 300, and 320 K. **c** The *bottom panel* shows the tail of the *top right* panel for the microwave domain on a logarithmic scale and for a temperature of 300 K

A black body at a higher temperature emits more radiation than a black body at lower temperature at all wavelengths. Furthermore, the maximum of emission shifts towards shorter wavelengths as the temperature increases. As a consequence, there is a clear distinction between the spectral domain corresponding to the emission by the Sun ($T \simeq 5750$ K, Fig. 5.11a) and the spectral domain corresponding to the emission by the Earth ($T \simeq 220$–320 K, Fig. 5.11b), although there is some overlap within the so-called near infrared domain, between 3 and 4 µm. Figure 5.11c shows, on a logarithmic scale, the tail of the distribution of terrestrial radiation emitted at very long wavelengths (i.e. the domain of microwaves which is useful for passive remote sensing of the atmosphere and surface). For these very long wavelengths, Planck's function can be well-approximated by the *Rayleigh–Jeans* approximation that is proportional to λ^{-4} and to the temperature, which is convenient to express the radiation directly in the form of a brightness temperature.

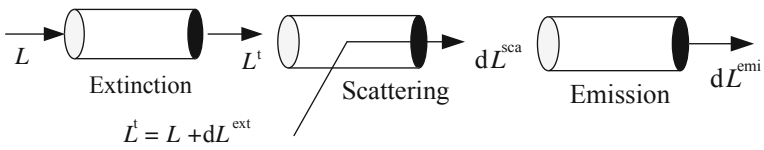


Fig. 5.12 The three main radiative processes: extinction, scattering, and emission

5.5 Atmospheric Radiative Transfer

The equation of radiative transfer describes the processes of scattering, absorption, and emission through an inhomogeneous medium such as the atmosphere. Analytic solutions exist in simple cases or if simplifying approximations are made. However, an accurate solution requires a numerical integration of the equation of radiative transfer.

5.5.1 Equation of Radiative Transfer

Let us note $L_\lambda(M, \lambda, \mathbf{s})$ the monochromatic radiance at wavelength λ at a point M in the solid angle $\mathrm{d}\omega$ and centred on the direction $\mathbf{s}$. For the sake of clarity we omit to mention the dependence on λ in the following. Let us consider a cylindrical volume element along the $\mathbf{s}$-axis with a surface $\mathrm{d}S$ and a length $\mathrm{d}l$ located at the abscissa l. The radiances at the entrance and exit of the cylinder are hereafter noted $L(l, \mathbf{s})$ and $L(l + \mathrm{d}l, \mathbf{s})$. Photons that come out the cylinder in the $\mathbf{s}$ direction can have three origins as illustrated on Fig. 5.12.

1. First some of the photons coming from the $\mathbf{s}$-direction and contributing to $L(l, \mathbf{s})$ have been absorbed or scattered in a different direction within the volume element. These processes (absorption and scattering) cause the extinction of the incident radiation, and the fraction of radiation that is transmitted through the cylinder is

$$L(l, \mathbf{s}) + \mathrm{d}L^{\mathrm{ext}} \text{ with } \mathrm{d}L^{\mathrm{ext}} < 0. \tag{5.32}$$

2. Second photons that are incident in the volume element from a direction $\mathbf{s}'$ different from $\mathbf{s}$ can be scattered in the direction $\mathbf{s}$. Let us note $\mathrm{d}L^{\mathrm{sca}}$ the corresponding radiance.
3. Third photons can be emitted in the volume element itself and come out in the $\mathbf{s}$ direction. Let us note $\mathrm{d}L^{\mathrm{emi}}$ the corresponding radiance.

The budget of radiance can, therefore, be written as

$$L(l + \mathrm{d}l, \mathbf{s}) = L(l, \mathbf{s}) + \mathrm{d}L^{\mathrm{ext}} + \mathrm{d}L^{\mathrm{sca}} + \mathrm{d}L^{\mathrm{emi}}. \tag{5.33}$$

Extinction

Absorption and scattering being linear processes that are statistically independent from each other, the extinction term can be rewritten as

$$\mathrm{d}L^{\mathrm{ext}} = -\sigma^{\mathrm{ext}} L(l,\mathbf{s})\,\mathrm{d}l \quad \text{with} \quad \sigma^{\mathrm{ext}} = \sigma^{\mathrm{abs}} + \sigma^{\mathrm{sca}}, \tag{5.34}$$

where σ^{abs} et σ^{sca} are, respectively, the absorption and scattering coefficients (in units of m^{-1}) defined previously.

Emission

According to Kirchhoff's law, a body emits a fraction of the radiation it can absorb:

$$\mathrm{d}L^{\mathrm{emi}} = B(T)\,\sigma^{\mathrm{abs}}\,\mathrm{d}l = B(T)\,d\varepsilon_\lambda, \tag{5.35}$$

where $\mathrm{d}\varepsilon_\lambda = \sigma^{\mathrm{abs}}\mathrm{d}l$ is the elementary monochromatic emissivity.

Scattering

The radiance due to the scattering process can be expressed as an integral over all possible incident directions weighted by the phase function and the scattering coefficient σ^{sca}:

$$\mathrm{d}L^{\mathrm{sca}} = \frac{\sigma^{\mathrm{sca}}}{4\pi}\left\{\iint_{4\pi} \mathrm{p}(\mathbf{s},\mathbf{s}_i)\; L(l,\mathbf{s}_i)\;\mathrm{d}\omega_i\right\}\mathrm{d}l, \tag{5.36}$$

where the 4π factor is here to normalize the phase function to the solid angle corresponding to a full sphere.

Equation 5.33 then becomes:

$$\begin{aligned}\mathrm{d}L(l,\mathbf{s}) &= L(l+\mathrm{d}l,\mathbf{s}) - L(l,\mathbf{s})\\ &= -\sigma^{\mathrm{ext}}\,L(l,\mathbf{s})\,\mathrm{d}l + \sigma^{\mathrm{abs}}\,B(T)\,\mathrm{d}l\\ &+\frac{\sigma^{\mathrm{sca}}}{4\pi}\left\{\iint_{4\pi} \mathrm{p}(\mathbf{s},\mathbf{s}_i) L(l,\mathbf{s}_i)\;\mathrm{d}\omega_i\right\}\mathrm{d}l,\end{aligned}$$

which can be broken down in an extinction term and a source term (emission and gain by scattering from other directions):

$$\mathrm{d}L(l,\mathbf{s}) = -\sigma^{\mathrm{ext}} L(l,\mathbf{s}) + J(l,\mathbf{s})\,\mathrm{d}l, \tag{5.37}$$

where the source function is defined as

$$J(l,\mathbf{s}) = \sigma^{\mathrm{abs}} B(T) + \frac{\sigma^{\mathrm{sca}}}{4\pi}\left\{\iint_{4\pi} \mathrm{p}(\mathbf{s},\mathbf{s}_i)\; L(l,\mathbf{s}_i)\;\mathrm{d}\omega_i\right\}. \tag{5.38}$$

5.5.2 *Extinction Only*

If only the extinction term is considered, which is the case in a medium where there are contributions from scattering and emission, the equation of radiative transfer reduces to

$$\mathrm{d}L(l,\mathbf{s}) = -\sigma^{\mathrm{ext}}(l)\, L(l,\mathbf{s})\, \mathrm{d}l. \tag{5.39}$$

The integration of $\mathrm{d}L/L$ between the two positions l_o and l_1 leads to

$$L(l_1,\mathbf{s}) = L(l_o,\mathbf{s})\, \exp\left(-\int_{l_o}^{l_1} \sigma^{\mathrm{ext}}(l')\, \mathrm{d}l'\right). \tag{5.40}$$

5.5.2.1 Optical Depth and Transmittance

The optical depth along a path $[l_o, l_1]$ is defined as the quantity:

$$\delta(l_o,l_1) = \int_{l_o}^{l_1} \sigma^{\mathrm{ext}}(l')\, \mathrm{d}l' \tag{5.41}$$

and the transmittance through the optical depth is

$$t(l_o,l_1) = e^{-\delta(l_o,l_1)}, \tag{5.42}$$

so that

$$L(l_1,\mathbf{s}) = L(l_o,\mathbf{s})\, e^{-\delta(l_o,l_1)} = L(l_o,\mathbf{s})\, t(l_o,l_1). \tag{5.43}$$

The transmittance $t(l_o,l_1)$ expresses an exponential decrease of the initial radiance, $L(l_o,\mathbf{s})$, due to extinction in the medium.

5.5.2.2 Gaseous Transmission

σ^{abs} being the product of the mass absorption coefficient k^{abs} with the amount of absorbing gas, its variation with altitude results from both variations in k^{abs} (due to variations in density and temperature) and variations in the amount of gas on the vertical. Using water vapour as an example, we compute the transmittance through a 1-km-thick layer situated at altitudes of 15, 8, and 1 km with varying degree of water vapour (Fig. 5.8). The intensity of the absorption line is stronger at 15 km but the small amount of water vapour at this altitude implies a transmittance close to 1. Most of the absorption comes from the lower layers in the atmosphere.

More generally, the exponential function in Eq. 5.43 implies that if absorption is large (i.e. because the wavelength corresponds to a strong absorption line, or because there is a large quantity of absorbing gas), there is little sensitivity to an increase

in absorbing gas as the transmittance saturates to zero. In contrast, if absorption is small (i.e. because the wavelength corresponds to a weak absorption line, because it is located in the wings of a stronger absorption line, or because the absorbing gas is in very small concentration), there is more sensitivity of the transmittance to an increase in concentration. For the same reason, if an absorption line is saturated, the profile of the wings (due to pressure and Doppler broadening) has more importance away than close to the line itself.

5.5.3 Scattering Medium

Let us now consider the substitution

$$\mathrm{d}\delta = \sigma^{\mathrm{ext}}\,\mathrm{d}l \tag{5.44}$$

and introduce a coordinate called *optical path* $\delta(l)$ verifying:

$$\delta(l) = \int_x^l \sigma^{\mathrm{ext}}(l')\,\mathrm{d}l'. \tag{5.45}$$

With this convention, the optical path is measured from the value of x for which $\delta(x) = 0$ (no extinction), to a value corresponding to l. The equation of radiative transfer, that we established earlier, can be expressed in terms of the optical path δ instead of l and becomes:

$$\mathrm{d}L(\delta, \mathbf{s}) = -L(\delta, \mathbf{s})\,\mathrm{d}\delta + J(\delta, \mathbf{s})\,\mathrm{d}\delta \tag{5.46}$$

with

$$J(\delta, \mathbf{s}) = J(l, \mathbf{s})/\sigma^{\mathrm{ext}}. \tag{5.47}$$

The previous equation can be multiplied by $e^{-(\delta_1-\delta)}$ and integrated between δ_o and δ_1 using integration by parts, in order to obtain:

$$L(\delta_1) = L(\delta_o)\,e^{-(\delta_1-\delta_o)} + \int_{\delta_o}^{\delta_1} J(\delta, \mathbf{s})\,e^{-(\delta_1-\delta)}\,\mathrm{d}\delta. \tag{5.48}$$

The radiance at the point of observation results from the initial radiance $L(\delta_o, \mathbf{s})$ attenuated along the optical path by the transmittance $\exp(-(\delta_1 - \delta_o))$ and the sum of the source functions from each of the volume elements themselves attenuated by the corresponding transmittances.

5.5.4 Plane-Parallel Atmosphere

Here the atmosphere is assumed to be homogeneous on the horizontal and the effect of the Earth's sphericity is neglected. The atmosphere can be reduced to one dimension on the vertical, which can be characterized by the z-coordinate oriented upwards ($z = 0$ at the surface). Such an atmosphere is referred to as a *plane-parallel* atmosphere. This assumption neglects three-dimensional effects (e.g. from clouds or from the orography) and is not valid at large solar zenith angles for which the Earth's sphericity cannot be neglected. The optical path δ defined above is oriented downwards ($\delta = 0$ at the top of atmosphere) is also called *optical depth*[4]. The zenith angle is measured against the z vertical axis. The sign of $\mu = \cos(\theta)$ can differentiate downwelling radiation ($\mu < 0$ for $\theta > 90°$) from upwelling radiation ($\mu > 0$ for $\theta < 90°$). An element of optical path in the **s**-direction referred by μ, φ is, therefore, equal to $\mathrm{d}l = \mathrm{d}z/\cos(\theta) = \mathrm{d}z/\mu$ in absolute value. With these conventions, the integration of the equation of radiative transfer for downwelling radiation becomes:

$$L^{\downarrow}(\delta, \mu, \varphi) = L^{\downarrow}(0, \mu, \varphi)\ \exp\left(-\frac{\delta}{-\mu}\right) + \int_0^{\delta} J(\delta', \mu, \varphi)\ \exp\left(-\frac{\delta - \delta'}{-\mu}\right)\ \frac{\mathrm{d}\delta'}{-\mu} \tag{5.49}$$

and for upwelling radiation:

$$L^{\uparrow}(\delta, \mu, \varphi) = L^{\uparrow}(\delta_s, \mu, \varphi)\ \exp\left(-\frac{\delta_s - \delta}{\mu}\right) + \int_{\delta}^{\delta_s} J(\delta', \mu, \varphi)\ \exp\left(-\frac{\delta' - \delta}{\mu}\right)\ \frac{\mathrm{d}\delta'}{\mu}. \tag{5.50}$$

In these expressions, $L^{\downarrow}(0, \mu, \varphi)$ is the downwelling radiance in the (μ, φ) direction at the top of the atmosphere where the origin of the optical depth coordinate lies, i.e. $\delta_{\mathrm{TOA}} = 0$. Conversely, $L^{\uparrow}(\delta_s, \mu, \varphi)$ is the upwelling radiance at the surface and δ_s is the optical depth at the surface, or equivalently, the total optical depth of the atmosphere.

The physical interpretation of these solutions is straightforward: the downwelling (upwelling) radiance that reaches the optical depth δ is the radiation at the upper (lower) boundary attenuated along its path (hence the exponential term) plus the sum of all source functions $J\,\mathrm{d}\delta/\mu$ from each effective optical depth $\mathrm{d}\delta/\mu$, themselves attenuated along their path.

[4] The difference between optical thickness and optical depth should be noted: the optical thickness is for a layer contained between l_o and l_1. It is, therefore, the difference between optical depths $\delta(l_o)$ and $\delta(l_1)$.

5.5.5 *Resolution of the Equation of Radiative Transfer*

To calculate the radiation transfer in the atmosphere requires the knowledge of atmospheric radiative properties on a discretized vertical grid: the amount of absorbing species and the characteristics of atmospheric scattering (δ, ϖ_λ, $\mathrm{p}(\theta, \varphi, \theta_i, \varphi_i)$) must be known at all layers. If several scatterers (i.e. molecules, aerosols, and cloud particles) are present, their respective phase functions should be combined as a weighted sum with their respective (elementary) scattering optical depths. It is often required to solve the equation of radiative transfer in the atmosphere. Depending on the applications, it may or may not be required to integrate spectrally over the wavelengths and over the directions to calculate downwelling and upwelling fluxes. Various techniques exist to resolve the equation of radiative transfer according to the precision required and the computational requirement.

5.5.5.1 Direct (Transmitted) Radiation

A prerequisite to solve the equation of radiative transfer is to compute the absorption and scattering optical thicknesses of each layer, which allows to calculate the direct (i.e. transmitted) radiance at each layer from the top-of-atmosphere incident radiance (for downwelling radiation) and from the surface reflected radiance (for upwelling radiation) and the attenuation terms ($e^{-\delta/\mu}$). The absorption coefficient must take into account the quantity and absorption cross section of each absorbing species in each layer. To compute these absorption coefficients, the *line-by-line* method is the most accurate but can be computationally demanding. It consists in summing up the contributions from all absorbing gases at each wavelength. As lines are broadened by the pressure and Doppler effects, some lines that are far away from the wavelength of calculation need to be accounted for, which increases the computational cost of the method. Other methods can be used to approximate the calculation of the transmittance on intervals of wavelengths (i.e. wavebands) such as band models, sum of exponentials and k-sum in order to decrease the computational time.

5.5.5.2 Diffuse Radiation

Successive Orders of Scattering

Scattering is the dominant process at the shorter wavelengths where absorption is small, in which case the source function can be expressed as

$$J(\delta, \mu, \varphi) = \frac{\varpi_\lambda}{4\pi} \int_{-1}^{1} \int_{0}^{2\pi} \mathrm{p}_\lambda(\mu, \varphi, \mu', \varphi')\, L(\delta, \mu', \varphi')\, \mathrm{d}\mu'\, \mathrm{d}\varphi'. \tag{5.51}$$

A physical solution consists in developing the source function according to successive orders of scattering. The source function for photons that undergo their n^{th} scattering is

$$J^{(n)}(\delta,\mu,\varphi)=\frac{\varpi_\lambda}{4\pi}\int_{-1}^{1}\int_{0}^{2\pi}\mathrm{p}_\lambda(\mu,\varphi,\mu',\varphi')\,L^{(n-1)}(\delta,\mu',\varphi')\,\mathrm{d}\mu'\,\mathrm{d}\varphi', \tag{5.52}$$

where $L^{(n-1)}(\delta,\mu',\varphi')$ corresponds to photons that have been scattered $n-1$ times. The 0^{th} order corresponds to direct radiation that has not been scattered at all. It can be computed from the Beer–Lambert law (also known as Beer's law):

$$L^{(0)}(\delta,\mu,\varphi)=F_\lambda^0\exp\left(-\frac{\delta}{-\mu_o}\right)\delta(\mu-\mu_o)\,\delta(\varphi-\varphi_o), \tag{5.53}$$

where δ is the Dirac delta function ($\int_{-\infty}^{\infty}\delta(x)\,\mathrm{d}x=1$ and $\delta(x)=0$ if $x\neq 0$) and F_λ^0 is the incident solar radiation flux at the top of the atmosphere in the (μ_o,φ_o) direction. Physically, the radiance at optical depth δ comes from all depths where it has been scattered from the (μ_o,φ_o) direction to the (μ,φ) direction[5]. Keeping the convention introduced previously where the cosine of the solar zenith angle, μ_o, is negative, the upward radiance at optical depth δ, corresponding to the first order of scattering, is

$$L^{\uparrow(1)}(\delta,\mu,\varphi)=\int_{\delta}^{\delta_s}F_\lambda^0\,\exp\left(-\frac{\delta'}{-\mu_o}\right)\frac{\varpi_\lambda}{4\pi}\,\mathrm{p}_\lambda(\mu,\varphi,\mu_o,\varphi_o)\,\exp\left(-\frac{\delta'-\delta}{\mu}\right)\frac{\mathrm{d}\delta'}{\mu}. \tag{5.54}$$

This expression can be extended to the downward radiance, to the second order of scattering, and the higher orders of scattering.

Single Scattering Approximation

The single scattering approximation consists in neglecting the contributions from multiple scattering. This approximation is justified in clear sky when the amount of scattering aerosols in the atmosphere is small, that is when the total (vertical) optical depth of the atmosphere is small. In this case, the radiance that emerges at the top of the atmosphere can be approximated by

$$\begin{aligned}L^{\uparrow(1)}(\delta=0,\mu,\varphi)=\;&\frac{\varpi_\lambda}{4\pi}\,F_\lambda^0\,\mathrm{p}_\lambda(\mu,\varphi,\mu_o,\varphi_o)\frac{1}{\mu}\left(\frac{1}{\mu}-\frac{1}{\mu_o}\right)^{-1}\times\\&\left(1-\exp\left[-\left(\frac{1}{\mu}-\frac{1}{\mu_o}\right)\delta_s\right]\right)\end{aligned} \tag{5.55}$$

and if $\delta_o\ll 1$

$$L^{\uparrow(1)}(0,\mu,\varphi)\approx\frac{\varpi_\lambda}{4\pi}\,F_\lambda^0\,\mathrm{p}_\lambda(\mu,\varphi,\mu_o,\varphi_o)\,\frac{\delta_s}{\mu}. \tag{5.56}$$

[5] We neglect the fact that incident solar radiation is not completely collimated but is included in a cone with half-aperture of 0.266° that corresponds to the solar disk.

Discrete-Ordinate Method

An accurate resolution of multiple scattering in the atmosphere requires more sophisticated methods. One of them is the so-called *discrete-ordinate method*, whereby the radiance is decomposed in a Fourier series in the azimuthal angle variable, φ:

$$L(\delta, \mu, \varphi) = \sum_{m=0}^{2N-1} L^{(m)}(\delta, \mu) \cos m\varphi. \tag{5.57}$$

The scattering function is itself decomposed in a series of associated Legendre functions:

$$\mathrm{p}_\lambda(\mu, \varphi, \mu', \varphi') = \sum_{m=0}^{2N-1} (2 - \delta_{0m}) \left(\sum_{l=m}^{2N-1} (2l+1) g_l P_l^m(\mu) P_l^m(\mu') \right) \cos m(\varphi - \varphi'), \tag{5.58}$$

where δ_{0m} is 1 if $m = 0$ and 0 otherwise. The equation of radiative transfer is, therefore, decomposed into $2N-1$ integro-differential equations that are independent of the azimuthal angle φ and correspond to each of the terms m of the decomposition. The integrals over the cosine of the solar zenith angle are then approximated using Gaussian quadratures and the new system of equations can be resolved analytically. We obtain for each of these equations a system of $2N$ linear differential equations, N being the order of the Gaussian quadrature. Practically, N is very dependent on the degree of anisotropy of the phase function.

The Adding–Doubling Method

The *adding–doubling* method consists in doubling and then adding a number of scattering layers by computing the multiple reflections at the interfaces. The computation is initialized for each atmospheric layer with a very thin layer for which the single scattering approximation is fully valid (typically $\delta_0 = 2^{-30}$). The reflectances and transmittances of the scattering layers can then be doubled as many times as needed in order to reach the scattering optical depth δ_i of layer i. The scattering layers can then be added according to the same principle to resolve the total reflectance and transmittance of the atmosphere. The principle of the method is indicated in Fig. 5.13. Here T_i and R_i represent the transmittance (in fact the sum of the transmittances for direct and diffuse radiation) and the reflectance of layer i, respectively. The combination of the contributions from multiple reflections at the interface between two layers results in a geometric series that converge rapidly for thin layers but more slowly for thicker and/or weakly absorbing layers.

Other methods exist to solve the equation of radiative transfer such as the successive orders of scattering method or Monte Carlo method.

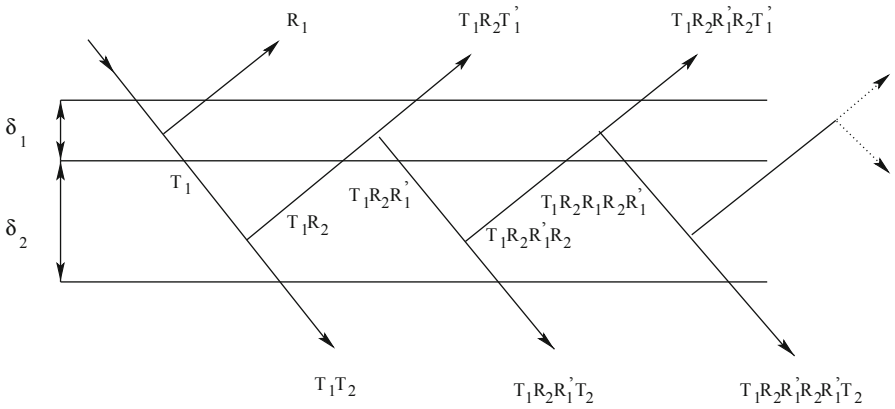


Fig. 5.13 Schematic of the adding–doubling method for resolving the radiative transfer equation. T_1 and T_2 are the transmittances of layers 1 and 2. R_1 and R_2 are the reflectances of layers 1 and 2

5.6 Absorption Bands, Energy, and Actinic Fluxes

We now discuss other aspects of atmospheric radiative transfer, in particular energetic aspects.

5.6.1 Main Molecular Absorption Bands in the Atmosphere

Because of its complex chemical composition, the atmosphere absorbs at many wavelengths. Atmospheric transmittances in the solar and infrared domains are shown in Figs. 5.14 and 5.15 for a range of trace gases. Atmospheric absorption occurs in the form of bands, with rapid variations of the absorption with the wavelength, or in the form of a continuum, where the spectral variations are much smoother. Spectral domains where there is little absorption are known as *atmospheric windows*.

At short wavelengths, the main absorbing gases are atomic oxygen and ozone, which are responsible for the absorption of ultraviolet radiation. Solar radiation below 310 nm does not reach the Earth's surface. Ozone also absorbs slightly between 0.4 and 1 μm (Chappuis bands) whereas molecular oxygen (or dioxygen) absorbs around 0.69 and 0.76 μm. Water vapour has many absorption bands around 0.93, 1.87, 2.7 μm, etc. In the troposphere, absorption is very weak in the visible window that extends from 0.4 to 0.67 μm. There are other atmospheric windows of varying widths at 1.06, 1.22, 2.2, and 3.7 μm.

Atmospheric gases that absorb in the infrared domain are known as greenhouse gases. The main greenhouse gas in the Earth's atmosphere is water vapour, H_2O, for which the rotational band extends from about 12 to 13 μm all the way to millimetric wavelengths. Carbon dioxide, CO_2, is the second most important greenhouse gas.

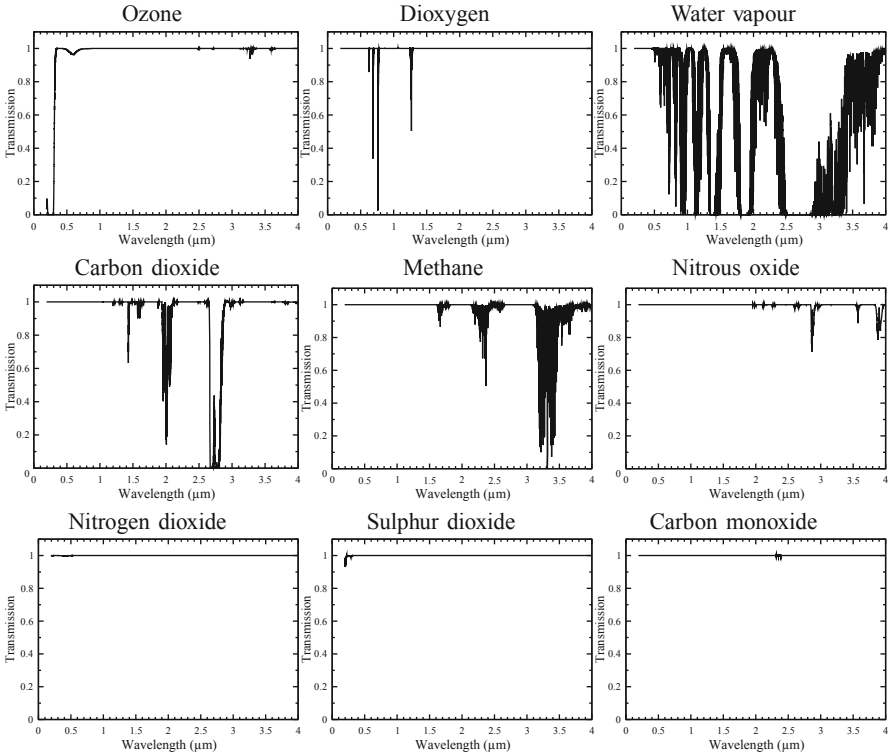


Fig. 5.14 Atmospheric transmittances in the solar domain for dioxygen and various atmospheric trace gases

Absorption by the CO_2 molecule around 15 μm is due to the fundamental vibrational–rotational transition upon which are superposed many transitions between vibrational states where the differences in energies correspond to similar wavelengths (i.e. those of isotopic molecules and those from transitions between vibrational states 1 and 2, 2 and 3 or any other combination). A vibrational–rotational band of water vapour is centred at 6.3 μm and extends approximately from 4 to 8 μm. Between the two water vapour bands, absorption is much weaker and the atmosphere is relatively but not completely transparent. This atmospheric window from 8 to 12 μm is the only significant atmospheric window in the infrared domain. Its role in radiative exchange in the atmosphere and between the Earth's surface and the outer space is particularly important because it corresponds to the maximum of thermal emission by a black body at temperatures typical of those of the surface–atmosphere system (see Fig. 5.11b). The ozone rotational band is located in the centre of the atmospheric window (9.6 μm). Other greenhouse gases, such as CO_2, chlorofluorocarbons (CFCs) and methane (CH_4) also present absorption bands in this window. Even though these absorption bands are weak, they are important and contribute significantly to the greenhouse effect because the atmosphere is fairly transparent and because they

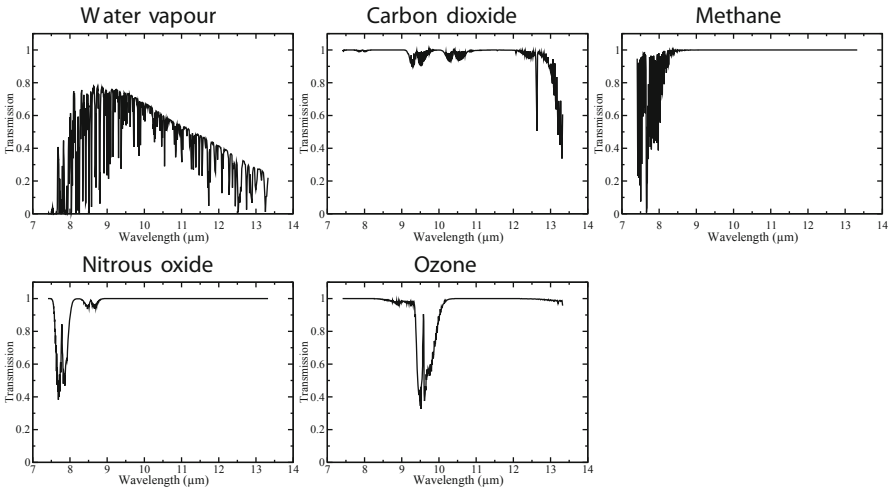


Fig. 5.15 Atmospheric transmittances in the infrared domain for various atmospheric trace gases. A typical vertical profile of temperature and humidity was used

correspond to the maximum of emission by the surface and atmosphere. An obvious consequence of the strategic position of these absorption bands is the rather large sensitivity of the greenhouse effect to non-CO_2 greenhouse gases such as methane and CFCs. Most of the absorption in the infrared atmospheric window is nevertheless due to the absorption continuum of water vapour due to the superposition of absorption in the wings of the many absorption lines that lie outside the window.

Atmospheric windows are used for remote sensing of the Earth's surface, aerosols and clouds. In contrast, remote sensing of gaseous species is made at the wavelengths where absorption takes place. Atmospheric sounding of temperature and humidity profiles can be performed through remote sensing at wavelengths where water vapour and another gas whose vertical profile is known (typically CO_2) absorb radiation.

5.6.2 Radiative Flux

For energetic aspects, the calculation of radiative fluxes is more relevant than that of radiances. The monochromatic radiative flux through a surface dA is equal to the component of the radiance that is perpendicular to the surface integrated over the whole spherical solid angle:

$$F_\lambda = \int_\omega L_\lambda \cos\theta \, \mathrm{d}\omega. \tag{5.59}$$

It is expressed in $\mathrm{Wm}^{-2}\ \mu\mathrm{m}^{-1}$.

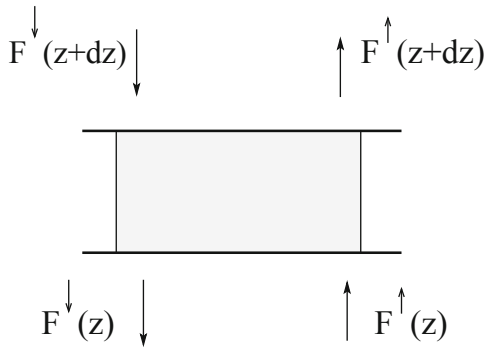


Fig. 5.16 Schematic of resolving the radiative transfer equation. $F^{\uparrow}$ and $F^{\downarrow}$ are the upward and downward vertical fluxes, respectively

In the atmosphere, it is usual to distinguish upward and downward radiative fluxes. They can easily be written as a function of the upward and downward radiances in spherical coordinates:

$$F_{\lambda}^{\uparrow} = \int_0^{2\pi} \int_0^{\pi/2} L_{\lambda}^{\uparrow}(\theta, \varphi) \cos\theta \, \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi \tag{5.60}$$

$$F_{\lambda}^{\downarrow} = \int_0^{2\pi} \int_0^{\pi/2} L_{\lambda}^{\downarrow}(\theta, \varphi) \cos\theta \, \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi. \tag{5.61}$$

If radiation is isotropic, it can easily be shown that the radiative flux can be written as

$$F_{\lambda}^{\uparrow} = \pi \, L^{\uparrow}. \tag{5.62}$$

The monochromatic radiative flux can be integrated over a range of wavelengths (e.g. the solar spectrum from roughly 0.2 to 4 μm, the infrared spectrum from roughly 4 to 100 μm, the spectral response function of a radiometer, etc.) to provide a radiative flux in unit Wm^{-2}.

5.6.3 Two-Flux Method

Two-flux methods are rapid and well-adapted to compute average quantities such as vertical fluxes in a plane-parallel atmosphere (Fig. 5.16). However, they are not suited to compute the angular distribution of (downward and upward) radiances in the case of a scattering atmosphere. There exist many variants of two-flux methods, with typical accuracies within 5 % in comparison to more accurate methods that explicitly compute the field of atmospheric radiances. Methods differ in their approximations of the angular dependence of the radiance field. In pure two-flux methods, it is assumed that:

$$L(\delta, \pm\mu) = \int_0^{2\pi} L(\delta, \mu, \varphi) \, \mathrm{d}\varphi = L^{\pm}(\delta) \qquad \forall \mu. \tag{5.63}$$

In the Eddington approximation, the following approximation is made:

$$L(\delta, \mu) = \int_0^{2\pi} L(\delta, \mu, \varphi)\, \mathrm{d}\varphi = L_o(\delta) + \mu L_1(\delta). \tag{5.64}$$

The angular integration of the equation of radiative transfer leads to rather simple analytical solutions. In the delta-Eddington approximation, the forward scattering peak is truncated before the Eddington approximation is applied. While two-flux methods are useful to understand atmospheric radiation, they are increasingly replaced with more accurate methods as more computational power becomes available.

5.6.4 Stefan–Boltzmann Law

In the case of a black body emitting radiation, the radiance integrated over the whole spectrum of wavelength is equal to:

$$B(T) = \int_0^{+\infty} B_\lambda(T)\, \mathrm{d}\lambda = \int_0^{+\infty} B_\nu(T)\, \mathrm{d}\nu = \frac{2\,\pi^4\,k_B^4}{15\,c^2\,h^3}\,T^4 \tag{5.65}$$

and can be obtained by the following substitution $x = h\,c/k_B\,\lambda\,T$ in the first integral knowing that $\int_0^{+\infty} x^3/(e^x - 1)\,\mathrm{d}x = \pi^4/15$. The emitted radiation being isotropic, the angular integration of the Planck function on a half-space provides the spectral emittance:

$$M_\lambda(T) = \int_{2\pi} B_\lambda(T)\, \cos\theta\, \mathrm{d}\omega = \pi\, B_\lambda(T). \tag{5.66}$$

Integrating over all wavelengths provides the emittance:

$$M(T) = \int_0^{+\infty} M_\lambda(T)\, \mathrm{d}\lambda = \sigma\, T^4, \tag{5.67}$$

where $\sigma = 2\pi^5 k_B^4/15c^2h^3 = 5.67\ 10^{-8}\ \mathrm{Wm^{-2}\ K^{-4}}$. This equation is known as the *Stefan–Boltzmann law* and σ is the Stefan–Boltzmann's constant.

5.6.5 Radiative Budget

The simplest climate model that one can think of consists in equilibrating the sources and sinks of energy for the surface–ocean–atmosphere system taken as a whole. In fact the quasi-totality of the energy exchanges between the climate system and its environment (i.e. the Earth's interior and the outer space) takes place in the form of radiative energy (geothermal energy that is slowly diffused at the bottom of the Earth's surface can be neglected). Radiative equilibrium implies that the Earth radiates as

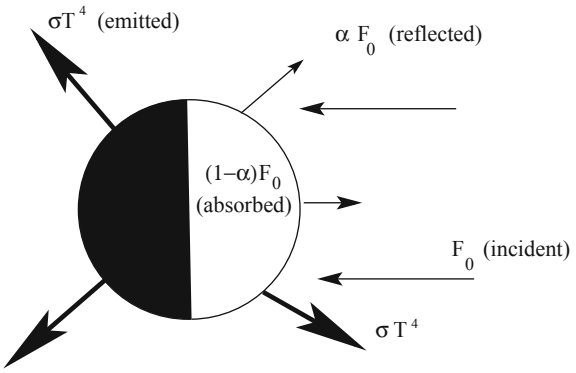


Fig. 5.17 Simplified energy budget in the terrestrial atmosphere

much terrestrial energy to space than it absorbs solar radiation. The source of energy occurs in the solar spectrum (also called shortwave radiation) while the sink of energy occurs in the infrared spectrum (also called longwave radiation).

The total amount of radiative energy received by the Earth is $\pi\ R^2\ F_0$, where R is the Earth's radius and F_0 is the total energy received per unit surface from the Sun (Fig. 5.17). Specifically this quantity is the solar flux received by a unit surface that is perpendicular to the Earth–Sun direction and roughly located at the mean Earth–Sun distance. It is referred to as the *solar constant* and its value has been estimated from atmospheric measurements and satellite instruments ($F_0 \approx 1362 \pm 3\ \mathrm{Wm}^{-2}$ according to the latest estimates). A fraction of this solar energy is reflected by the Earth's system. With α denoting the planetary albedo in the solar spectrum ($\alpha \approx 0.30$), the absorbed solar energy amounts to:

$$Q_{\mathrm{abs}} = \pi\ R^2\,(1-\alpha)\,F_0. \tag{5.68}$$

The amount of energy emitted by the Earth's system, assuming it behaves as a black body with a uniform temperature T, is

$$F^{\uparrow} = 4\,\pi\ R^2\,\sigma\ T^4. \tag{5.69}$$

Radiative equilibrium implies that:

$$Q_{\mathrm{abs}} - F^{\uparrow} = 0 \tag{5.70}$$

that is

$$4\,\sigma\ T^4 = (1-\alpha)\,F_0. \tag{5.71}$$

This simple model predicts an equilibrium temperature of 255 K, which is too small in comparison to the observed average Earth's surface temperature of 288 K. The difference is due to the greenhouse effect related to the absorption and emission of infrared radiation in the atmosphere by greenhouse gases including water vapour.

Figure 1.1 in Chap. 1 illustrates the complexity of the Earth's radiative budget. Incident solar radiation is $F_0/4 = 342$ Wm^{-2} (corresponding to an earlier estimate of the solar constant of 1368 Wm^{-2}). Reflection of solar radiation is due to the atmosphere (molecules, aerosols, and clouds) for 77 Wm^{-2} and the surface for 30 Wm^{-2}. The rest of solar radiation is absorbed in the atmosphere for 67 Wm^{-2} and at the surface of oceans and continents for 168 Wm^{-2}. Outgoing longwave radiation (235 Wm^{-2}) compensates for the absorption of solar radiation[6]. It originates from longwave radiation emitted by the atmosphere itself (165 Wm^{-2} by gases and 30 Wm^{-2} by clouds). Only a small fraction of the outgoing longwave radiation that is emitted by the surface (40 Wm^{-2} out of 390 Wm^{-2}) goes through the atmosphere without being absorbed. This is due to the fact that, except for the atmospheric windows between 8 and 12 μm, the atmosphere is very absorbing in the longwave spectrum. The greenhouse effect is substantial because the downward infrared radiation at the surface (324 Wm^{-2}) is only slightly smaller than the upward infrared radiation at the surface (390 Wm^{-2}). The energy budget would not be closed at the surface and in the atmosphere without considering the sensible heat flux (24 Wm^{-2} abandoned by the surface to the atmosphere) and the latent heat flux (78 Wm^{-2} corresponding to the evaporation of solid and liquid water at the Earth's surface and deposited in the atmosphere upon the condensation of water vapour).

5.6.6 Actinic Fluxes

Diffuse radiation caused by the scattering of solar radiation by the Earth's surface, clouds, aerosols and molecules can increase substantially the amount of shortwave radiation available for photochemical reactions. As we are interested in molecular absorption, it is the amount of incident radiation on a point rather than the flux through a surface that matters. This quantity is called *actinic flux*.

Let us consider first of a nonscattering atmosphere. In this case, solar radiation is completely collimated. If F_o is the irradiance on a surface perpendicular to the Sun's direction, i.e. it is the flux from the incoming solar beam, then the radiance is

$$L_o = \frac{F_o}{\Delta\omega}\,\delta(\mu - \mu_o)\,\delta(\varphi - \varphi_o), \tag{5.72}$$

so that $L_o = 0$ outside the direction of the solar beam of solid angle $\Delta\omega$. The actinic flux is

$$\Phi_o = \int_{2\pi} L_o\, d\omega = L_o\, \Delta\omega = F_o. \tag{5.73}$$

[6] In a warming climate, as we currently experience, outgoing longwave radiation slightly lags behind absorbed solar radiation because some energy is stored in the ocean. Hence, outgoing longwave radiation is slightly smaller than absorbed solar radiation once interannual variability is smoothed out.

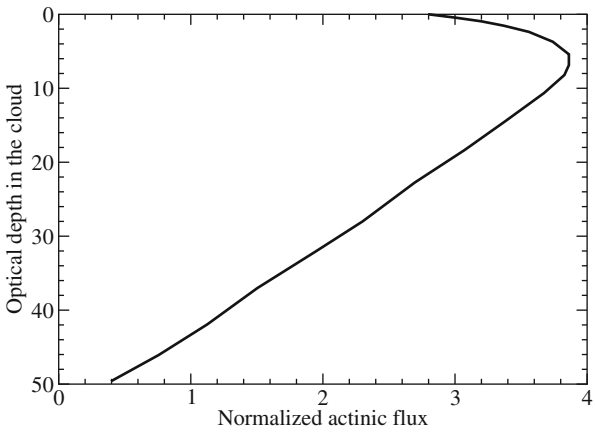


Fig. 5.18 Amplification of the actinic flux inside a typical stratocumulus cloud. Calculations have been made using the discrete-ordinate method using a cloud optical depth $\tau = 50$ and a solar zenith angle of 0°

Let us assume now that the same amount of downward solar radiation is isotropic:

$$L(\mu, \varphi) = L_d \quad \forall \mu, \forall \varphi. \tag{5.74}$$

The same quantity of energy, F_o, is split over all directions and the irradiance is

$$F_o = \int_{2\pi} L_d \, \cos i \, \mathrm{d}\omega = \pi \, L_d \tag{5.75}$$

and the actinic flux is

$$\phi_d = \int_{2\pi} L_d \, \mathrm{d}\omega = 2\,\pi \, L_d = 2 \, F_o. \tag{5.76}$$

Diffuse isotropic radiation is, therefore, twice as efficient for photolysis and photochemistry than collimated radiation. As a consequence, in a cloud-free atmosphere, reflection by a surface can increase the actinic flux by a factor of up to three (direct solar radiation and reflected radiation which, if it is isotropic, is twice as efficient than the direct incoming radiation). This factor of three increase is of course an absolute upper limit because surface albedo is less than 1 and reflected radiation is not isotropic. Specular reflection, for which the reflection of a collimated beam is itself collimated, can only increase the actinic flux by a factor of two (the incident and the reflected solar radiation). In cloudy situations, downward solar radiation is itself diffuse. The amplification can in principle reach a factor of four, in the theoretical case of isotropically diffuse radiation and no absorption in the cloud. Figure 5.18 shows the variation of the actinic flux normalized to the incident flux inside a stratocumulus. In the upper part of the cloud the amplification of the actinic flux approaches the upper limit of four.

5.6.7 Polarization of Radiation

We have left out until this point an additional property of monochromatic electromagnetic waves, that is the state of polarization. The electric field of a plane monochromatic wave with wavelength λ and frequency f propagating in the direction z can be written as

$$\mathbf{E} = \mathbf{A}\,\cos\left(\frac{2\pi\,z}{\lambda} - 2\pi\,f\,t\right) + \mathbf{B}\,\sin\left(\frac{2\pi\,z}{\lambda} - 2\pi\,f\,t\right),$$

where **A** and **B** are two vectors and t is the time. The electric field describes an ellipse at any given z and the wave is said to be elliptically polarized. If the two vectors **A** and **B** are orthogonal with the same magnitude, then **E** oscillates on a circle and the wave is said to be circularly polarized. If one of the two vectors **A** or **B** is zero, then **E** oscillates on a straight line and the wave is said to be linearly polarized.

There are many ways to decompose the electric field of a polarized wave. For instance, the electric field (represented by complex numbers) can be expressed as

$$\mathbf{E} = E_x\,\mathbf{e_x} + E_y\,\mathbf{e_y},$$

where $\mathbf{e_x}$ and $\mathbf{e_y}$ are orthogonal unit vectors (referred to as horizontal and vertical) and E_x and E_y are the complex amplitudes. It is customary to define the polarized states of light using four observable parameters, called the Stokes parameters and noted I, Q, U, and V:

- $I = E_x\,E_x^* + E_y\,E_y^*$ corresponds to the irradiance measured in the absence of polarizers.
- $Q = E_x\,E_x^* - E_y\,E_y^*$ corresponds to the difference in irradiances measured through a horizontal and a vertical polarizer.
- $U = E_x\,E_y^* + E_x^*\,E_y$ corresponds to the difference in irradiances measured through $+45°$ and $-45°$ polarizers.
- $V = (E_x\,E_y^* - E_x^*\,E_y)\,i$ corresponds to the difference in irradiances measured through two right-handed and left-handed circular polarizers.

The Stokes parameters are not independent from each other: for a given monochromatic wave $I^2 = Q^2 + U^2 + V^2$. It should be noted that $V = 0$ for linearly polarized radiation. For quasi-monochromatic light, the complex amplitudes vary in time and depending on the degree of correlation between E_x and E_y, the beam is said to be unpolarized, partially polarized or polarized. In this case, the Stokes parameters are defined as the time average of instantaneous I, Q, U, and V parameters over an interval that is long enough compared to the period of the wave. It can be shown that $I^2 \geq Q^2 + U^2 + V^2$. If the light is unpolarized, then $Q = U = V = 0$. The degree of polarization and degree of linear polarization are defined as $\sqrt{Q^2 + U^2 + V^2}/I$ and $\sqrt{Q^2 + U^2}/I$, respectively.

Modelling polarization therefore requires to represent the radiation not just by its intensity (noted L above, but I here for consistency with usual notations) but

by a vector, known as the Stokes vector **I**, of dimension four and comprised of the four Stokes parameters. Scattering in a given direction is then expressed by a 4×4 matrix. The Stokes vector of the scattered radiation at a distance R from the scatterer can be obtained from the incident Stokes vector according to the following matrix multiplication:

$$\begin{pmatrix} I_s \\ Q_s \\ U_s \\ V_s \end{pmatrix} = \frac{\lambda^2}{4\,\pi^2\,R^2} \begin{pmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{21} & S_{22} & S_{23} & S_{24} \\ S_{31} & S_{32} & S_{33} & S_{34} \\ S_{41} & S_{42} & S_{43} & S_{44} \end{pmatrix} \begin{pmatrix} I_i \\ Q_i \\ U_i \\ V_i \end{pmatrix}. \tag{5.77}$$

The elements of the scattering matrix **S** depend on the geometry and properties of the scattering medium. For a homogeneous sphere, the scattering matrix simplifies to

$$\mathbf{S} = \begin{pmatrix} S_{11} & S_{12} & 0 & 0 \\ S_{12} & S_{11} & 0 & 0 \\ 0 & 0 & S_{33} & S_{34} \\ 0 & 0 & -S_{34} & S_{33} \end{pmatrix}, \tag{5.78}$$

where it can be shown that $S_{11}^2 = S_{12}^2 + S_{33}^2 + S_{34}^2$. The corresponding S coefficients are predicted by the Mie theory and the reader is referred to Appendix C for further details.

Finally, we can relate the Stokes matrix to the phase function introduced previously. We can rewrite Eq. 5.77 as

$$\mathbf{I_s} = \frac{\lambda^2}{4\,\pi^2\,R^2}\,\mathbf{S}\,\mathbf{I_i} = \frac{1}{4\,\pi\,R^2}\,\frac{\lambda^2\int_{4\,\pi} S_{11}\,\mathrm{d}\omega}{4\,\pi^2}\,\frac{4\,\pi\,\mathbf{S}}{\int_{4\,\pi} S_{11}\,\mathrm{d}\omega}\,\mathbf{I_i} = \frac{s^{\mathrm{sca}}}{4\,\pi\,R^2}\,\mathbf{P}\,\mathbf{I_i}, \tag{5.79}$$

where we have used the normalized phase function for polarized radiation

$$\mathbf{P} = \frac{4\,\pi\,\mathbf{S}}{\int_{4\,\pi} S_{11}\,\mathrm{d}\omega} \tag{5.80}$$

and the definition of the scattering cross section

$$s^{\mathrm{sca}} = \int_{4\,\pi} I_s\,R^2\,\mathrm{d}\omega/I_i = \frac{\lambda^2}{4\,\pi^2}\int_{4\,\pi} S_{11}\,\mathrm{d}\omega. \tag{5.81}$$

We show on Fig. 5.19 the P_{11} and P_{12} elements of the phase function as well as the polarization ratio of the radiation resulting from the scattering of unpolarized incident

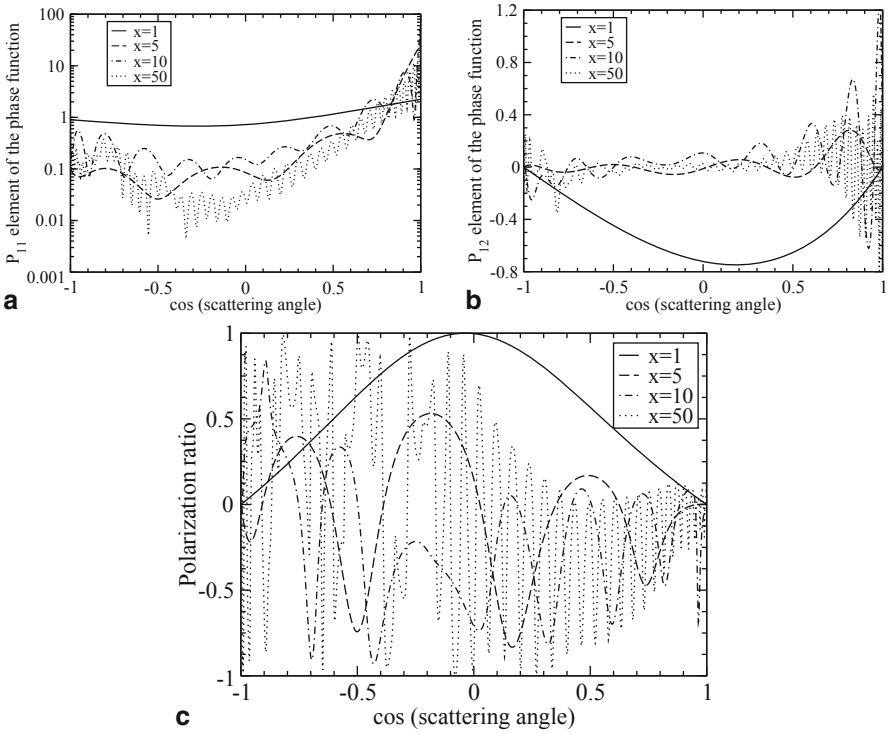


Fig. 5.19 **a** P_{11} and **b** P_{12} elements of the phase function, and **c** polarization ratio of scattered unpolarized radiation by spherical homogeneous particles with a range of Mie parameters and a fixed imaginary refractive index $m = 1.33 - 10^{-8}i$

radiation by spherical homogeneous particles. The polarization ratio shows large and multiple variations with the scattering angle at large Mie parameters. When averaged over a realistic particle size distribution, the polarization ratio approaches zero for a coarse mode aerosol, with the notable exception of some oscillations observed at large scattering angles in the case of coarse spherical particles such as coarse sea salt particles. In contrast it can be seen that small particles compared to the wavelength (i.e. typically accumulation mode aerosols) have a clear signature on polarized light.

All of this is highly relevant to the atmosphere because atmospheric radiation can be partially linearly polarized. Incoming solar radiation is not polarized because it results from incoherent processes. However, the unpolarized incident radiation can become polarized through scattering in the atmosphere and reflection at the surface. Resolution of the equation of radiative transfer needs in principle to account for the polarization of radiation in the atmosphere, which is computationally more demanding but necessary for some applications. Accounting for polarization only has a small effect on radiative fluxes themselves (i.e. the intensity of the radiation field integrated over a hemisphere) but can be substantial on the directional radiances (up to 10 % in some directions at 450 nm). The effect is larger at shorter wavelengths especially in

the ultraviolet. It can also be interesting to observe and model polarization when it comes to remote sensing of the atmosphere (i.e. to retrieve aerosol and cloud properties in the visible spectrum, see Chap. 6). This is because aerosols and cloud droplets contribute to polarize radiation upon scattering. Multiple scattering, however, tends to depolarize radiation. Because of the interesting properties of polarization, a number of active remote sensing techniques, such as the lidar, also rely on the emission and detection of polarized radiation.

Exercises

1a. The Sun diameter is 1.4×10^6 km. The average Earth–Sun distance is 150×10^6 km. At this distance a surface perpendicular to the direction of incoming solar radiation receives a radiative flux $S = 1370\ \mathrm{Wm}^{-2}$. Estimate the temperature at which the Sun emits its radiation assuming it behaves as a black body.

1b. The Earth is assumed to be in radiative equilibrium. Its average albedo is $\alpha = 0.30$. Compute the average temperature at which the Earth emits its radiation assuming it behaves as a black body and neglecting the presence of an atmosphere. Compare this temperature to the Earth's observed average surface temperature which is 288 K. Provide an explanation for the discrepancy.

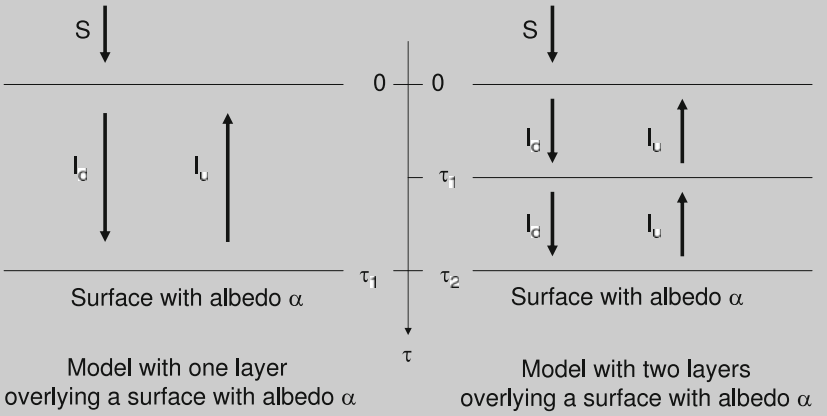
1c. The Earth's greenhouse effect is represented with a simplified model which is comprised of the Earth's surface and a single homogeneous atmospheric layer. T_1 is the temperature of the atmosphere which is assumed transparent to solar radiation, but behaves like a black body in the infrared. Calculate the temperature of the Earth's surface T_0.

1d. Same question as in 1c above, but considering that a fraction $\beta = 0.2$ of the absorbed solar radiation is actually absorbed by the atmosphere.

1e. Same question as 1d above, but considering that the atmospheric layer has an absorptivity (and therefore an emissivity ϵ) equal to 0.90.

2. Integrate Eq. 5.60 to derive Eq. 5.62.

3. Let us consider a classical two-stream radiative transfer model. The atmosphere is assumed to be plane-parallel. Radiation can be either downward (I_d) or upward (I_u) and no other direction of propagation is considered. Both I_u and I_d are positive by definition. The optical depth τ is used as the vertical coordinate and it taken equal to 0 at the top of the atmosphere and directed downwards. The surface albedo is α as shown on the schematic below.



We note β the backscatter fraction and ϖ the single scattering albedo. Indices 1 and 2 are used to refer to the top and bottom layers in the two-layer case.

3a. We now consider one of the two layers. Express the Beer–Lambert law for I_u and I_d under the differential form taking into account the two-flux approximation (i.e. scattering in the forward direction does not contribute to attenuate downward or upward radiation, only scattering in the backward direction is considered). This makes a system of two equations for each layer.

3b. We set $E = I_u + I_d$ (total flux) and $F = I_u - I_d$ (net flux). Rewrite the system of equations for variables E and F. Verify that

$$\begin{cases} E = 2\,\beta\,F_0\,\tau + E_0 \\ F = F_0 \end{cases}$$

is a solution in the conservative case (i.e. $\varpi = 1$) and that

$$\begin{cases} E = E_0 \exp(\gamma\tau) + E'_0 \exp(-\gamma\tau) \\ F = \frac{1-\varpi}{\gamma}\,E_0\,\exp(\gamma\tau) - \frac{1-\varpi}{\gamma}\,E'_0\,\exp(-\gamma\tau), \end{cases}$$

where $\gamma = \sqrt{(1-\varpi)(1-\varpi+2\,\varpi\,\beta)}$ is a solution in the nonconservative case ($\varpi < 1$).

3c. We first consider a one-layer model (left panel). Express the boundary conditions at the top and bottom of the layer and deduce the constants E_0 and E'_0.

3d. Compute the albedo of the surface–atmosphere system $\alpha' = I_u(0)/I_d(0)$

and compare it to the surface albedo α.

3e. Repeat questions 3c and 3d for the two-layer model (right panel) when only the bottom layer is absorbing. Write the boundary conditions at the top and bottom of each of the two layers for the solutions provided above.

3f. Same question as above, but only the upper layer is absorbing.

Solutions

1a. $T_{\text{Sun}} = 5770$ K.

1b. $T_{\text{Earth}} = 255$ K. This estimate neglects the atmospheric greenhouse effect.

1c. $T_1 = 255$ K and $T_0 = 303$ K.

1d. $T_1 = 255$ K and $T_0 = 295$ K.

1e. $T_1 = \sqrt[4]{\dfrac{S\,(1-\alpha)(\beta+\epsilon-\beta\epsilon)}{4\,\epsilon\,\sigma\,(2-\epsilon)}} = 250$ K

$T_0 = \sqrt[4]{\dfrac{S\,(1-\alpha)(2-\beta)}{4\,\sigma\,(2-\epsilon)}} = 288$ K.

References

Bohren CF, Huffman DR (1998) Absorption and scattering of light by small particles. Wiley Science Paperback Series, New York, 544 pp

Fouchet T (2000) Physico-chimie de l'atmosphère jovienne à partir de l'analyse des données du satellite infrarouge ISO, Thèse de doctorat de l'Université Paris 6, 274 pp

López-Puertas M, Taylor FW (2001) Non-LTE radiative transfer in the atmosphere. World Scientific Publishing Co Pte Ltd, 504 pp

van de Hulst HC (1982) Light scattering by small particles. Dover Publications, New York

Further Reading (Textbooks and Articles)

Bohren CF, Clothiaux EE (2006) Fundamentals of atmospheric radiation: an introduction with 400 problems. Wiley-VCH, 490 pp

Goody RM, Yung YL (1989) Atmospheric radiation, theoretical basis, 2nd edn. Oxford University Press, Oxford, 519 pp

Hansen JE, Travis LD (1974) Light scattering in planetary atmospheres. Sp Sci Rev 16:527–610
Lenoble J (1993) Atmospheric radiative transfer. A. Deepak Publishing, Hampton, 532 pp
Liou K-N (2002) An introduction to atmospheric radiation, vol 84, 2nd edn. International Geophysics Series, Academic, San Diego, California, 583 pp
Madronich S (1987) Photodissociation in the atmosphere. 1. Actinic flux and the effects of ground reflections and clouds. J Geophys Res 92:9740–9752
Mishchenko MI, Travis LD, Mackowski DW (1996) T-matrix computations of light scattering by nonspherical particles: a review. J Quant Spectrosc Radiat Transf 55:535–575

Chapter 6
In Situ and Remote Sensing Measurements of Aerosols

Abstract This chapter builds on the radiative transfer chapter to explain the foundations for aerosol remote sensing in the atmosphere. It categorizes the various passive remote sensing techniques according to the process involved (scattering, extinction or emission), the viewing geometry and the wavelength considered. The principles for aerosol remote sensing from the ground using sunphotometry and from space are then introduced with a highlight on the importance of characterizing the surface for spaceborne retrieval. This is complemented by a discussion of active remote sensing using the lidar and Raman lidar techniques. A range of instruments for in situ measurements are then presented such as the cascade impactor, particle counters, differential mobility analyzers or the single particle soot photometer.

Keywords Remote sensing · Aerosol retrieval · Scattering · Particle counter · Impactor · Aerosol measurement

Aerosol measurement is a very broad topic that would require a textbook in itself. This chapter is therefore limited to a presentation of the principles of aerosol *remote sensing* and in situ measurements, but also provide concrete examples of measurement techniques when possible.

The choice was made here to present well-established techniques, and only occasionally more novel or less well-established techniques, when such techniques have the potential to provide complementary information on aerosols. A rather exhaustive set of references is also provided so that the interested reader can go back to the original description of a given technique. A number of satellite instruments are mentioned, among which some sensors that have been specifically designed for the characterization of the aerosols, and other sensors for which aerosols may not have been the primary target but an application to aerosols has been found.

6.1 Introduction to Aerosol Remote Sensing

The concept of *remote sensing* includes all measurements that are made remotely, that is at some distance of the object being observed. The information between the

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_6

observable and the *observer* is carried out by electromagnetic waves (Fig. 5.1)[1]. We often refer to the aerosol influence on the radiation as the *signal* that we seek to isolate from other influences. If the electromagnetic radiation is initially emitted artificially, the remote sensing is said to be *active*, with the best known applications being the *lidar*[2] in the optical domain and the *radar*[3] in the radiowave domain. In the opposite case where the observation is based on natural electromagnetic radiation, for instance shortwave radiation emitted by the Sun or longwave and microwave radiation emitted by the Earth's system, the remote sensing is said to be *passive*.

The distinction between shortwave and longwave radiation introduced previously remains useful when describing and categorizing different aerosol remote sensing techniques. Atmospheric aerosols being relatively small in size, there is no benefit of using radiation with longer wavelength, such as radiowaves, to which aerosols are largely transparent. Radiative processes of importance, when it comes to aerosol remote sensing, are scattering and absorption of solar radiation, and scattering, absorption and emission of terrestrial radiation. These considerations make it relevant to classify passive remote sensing methods in three categories that rely respectively on (1) extinction of solar radiation, (2) scattering of solar radiation, and (3) absorption and emission of terrestrial radiation. More sophisticated approaches can make a combined use of these methods.

Remote sensing methods that rely on the measurement of the extinction of the incident solar radiation can take place at the ground using sunphotometry (Fig. 6.1a) and from satellite using occultation techniques (Fig. 6.1b). This method is well suited to retrieve the extinction aerosol optical depth along a given path. Observations of the diffused component of the solar radiation form the basis of a wide variety of methods that measure scattered radiances from the ground (Fig. 6.1c and d) or from space by pointing either downward (Fig. 6.1e) or in a direction tangent to the upper atmosphere (Fig. 6.1f). Those are termed nadir and limb observations, respectively. Methods that rely on the observations of solar radiation are the most numerous and the most popular because they are very sensitive to the presence of aerosols. We will also show that there is also information on the aerosol in the polarized state of the reflected solar radiation.

Methods based on the emission of terrestrial radiation and its measure in the atmospheric window are useful to detect the presence and properties of desert dust whose quantities emitted and transported in the atmosphere can be substantial (Fig. 6.1g) and stratospheric aerosols in particular through limb measurements (Fig. 6.1h).

The different techniques are summarized in Table 6.1.

---

[1] Remote sensing can also use other types of waves, such as acoustic waves or gravity waves. However for the purpose of observing aerosols, only electromagnetic waves are used, hence the somewhat restrictive definition provided here.

[2] Acronym for *LIght Detection And Ranging*.

[3] Acronym for *RAdiowave Detection And Ranging*. Aerosols being transparent to radiowaves, radar is not a suitable technique to measure them, which is why this technique is not discussed further.

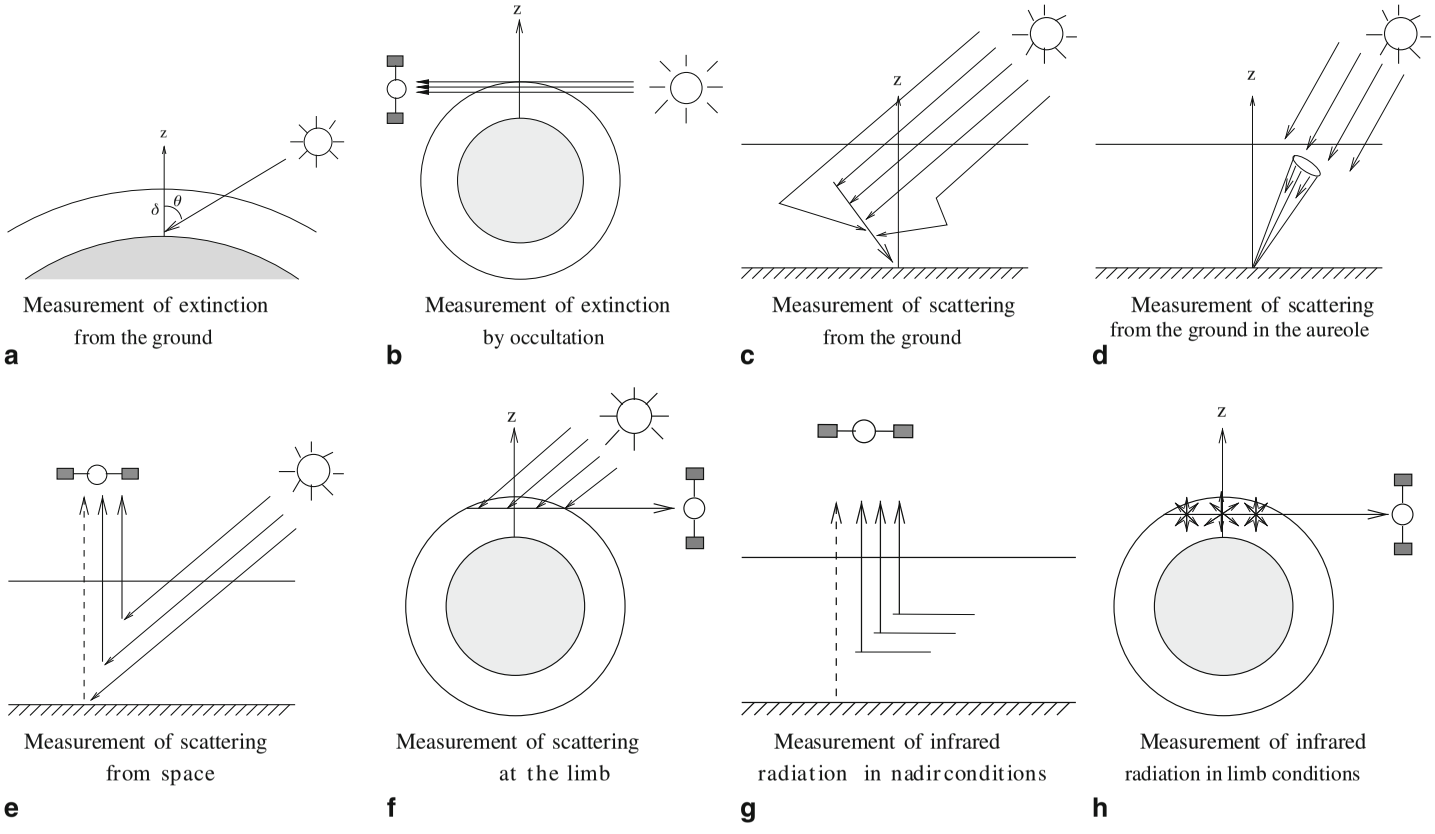


Fig. 6.1 Schematic of various passive remote sensing techniques for the observation of atmospheric aerosols. Techniques **a**–**f** rely on the measurement of solar radiation, while techniques **g**–**h** rely on the measurement of infrared radiation. The *arrows* that point towards the satellite indicate the type of radiation that is measured. The *dashed arrows* indicate contributions that need to be corrected for measuring aerosol properties

Table 6.1 Summary of the various passive remote sensing techniques and relevant satellite instruments and missions for the observation of atmospheric aerosols. Some key references are also provided for each technique. A few forthcoming instruments are indicated in italic. Refer to the acronym table at the beginning of this book for the meaning of the acronyms

Remote sensing principle	Section	Examples of instruments	References
Measurement of the extinction of solar radiation from the ground	Figure 6.1a Section 6.2.2	AERONET, PHOTONS	Shaw (1983) Holben et al. (1998) Dubovik et al. (2000)
Measurement of the extinction of solar radiation from space	Figure 6.1b Section 6.2.3	SAGE, OSIRIS, GOMOS, MAESTRO	Kent et al. (1991)
Measurement of the scattering of solar radiation from the ground	Figure 6.1c, d Section 6.3.2	AERONET, PHOTONS	Twitty (1975); Nakajima et al. (1983); Dubovik et al. (2000)
Measurement of the scattering of solar radiation from space	Figure 6.1e Section 6.3.3 Figure 6.3	AVHRR, MERIS, MODIS, MISR, SCIAMACHY, POLDER, ATSR, AATSR, SEVIRI, PARASOL, VIIRS, *3MI*	Rao et al. (1989); Herman et al. (1997); Remer et al. (2005); Tanré et al. (2011)
Measurement of infrared radiation at nadir	Figure 6.1g Section 6.4.2	METEOSAT, SEVIRI, AIRS, IASI, *IASI-NG*	Ackerman (1997); Legrand et al. (1989); Pierangelo et al. (2004), (2013)
Measurement of infrared radiation at the limb	Figure 6.1h Section 6.4.3		Thomas et al. (1983)

The radiation received at the surface is almost exclusively a signature of the atmosphere, while the radiation received by a satellite instrument, exception made of limb measurements, contains both surface and atmospheric contributions. In the latter case, the measurement strategy will be to look for observation scenes for which the surface contribution is either small or easy to model and correct. The surface contribution can then be subtracted from the observed radiance to isolate the aerosol signal. As we are interested in aerosol properties, it is appropriate to measure radiation in wavelengths corresponding to atmospheric windows (0.35–0.80 μm in the visible and 8–12 μm in the infrared). Atmospheric windows in the visible and near infrared are illustrated on Fig. 6.2.

Generally speaking remote sensing is an inverse problem, which is more difficult to solve than a direct radiative transfer calculation in an atmosphere whose chemical composition is known. Different inversion techniques have been developed but many share similar characteristics. To be successful an inversion technique has to be both physically sound and mathematically robust. We emphasize here the physical basis of the aerosol inversion techniques and the reader is directed to other textbooks to gain more understanding of the mathematical tools used in remote sensing.

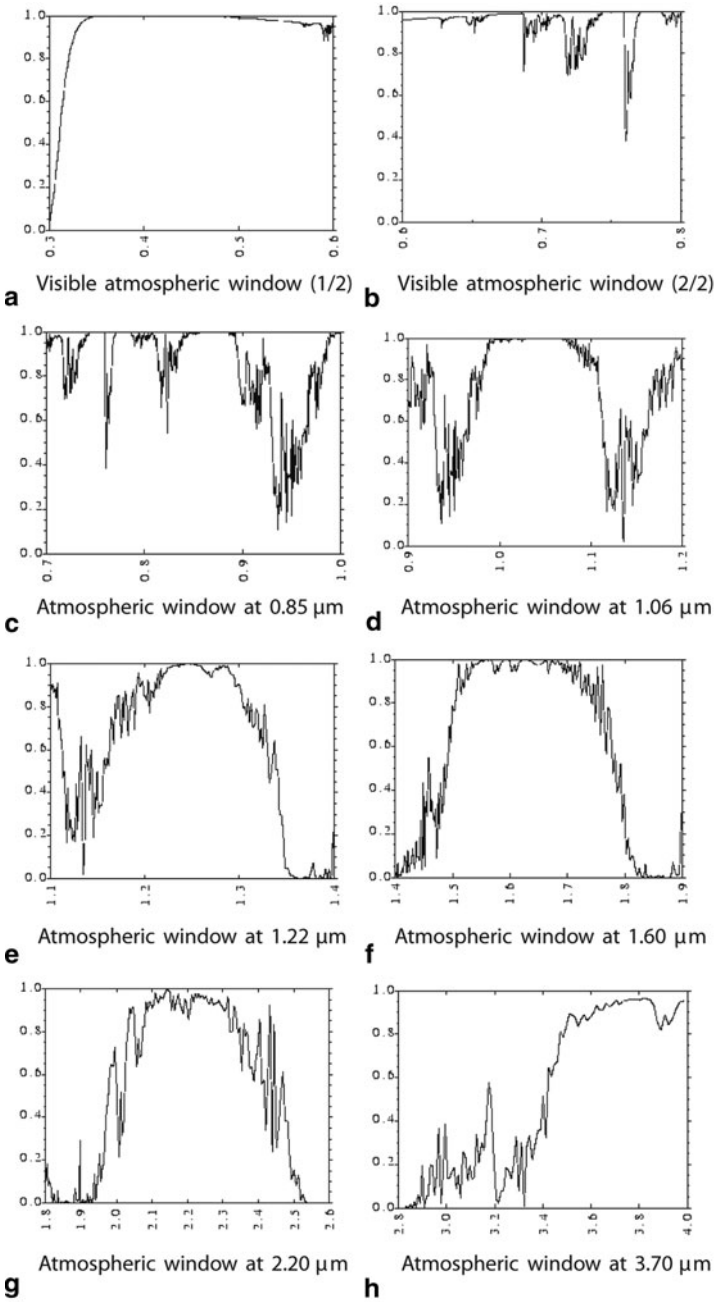


Fig. 6.2 Transmittance along a vertical optical path for a standard atmosphere (referred to as US62) as a function of wavelength (in µm). The simulation has been performed with the 6S radiative transfer model (Vermote et al. 1997). Atmospheric windows correspond to spectral intervals where the transmittance is close to unity

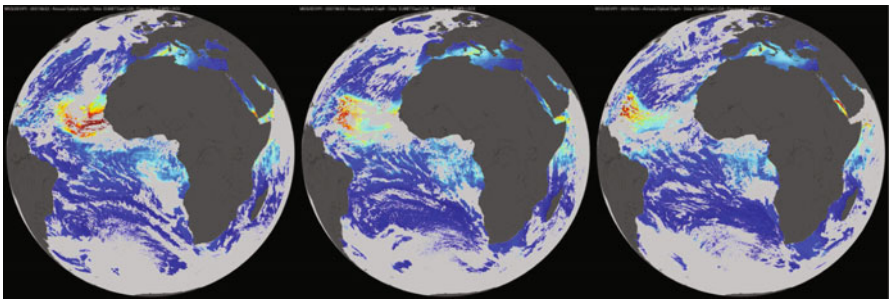


Fig. 6.3 Desert dust plume across the Atlantic Ocean on 22, 23, and 24 June 2007. The graphic shows the aerosol optical depth at 550 nm as retrieved from the SEVIRI instrument onboard the METEOSAT geostationary satellite. (© ICARE/LSCE)

6.2 Passive Remote Sensing: Measurement of the Extinction

6.2.1 General Principles

Measuring the extinction caused by the aerosols forms the basis of a technique that can be applied either from the ground, with a sunphotometer, or from space, with a spectrometer. A rather large range of wavelengths, from ultraviolet to infrared radiation, can be used depending on the amount and type of aerosols considered.

Solar radiation measured by the photometer or spectrometer can be simply written as:

$$E_\lambda = f_\lambda \ \exp(-\tau_\lambda) = f_\lambda \ \exp(-\tau_\lambda^{\nu}/\cos(\theta_0)) \tag{6.1}$$

where f_λ is the solar irradiance at the top of the atmosphere, τ_λ is the total extinction optical depth due to molecules, aerosols, and clouds along the slanted optical path, θ_0 the solar zenith angle, and τ_λ^{ν} the optical depth on the vertical. In practice, the measure of extinction is performed in different channels that correspond to narrow spectral bands. A first step is to eliminate measurement scenes that are contaminated by clouds. This is usually done through a number of tests that apply some threshold on the measured radiances and screen the scenes for those the temporal consistency of the measurements is low. The former test is because aerosol layers are usually less dense than cloud layers. The latter test is because the properties of clouds generally vary on much shorter timescales than those of aerosol layers that tend to be more homogeneous in space and time. The effect of scattering from molecules can then be corrected if the surface pressure is known. If the measurement is made at a wavelength where there is absorption by a trace gas, then the gaseous absorption optical depth must also be corrected for.

6.2.2 Ground-Based Photometry

The principle of sunphotometry from the ground is illustrated on Fig. 6.1a. There is a long history of such measurements with the pioneering work of Pierre Bouguer whose initial objective was to measure the *turbidity* or haziness of the atmosphere (Shaw 1983). Sunphotometry is now a well-established technique that can provide high quality aerosol optical depth measurements with an absolute accuracy better than 0.001. In the case of ground-based measurements with a solar zenith angle θ_0, the optical depth τ of Eq. 6.1 is linked to the optical depths on the vertical due to molecules, τ_m, and aerosols, τ_a :

$$\tau = (\tau_m + \tau_a)/\cos\theta_0 = m\,(\tau_m + \tau_a) \tag{6.2}$$

where $m = 1/\cos\theta$ is the air mass or air mass factor. It should be noted that this expression is only valid for solar zenith angles that are not too large. When the sun is low on the horizon, it is important to account for the index of refraction of the atmosphere that deviates from unity and bends slightly the incident solar radiation. In the absence of molecular absorption, the knowledge of surface pressure is sufficient to estimate τ_m as a function of the wavelength and retrieve the aerosol extinction optical depth τ_a at several wavelengths in the atmospheric windows. The inversion gets more complicated if the measurement is performed in a channel which is sensitive to molecular absorption by gases such as ozone or water vapour. If absorption is weak, it is enough to assume a typical vertical profile for the absorbing gas and make a correction to the observed extinction optical depth. If the absorption is more substantial, it is necessary to use additional (ancillary) information or incorporate more channels in the retrieval to estimate simultaneously the aerosol optical depth and the gaseous absorption.

In reality sunphotometers measure a numerical count $V(\lambda)$, and the equation takes the form:

$$V(\lambda) = V_0(\lambda)\ \exp(-m\ \tau_\lambda) \tag{6.3}$$

where $V_0(\lambda)$ is the numerical count that would be measured at the top of atmosphere. The inversion of Eq. 6.1 thus requires the instrument to be well calibrated, that is to effectively measure $V_0(\lambda)$ if it was placed at the top of atmosphere. Calibration is therefore critical, especially for observations made within a network, and for instruments located in background regions where a small error on the calibration translates into a large relative error on the aerosol optical depth. Calibration is usually done in a "clean" location and/or in altitude. It consists in measuring the numerical count $V(\lambda)$ for various air mass factors (i.e. at different times of the day). The quantity $V_0(\lambda)$ can then be estimated from a linear regression of $V(\lambda)$ against the air mass factor $m = 1/\cos\theta$, which is known as the *Bouguer–Langley method*. Calibration can also be performed in an absolute way using a master instrument traceable to a standard. As this is a heavy and long procedure that requires bringing all instruments to the laboratory, it is usual to calibrate one or several instruments in an absolute way and calibrate other instruments against these.

There exist several networks of sunphotometers in the world. The AERONET[4] network from NASA, coupled to the French network PHOTONS[5], has a wide coverage and provides accurate raw measurements and high quality inversions in cloud-free conditions. Other stations contribute to the Global Atmospheric Watch programme of the World Meteorological Organization[6], including a network of precision filter radiometers. These networks play a major role in the monitoring of atmospheric aerosols, the evaluation of satellite aerosol measurements and our understanding of the regional variations in aerosol properties.

6.2.3 Spaceborne Occultation Measurements

The measurement principle behind Eq. 6.1 can also be used to retrieve aerosol optical depth from space in limb geometry by pointing the instrument towards the Sun (see Fig. 6.1b). Successive measurements performed as the satellite moves along its orbit sample an increasing or decreasing range of heights, which allows the retrieval of the vertical profile of the aerosol extinction coefficient under the assumption that aerosol concentrations and properties do not vary much along the horizontal. This technique is appropriate to measure aerosols in the stratosphere but cannot penetrate much in the troposphere. This is because the optical path increases rapidly as the instrument scans lower down in the troposphere, which also increases the probability of the line of sight being contaminated by the presence of a cloud.

The solar occultation measurement technique has been used by the series of SAGE instruments onboard the SAM II, SAGE I, SAGE II, and SAGE III missions, which have measured the vertical profile of stratospheric aerosols for several decades. This series of mission is now discontinued but similar measurements are conducted with the MAESTRO instrument. The same principle can be used by pointing towards stars instead of the Sun if the spectrometer has enough sensitivity, as it is the case with the GOMOS instrument.

6.2.4 Retrieval of Aerosol Size Distribution

The spectral variation of aerosol optical depth contains information on the aerosol size distribution because the smallest particles (a few tenths of a micrometre in radius) scatter solar radiation more effectively than the largest particles (a few micrometres in radius). The optical depth of an aerosol population with homogeneous properties

[4] http://aeronet.gsfc.nasa.gov.

[5] http://loaphotons.univ-lille1.fr.

[6] http://www.gaw-wdca.org.

on the vertical filling a layer with thickness Δz is:

$$\tau_a(\lambda) = \Delta z \int \pi r^2 \, Q^{\text{ext}}(r, \lambda, m) \, n(r) \, \mathrm{d}r \tag{6.4}$$

where $Q^{\text{ext}}(r, \lambda, m)$ is the extinction factor—function of the aerosol radius, r, the wavelength, λ, and the complex refractive index, m, itself a function of λ—and $n(r)$ is the size distribution that we seek to inverse. This type of equation, where the unknown is under the integral sign, is a Fredholm equation of the first type and can be solved mathematically. If $\tau_a(\lambda)$ is measured at several wavelengths in the visible and near-infrared domain, an approximative aerosol size distribution can be inverted (King et al. 1978). The inversion does not allow to retrieve particles with a radius larger than 3–4 μm. A more accurate inversion of the aerosol size distribution requires to exploit the information contained in the scattered solar radiation rather than just in the extinction, which is the subject of the next subsection.

6.3 Passive Remote Sensing: Measurement of the Scattering

6.3.1 General Principles

The aerosol phase function depending strongly on the aerosol size distribution, it is judicious to use the information contained in the angular distribution of the scattered radiation to estimate the aerosol physical properties. Assuming single scattering (i.e. neglecting multiple scattering) and a black surface (i.e. there is no radiation reflected by the surface), the radiance scattered in a direction making an angle Θ with the incident solar radiation can be written:

$$L_\lambda(\Theta) = \frac{\varpi_\lambda \, \tau_\lambda \, P_\lambda(\Theta)}{4 \, \pi \, \mu} f_\lambda \tag{6.5}$$

where ϖ_λ, τ_λ, and $P_\lambda(\Theta)$ are the single scattering albedo, the optical depth and the phase function of the mixture of molecules and aerosols at the wavelength λ, and μ is the cosine of the viewing zenith angle. The aerosol phase function can be approximated from the measure of the radiance corrected from the effect of molecules. The equation to be solved now becomes:

$$P_\lambda^{\text{aerosol}}(\Theta) = \int P(\Theta, r, m) \, n(r) \, \mathrm{d}r \tag{6.6}$$

where $P(\Theta, r, m)$ is the phase function of a particle with radius r and refractive index m. The equation above is another Fredholm equation of the first type. In practice, the inversion is complicated by the fact that it is necessary to account for multiple scattering, molecular absorption, contributions from the surface, and to some extent by the aerosol vertical profile, at least in the shorter wavelengths of solar radiation.

6.3.2 Ground-Based Measurement of Scattered Radiation

There are several advantages to measure the spectral dependence of the scattered solar radiation from the ground. First, it is possible to sample a large range of scattering angles from very small angles in the aureole (Fig. 6.1c) to large scattering angles when the Sun is low on the horizon. Second, the surface contribution to scattered (diffuse) downward radiation is generally weak even when the surface albedo is relatively large. This is why most sunphotometers discussed earlier in this chapter also include sky radiance measurements in addition to direct measurements pointing in the direction of the Sun. Measurements are generally performed in the *almucantar*, that is to say in a cone of zenith angle equal to the solar zenith angle. Some instruments are capable of measuring with high accuracy the scattered radiation in a small cone around the direction of the Sun, a region called the *aureole* which is characterized by small scattering angles and a high sensitivity to large aerosol particles. Retrieval algorithms for sunphotometers, such as those used for the AERONET network, combine the information from the Sun (i.e. direct) and sky (i.e. diffuse) radiances at several wavelengths. This is an effective way to retrieve an aerosol size distribution and an aerosol single scattering albedo that are representative of the aerosol burden when the aerosol optical is typically larger than 0.4 (Dubovik et al. 2000, 2002).

6.3.3 Spaceborne Measurements of Scattered Radiation

6.3.3.1 Contribution from the Surface

Aerosol remote sensing from space is complicated by the fact that the Earth's surface contaminates the atmospheric signal that contains an aerosol signature. Figure 6.4 shows in a schematic way how the satellite signal can be decomposed between an atmospheric signal, a contribution from the surface, and other contributions due in particular to the presence of clouds. The latter contributions are generally minimized by screening out not only cloudy observation scenes but also clear-sky scenes that are located nearby cloudy scenes. Regarding the surface contribution, the challenge is to choose viewing geometries and wavelengths for which this contribution can be minimized or to build a surface reflectance model that is accurate enough so that the effect of the surface reflectance can be corrected for in the observed satellite signal.

In the case of a Lambertian surface that reflects incoming solar radiation isotropically, it can be shown that the reflectance coming out of the atmosphere can be written as:

$$\rho^*(\theta_0, \theta_v, \varphi_v) = \rho_a(\theta_0, \theta_v, \varphi_v) + \frac{\rho}{1 - \rho\, S}\, T(\theta_0)\, T(\theta_v) \tag{6.7}$$

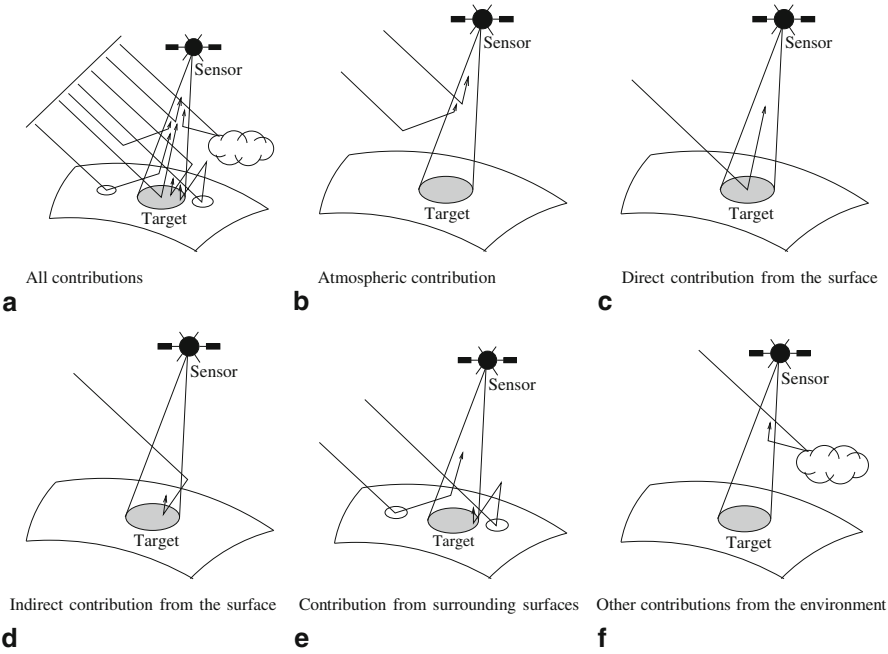


Fig. 6.4 Schematic description of the various contributions to the solar radiances measured by a spaceborne radiometer. (Adapted from Tanré et al. (1979) for spaceborne remote sensing)

where ρ_a is the intrinsic radiance of the atmosphere expressed in the form of a reflectance, ρ is the surface reflectance, S is the spherical albedo, $T(\theta)$ is the total atmospheric transmittance, θ_0 and θ_v are the solar and viewing zenith angles and φ_v is the difference in azimuthal angle between incident and reflected solar radiation. The information on aerosol is essentially contained in the first term of the right-hand side of the equation even though aerosols also affect the atmospheric transmittances.

6.3.3.2 Analysis of the Atmospheric Signal

Neglecting multiple scattering in the atmosphere, the equation of radiative transfer (without any emission source) can be simplified into:

$$\frac{\mu \, \mathrm{d}L_1(\tau, \mu, \varphi)}{\mathrm{d}\tau} = L_1(\tau, \mu, \varphi) - \frac{\varpi}{4\pi} \, \mathrm{p}(\tau, \mu, \varphi, \mu_0, \varphi_0) \, f_\lambda \, \exp(\tau/\mu_0) \qquad (6.8)$$

where L_1 is the radiance for the first order of scattering, τ is the optical depth which is used as the vertical coordinate with the top of atmosphere as origin, ϖ is the single scattering albedo p is the phase function, f_λ is the solar spectral irradiance and μ is the cosine of the zenith angle counted negatively for downward radiation. If the surface is black (that is if the surface reflectance is zero), the equation can be solved

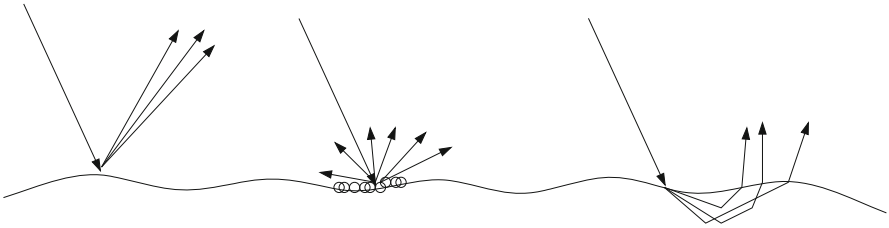


Fig. 6.5 Contributions to the ocean surface reflectance. From *left* to *right*: specular (or Fresnel) reflection, scattering by ocean foam, and underlight that comes from scattering of solar radiation entering the ocean

analytically and the radiance coming out of the atmosphere is obtained as:

$$L^{\uparrow}_{\tau=0} = \frac{\varpi\ \tau\ P(\Theta)\ f_{\lambda}}{4\ \pi\ \mu_{v}} \tag{6.9}$$

which gives in terms of reflectance:

$$\rho_{a} = \frac{\pi\ L^{\uparrow}_{\tau=0}}{|\mu_{0}|\ f_{\lambda}} = \frac{\varpi\ \tau\ P(\Theta)}{4\ |\mu_{0}|\ \mu_{v}} \tag{6.10}$$

where Θ is the scattering angle. In practice, it is important to account for multiple scattering, molecular absorption and the (non-Lambertian) surface reflectance.

6.3.3.3 Modelling the Oceanic Reflectance

Over the oceans, the surface reflectance comes from specular reflection (also called Fresnel reflection), reflection by sea foam, and underwater light that is back scattered by water molecules and particles in suspension (Fig. 6.5). Specular reflection is the mirror-like reflection of solar radiation over the ocean and is responsible for a glint. It occurs around the direction $\theta_{v} = \theta_{0}$, $\varphi_{v} = \varphi_{0} + 180°$ and its angular extent depends on the direction of the waves at the surface of the ocean and the distribution of their slopes. As these parameters are difficult to predict or measure, aerosol retrieval will avoid viewing geometries that correspond to specular reflection and scenes where the presence of sea foam is too important. This is usually diagnosed from the near-surface (10-m) wind speed which is known as ancillary data from a numerical weather prediction model or from measurements made with a scatterometer. In other situations, the surface reflectance depends on the concentration of plankton in the surface ocean water (Fig. 6.6) and in the case of coastal waters the quantity of sediments in suspension. It can be seen from Fig. 6.6 that surface reflectances are smaller in the green (wavelengths 500–600 nm) than in the blue (wavelengths 400–500 nm).

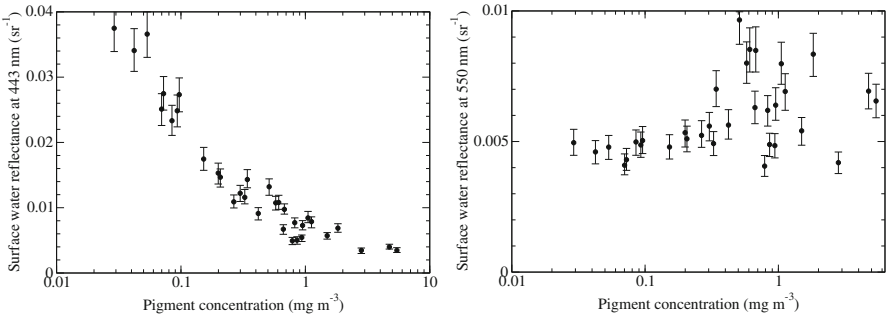


Fig. 6.6 Reflectances of the ocean surface as a function of the phytoplankton pigment concentration (in mg m^{-3}). (From Gordon et al. (1988). © American Geophysical Union)

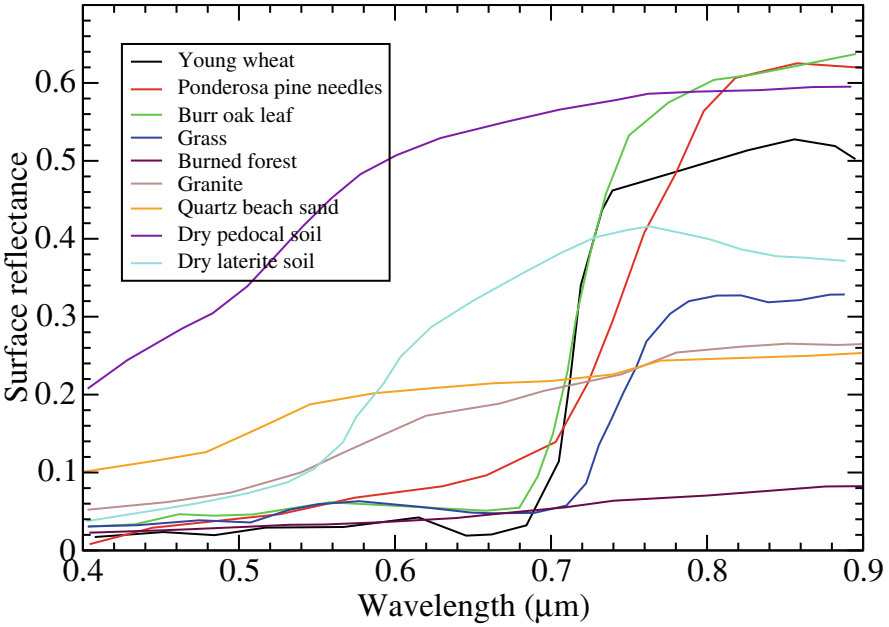


Fig. 6.7 Reflectance for different types of continental surfaces. (© 1992, IEEE. Reproduced with permission from Kaufman and Tanré (1992))

6.3.3.4 Modelling the Continental Surface Reflectance

The situation is different over terrestrial surfaces and several techniques exist to minimize or correct for the impact of the surface on the measured satellite signal. Some techniques seek to select the darkest targets, such as lakes or vegetation, especially forested areas, around 650 nm and at wavelengths less than 500 nm (Fig. 6.7). The retrieval principle is thus similar to that over oceanic surfaces (Soufflet et al. 1997).

Other techniques consist in measuring radiances in several viewing geometries and assuming a particular model for the shape of the bidirectional reflectance function (often abbreviated BDRF). The BDRF can then be scaled up or down to fit the measurements and the surface albedo can be retrieved simultaneously with the aerosol properties. This can be done with a dual view instrument such as ATSR and AATSR (North et al. 1999) or with a multi-angular instrument such as MISR (Diner et al. 2005), POLDER (Deuzé et al. 2000, 2001) or PARASOL (Tanré et al. 2011). A variant to this technique that can be applied to geostationary satellites consists in using a number of consecutive measurements during the day while assuming the aerosol optical depth and aerosol properties are constant or vary slowly during that period (Govaerts et al. 2010; Wagner et al. 2010). The consecutive measurements of a particular pixel share the same viewing zenith and azimuthal angles but correspond to different solar zenith angles.

It is also possible to make the assumption of a constant ratio between surface reflectances at two different wavelengths. Kaufman and Tanré (1996) proposed to use the surface reflectance at 2.2 μm (where accumulation mode aerosols are almost transparent) in order to estimate the surface reflectance in the visible. A last technique consists in constructing a surface bidirectional reflectance model from measurements made on days where the aerosol concentrations are smallest and use these surface properties for the retrieval on aerosol-laden days. As surface properties vary with the season—-in particular because of the seasonal evolution of the vegetation—-these properties need to be updated regularly for the retrieval to be correct.

6.3.3.5 Use of Polarized Radiances

The analysis of the polarized reflected solar radiances provides an interesting, yet little used, complement to the above-mentioned techniques. As discussed in Sect. 5.6.7, radiation emitted by the Sun is not polarized but gets partly polarized upon reflection by the surface and scattering by molecules and aerosols. The contribution of the surface to the polarized reflectance can be easier to correct for than for the total reflectance, especially over land. The measure of the polarized radiances in the solar spectrum thus brings information on the amount and properties of accumulation mode aerosols (Herman et al. 1997; Mishchenko and Travis 1997). It can also be used to differentiate spherical from nonspherical coarse mode aerosols. Figure 6.8 shows the accumulation mode aerosol optical depth as retrieved by the POLDER instrument which exploits the information content of polarized radiances at the 670 and 865 nm wavelengths.

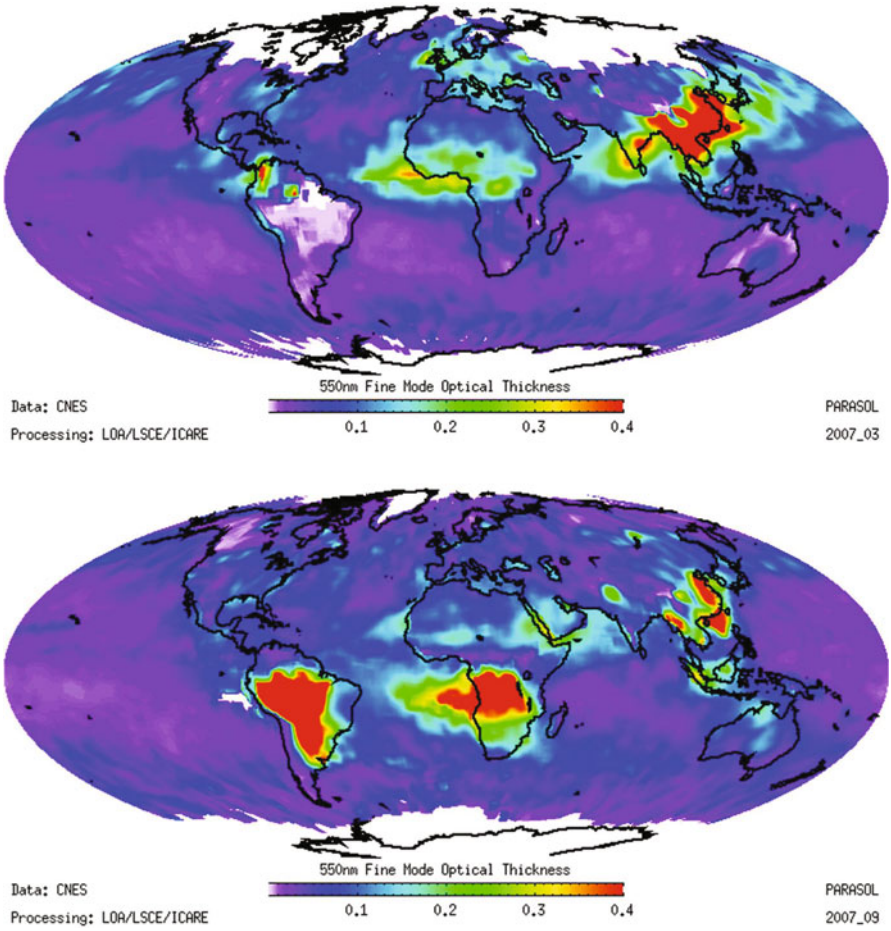


Fig. 6.8 Optical depth of the aerosol accumulation mode as retrieved from the POLDER instrument onboard the PARASOL satellite for the months of March 2007 and September 2007. (© CNES/LOA/LSCE/ICARE)

6.4 Measurement of Infrared Radiation

6.4.1 General Principles

Remote sensing in the infrared spectrum exploits the fact that the atmosphere absorbs and emits infrared radiation in different amounts along the vertical. Outside the atmospheric window, the atmosphere is opaque so that the energy radiated by the surface is typically absorbed and re-emitted many times by greenhouse gases and clouds and to a fairly small extent by aerosols. Inside the atmospheric window

and in the absence of clouds and aerosols, a significant part of the energy radiated by the surface go through the atmosphere and can escape to space. As clouds and aerosols can absorb and re-emit such radiation, the outgoing infrared radiation contains information on the amount and properties of clouds and aerosols. In fact the outgoing infrared radiation measured by a satellite instrument will come from a range of altitudes depending on the wavelength, which allows a vertical profiling of the atmosphere, at least for temperature and humidity, but much less so for clouds and aerosols.

Infrared radiometers measure a radiance in spectral bands (also called channels) which can be large (broadband radiometer) or narrow (narrow band radiometer, infrared sounder, Fourier interferometer). For the sake of simplicity, the radiance B_λ is expressed by means of a brightness temperature $T_b(\lambda)$ which corresponds to the temperature of a black body emitting the same amount of radiation at that wavelength. The measured brightness temperature therefore verifies:

$$B_\lambda = \frac{2\,h\,c^2}{\lambda^5\,\left[\exp\left(h\,c/k_B\,T_b(\lambda)\,\lambda\right)-1\right]}. \tag{6.11}$$

6.4.2 *Spaceborne Nadir Measurement of Infrared Radiation*

The atmosphere does not behave as a black body however. Clouds and aerosol layers are also imperfect black bodies. The degree to which clouds and aerosols behave as a black body depends on the concentration of water droplets, ice crystals, as well as on the spectral properties of liquid water, ice water or the aerosol material. The observed brightness temperature at the top of atmosphere is therefore a function of the wavelength. In any case the brightness temperature is smaller for an aerosol-laden atmosphere as compared to a clear atmosphere.

New generation infrared sounders, such as AIRS and IASI, can be used to retrieve more detailed and more accurate information on desert dust. Once the temperature and humidity profiles are retrieved, other channels can be used to retrieve a typical dust aerosol size distribution and altitude. This is done by minimizing a cost function that depends on these parameters assuming the dust refractive index in the longwave spectrum is known (Pierangelo et al. 2004; Peyridieu et al. 2013).

Radiances and Reflectances

Both radiances and reflectances are used in the scientific literature and it is important to link the two quantities. Let $\Phi = \mathrm{d}Q/\mathrm{d}t$ be the radiative flux, i.e. the radiative energy emitted, transported or received per unit time. The radiance L is defined as

$$L = \frac{\mathrm{d}^2\Phi}{\mathrm{d}\omega\,\mathrm{d}\Sigma}$$

where $\mathrm{d}^2\Phi$ is the radiative flux across a surface $\mathrm{d}\Sigma$ perpendicular to the radiation beam in a solid angle $\mathrm{d}\omega$. The radiance is therefore expressed in units of $\mathrm{W\,m^{-2}\,sr^{-1}}$. The angular distribution of L (and its polarization state) at a given point defines the radiative field at that point. The radiative flux density on an oriented surface $\mathrm{d}S$ is given by

$$E = \int\int_{2\pi} L(s)\,\cos\theta\,\mathrm{d}\omega.$$

If the direction of propagation of the radiation is characterized by the angles (θ, φ), the elementary solid angle can be written as $\mathrm{d}\omega = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi$ and the flux density becomes

$$E = \int_0^{2\pi}\int_0^{\pi/2} L(\theta,\varphi)\,\cos\theta\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi.$$

It should be noted that in the case of a isotropic radiation field with a uniform radiance L_0, the previous equation simplifies to:

$$E = \pi\,L_0.$$

These notations can equally be applied to an emitted and receiving surface $\mathrm{d}S$. In the same way a spectral radiance (in units $\mathrm{W\,m^{-2}\,sr^{-1}\,\mu m^{-1}}$) can be defined.

Bidirectional reflectances are defined from radiances by

$$\rho(\theta_0,\theta,\varphi) = \frac{\pi\,L(\theta_0,\theta,\varphi)}{E_0\,\cos\theta_0}$$

where θ_0 is the solar zenith angle and E_0 is the solar spectral irradiance. It should be noted that a bidirectional reflectance can take values larger than unity in some directions.

6.4.3 Spaceborne Limb Measurement of Infrared Radiation

Spaceborne measurement of infrared radiation at the limb is an uncommon technique to retrieve atmospheric aerosols but is mentioned here for completeness. The technique can work in the case of large concentrations of stratospheric aerosols (Fig. 6.1h; Thomas et al. 1983).

6.5 Active Remote Sensing: Lidar

6.5.1 General Principles

The principle of the lidar measurement is to emit a short light pulse with a laser and to measure the intensity (and sometimes the wavelength shift and the polarization) of the light backscattered by the atmosphere (i.e. molecules, aerosols, and hydrometeors). The so-called *backscattered signal* is recorded and analyzed as a function of time after the laser pulse, so that the altitude of the measurement can be estimated with precision given the speed of light in the atmosphere. The lidar can be used at the surface in which case it is pointing upwards to the atmosphere, from a satellite in which case it is pointing downwards, or from an aircraft where it can point downwards or upwards depending on what part of the atmosphere is targeted. However the lidar does not have to point vertically: some surface-based lidar systems can be used in zenith or slant-angle mode. Lidars usually use wavelengths in the ultraviolet to infrared part of the spectrum, with most aerosol lidars operating at visible to near-infrared wavelengths where the aerosol optical depth is large enough. The lidar is therefore a unique technique to retrieve the vertical profile of atmospheric properties related to scattering and absorption by molecules and aerosols.

6.5.2 The Lidar Equation

The signal measured by the lidar can be written as:

$$LS(z,\lambda) = C_0(\lambda)\,\frac{A}{z^2}\,\beta(z,\lambda)\,\exp\left(-2\int_0^z \sigma(z',\lambda)\,\mathrm{d}z'\right) \tag{6.12}$$

where z is the distance to the lidar, $C_0(\lambda)$ is the intensity of the light emitted by a laser at wavelength λ, A is the telescope area, so that A/z^2 is the solid angle of the signal receiver, $\beta(z,\lambda)$ is the backscattered coefficient (in $\mathrm{m}^{-1}\,\mathrm{sr}^{-1}$) and $\sigma(z,\lambda)$ is the extinction coefficient (in m^{-1}). In the absence of clouds, the backscattered coefficient is equal to:

$$\beta(z,\lambda) = P_{180}\,\varpi\,\sigma(z,\lambda) = P^R_{180}\,\sigma_R + P^a_{180}\,\varpi_a\,\sigma_a \tag{6.13}$$

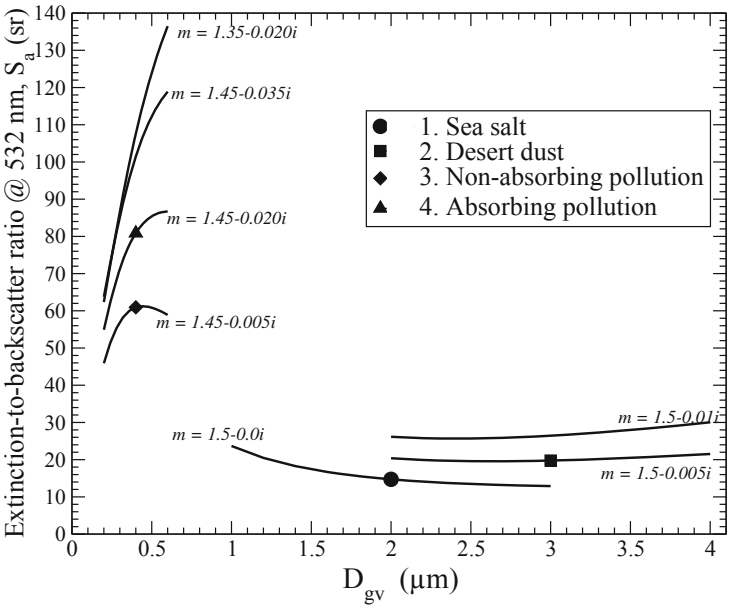


Fig. 6.9 Extinction-to-backscatter ratio (sr) as a function of the mean geometric volume diameter, D_{gv}, for different aerosol refractive indices. Four points corresponding to four typical aerosol types have been superposed on top of four lines where the aerosol size is varied. The size distributions are log-normal with a geometric standard deviation $\sigma_g = 2$

where the index R refers to the molecule scattering term (i.e. Rayleigh scattering) and the index a refers to the aerosol extinction term. The exponential term in Eq. 6.12 expresses the extinction of the light pulse along its round trip from the laser to the target and back to the telescope. Practically, the lidar signal has to be calibrated so that the lidar equation can be rewritten as:

$$LS(z,\lambda) = \frac{C_1}{z^2}\left(P^R_{180}\,\sigma_R(z,\lambda) + P^a_{180}\,\varpi_a\,\sigma_a(z,\lambda)\right) \times \exp\left(-2\int_0^z (\sigma_R(z',\lambda) + \sigma_a(z',\lambda))\,\mathrm{d}z'\right). \tag{6.14}$$

Equation 6.14 can be resolved step by step, the inversion of $\sigma_a(z,\lambda)$ between the lidar and altitude z being then available to retrieve the backscattering at the altitude $z + \mathrm{d}z$. This is only possible though if the value of the phase function at 180°, P^a_{180}, is known. This quantity, equal to the ratio between backscattering and total extinction, is a key parameter for resolving the lidar equation. Unfortunately it varies significantly with the aerosol type. Figure 6.9 shows such variations for the extinction-to-backscatter ratio (i.e. the inverse of P^a_{180}) as a function of aerosol size and refractive index. This parameter can be assumed using *a priori* knowledge on

aerosol type, either uniform or varying along the vertical, but the price to pay is a large uncertainty in the retrieved aerosol extinction coefficient. This uncertainty grows larger with the range of the retrieval. It is also possible to inverse the extinction-to-backscatter ratio when the lidar measurements are made at several wavelengths and/or along different directions that sample a same height for different optical paths (Althausen et al. 2000). As lidars operate in the solar spectrum, nighttime measurements are less noisy than daytime measurements.

As an illustration, Fig. 6.10 shows raw data from the CALIPSO spaceborne lidar along a vertical transect over the Mediterranean Sea during an episode of intense forest fires.

6.5.3 Raman Lidar

As discussed previously, it is difficult to solve the lidar equation without further information on the aerosol type. Raman lidars form a class of lidar systems that try to circumvent this issue by exploiting the Raman effect. This effect is an optical phenomenon discovered by Indian physicist Chandrashekhara Venkata Raman in 1928, also described in Sect. 5.3.4. Scattering of radiation is generally elastic, that is it occurs without any change in the wavelength of the radiation, as it is the case for Rayleigh scattering. The Raman effect, however, results from inelastic scattering by molecules during which the wavelength of the scattered radiation is slightly modified. The small shift in wavelength is associated with a small exchange of energy between the incident photon and an air molecule whose energy level is altered. The intensity of the Raman lines depends on the distribution of molecules between the various vibrational modes. The intensity of Raman scattering is very small in comparison to Rayleigh scattering, which necessitates a high-precision measurement if it is to be exploited in the aerosol retrieval.

The Raman lidar measures the radiation that has been backscattered with a wavelength shift, which allows the detection of the vertical profile of molecular species (N_2, O_2, and H_2O). The lidar equation for the Raman component of the lidar signal can be expressed as:

$$LSR(z,\lambda_r) = C_0(\lambda)\,\frac{A}{z^2}\,\beta_{\text{Raman}}(z,\lambda_r)\,\exp\left(-\int_0^z (\sigma(z',\lambda)+\sigma(z',\lambda_r))\,\mathrm{d}z'\right) \tag{6.15}$$

where β_{Raman} is the molecular backscatter coefficient corresponding to the Raman effect. The integral in the exponential corresponds to the original wavelength λ of the emitted radiation and to the modified wavelength λ_r of the returning radiation. As β_{Raman} is only affected by molecular scattering and not by aerosol extinction, it becomes possible to separate the molecular from the aerosol contributions. The Raman lidar thus allows measuring the aerosol backscatter and extinction independently, which makes a more accurate inversion of the lidar signal equation possible

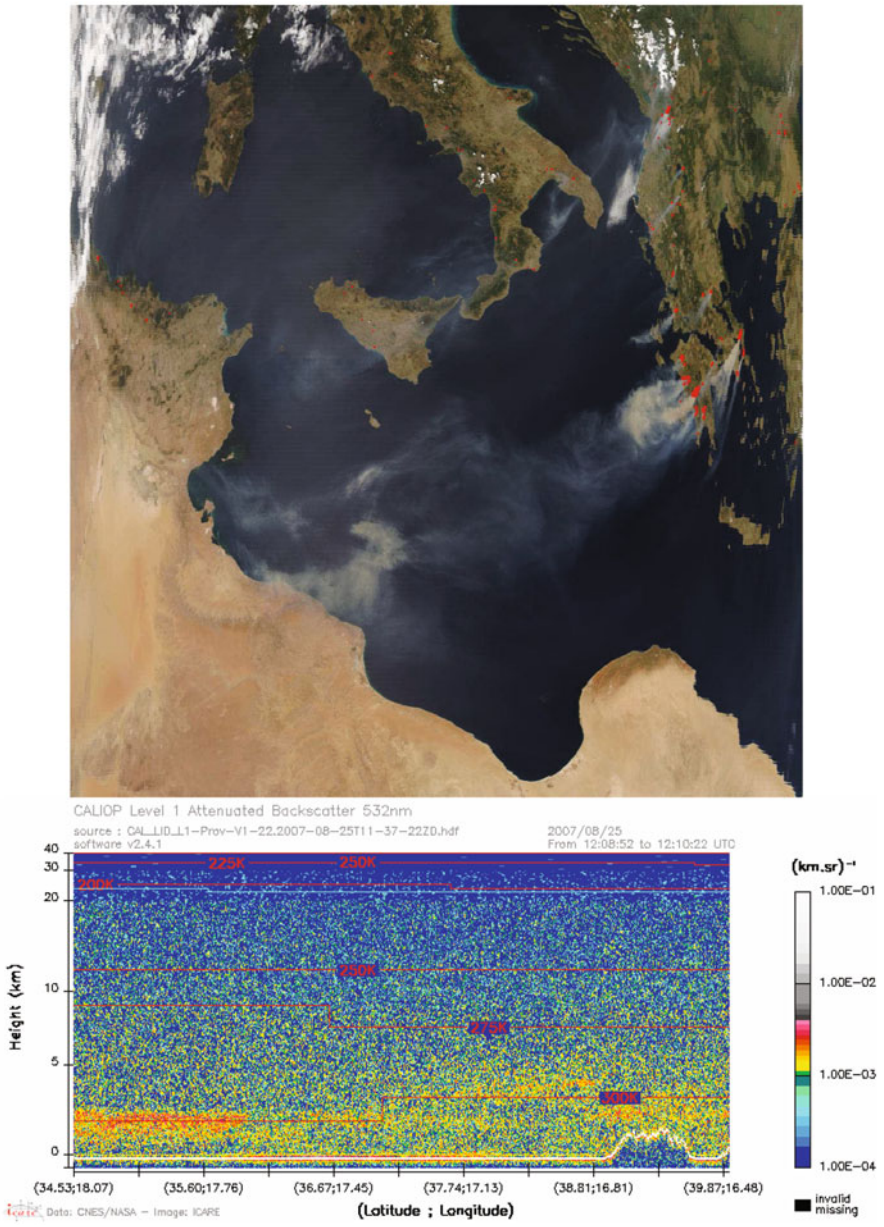


Fig. 6.10 Satellite images of August 2007 forest fires over the Mediterranean region. The top panel shows a colour composite from the MODIS instrument onboard the Aqua satellite with superposed fire pixels in red. The bottom panel shows the attenuated backscatter coefficient from the CALIOP lidar onboard the CALIPSO satellite along a vertical transect ranging from (34.5°N, 18.1°E) to (39.9 °N, 16.5 °N). (© NASA/ICARE)

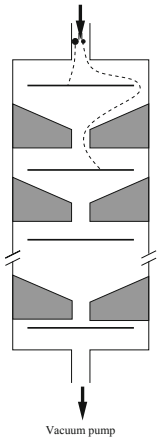


Fig. 6.11 Schematic description of a cascade impactor. The successive plates are located closer and closer to the nozzles where the air stream comes from. The trajectories of a small and a large particle are shown with *dashed curves*

(Ansmann et al. 1990; Ferrare et al. 1998). There exist different possible configurations for Raman lidars but they generally operate at several wavelengths with only one wavelength equipped with the capability to detect the Raman effect for one or several atmospheric molecules whose vertical profile is known.

6.6 In Situ Aerosol Measurements

This section is a brief introduction to in situ aerosol measurement techniques. Its aim is to provide a broad description of the principles of the main techniques without going into details of the instrumentation. As the size and concentrations of aerosols can be fairly small, and the range in aerosol sizes is very large, making accurate in situ measurements is always a challenge and requires a lot of high-precision instrumentation and know-how. We distinguish in what follows techniques that rely on the collection of aerosols on filters (thus requiring a rather long integration time to accumulate enough aerosol particles) from so-called *online techniques* that can sample the atmospheric aerosols on a continuous basis.

6.6.1 Measurement of Aerosol Concentrations

6.6.1.1 Cascade Impactor

A cascade impactor is an instrument that is used to measure the aerosol mass granulometry (i.e. mass size distribution). It relies on the fact that the larger the particle in an air stream, the larger the inertia because of its larger mass. A vacuum pump is used to force an aerosol-laden air stream into a device, called an impactor. It is characterized by several stages linked to each other through a pipe with a diameter that decreases as the air goes from the first to the last stages (see Fig. 6.11).

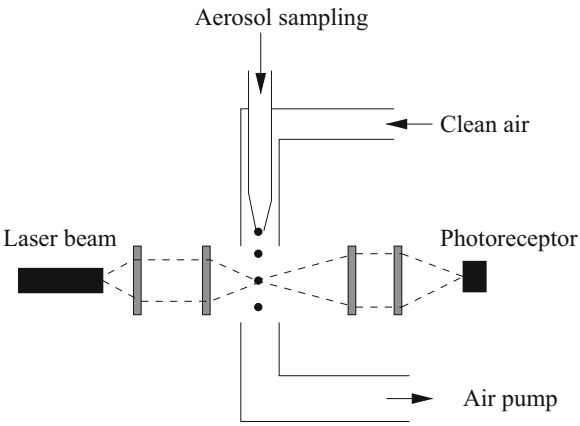


Fig. 6.12 Schematic description of an optical particle counter based on the detection of scattered light. There are various geometries and systems to focus the light emitted by the laser beam onto the particle and the light scattered by the aerosol particle onto the photoreceptor

At each stage the air stream is deviated by an impaction plate so that the heaviest (and largest) particles collide with the plate. The succession of the stages (typically from 4 to 10) enables to separate and class the impacted particles according to their aerodynamic diameter (i.e. the diameter of an equivalent spherical particle with a density of 1000 kg m^{-3} that would have the same aerodynamic behaviour in the air flow). The mass of aerosol particles impacted on the collection membranes placed at each stage provides a measure of the aerosol size distribution. A cascade impactor is a relatively simple instrument with a good accuracy if used appropriately. It can be coupled with measurement of the chemical composition of the bulk aerosol deposited at each stage.

6.6.1.2 Optical Particle Counter

An optical particle counter is an instrument that can measure the number of aerosol particles in a continuous flow that samples atmospheric air. The particles are diluted with clean filtered air so that the system can detect and count particles one by one (see Fig. 6.12). The particle counting can rely on the obscuration or on the scattering of a concentrated light beam. A particle counter can be coupled with a humidifier so that aerosol particles are humidified and grow to a sufficient size to be detected easily. The air flow has to be adjusted to the aerosol atmospheric concentrations.

Standard optical particle counters cannot measure particles much below 50 nm. Condensation particle counters are used to count smaller particles down to a size of a few nanometres. The principle is to control supersaturation to a value of 100–200 %, large enough to activate most aerosols, but below the supersaturation for homogeneous nucleation of water vapour which is about 300 %. Aerosols that get activated at such large supersaturations and can be measured by a condensation particle counter are called *condensation nuclei* or CN.

6.6.1.3 Differential Mobility Analyzer

The principle of a differential mobility analyzer is to separate charged aerosol particles according to their mobility in an electric field prior to optical detection. Aerosol particles are first sampled in the atmosphere and charged before being introduced in an electric field. Aerosol particles are accelerated by the electric field proportionally to their mass. Only particles within a small mass range can get out of the analyzer and be directed towards an optical counter. It is possible to span and measure the entire aerosol size distribution by varying the electric field. It is important for the technique to be effective to apply the same charge on each aerosol particle, which can be achieved with a charge neutralizer. Unfortunately the largest particles can retain several charges and it may be easier to remove them from the air flow with an impactor. For the smallest particles (with typical sizes of 10 nm and below), the Brownian motion of the aerosol particles in the analyzer need to be accounted for, which complicates the task of inverting the aerosol size distribution. Differential mobility analyzers can be used in tandem in order to measure the aerosol hygroscopicity. In such a setup, each one of the two differential mobility analyzers are preceded by a device to control the humidity in order to dry one aerosol line while keeping the second one moist. The growth factor, defined as the ratio of the moist and dry particle diameters, D/D_0, can be estimated as a function of particle size (by varying the electric field) and relative humidity.

6.6.2 Measurement of Aerosol Chemical Composition

6.6.2.1 Filter Measurements

Aerosols can be collected on filters and the bulk aerosol deposited on the filter can then be weighted. Chemical analysis of the aerosol particles deposited on the filters provides information on the chemical composition of an aerosol population. It is also possible to restrict the analysis to a range of aerosol sizes if the collection of aerosols on the filters is done using an impactor. Individual aerosol particles collected on filters can also be visualized and identified using an electronic microscope.

The total quantity of carbon in the aerosol can be estimated by volatilizing and burning the aerosol material deposited on a filter, and then by measuring the amount of CO_2 that is emitted. Generally the filter is first heated in the absence of oxygen, which leads to the volatilization of the organic carbon in the aerosols, before it is heated again in the presence of oxygen in order to volatilize the black carbon (or elemental carbon) which is more refractory. The filter can also be heated progressively in order to measure the aerosol carbon by class of volatility. It is important to ensure that the organic carbon does not oligomerize or polymerize during the heating process, as it would artificially increase the refractory character of the aerosol.

6.6.2.2 Aerosol Mass Spectrometry

Aerosol mass spectrometry is an online measurement technique whereby aerosol particles entering the system are bombarded with a laser, which breaks them down in small pieces that are then measured through standard mass spectrometry. By detecting the cations and anions that are present in the particles, aerosol mass spectrometry provides a semi-quantitative measurement of the aerosol chemical composition but does not provide information on the state of mixture.

6.6.3 Measurement of Aerosol Scattering

6.6.3.1 Passive Cavity Aerosol Spectrometer

The Passive Cavity Aerosol Spectrometer belongs to the class of optical counters, but measures visible radiation scattered by the aerosol particles for scattering angles ranging between 35 and 165°. If the refractive index of the aerosol is known it is possible to estimate an equivalent size distribution in the approximate range 0.1–30 μm.

6.6.3.2 Nephelometry

A *nephelometer* is an instrument that measures the radiation that is scattered or backscattered by aerosols which are directed into a tube. Only a fraction of the scattered or backscattered radiation can actually be measured because of the geometry of the instrument. The range of scattering angles that can be integrated is typically 7–170° for scattered radiation and 90–170° for backscattered fraction. A correction needs to be applied to account for the missing ranges of scattering angles in the forward and backward directions. Other correction terms are required to remove the contributions from molecular scattering, scattering from the instrument walls and parasite light. Several instruments are commercialized, e.g. by TSI™ and Radiance Research™.

6.6.4 Measurement of Aerosol Absorption

6.6.4.1 Filter-Based Optical Measurement

Aerosol absorption can be measured as the difference in the transmission of light through a filter before and after aerosols are deposited on the filter. The issue is that aerosol scattering also decreases the transmission of light through the filter, which needs to be corrected. Absorption is therefore estimated as the difference

between extinction and scattering. Several instruments are available on the market and have been widely used such as the Particle Soot Absorption Photometer™ or the Aethalometer™.

6.6.4.2 Photo-Acoustic Absorption Measurement

The principle of photo-acoustic absorption measurements is to illuminate aerosol particles flowing in a small volume with a laser beam. Absorbing aerosol particles get heated and emit an acoustic wave that can then be measured. The instrument can be calibrated with a gas whose concentration and absorbing cross section at the wavelength of the laser beam are well known. This technique is of interest in that it allows a direct measure of the aerosol absorption.

6.6.4.3 Single Particle Soot Photometer (SP2)

The Single Particle Soot Photometer (SP2) is a recent device, which can measure mass and mixing state (typically coating thickness) of individual black carbon (BC) particles (Moteki and Kondo 2007, 2008). A laminar air flow containing particles is led into a small volume and illuminated with a laser beam. Particles that sufficiently absorb laser energy are heated to several thousand kelvins, and emit black body radiation in the visible spectrum, which is detected by photomultiplier tubes. The radiation emitted by the heated particles is typically measured at two different wavelengths to estimate the black body temperature at the time of vaporization. The vaporization temperature of BC particles is about 4000 K and can be discriminated from other types of particles with different vaporization temperatures, such as metallic particles. It is then possible to measure the number and mass concentration of BC particles (BC cores). The coating thickness of large enough BC-containing particles can also be estimated by an analysis, based on Mie theory, of the time-dependent scattering intensity during the course of travel of the particle in the laser beam. Further, it has now become possible to discriminate attached type (bare BC on the surface of non-BC particles) and coated type BC (BC embedded within or coated by non-BC compounds) particles.

6.7 Conclusions

In situ measurements from the ground or from aircraft provide accurate local information on aerosol concentrations and properties but cannot sample properly the atmosphere to characterize the spatial and temporal variability of the global aerosol. Remote sensing of aerosols from the ground and from space is thus indispensable for monitoring and understanding atmospheric aerosols and their role on the climate system. Only satellite measurements can observe aerosols at the global scale, although

they provide a limited characterization of the aerosols. In particular, it is difficult to retrieve aerosol chemical composition and absorption from satellite observations. The vertical profile is now better observed at the global scale with the advance of spaceborne lidars.

Mathematical techniques exist that allow to combine the (imperfect) information from models with the (incomplete and also imperfect) information from measurements. These techniques are known as *data assimilation* methods. Figure 2.5 shows the aerosol optical depth simulated by a numerical weather prediction (NWP) model in which twice-daily measurements from the MODIS instrument onboard the Terra and Aqua satellites are assimilated (Morcrette et al. 2009; Benedetti et al. 2009). Data assimilation techniques are introduced in Chap. 7.

Exercises

1a. The IASI instrument measures radiances of 0.0700 and 0.0913 W m^{-2} sr^{-1} cm at wavenumbers 711.0 and 965.5 cm^{-1}. Compute the wavelengths corresponding to these wavenumbers. Compute the brightness temperatures corresponding to the measured radiances at these two wavenumbers. Why do they differ from each other?

1b. In the presence of desert dust, the measured radiances are now 0.0699 and 0.0881 W m^{-2} sr^{-1} cm. Compute the brightness temperatures corresponding to the measured radiances. Provide an explanation for the measurements.

2. A satellite instrument measures a radiance of 0.35 W m^{-2} sr^{-1} nm^{-1} above a "black" surface at a wavelength of 500 nm and for a scattering angle of 120°. The viewing zenith angle is 30°. The average phase function of the aerosols (which are nonabsorbing) is 0.2 sr^{-1}. The solar irradiance at the top of the atmosphere at 500 nm equals 2 W m^{-2} nm^{-1}. The surface pressure is taken equal to 1013 hPa. Estimate the optical depth and phase function by atmospheric molecules with the help of Chap. 5. Estimate the vertical aerosol optical depth.

3. This exercise aims to find an analytical solution to the lidar signal equation in a way similar to Fernald et al. (1972). We consider a lidar on the ground and pointing to the zenith. Scattering by air molecules are neglected for now and the aerosol type is assumed to be known and uniform on the vertical. In particular the aerosol extinction-to-backscatter ratio is $S_a = \sigma_a(z)/\beta_a(z)$ for all height z. The lidar signal equation at a given height can be written as:

$$LS(z) = \frac{C}{z^2} \frac{\sigma_a(z)}{S_a} T^2(z)$$

where C is a constant and $T^2(z) = \exp\left(-2 \int_0^z \sigma_a(z')\, dz'\right)$.

3a. Derive an expression for $T^2(z)$ as a function of z.

3b. Inject the equation obtained at question 3a in the lidar signal equation and express $\mathrm{d}\, T^2(z)/\mathrm{d}z$ as a function of z and $LS(z)$.

3c. Integrate the previous equation between altitudes z=0 and z. Combine this new equation with that of the lidar signal equation to get an expression of $\sigma_a(z)$ as a function of z and $LS(z)$.

3d. We now consider scattering by molecules, so that the lidar signal equation can now be written:

$$LS(z) = \frac{C}{z^2} \left(\frac{\sigma_R(z)}{S_R} + \frac{\sigma_a(z)}{S_a} \right) T_R^2(z)\, T_a^2(z)$$

Derive the expression of $T_a^2(z)$ with respect to z and reinject this equation into the lidar signal equation in order to obtain a differential equation for $T_a^2(z)$. Derive an analytical solution for $\sigma_a(z)$. It is reminded that a solution to the differential equation $y' + a(x)\, y = b(x)$ is

$$f(x) = \exp\left(- \int_0^x a(x')\, \mathrm{d}x' \right) \left\{ f(0) + \int_0^x b(x')\, \exp\left(\int_0^{x'} a(x")\mathrm{d}x" \right) \mathrm{d}x' \right\}.$$

Solutions

1a. $T_b(711.0\ \mathrm{cm}^{-1}) = 247.7$ K, $T_b(965.5\ \mathrm{cm}^{-1}) = 291.0$ K. The difference is due to the fact that the first channel is sensitive to the presence of water vapour and ozone.

1b. $T_b(711.0\ \mathrm{cm}^{-1}) = 247.6$ K, $T_b(965.5\ \mathrm{cm}^{-1}) = 288.8$ K. Only the second channel is sensitive to the presence of desert dust.

2. $\delta_{\mathrm{mol}} = 0.16$, $P_{\mathrm{mol}}(120°) = 0.937\ \mathrm{sr}^{-1}$

$$\delta_{\mathrm{mol}} P_{\mathrm{mol}}(120°) + \delta_{\mathrm{aer}} P_{\mathrm{aer}}(120°) = \frac{4\, \pi\, \mu_v\, L^{\uparrow}}{\varpi\, f_\lambda} = 0.19$$

$\delta_{\mathrm{aer}} = 0.2$.

3a. It can be obtained upon differentiation:

$$\frac{\mathrm{d}\,T^2(z)}{\mathrm{d}z} = -2\,\sigma_a(z)\,T^2(z)$$

3b. After combining the previous equation with the lidar signal equation and rearranging some terms, one gets:

$$\frac{\mathrm{d}\,T^2(z)}{\mathrm{d}z} = \frac{-2\,S_a}{C}\,LS(z)\,z^2$$

3c. After integration, one gets:

$$T^2(z) - 1 = \frac{-2\,S_a}{C}\int_0^z LS(z')\,z'^2\,dz'$$

It can be noted that this equation can be used to deduce S_a if one knows $T^2(z)$ for a particular value of z (such as the top of the atmosphere). The equation can be used as well to get an analytical solution:

$$\sigma_a(z) = \frac{S_a\,z^2\,LS(z)}{C - 2\,S_a\int_0^z LS(z')\,z'^2\,dz'}$$

3d. The quantity $T_a^2(z)$ follows a differential equation

$$\frac{\mathrm{d}\,T_a^2(z)}{\mathrm{d}z} - \frac{2\,S_a}{S_r}\,\sigma_R(z)\,T_a^2(z) = -\frac{2\,S_a\,z^2\,LS(z)}{C\,T_R^2(z)}$$

whose solution allows to extract

$$\sigma_a(z) = -\frac{S_a}{S_r}\,\sigma_R(z) + \frac{S_a}{C}\,LS(z)\,z^2\,T_R(z)^{2\,S_a/S_r-2}\times \left\{1 - \frac{2\,S_a}{C}\int_0^z LS(z)\,z^2\,T_R(z)^{-2S_a/S_r-S_R}\,dz'\right\}^{-1}.$$

References

Ackerman SA (1997) Remote sensing aerosols using satellite infrared observations. J Geophys Res 102:17069–17079. doi:10.1029/96JD03066

Althausen D, Müller D, Ansmann A, Wandinger U, Hube H, Clauder E, Zörner S (2000) Scanning 6-wavelength 11-channel aerosol lidar. J Atmos Ocean Technol 17:1469–1482

Ansmann A, Riebesell M, Weitkamp C (1990) Measurement of atmospheric aerosol extinction profiles with a Raman lidar. Opt Lett 15:746–748

Benedetti A, Morcrette J-J, Boucher O, Dethof A, Engelen RJ, Fischer M, Flentjes H, Huneeus N, Jones L, Kaiser JW, Kinne S, Mangold A, Razinger M, Simmons AJ, Suttie M, The GEMS-AER Team (2009) Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part II: data assimilation. J Geophys Res 114:D13205. doi:10.1029/2008JD011115

Deuzé J-L, Goloub P, Herman M, Marchand A, Perry G, Susana S, Tanré D (2000) Estimate of the aerosol properties over the ocean with POLDER. J Geophys Res 105:15329–15346

Deuzé J-L, Bréon F-M, Devaux C, Goloub P, Herman M, Lafrance B, Maignan F, Marchand A, Nadal F, Perry G, Tanré D (2001) Remote sensing of aerosols over land surfaces from POLDER-ADEOS-1 polarized measurements. J Geophys Res 106(D5):4913–4926

Diner DJ, Martonchik JV, Kahn RA, Pinty B, Gobron N, Nelson DL, Holben BN (2005) Using angular and spectral shape similarity constraints to improve MISR aerosol and surface retrievals over land. Remote Sens Environ 94:155–171

Dubovik O, King MD (2000) A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements. J Geophys Res 105:20673–20696

Dubovik O, Smirnov A, Holben BN, King MD, Kaufman YJ, Eck TF, Slutsker I (2000) Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network (AERONET) Sun and sky radiance measurements. J Geophys Res 105:9791–9806. doi:10.1029/2000JD900040

Dubovik O, Holben B, Eck TF, Smirnov A, Kaufman YJ, King MD, Tanré D, Slutsker I (2002) Variability of absorption and optical properties of key aerosol types observed in worldwide locations. J Atmos Sci 59:590–608

Fernald FG, Herman BM, Reagan JA (1972) Determination of aerosol height distribution by lidar. J Appl Meteorol 11:482–489

Ferrare RA, Melfi SH, Whiteman DN, Evans KD, Leifer R (1998) Raman lidar measurements of aerosol extinction and backscattering 1. Methods and comparisons. J Geophys Res 103:19663–19672

Gordon HR, Brown OB, Evans RH, Brown JW, Smith RC, Baker KS, Clark DK (1988) A semianalytic radiance model of ocean color. J Geophys Res 93:10909–10924. doi:10.1029/JD093iD09p10909

Govaerts YM, Wagner S, Lattanzio A, Watts P (2010) Joint retrieval of surface reflectance and aerosol optical depth from MSG/SEVIRI observations with an optimal estimation approach: 1. Theory. J Geophys Res 115:D02203. doi:10.1029/2009JD011779

Herman M, Deuzé JL, Devaux C, Goloub P, Bréon FM, Tanré D (1997) Remote sensing of aerosols over land surfaces including polarization measurements and application to POLDER measurements. J Geophys Res 102:17039–17049

Holben BN, Eck TF, Slutsker I, Tanré D, Buis JP, Setzer A, Vermote E, Reagan J, Kaufman YJ, Nakajima T, Lavenu F, Jankowiak I, Smirnov A (1998) AERONET—a federated instrument network and data archive for aerosol characterisation. Remote Sens Environ 66:1–16

Kaufman YJ, Tanré D (1992) Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. IEEE Trans Geosci Remote Sens 30:261–270

Kaufman YJ, Tanré D (1996) Strategy for direct and indirect methods for correcting the aerosol effect on remote sensing: from AVHRR to EOS-MODIS. Remote Sens Environ 55:65–79

Kent GS, McCormick MP, Schaffner SK (1991) Global optical climatology of the free tropospheric aerosol from 1.0-μm satellite occultation measurements. J Geophys Res 96:5249–5267

King MD, Byrne DM, Herman BM, Reagan JA (1978) Aerosol size distributions obtained by inversions of spectral optical depth measurements. J Atmos Sci 35:2153–2167

Legrand M, Bertrand JJ, Desbois M, Menenger L, Fouquart Y (1989) The potential of infrared satellite data for the retrieval of Saharan dust optical depth over Africa. J Appl Meteorol 28:309–321

Mishchenko MI, Travis LD (1997) Satellite retrieval of aerosol properties over the ocean using polarization as well as intensity of reflected sunlight. J Geophys Res 102:16989–17013

Morcrette J-J, Boucher O, Jones L, Salmond D, Bechtold P, Beljaars A, Benedetti A, Bonet A, Kaiser JW, Razinger M, Schulz M, Serrar S, Simmons AJ, Sofiev M, Suttie M, Tompkins AM, Untch A (2009) Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Forward modelling. J Geophys Res 114:D06206. doi:10.1029/2008JD011235

Moteki N, Kondo Y (2007) Effects of mixing state on black carbon measurements by laser-induced incandescence. Aerosol Sci Technol 41:398–417

Moteki N, Kondo Y (2008) Method to measure time-dependent scattering cross sections of particles evaporating in a laser beam. J Aerosol Sci 39:348–364

Nakajima T, Tanaka M, Yamauchi T (1983) Retrieval of the optical properties of aerosols from aureole and extinction data. Appl Opt 22:2951–2959

North PRJ, Briggs SA, Plummer SE, Settle JJ (1999) Retrieval of land surface bidirectional reflectance and aerosol opacity from ATSR-2 multiangle imagery. IEEE Trans Geosci Remote Sens 37:526–537

Peyridieu S, Chédin A, Capelle V, Tsamalis C, Pierangelo C, Armante R, Crevoisier C, Crépeau L, Siméon M, Ducos F, Scott NA (2013) Characterisation of dust aerosols in the infrared from IASI and comparison with PARASOL, MODIS, MISR, CALIOP, and AERONET observations. Atmos Chem Phys 13:6065–6082

Pierangelo C, Chédin A, Heilliette S, Jacquinet-Husson N, Armante R (2004) Dust altitude and infrared optical depth from AIRS. Atmos Chem Phys 4:1813–1822

Rao CRN, Stowe LL, McLain EP (1989) Remote sensing of aerosols over the oceans using AVHRR data: theory, practice and applications. Int J Remote Sens 10:743–749

Remer LA, Kaufman YJ, Tanré D, Mattoo S, Chu DA, Martins JV, Li R-R, Ichoku C, Levy RC, Kleidman RG, Eck TF, Vermote E, Holben BN (2005) The MODIS aerosol algorithm, products, and validation. J Atmos Sci 62:947–973

Shaw GE (1983) Sun photometry. Bull Am Meteorol Soc 64:4–10

Soufflet V, Tanré D, Royer A, O'Neill NT (1997) Remote sensing of aerosols over boreal forest and lake water from AVHRR data. Remote Sens Environ 60:22–34

Tanré D, Herman M, Deschamps P-Y, De Leffe A (1979) Atmospheric modeling for space measurements of ground reflectances including bidirectional properties. Appl Opt 18:3587–3594

Tanré D, Bréon F-M, Deuzé J-L, Dubovik O, Ducos F, François P, Goloub P, Herman M, Lifermann A, Waquet F (2011) Remote sensing of aerosols by using polarized, directional and spectral measurements within the A-Train: the PARASOL mission. Atmos Meas Tech 4:1383–1395

Thomas GE, Jakosky BM, West RA, Sanders RW (1983) Satellite limb-scanning thermal infrared observations of the El Chichón stratospheric aerosol: First results. Geophys Res Lett 10:997–1000

Twitty JT (1975) The inversion of aureole measurements to derive aerosol size distributions. J Atmos Sci 32:584–591

Vermote EF, Tanré D, Deuzé J-L, Herman M, Morcrette J-J (1997) Second simulation of the satellite signal in the solar spectrum: an overview. IEEE Trans Geosci Remote Sens 35:675–686

Wagner SC, Govaerts YM, Lattanzio A (2010) Joint retrieval of surface reflectance and aerosol optical depth from MSG/SEVIRI observations with an optimal estimation approach: 2. Implementation and evaluation. J Geophys Res 115:D02204. doi:10.1029/2009JD011780

Further Reading (Textbooks and Articles)

Hinds WC (1999) Aerosol technology: properties, behavior, and measurement of airborne particles, 2nd edn. Wiley, New York, 504 p

Kaufman YJ, Tanré D, Boucher O (2002) A satellite view of aerosols in the climate system. Nature 419:215–223

Kokhanovsky A, de Leeuw G (eds) (2008) Satellite aerosol remote sensing over land. Springer-Verlag, Berlin and Heidelberg, 405 p

Lenoble J, Remer L, Tanré D (eds) (2012) Aerosol remote sensing. Springer-Verlag, Berlin, 428 pp

Stephens GL (1994) Remote sensing of the lower atmosphere, an introduction. Oxford University Press, Oxford, 523 p

Chapter 7
Aerosol Data Assimilation

Abstract Data assimilation aims to combine information from observations describing the real world, some prior information from a model, and in some cases the physical and chemical laws governing the evolution of the system in consideration. This chapter provides a short description of the mathematical principles of data assimilation from the best linear unbiased estimator to variational techniques and Kalman filters. It introduces the vocabulary (background, control vector, analysis) and highlights the importance of error covariance matrices in data assimilation. Finally, the chapter reviews past applications of data assimilation for atmospheric aerosols.

Keywords Data assimilation · Error · Uncertainty · Minimization · Variational · Kalman filter

7.1 Introduction

Uncertainties associated with aerosol emissions, both in terms of their intensity and distribution pattern, atmospheric processes, and optical properties, represent a significant part of the uncertainty associated with the quantification of the impact of aerosols on climate and air quality in regional and global models. As it proves difficult to decrease these uncertainties, an interesting approach is to develop approaches that constrain models with available information from observations. Such approaches are known generically under the term of *data assimilation*, whereby observational data and numerical models are combined together to find a statistically-optimal solution that represents the best compromise between an *a priori* or *background* information and the observations.

Different assimilation techniques exist, which differ in their level of complexity, validity assumptions and numerical cost, and can be implemented and applied to various geophysical problems (Bouttier and Courtier 1999). While such techniques have been applied to numerical weather prediction for a long time, they have only started recently to be used in atmospheric chemistry both to provide optimal initial conditions for the subsequent prediction as well as to constrain emissions. Assimilation techniques include optimal interpolation, Kalman filters (KF), and the

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_7

three-dimensional and four-dimensional variational (3D-Var and 4D-Var) assimilation schemes. In spite of their differences, all these schemes seek to combine information from different sources: observations describing the real world, some prior information from a model, and in some cases the physical and chemical laws governing the evolution of the system in consideration (as described in a model). We provide here a short description of the mathematical principles of data assimilation and a review of existing applications for atmospheric aerosols.

7.2 Basic Principles of Data Assimilation

Understanding data assimilation requires the introduction of some algebra. We seek to find the optimum, $\mathbf{x^a}$, of a state vector, $\mathbf{x}$, given *a priori* knowledge on the state vector, $\mathbf{x^b}$, and a vector of observations, $\mathbf{y}$. Such a problem for aerosols is often undetermined, because there is not enough data or because the data only relates indirectly to the model variables. The optimum is called the *analysis* and the *a priori* state vector is also known as the *background*. The principle of data assimilation relies on the minimization of a cost function (Fig. 7.1a) in order to find the optimum $\mathbf{x^a}$. This cost function, $J(x)$, is defined as the sum of the departures of a potential solution $\mathbf{x}$ and of the corresponding simulated observations to the *a priori* information $\mathbf{x}^b$ and to the given observations $\mathbf{y}$, which leads to the following general expression:

$$J(\mathbf{x}) = J_b(\mathbf{x}) + J_o(\mathbf{x}) = \frac{1}{2}\,(\mathbf{x}-\mathbf{x^b})^T\,\mathbf{B}^{-1}\,(\mathbf{x}-\mathbf{x^b}) \\ +\frac{1}{2}\,(H(\mathbf{x})-\mathbf{y})^T\,\mathbf{R}^{-1}(H(\mathbf{x})-\mathbf{y}) \tag{7.1}$$

where H is the (nonlinear) observation operator that computes the observation vector that would correspond to the state vector $\mathbf{x}$, $\mathbf{B}$ is the covariance matrix of the error statistics of the background state vector $\mathbf{x}^b$, and $\mathbf{R}$ is the covariance matrix of the error statistics of the observation vector $\mathbf{y}$. The superscript T denotes the transpose operator for a matrix. More specifically, $\mathbf{B}$ can be built from the statistics of the error vector $\mathbf{e}$ of the background vector against its true value

$$\mathbf{B} = (B_{ij}) = \left(\rho(e_i, e_j)\sqrt{\mathrm{var}(e_i)}\sqrt{\mathrm{var}(e_j)}\right) \tag{7.2}$$

where $\rho(e_i, e_j)$ designates the correlation coefficient between e_i and e_j (1 when $i = j$) and $\mathrm{var}(e_i)$ is the variance of e_i. The $\mathbf{R}$ matrix represents not only the errors in the observation, but also the errors of the observation operator, and representativeness errors associated with the model discretization and observation sampling.

It can be understood intuitively that the larger the errors on the background state vector (i.e. the larger the diagonal terms in the $\mathbf{B}$ matrix), the more weight is given to the observations, and vice versa. The $\mathbf{B}$ and $\mathbf{R}$ matrices are composed of the error variances on the individual components of the state and observation vectors (diagonal terms) and the error covariances (off-diagonal terms). Although data assimilation is

quite sensitive to the way these covariance error matrices are defined, it is often difficult to estimate them rigorously so that a number of approximations are often made. For instance, error covariance terms are often neglected for the observations and a diagonal $\mathbf{R}$ matrix is used. There are rules to verify the statistical consistency of the $\mathbf{R}$ and $\mathbf{B}$ matrices and the actual model-observation misfits (Desroziers et al. 2005). Finally, it should be stressed that only if the errors are Gaussian and unbiased does the analysis $\mathbf{x^a}$ correspond to the optimum in a Bayesian sense (in this case the method is also called the *best linear unbiased estimator* or BLUE).

The method used to minimize the cost function J depends among other aspects on the size of the state vector as well as the relative sizes of the $\mathbf{B}$ and $\mathbf{R}$ matrices and the difficulties associated with their inversion (Chevallier et al. 2005). When the size of the state vector is small, or the number of observations is small (typically less than a few 10,000s), assuming the problem is not too nonlinear, and the errors are unbiased, the analysis or *a posteriori* can be computed analytically:

$$\begin{aligned} \mathbf{x}^a &= \mathbf{x}^b - (\mathbf{H}^T\,\mathbf{R}^{-1}\,\mathbf{H} + \mathbf{B}^{-1})^{-1}\,\mathbf{H}^T\,\mathbf{R}^{-1}\,(\mathbf{H}\,\mathbf{x}^b - \mathbf{y}) \\ &= \mathbf{x}^b - \mathbf{B}\,\mathbf{H}^T\,(\mathbf{H}\,\mathbf{B}\,\mathbf{H}^T + \mathbf{R})^{-1}\,(\mathbf{H}\,\mathbf{x}^b - \mathbf{y}) \\ &= \mathbf{x}^b + \mathbf{K}\,(\mathbf{y} - \mathbf{H}\,\mathbf{x}^b) \end{aligned} \tag{7.3}$$

where $\mathbf{H}$ is the linearized operator of H, also called *Jacobian matrix*, and $\mathbf{K}$ is known as the *gain matrix*. Each element of the Jacobian matrix is given by the partial derivative of one element of observation vector ($\mathbf{y}$) with respect to one element of the state vector ($\mathbf{x}$). Each column of the Jacobian matrix can be computed from finite differences or an explicit computation with a linearized version of the H code, also called *tangent linear* code, which provides the sensitivities of all output parameters with respect to one input parameter. The uncertainty in the analysis is described by the analysis error covariance matrix

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\,\mathbf{H})\,\mathbf{B} \tag{7.4}$$

where $\mathbf{I}$ is the identity matrix.

When the size of the state vector is large, the idea is to avoid computing the gain matrix $\mathbf{K}$. The analysis $\mathbf{x}^a$ is computed instead iteratively through a suitable gradient descent algorithm. A convenient way to do this is through variational approaches, where the gradient of the cost function

$$\nabla J(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) - \mathbf{H}^T\,\mathbf{R}^{-1}\,(\mathbf{y} - H(\mathbf{x})) \tag{7.5}$$

can be computed directly at a relatively low computational cost once the $\mathbf{H}^T$ operator, also called the *adjoint* of the observation operator, has been coded. The adjoint model provides the sensitivities of the cost function to perturbations in the input parameters ($\mathbf{x}$). If the system is well behaved, then only a small number of iterations, typically less than 100, are performed to minimize the J function (Fig. 7.1b). A disadvantage of the variational approach is that it does not easily provide an estimate of the uncertainty of the analysis. This is not impossible though as the $\mathbf{A}$ matrix can be expressed as the

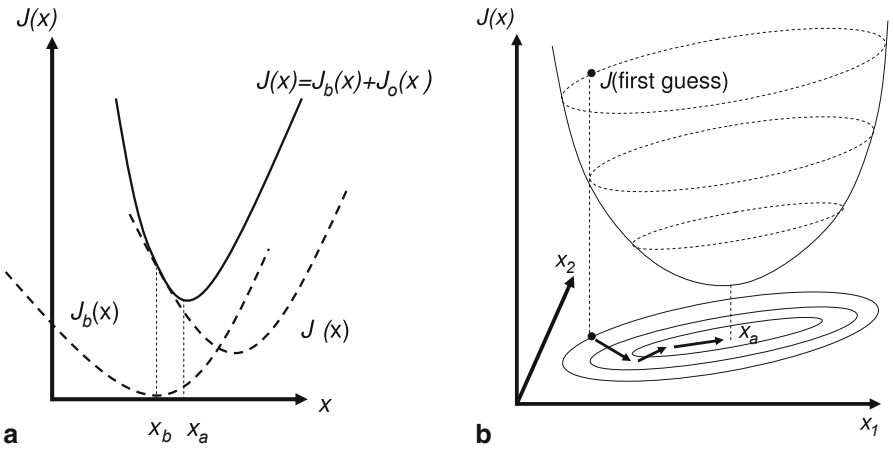


Fig. 7.1 Schematic representation of **a** the cost function (here in a one-variable model space) and **b** the minimization of the variational cost function (here in a two-variable model space). The minimization works by going down the gradient in some iterative steps. (Adapted from Bouttier and Courtier (1999))

inverse of the Hessian of the cost function (i.e. a square matrix of the second partial derivatives of J).

Variational assimilation methods can be 3-D (as in Eq. 7.5), in which case the minimization is done within a time window but without constraining explicitly the temporal trajectory of the state vector, or 4-D, in which case the time evolution of the state vector as simulated by the model comes as an additional constraint. The extension to 4D-Var requires the tangent linear and adjoint of the chemistry-transport model or a simple version of it, but is also more accurate.

The data assimilation can be *sequential*—when observations are assimilated in consecutive sequences—or *nonsequential*—when all available observations are assimilated at once. A sequential data assimilation scheme, therefore, requires some cycling over time in order to update the state vector and its error statistics from one sequence to the next. A *filter* is made of this cycling of the error statistics and of a Bayesian analysis through Eqs. 7.3 and 7.4. This is typically done through a predict and update approach in the so-called *Kalman filter*:

$$\begin{cases} \mathbf{x}_k^b = \mathbf{M}_k\, \mathbf{x}_{k-1}^a \\ \mathbf{B}_k = \mathbf{M}_k\, \mathbf{A}_{k-1}\, \mathbf{M}_k^T + \mathbf{Q}_k \\ \mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k\, (\mathbf{y} - \mathbf{H}\, \mathbf{x}_k^b) \\ \mathbf{A}_k = (\mathbf{I} - \mathbf{K}_k\, \mathbf{H}_k)\, \mathbf{B}_k \end{cases} \tag{7.6}$$

where $\mathbf{x}_k^b$ is the predicted (*a priori*) estimate, $\mathbf{x}_k^a$ the updated (*a posteriori*) estimate, $\mathbf{B}_k$ the predicted error covariance matrix for $\mathbf{x}_k^b$, $\mathbf{A}_k$ the updated error covariance matrix for $\mathbf{x}_k^a$, and $\mathbf{Q}_k$ the error of the model operator $\mathbf{M}_k$ for sequence k. An ensemble

approach can replace some or all of the analytical computations above with the statistics of a (generally rather small) discrete ensemble of model runs. The *ensemble Kalman filter* (EnKF) is one example of the ensemble approach based on the Kalman filter but there are other ones, such as *particle filters* or variations of the EnKF itself.

7.3 Applications of Data Assimilation for Aerosols

Quite logically the first applications of data assimilation to aerosols were to determine an optimal state for the distribution of aerosol concentration, burden or optical depth given imperfect and patchy satellite aerosol observations and equally—if not more—imperfect global aerosol models. In this case the state vector $\mathbf{x}$ consists of the aerosol concentrations. Collins et al. (2001) and Rasch et al. (2001) pioneered data assimilation for aerosols by using an optimal interpolation scheme to assimilate aerosol optical depth (AOD) from the Advanced Very High Resolution Radiometer (AVHRR) during the Indian Ocean Experiment (INDOEX). The technique was used both to forecast aerosol fields during the experiment to best deploy ships and airplanes and to help explain the aerosol distribution after the field campaign. Weaver et al. (2007) used an assimilation method based on the Kalman filter to assimilate radiances from the Moderate Resolution Imaging Spectroradiometer (MODIS).

Furthermore, assimilation of aerosol observations has also started to be incorporated into operational systems to improve aerosol forecasting capabilities. The first attempt was done by the Naval Research Laboratory (NRL) where an AOD assimilation scheme was implemented in the Navy Aerosol Analysis and Prediction System (NAAPS) (Zhang et al. 2008). Niu et al. (2008) developed an assimilation system for the Chinese Unified Atmospheric Chemistry Environment—Dust (CUACE/Dust) forecast system for the operational forecasting of dust aerosols. Finally, the European Centre for Medium range Weather Forecast (ECMWF) included an assimilation scheme for aerosol optical depth from satellite sensors into its Integrated Forecasting System (IFS) to produce aerosol forecasts and reanalyses of aerosol fields (Benedetti et al. 2009; Fig. 2.5).

Data assimilation can also be applied to estimate the emission intensity of different aerosol species and/or gaseous precursor on both the global and regional scale. In this case the state vector is composed of the emissions either at the model grid-box resolution or as averages over some predefined regions. In a variational-like method, Dubovik et al. (2008) assimilated MODIS AOD at 550 nm to estimate the location and intensity of fine and coarse mode aerosol emissions. Hakami et al. (2005) used a 4D-Var approach to estimate the emissions and initial and boundary conditions of black carbon aerosols by assimilating surface concentration measurements. Huneeus et al. (2012, 2013) estimated natural and anthropogenic emissions of primary aerosols and sulphur dioxide (SO_2) at the monthly resolution and for the year 2010 by assimilating daily total and fine mode AOD at 550 nm from MODIS into a global aerosol model of intermediate complexity. Schutgens et al. (2012) applied a fixed-lag ensemble Kalman smoother to estimate the emissions of dust, sea salt and carbon aerosol

and, the precursor gas SO_2 by assimilating AOD and the Ångström coefficient. This fixed-lag Kalman smoother is in essence a Kalman filter that iteratively estimates emissions assuming a linear model and Gaussian errors.

Models of different levels of complexity are being used in such data assimilation studies. While models of high complexity faithfully represent the known physical and chemical processes, simplified models and those of intermediate complexity have the advantage of focusing on important processes and of making simulations computationally efficient and conceptually easier to understand. Henze et al. (2004) and Sandu et al. (2005) developed inverse box models of aerosol dynamics that focus on the physical particle dynamics with limited chemical and thermodynamical transformations. Zhang et al. (2008) used an aerosol model with a simplified aerosol representation in the operation assimilation system from the NRL numerical weather prediction (NWP) system. An aerosol model of intermediate complexity was developed and introduced into the IFS of the ECMWF but an even simpler one-variable aerosol model is used in the 4D-Var assimilation loop (Benedetti et al. 2009). Most of the aerosol assimilation systems rely on MODIS AOD product (e.g. Zhang et al. 2008; Dubovik et al. 2008; Benedetti et al. 2009; Kaiser et al. 2012; Huneeus et al. 2012, 2013). However, satellite AOD products have also been assimilated from other instruments such as the Polarization and Directionality of the Earth's Reflectances (POLDER) instrument (Generoso et al. 2007) and AVHRR. Studies also exist that instead of assimilating AOD have directly assimilated radiances from MODIS (Weaver et al. 2007; Wang et al. 2012; Xu et al. 2013). Additional satellite products that have been assimilated are the aerosol index (AI) from the Total Ozone Mapping Spectrometer (TOMS) (Zhang et al. 2005) and the vertical profiles of extinction coefficients from the Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP) onboard the Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) platform (Sekiyama et al. 2010). There are a few studies that have assimilated ground-based observations. For instance Fu et al. (2012) estimated the emissions of carbonaceous aerosols in China by constraining the fluxes with surface concentration measurements. Hakami et al. (2005) assimilated in situ measurements of BC mass concentration to estimate BC emissions and initial and boundary conditions. Yumimoto et al. (2007, 2008) used a 4D-Var data assimilation system for a regional dust model vertical profiles of dust extinction coefficient from the NIES lidar network. Finally, systems assimilating both ground-based and satellite data also exist. Schutgens et al. (2010, 2012) combined simultaneously AOD and Ångström coefficient from the Aerosol Robotic Network (AERONET) and AOD from MODIS. The CUACE/Dust forecast system assimilates surface visibility and dust loading retrieval from the Chinese geostationary satellite FY-2C (Niu et al. 2008).

Exercises

1. A model predicts a burden of dry aerosol of 5 ± 2 mgm^{-2}. The aerosol is estimated to have a mass extinction efficiency of 4 ± 2 m^{2}g^{-1}, which includes the effect of hygroscopicity. The aerosol optical depth (AOD) is observed to be 0.025 ± 0.005. The $\pm$ indicates one standard deviation uncertainty for each modelled or observed parameter. Assuming all estimates are unbiased, compute the optimal dry aerosol burden and its uncertainty given the modelled value, the aerosol mass extinction efficiency, the observed AOD, and their uncertainties.

2. This exercise aims to develop the adjoint of a model that simulates the signal measured by a spaceborne single-wavelength lidar for an atmosphere that contains a single aerosol type. The lidar signal equation can be discretized on the vertical:

$$LS_k = \eta_k \, \sigma_k \, \exp\left(-2 \sum_{j=k}^{k_{\max}} \sigma_j \, \Delta z_j\right)$$

where $\eta_k = C_0 \, \frac{A}{R_k^2} \, \beta_k$ brings together all the constant terms. The levels are numbered from the surface and $k_{\max}$ represents the index of the uppermost atmospheric level. The control variable in this case is made of the vertical profile of the extinction coefficient σ_k at the given wavelength over the $k_{\max}$ levels of the atmosphere. A system of equations is created by combining the identity equation for the extinction coefficient and the lidar signal equation.

$$\begin{cases} \sigma_k = \sigma_k \\ LS_k = \eta \, \sigma_k \, \exp\left(-2 \sum_{j=k}^{k_{\max}} \sigma_j \, \Delta z_j\right) \end{cases}$$

2a. Linearize the lidar signal equation with respect to the extinction coefficient so that to obtain the tangent linear model.

2b. Rewrite the system in matrix form with a dimension $2 \times k_{\max}$. The first $k_{\max}$ lines represent the perturbed identity equation for the extinction coefficient at each level and the next $k_{\max}$ lines correspond to the perturbed lidar signal equation at each atmospheric layer.

2c. Derive the adjoint model. The adjoint variables can be noted with a * symbol so as to distinguish them from those of the tangent linear system.

Solutions

1. We can apply the best linear unbiased estimator where all the vectors and error matrices reduce to scalar values. The H operator consists of the multiplication by the mass extinction efficiency. $\mathbf{x^b} = (5\ 10^{-3})$, $\mathbf{B} = (4\ 10^{-6})$, and $\mathbf{y} = (0.025)$. The $\mathbf{R}$ matrix combines the uncertainty on the observation with that of the observation operator transformed in AOD space: $\mathbf{R} = (0.005^2 + 5\ 10^{-3} \times 4 \times 5\ 10^{-3}) = (1.25\ 10^{-4})$. Applying Eqs. 7.3 and 7.4 leads to a best estimate for the dry aerosol burden of 5.4 mgm^{-2} and a (reduced) standard deviation of 1.6 mgm^{-2}.

2a. The system of linearized equations can be written in the following way:

$$\begin{cases} \delta\sigma_k = & \delta\sigma_k \\ \delta LS_k = & \eta_k\, \delta\sigma_k\, \exp\left(-2 \sum_{j=k}^{k_{\max}} \sigma_j\, \Delta z_j\right) \\ & -2\, \eta_k\, \sigma_k\, \exp\left(-2 \sum_{j=k}^{k_{\max}} \sigma_j\, \Delta z_j\right) \sum_{i=k}^{k_{\max}} \delta\sigma_i\, \Delta z_i \end{cases}$$

2b. The system of equations can be written in matrix form with a matrix of dimension $2\,k_{\max} \times 2\,k_{\max}$.

$$\begin{pmatrix} \delta\sigma_1 \\ \vdots \\ \delta\sigma_k \\ \vdots \\ \delta\sigma_{k_{\max}} \\ \delta LS_1 \\ \vdots \\ \delta LS_k \\ \vdots \\ \delta LS_{k_{\max}} \end{pmatrix} = \mathbf{H} \begin{pmatrix} \delta\sigma_1 \\ \vdots \\ \delta\sigma_k \\ \vdots \\ \delta\sigma_{k_{\max}} \\ \delta LS_1 \\ \vdots \\ \delta LS_k \\ \vdots \\ \delta LS_{k_{\max}} \end{pmatrix}$$

where

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots & \vdots & & \vdots \\ \vdots & 0 & 1 & 0 & \vdots & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \alpha^{11} & \cdots & \cdots & \cdots & \alpha^{1k_{\max}} & 0 & \cdots & 0 \\ 0 & \ddots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & 0 & \alpha^{kk} & \alpha^{kl} & \vdots & \vdots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & \alpha^{k_{\max}k_{\max}} & 0 & \cdots & 0 \end{pmatrix}$$

and α^{kl} is defined as:

$$\alpha^{kl} = \begin{cases} \eta_k \exp\left(-2\sum_{j=k}^{k_{\max}} \sigma_j \Delta z_j\right) - 2\eta_k \sigma_k \Delta z_l \exp\left(-2\sum_{j=k}^{k_{\max}} \sigma_j \Delta z_j\right) & \text{if } k = l \\ -2\eta_k \sigma_k \Delta z_l \exp\left(-2\sum_{j=k}^{k_{\max}} \sigma_j \Delta z_j\right) & \text{if } k < l \end{cases}$$

2c. The adjoint system in matrix form can be obtained by transposing the matrix of the tangent linear model, that is

$$\begin{pmatrix} \delta\sigma_1^* \\ \vdots \\ \delta\sigma_k^* \\ \vdots \\ \delta\sigma_{k_{\max}}^* \\ \delta LS_1^* \\ \vdots \\ \delta LS_k^* \\ \vdots \\ \delta LS_{k_{\max}}^* \end{pmatrix} = \mathbf{H}^T \begin{pmatrix} \delta\sigma_1^* \\ \vdots \\ \delta\sigma_k^* \\ \vdots \\ \delta\sigma_{k_{\max}}^* \\ \delta LS_1^* \\ \vdots \\ \delta LS_k^* \\ \vdots \\ \delta LS_{k_{\max}}^* \end{pmatrix}$$

where

$$\mathbf{H}^T = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 & \alpha^{11} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots & & \ddots & \ddots & & \vdots \\ \vdots & \ddots & 1 & \ddots & \vdots & \vdots & & \alpha^{kk} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 & & & \alpha^{kl} & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 & \alpha^{1k_{\max}} & \cdots & \cdots & & \alpha^{k_{\max}k_{\max}} \\ 0 & & \cdots & & 0 & 0 & & \cdots & & 0 \\ \vdots & & & & \vdots & \vdots & & & & \vdots \\ 0 & & & & 0 & 0 & & & & 0 \\ \vdots & & & & \vdots & \vdots & & & & \vdots \\ 0 & & \cdots & & 0 & 0 & & \cdots & & 0 \end{pmatrix}$$

which gives the following equation for a line k:

$$\begin{cases} \delta\sigma_k^* = \delta\sigma_k^* + \sum_{i=1}^{k} \alpha^{ik}\, \delta LS_i^* \\ \delta LS_k^* = 0 \end{cases}$$

Combining these equations with previous ones, the system of equations for the adjoint model of the lidar signal is then obtained:

$$\begin{cases} \delta\sigma_k^* = & \delta\sigma_k^* + \eta\ \exp\left(-2\sum_{j=k}^{k_{\max}} \sigma_j \Delta z_j\right) \delta LS_k^* \\ & -2\eta\ \Delta z_k \sum_{i=1}^{k} \sigma_i\ \exp\left(-2\sum_{j=i}^{k_{\max}} \sigma_j\ \Delta z_j\right)\ \delta LS_i^* \\ \delta LS_k^* = & 0 \end{cases}$$

References

Benedetti A, Morcrette J-J, Boucher O, Dethof A, Engelen RJ, Fischer M, Flentjes H, Huneeus N, Jones L, Kaiser JW, Kinne S, Mangold A, Razinger M, Simmons AJ, Suttie M, The GEMS-AER team (2009) Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part II: Data assimilation. J Geophys Res 114:D13205. doi:10.1029/2008JD011115

Bouttier F, Courtier P (March 1999) Data assimilation concepts and methods. Meteorological Training Course Lecture Series, 58 pp

Chevallier F, Fisher M, Peylin P, Serrar S, Bousquet P, Bréon F-M, Chedin A, Ciais P (2005) Inferring CO_2 sources and sinks from satellite observations: method and application to TOVS data. J Geophys Res 110:D24309. doi:10.1029/2005JD006390

Collins WD, Rasch PJ, Eaton BE, Khattatov BV, Lamarque J-F, Zender CS (2001) Simulating aerosols using a chemical transport model with assimilation of satellite aerosol retrievals: methodology for INDOEX. J Geophys Res 106:7313–7336

Desroziers G, Berre L, Chapnik B, Poli P (2005) Diagnosis of observation, background and analysis-error statistics in observation space. Q J Royal Meteorol Soc 131:3385–3396

Dubovik O, Lapyonok T, Kaufman YJ, Chin M, Ginoux P, Kahn RA, Sinyuk A (2008) Retrieving global aerosol sources from satellites using inverse modeling. Atmos Chem Phys 8:209–250

Fu TM, Cao JJ, Zhang XY, Lee SC, Zhang Q, Han YM, Qu WJ, Han Z, Zhang R, Wang YX, Chen D, Henze DK (2012) Carbonaceous aerosols in China: top-down constraints on primary sources and estimation of secondary contribution. Atmos Chem Phys 12:2725–2746

Generoso S, Bréon FM, Chevallier F, Balkanski Y, Schulz M, Bey I (2007) Assimilation of POLDER aerosol optical thickness into the LMDz-INCA model: implications for the Arctic aerosol burden. J Geophys Res 112:D02311. doi:10.1029/2005jd006954

Hakami A, Henze DK, Seinfeld JH, Chai T, Tang Y, Carmichael GR, Sandu A (2005) Adjoint inverse modeling of black carbon during the Asian Pacific regional aerosol characterization experiment. J Geophys Res 110:D14301. doi:10.1029/2004JD005671

Henze DK, Seinfeld JH, Liao W, Sandu A, Carmichael GR (2004) Inverse modeling of aerosol dynamics: condensational growth. J Geophys Res 109:D14201. doi:10.1029/2004jd004593

Huneeus N, Boucher O (2007) One-dimensional retrieval of extinction coefficient from synthetic lidar and radiometric measurements. J Geophys Res 112:D14303. doi:10.1029/2006JD007625

Huneeus N, Chevallier F, Boucher O (2012) Estimating aerosol emissions by assimilating observed aerosol optical depth in a global aerosol model. Atmos Chem Phys 12:4585–4606

Huneeus N, Boucher O, Chevallier F (2013) Atmospheric inversion of SO_2 and primary aerosol emissions for the year 2010. Atmos Chem Phys 13:6555–6573

Kaiser JW, Heil A, Andreae MO, Benedetti A, Chubarova N, Jones L, Morcrette JJ, Razinger M, Schultz MG, Suttie M, van der Werf GR (2012) Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. Biogeosciences 9:527–554

Morcrette J-J, Boucher O, Jones L, Salmond D, Bechtold P, Beljaars A, Benedetti A, Bonet A, Kaiser JW, Razinger M, Schulz M, Serrar S, Simmons AJ, Sofiev M, Suttie M, Tompkins AM, Untch A (2009) Aerosol analysis and forecast in the ECMWF Integrated Forecast System: forward modelling. J Geophys Res 114:D06206. doi:10.1029/2008JD011235

Niu T, Gong SL, Zhu GF, Liu HL, Hu XQ, Zhou CH, Wang YQ (2008) Data assimilation of dust aerosol observations for the CUACE/dust forecasting system. Atmos Chem Phys 8:3473–3482

Rasch PJ, Collins WD, Eaton BE (2001) Understanding the Indian Ocean Experiment (INDOEX) aerosol distributions with an aerosol assimilation. J Geophys Res 106:7337–7355

Sandu A, Daescu DN, Carmichael GR, Chai T (2005) Adjoint sensitivity analysis of regional air quality models. J Comput Phys 204:222–252

Schutgens NAJ, Miyoshi T, Takemura T, Nakajima T (2010) Applying an ensemble Kalman filter to the assimilation of AERONET observations in a global aerosol transport model. Atmos Chem Phys 10:2561–2576

Schutgens NAJ, Nakata M, Nakajima T (2012) Estimating aerosol emissions by assimilating remote sensing observations into a global transport model. Remote Sens 4:3528–3543

Sekiyama TT, Tanaka TY, Shimizu A, Miyoshi T (2010) Data assimilation of CALIPSO aerosol observations. Atmos Chem Phys 10:39–49

Wang J, Xu X, Henze DK, Zeng J, Ji Q, Tsay S-C, Huang J (2012) Top-down estimate of dust emissions through integration of MODIS and MISR aerosol retrievals with the GEOS-Chem adjoint model. Geophys Res Lett 39:L08802. doi:10.1029/2012GL051136

Weaver C, da Silva A, Chin M, Ginoux P, Dubovik O, Flittner D, Zia A, Remer L, Holben B, Gregg W (2007) Direct insertion of MODIS radiances in a global aerosol transport model. J Atmos Sci 64:808–826

Xu X, Wang J, Henze DK, Qu W, Kopacz M (2013) Constraints on aerosol sources using GEOS-Chem adjoint and MODIS radiances, and evaluation with multisensor (OMI, MISR) data. J Geophys Res 118:6396–6413

Yumimoto K, Uno I, Sugimoto N, Shimizu A, Satake S (2007) Adjoint inverse modeling of dust emission and transport over East Asia. Geophys Res Lett 34:L08806. doi:10.1029/2006gl028551

Yumimoto K, Uno I, Sugimoto N, Shimizu A, Liu Z, Winker DM (2008) Adjoint inversion modeling of Asian dust emission using lidar observations. Atmos Chem Phys 8:2869–2884. doi:10.5194/acp-8-2869-2008

Zhang S, Penner JE, Torres O (2005) Inverse modeling of biomass burning emissions using Total Ozone Mapping Spectrometer aerosol index for 1997. J Geophys Res 110:D21306. doi:10.1029/2004jd005738

Zhang JL, Reid JS, Westphal DL, Baker NL, Hyer EJ (2008) A system for operational aerosol optical depth data assimilation over global oceans. J Geophys Res 113:D10208. doi:10.1029/2007JD009065

Further Reading (Textbooks and Articles)

Ide K, Courtier P, Ghil M, Lorenc AC (1997) Unified notation for data assimilation: operational, sequential and variational. J Meteorol Soc Jpn 75:181–189

Lahoz W, Swinbank R, Khattatov B, Menard R (2010) Data assimilation: making sense of observations. Springer-Verlag, Berlin, 475 pp

Rodgers CD (2000) Inverse methods for atmospheric sounding: theory and practice. World Science, Tokyo, 240 pp

Tarantola A (2004) Inverse problem theory and model parameter estimation. SIAM: Society for Industrial and Applied Mathematics, Philadelphia, 352 pp

Tipping E (2005) Inverse problems in atmospheric constituent transport. Cambridge Atmospheric and Space Science Series, Cambridge University Press, Cambridge, 425 pp

Chapter 8
Aerosol–Radiation Interactions

Abstract This chapter focuses on the radiative effects of aerosols in the solar part of the electromagnetic spectrum. The concepts are introduced using simplified expressions for scattering and absorbing aerosols, highlighting the importance of aerosol upscatter, single scattering albedo and surface albedo. Exact calculations are then presented for different aerosol types for top-of-atmosphere, surface and atmospheric radiative forcing. An assessment of the radiative forcing due to aerosol–radiation interactions is then presented and the role of rapid adjustments (also known as the aerosol semidirect effect) is discussed according to atmospheric conditions. Finally, the mechanisms of the aerosol radiative impacts on the snowpack are discussed and quantified.

Keywords Radiation · Direct effect · Radiative forcing · Scattering · Absorption · Snow

8.1 Introduction

As we have seen in Chap. 2, the radiative effect due to aerosol–radiation interactions results from the scattering and absorption of solar radiation (the parasol effect) and to a lesser extent from the absorption and reemission of terrestrial radiation (the greenhouse effect). The latter effect is negligible except for coarse mode aerosols (desert dust and sea spray) and stratospheric aerosols. This chapter therefore focuses on the radiative effect of aerosols in the solar part of the electromagnetic spectrum. Chapters 3 and 5 provide the theoretical basis to tackle this radiative effect in a quantitative manner, which is done here through simplified equations. While these simplified equations are useful for process understanding, they are not sufficient for an accurate calculation of these effects.

Aerosol–radiation interactions occur essentially in cloud free conditions[1] because aerosols are more effective at modifying the top-of-atmosphere flux of solar radiation in the absence of clouds that are themselves good scatterers. Although the presence of clouds generally suppresses the radiative effect of aerosols, it does not necessarily cancel it altogether, either because the cloud is sufficiently thin, or because the aerosol

[1] Cloud free conditions are often referred to as *clear sky* in contrast to *cloudy sky* or *all sky* conditions.

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_8

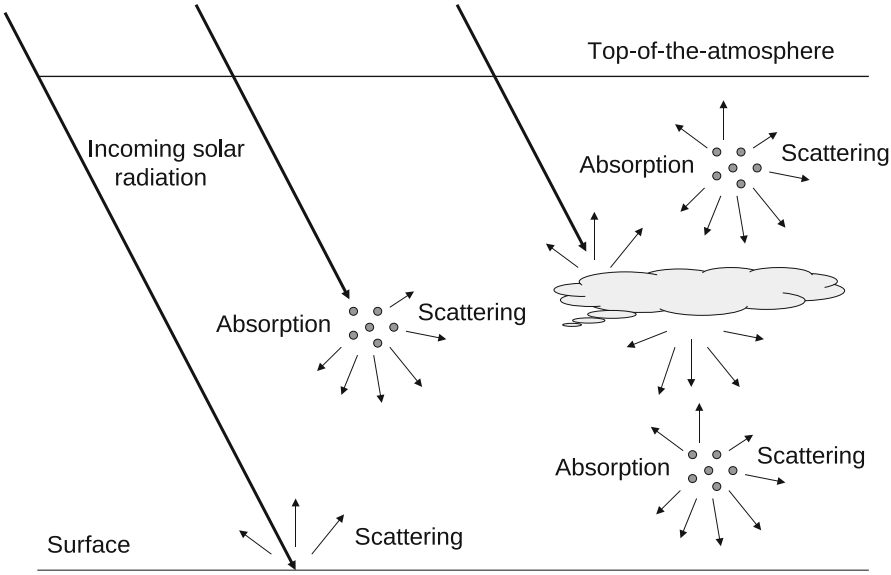


Fig. 8.1 Schematic of the aerosol–radiation interactions. Aerosols scatter solar radiation in all directions; they can also absorb solar radiation. The amount and properties of solar radiation that interact with the aerosols depend on the solar zenith angle, altitude, the surface properties, and the presence or absence of clouds

is absorbing. Figure 8.1 shows schematically the various components of aerosol–radiation interactions. This does not only depend on the amount or properties of the aerosols, but also on the properties of the solar radiation that interacts with them. This is why the solar zenith angle, the properties of the surface, and the presence of clouds are as important as aerosol optical properties.

It is important to distinguish the radiative effect from the radiative forcing due to aerosol–radiation interactions. The former refers to all (natural and anthropogenic) aerosols while the latter refers explicitly to anthropogenic aerosols. Calculating the radiative forcing due to aerosol–radiation interactions (RFari) requires to know not only aerosol and environmental properties, but also the fraction of the aerosol that is of anthropogenic origin and, in the case of an internal mixing, the properties that the aerosols would have had in the absence of anthropogenic emissions. Given the ubiquitous nature of anthropogenic aerosols, it is not easy to characterize the natural aerosols from observations alone. In particular there is a lot of uncertainty as to what the preindustrial continental aerosol actually is (Carslaw et al. 2013). This is why most estimates of the radiative forcing of aerosols require assumptions and involve some degree of modelling.

The radiative effect (or forcing) of aerosols is measured in unit of W m^{-2} and is generally estimated at the top of atmosphere with downward flux counted positive. A negative radiative forcing corresponds to an energy loss for the climate system, which leads to cooling. Inversely, a positive radiative forcing corresponds to a gain of energy,

and therefore to a warming of the system. It should be noted, and this is something we will come back to in Chap. 10, that for an inhomogeneous radiative forcing such as that of aerosols, the patterns of cooling and/or warming do not necessarily follow the patterns of radiative forcing. This chapter will only discuss the radiative impact of aerosols, and not the climate response to this radiative impact. Although the radiative impact of aerosols is usually estimated at the top of atmosphere, it is useful to estimate the radiative impact of aerosols at the surface and in the atmosphere, but such quantities need to be interpreted cautiously.

This chapter also discusses rapid adjustments associated with the radiative effects due to aerosol–radiation interactions (i.e. what was previously called the semi-direct effect) as well as the radiative effects due to aerosol–surface or aerosol–snow interactions (see Fig. 2.9).

8.2 Atmospheric Radiative Effects Due to Aerosols

8.2.1 Simplified Equation for Scattering Aerosols

Let us consider a layer of scattering (i.e. non-absorbing) aerosols, with (vertical) optical depth τ_a and upscatter fraction $\beta_a(\theta_0)$. We neglect for now the presence of a gaseous atmosphere. Several effects need to be taken into account to understand the dependence with the solar zenith angle, θ_0, of the interactions between aerosols and solar radiation. First of all, the amount of available solar radiation reaching a horizontal surface is equal to $S \cos\theta_0$ where $S \simeq 1362\,\mathrm{W\,m^{-2}}$ is the solar constant[2]. It decreases therefore with $\cos\theta_0$ as the solar zenith angle increases from 0 (the Sun is at the zenith) to 90° (the Sun is on the horizon). Second, the optical depth along the direction of propagation of solar radiation is $\tau_a/\cos\theta_0$ and increases therefore with $1/\cos\theta_0$ as the solar zenith angle increases. We can consider that, at least under the single scattering assumption, these two effects compensate each other. Of course, when θ_0 tends towards 90°, the single scattering assumption is no longer valid and the aerosol radiative effect has to tend towards 0 because there is no energy available to be lost. Third, we need to consider the dependence of the upscatter fraction with the solar zenith angle (see Sect. 3.6.3). For spherical aerosols of the accumulation mode, the upscatter fraction increases from a value of about 0.1 for a solar zenith angle of 0° to 0.5 for a solar zenith angle of 90° (see Fig. 3.8). It should be noted that the upscatter fraction for $\theta_0 = 0$ depends strongly on the aerosol size and refractive index. This value decreases for coarse mode aerosols that are larger and scatter more in the forward direction. However, the value of 0.5 for $\theta_0 = 90°$ is the same for all spherical (or randomly oriented) aerosols as it corresponds to the "downward–upward" symmetry for a particle illuminated along an horizontal axis.

[2] The solar constant is the flux of solar radiation outside the atmosphere crossing a surface perpendicular to the direction of propagation of the solar radiation.

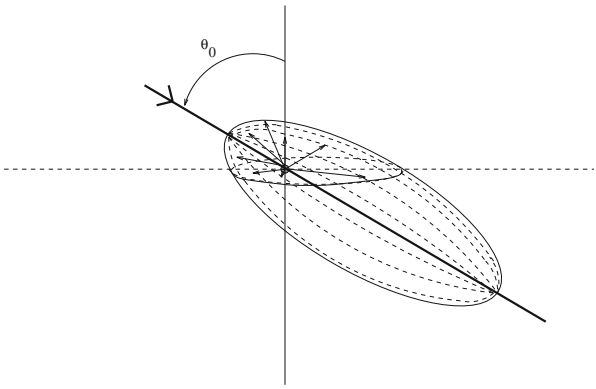


Fig. 8.2 Schematic illustrating upscatter by an aerosol. The bold line indicates the direction of the incident solar radiation (solar zenith angle θ_0). The phase function is represented by an ellipsoid which intercepts the horizontal plane. Thin arrows indicate diffuse radiation that participates to the upscatter. The larger the solar zenith angle, the more forward scattering contributes to the upscatter

Figure 8.2 illustrates the fact that as the solar zenith angle increases, a larger fraction of the forward scattering (scattering angles ranging between 0 and 90°) contributes to scattering in the upward direction. Combining these effects, we show that the radiative effect of a thin aerosol layer increases with the solar zenith angle up to the point where the combined effects of multiple scattering and decrease of available solar radiation prevail.

Let us now consider a layer of scattering aerosols in a cloud-free atmosphere and denote β_a the upscatter fraction, while deliberately ignoring for now its dependence with the solar zenith angle. This aerosol layer is placed between the surface with albedo R_s and an atmospheric layer with transmittance T (defined here as the fraction of solar radiation that goes through the atmosphere without being absorbed or backscattered). The solar zenith angle is noted θ_0 as before. The reflectance of the aerosol layer alone can be approximated by:

$$R_a \approx \beta_a \tau_a / \cos\theta_0 \tag{8.1}$$

and its transmittance

$$T_a = 1 - R_a \approx 1 - \beta_a \tau_a / \cos\theta_0. \tag{8.2}$$

The reflectance of the coupled surface–aerosol system can be obtained by adding the successive reflections that come out of the system:

$$\begin{aligned} R_{as} &= R_a + T_a\, R_s\, T_a + T_a\, R_s\, R_a\, R_s\, T_a + T_a\, (R_s\, R_a)^2\, R_s\, T_a + \ldots \\ &= R_a + T_a^2\, R_s \sum_{i=0}^{\infty} (R_s\, R_a)^i \\ &= R_a + \frac{T_a^2\, R_s}{1 - R_s\, R_a} \\ &= R_a + \frac{(1 - R_a)^2\, R_s}{1 - R_s\, R_a} \end{aligned} \tag{8.3}$$

where it has been assumed that the reflectance of the aerosol layer is the same for the

upward and downward solar radiation. The change in reflectance due to the aerosol layer is equal to:

$$\Delta R_{as} = R_{as} - R_s = R_a + \frac{(1 - R_a)^2 R_s}{1 - R_s R_a} - R_s \approx R_a (1 - R_s)^2 \tag{8.4}$$

under the assumption that $R_a \ll 1$.

The change in reflected solar radiation caused by the aerosol layer at a given place and time can therefore be approximated as:

$$\Delta F = -S \cos\theta_0 T^2 \Delta R_{as} \approx -S (1 - R_s)^2 T^2 \tau_a \beta_a. \tag{8.5}$$

It can be seen that the radiative effect increases linearly with the aerosol optical depth and with the upscatter fraction, and tends to zero when the surface albedo tends to 1 (i.e. a fully reflective surface). If we consider that in first approximation the radiative forcing due to aerosol–radiation interactions is nonexistent in cloudy conditions and that the various parameters involved in its calculations are not correlated, Eq. 8.5 can be integrated in space and time to obtain an estimate of the global-mean radiative forcing by aerosols:

$$\overline{\Delta F} = -\frac{1}{2} S \overline{T}^2 (1 - \overline{A_c}) (1 - \overline{R_s})^2 \overline{\beta_a} \, \overline{\tau} \tag{8.6}$$

where A_c is the cloud cover. Bars indicate an average over the range of atmospheric conditions that exist over the globe, except for $\overline{\beta_a}$ that represents the isotropic upscatter fraction as defined in Chap. 3. The factor ½ appears because every single point of the globe is in the night on average half of the year. This formulation, which is due to Charlson et al. (1991), is attractive for its simplicity but remains very approximative and neglects the potential correlations between the various parameters.

For the sulphate aerosol, which historically has been the first one to be studied for its link to climate and climate change, it is possible to link τ_a to the properties of the sulphate aerosol and sulphur cycle, so that:

$$\Delta F = \underbrace{-\frac{1}{2} S (1 - A_c) T^2 (1 - R_s)^2}_{\text{geophysical}} \underbrace{\overline{\beta} \, \alpha_{SO_4^{2-}} f(\text{RH})}_{\text{microphysical}} \underbrace{Q_{SO_2} Y_{SO_4^{2-}} \tau_{SO_4^{2-}}}_{\text{geochemical}} \tag{8.7}$$

where $\alpha_{SO_4^{2-}}$ is the mass scattering efficiency at 30 % relative humidity (i.e. in dry conditions), $f(\text{RH})$ a corrective factor that takes into account the growth of the aerosols with relative humidity, Q_{SO_2} (in moles S/year) the anthropogenic flux of SO_2 into the atmosphere, $Y_{SO_4^{2-}}$ the fraction of SO_2 that is converted into sulphate aerosol, and $\tau_{SO_4^{2-}}$ is the residence time (in years) of sulphate aerosols in the atmosphere. We have omitted the bars here (except for $\overline{\beta}$) for the sake of simplicity. The horizontal brackets isolate (in a nonexclusive way) the different types of parameters involved in this calculation: geophysical, microphysical, and geochemical. It is therefore important to be able to characterize with precision all these parameters and their spatial and temporal covariations in order to estimate the radiative forcing due to the interactions between sulphate aerosols and solar radiation.

8.2.2 Simplified Equation for Absorbing Aerosols

The simple Eqs. 8.5 and 8.6 are only valid for a non-absorbing aerosol. We know however that the atmospheric aerosol can be more or less absorbing, in particular because of the presence of black carbon in combustion aerosols.

The previous formulation can be generalized so as to include the effect of aerosol absorption (Haywood and Shine 1995). In this case, the reflectance and transmittance can be written as:

$$\begin{cases} R_a \approx \beta_a\, \varpi_a\, \tau_a / \cos\theta_0 \\ T_a = 1 - R_a - A_a \approx 1 - \beta_a\, \varpi_a\, \tau_a / \cos\theta_0 - (1 - \varpi_a)\, \tau_a / \cos\theta_0 \end{cases} \tag{8.8}$$

where A_a is the absorptivity in the aerosol layer and ϖ_a is the single scattering albedo of the aerosol. Equation 8.4 becomes:

$$\Delta R_{as} = R_a + \frac{T_a^2\, R_s}{1 - R_s\, R_a} - R_s \approx R_a \left((1 - R_s)^2 - 2 R_s\, \frac{A_a}{R_a} \right). \tag{8.9}$$

The instantaneous radiative forcing for a thin aerosol layer in clear-sky conditions can then be approximated in the following way:

$$\Delta F \approx -S\, T^2\, \varpi_a\, \beta_a\, \tau_a \left((1 - R_s)^2 - \frac{2\, R_s\, (1 - \varpi_a)}{\beta_a\, \varpi_a} \right). \tag{8.10}$$

It can be seen that this expression reduces to Eq. 8.5 when $\varpi_a = 1$.

It now appears that it is the triplet of parameters (β_a, ϖ_a, R_s) that determines the sign of the radiative effect due to aerosol–radiation interactions. The radiative effect of aerosols is more likely to be positive—-which induces a warming of the climate system—if the surface albedo R_s is large, the upscatter fraction β_a is small or the single scattering albedo ϖ_a is small. It can be seen that Eq. 8.10 produces a radiative forcing that is positive when the aerosol single scattering albedo is above a critical value:

$$\varpi_a < \varpi_a^c = \frac{2\, R_s}{\beta_a\, (1 - R_s)^2 + 2\, R_s}. \tag{8.11}$$

The critical aerosol single scattering albedo, ϖ_a^c, decreases when β_a increases and/or when R_s decreases. If $R_s = 0$, then $\varpi_a^c = 0$ and the aerosol can only increase the reflectance of the surface–atmosphere–aerosol system. It is easy to induce a positive aerosol radiative effect above a reflective surface such as a snow-covered surface, sea ice or if there is a cloud layer below the aerosol layer. The relative positions of the absorbing aerosol and cloud layers are therefore critical in the calculation of the aerosol radiative effect.

Embedded in the previous formulations are the assumptions that (i) transmittances of the atmospheric layer are the same for upward and downward solar radiation (ii) the reflectivities of the aerosol layer are the same for the upward and downward

solar radiation, and (iii) it is possible to consider averaged quantities over the solar spectrum. None of these assumptions is really justified. A more general formulation has to take into account the spectral variations of the aerosol, surface and atmospheric properties, and differentiate upward and downward transmittances and reflectances:

$$\Delta F(\mu_0) = -\int \mu_0 \, F_T(\lambda) \, T^{\downarrow}(\lambda, \mu_0) \, \Delta R_{as}(\lambda, \mu_0) \, T^{\uparrow}(\lambda) \, \mathrm{d}\lambda \tag{8.12}$$

where μ_0 is the cosine of the solar zenith angle, $\int F_T(\lambda)\,\mathrm{d}\lambda$ is the solar constant, $T^{\downarrow}(\lambda, \mu_0)$ and $T^{\uparrow}(\lambda)$ are the transmittances of the atmosphere located above the aerosol for downward solar radiation and for upward solar radiation scattered by the aerosol. The variation of the reflectance of the surface–atmosphere–aerosol system due to the presence of the aerosols is given by:

$$\Delta R_{as}(\lambda, \mu_0) = R_a(\mu_0) + \frac{T_a(\mu_0) \, R_s(\mu_0) \, \overline{T_a}}{1 - \overline{R_s} \, \overline{R_a}} - R_s(\mu_0) \tag{8.13}$$

where $R_a(\mu_0)$ is the reflectance of the aerosol layer for incident (direct) solar radiation which can be linked to the monodirectional upscatter fraction $\beta_a(\theta_0)$, $\overline{R_a}$ is the reflectance of the aerosol layer for diffuse solar radiation which can be linked to the isotropic upscatter fraction $\overline{\beta_a}$, $R_s(\mu_0)$ and $\overline{R_s}$ are the surface reflectivities for incident and diffuse solar radiation, respectively.

8.2.3 Radiative Transfer Calculations

8.2.3.1 Solar Radiation

The simplified formulations presented above have their limitations and an accurate numerical method, to resolve this the radiative transfer equation is required for calculating the radiative effects due to aerosol–radiation interactions. Figures 8.3 and 8.4 show how the aerosol radiative effect varies as a function of the solar zenith angle, surface albedo, and aerosol optical depth. These state-of-the-art calculations were made by Bellouin et al. (2005) for different representative aerosol types: sea spray, desert dust, pollution aerosol with little absorption, biomass burning aerosol.

It can be observed on Fig. 8.3a how the aerosol radiative effect, estimated at the top atmosphere, depends on the solar zenith angle. For a relatively dark surface such as the ocean, the radiative effect goes through a minimum (i.e. a maximum in reflectance) for a solar zenith angle of about 60–70°, which confirms the qualitative interpretation made earlier. For a brighter surface ($R_s = 0.5$), the minimum occurs for a larger solar zenith angle and the radiative effect gets positive for smaller solar zenith angles (i.e. when the sun is higher in the sky). It can also be noticed that for a given aerosol optical depth, sea spray and desert dust aerosols exert larger radiative effects than accumulation mode aerosols despite a lower upscatter fraction. This is because accumulation mode aerosols have an optical depth that decreases fairly

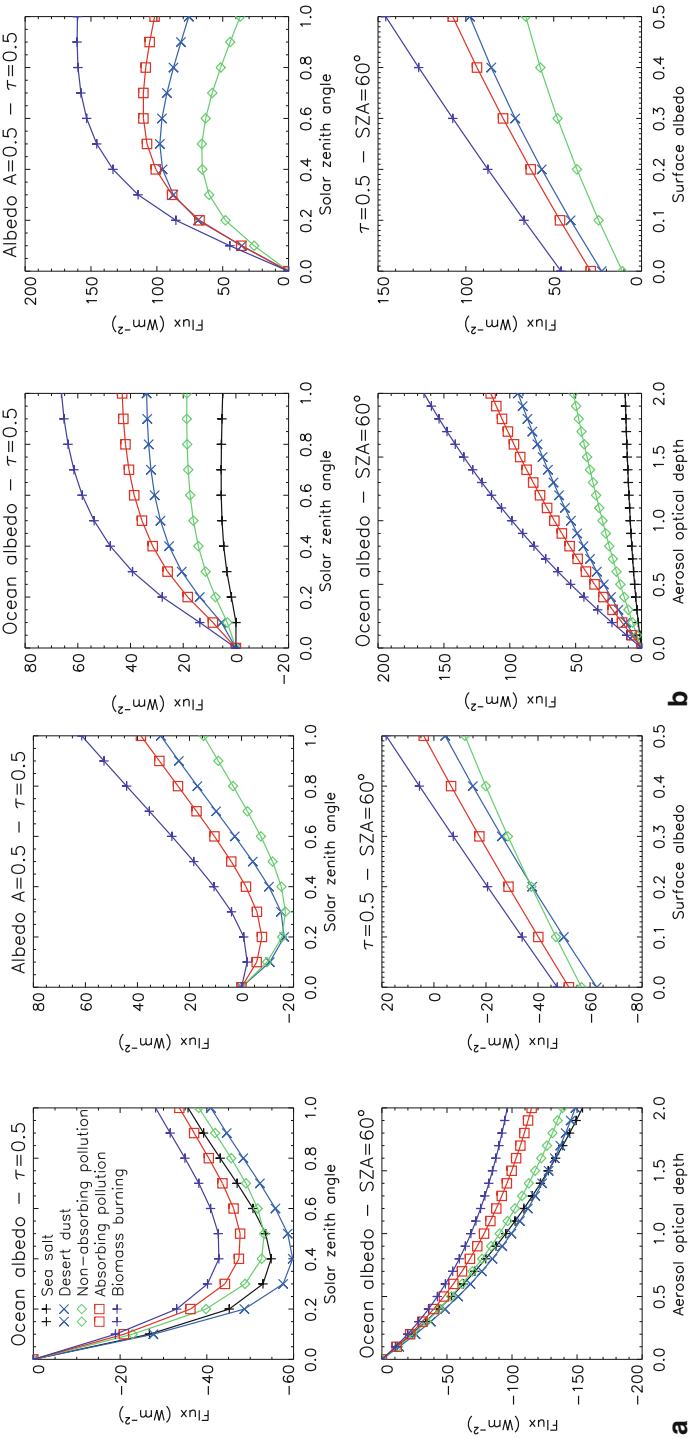


Fig. 8.3 Change in shortwave radiative fluxes (W m^{-2}) due to aerosol–radiation interactions as simulated by the STREAMER radiative transfer model for different aerosol types. The changes in radiative fluxes are for **a** the top of the atmosphere and **b** the atmosphere (i.e. the difference between the top and bottom of the atmosphere); they are plotted as a function of the solar zenith angle (SZA) for a given surface albedo and aerosol optical depth, as a function of the aerosol optical depth for a given surface albedo and SZA, and as a function of the surface albedo for a given aerosol optical depth and SZA. The aerosol optical depth is given at 550 nm but varies with the wavelength. The results for sea-salt aerosols are only showed for the surface albedo of an oceanic surface (that varies accordingly with SZA). (Drawn from calculations made by Bellouin et al. (2005))

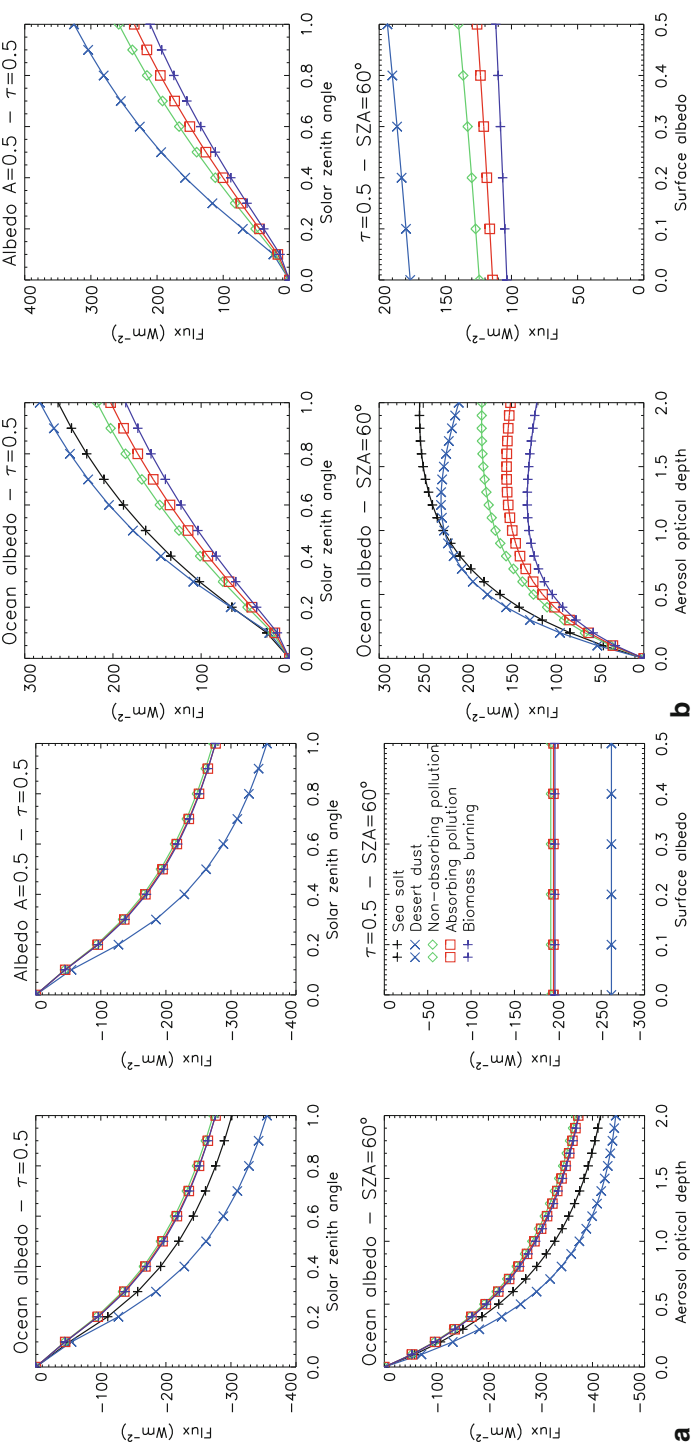


Fig. 8.4 Change in shortwave radiative fluxes (W m^{-2}) at the surface due to aerosol–radiation interactions as simulated by the STREAMER radiative transfer model for different aerosol types for **a** direct radiation and **b** diffuse radiation. The aerosol optical depth is given at 550 nm but varies with the wavelength. Results are presented as in Fig. 8.3. (Drawn from calculations made by Bellouin et al. (2005))

rapidly in the near infrared part of the spectrum in contrast to sea spray and desert dust aerosols whose optical depth exhibits much less spectral variation. The dependence on surface albedo is also consistent with that of the simplified formulations discussed previously. It can be seen however that the radiative effect does not respond linearly to the aerosol optical depth, which can be explained by the saturation effect due to multiple scattering.

Figures 8.3b, 8.4a and b further show the atmospheric and surface aerosol radiative effect for incident (direct) and diffuse solar radiation. Aerosols contribute to decrease direct solar radiation at the surface but increase diffuse solar radiation. Their net effect is to decrease the total (direct and diffuse) amount of solar radiation received at the surface. A fraction of the radiative energy that is not absorbed by the surface is in fact absorbed by the atmosphere. The atmospheric radiative effect is defined as the difference between the top-of-atmosphere and surface radiative effects:

$$\Delta F_{\text{atm}} = \Delta F_{\text{toa}} - \Delta F_{\text{surf}} = \Delta F_{\text{toa}} - \Delta F_{\text{surf direct}} - \Delta F_{\text{surf diffuse}} \tag{8.14}$$

It will be shown in Chap. 10 that the aerosol atmospheric radiative forcing (which is of positive sign) controls the response of the global water cycle to the aerosol forcing. The top-of-atmosphere radiative forcing (which can be of positive or negative signs) controls the global mean surface temperature change, as long as aerosols are located in the troposphere. This is because the surface and the atmosphere are in radiative-convective equilibrium and it is therefore the sum of the surface and atmospheric radiative forcings that matters for the temperature response.

8.2.3.2 Infrared Radiation

Aerosols are also responsible for a greenhouse effect whose strength depends on the concentrations of aerosols, their optical properties in the infrared part of the spectrum, their vertical profile and environmental conditions (temperature and emissivity of the surface, vertical temperature, and humidity profiles, cloud amount and properties). Large concentrations of aerosols are required for the infrared radiative impact of aerosols to be significant. It is also necessary that aerosols have an altitude that is high enough to cause a temperature difference between the surface and the aerosol layer. In practice, only the infrared radiative impacts of desert dust aerosols—-whose mass in the atmosphere is large—and stratospheric volcanic aerosols—whose altitude is high—are significant. For stratospheric volcanic aerosols, the negative radiative effect in the solar spectrum (parasol effect) is larger in magnitude than the positive radiative effect (greenhouse effect) in the infrared spectrum and dominates the net radiative effect. For desert dust, the greenhouse effect dominates over land, while the parasol effect can dominate over ocean where the surface albedo is rather small.

8.2.4 Global Estimates and Sources of Uncertainty

The radiative impact of aerosols depends on the aerosol properties but also the characteristics of the incident solar radiation. At first order, the incident solar radiation is characterized by the spatial and temporal distribution of solar zenith angle that varies as a function of latitude, time of day and day in the year. The solar zenith angle takes larger values on average at high latitudes than at low latitudes. As the radiative impact of aerosols goes through a minimum for solar zenith angles of 60–70°, this compensates for the somewhat smaller amount of radiative energy received at high latitudes. Figure 8.5a shows the evolution along the year of the radiative impact of a uniform distribution of aerosols in the northern hemisphere. One can visualize that the radiative impact has the largest magnitude at high latitudes during summertime, smallest magnitude during wintertime, and is relatively constant in the tropics. Figure 8.5b shows the case of an aerosol that has a latitudinal distribution that varies as $\sin(2\phi)$, where ϕ is the latitude, hence with concentration going through a maximum at a latitude of 45° and to zero at latitudes $\phi = 0$ and 90°. This distribution represents that of an idealized aerosol emitted in the northern hemisphere. Under these conditions, the magnitude of the radiative impact reaches a maximum between 30 and 60° depending on the season.

The Intergovernmental Panel on Climate Change (IPCC) uses available observations, global aerosol models, and expert judgment to estimate the radiative forcing due to aerosol–radiation interactions (RFari) and its uncertainty. We define RFari as the change in irradiance at the top of atmosphere due to changing aerosol atmospheric concentrations between the preindustrial period (typically 1750) and present-day prior to any rapid adjustments:

$$\begin{aligned} \Delta F_{\text{ari}} &= F_{\text{present-day aerosol}} - F_{\text{preindustrial aerosol}} \\ &= (F_{\text{present-day aerosol}} - F_{\text{no aerosol}}) \\ &\quad - (F_{\text{preindustrial aerosol}} - F_{\text{no aerosol}}). \end{aligned} \tag{8.15}$$

Such a quantity is typically calculated using double radiation calls in a climate or global aerosol model either with and without the anthropogenic aerosols or as a double difference against a no-aerosol reference, with tropospheric state held fixed. The IPCC assessment for RFari is summarized in Table 8.1 for the different anthropogenic aerosol species (Boucher et al. 2013). Sulphate, nitrate, and secondary organic aerosols contribute a negative RFari, while black carbon aerosols contribute a positive RFari. Biomass burning aerosols are partially absorbing so that their RFari can be either positive or negative. Anthropogenic activities may have contributed to the desertification in some regions such as the Sahel, which has probably led to an increase in desert dust emissions (Ginoux et al. 2012). There are also natural causes for variations in desert dust emissions and it is difficult to disentangle the two, which is why the RFari reported by the IPCC may include a contribution from climate feedbacks. Desert dust aerosols are responsible for a greenhouse effect which partly compensates for their negative radiative effect in the solar spectrum.

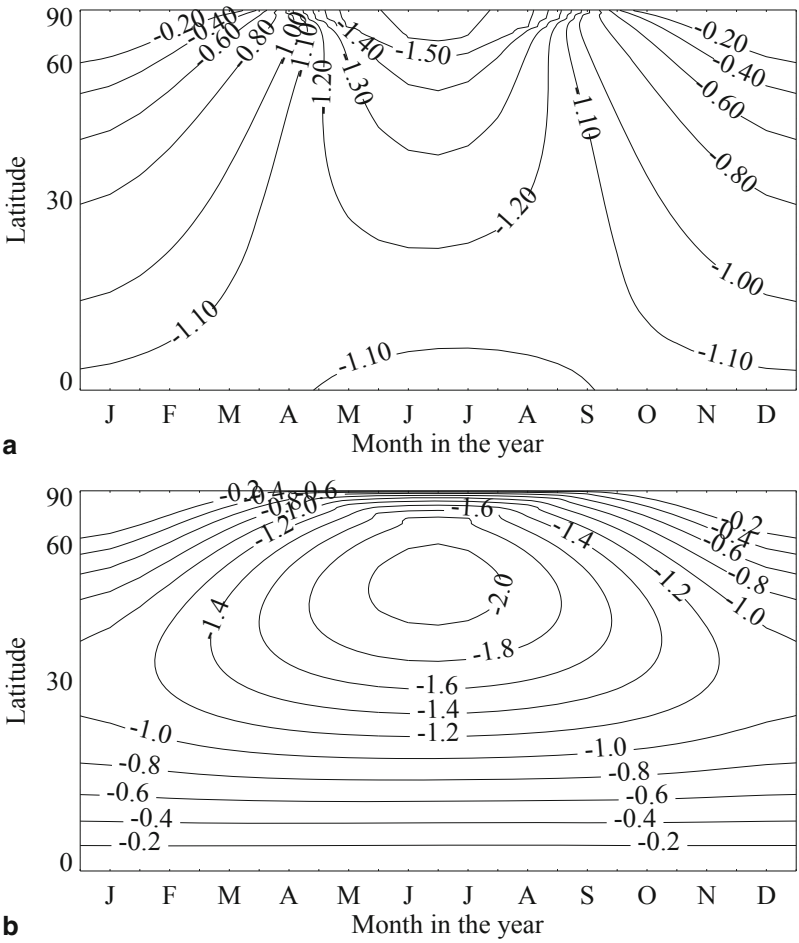


Fig. 8.5 Radiative forcing (W m^{-2}) due to aerosol–radiation interactions in clear sky as a function of the latitude and month in the year for **a** a uniform distribution of the aerosol optical depth and **b** a latitudinal distribution of the aerosol optical depth which is taken proportional to sin (2ϕ) where ϕ is the latitude. In both cases, the radiative forcing is normalized to an arbitrary northern hemisphere average of -1 W m^{-2}

Estimates presented in Table 8.1 are subject to relatively large uncertainties. These are dominated by uncertainties on the aerosol concentrations and optical properties. The relative positions of clouds and aerosols is also a major cause of uncertainty in the calculations of the radiative forcing by absorbing aerosols. Surprisingly, atmospheric radiative transfer is also a source of uncertainty because of assumptions on how to handle environmental conditions (surface albedo and clouds), the way to resolve the atmospheric radiative transfer equation, and to integrate it over the solar spectrum (Stier et al. 2013).

Table 8.1 Global-mean radiative forcing due to aerosol–radiation interactions. The uncertainties quoted span the 5–95 % confidence range. Note that the species breakdown is less certain than the total RFari and does not sum to the total exactly. (From Boucher et al. (2013))

Aerosol type	Net radiative forcing (W m^{-2})	Uncertainty range (W m^{-2})
Fossil-fuel sulphates	−0.40	−0.6 to −0.2
Fossil-fuel and biofuel black carbon	+0.40	+0.05 to +0.6
Fossil-fuel and biofuel organic aerosols	−0.12	−0.4 to +0.1
Nitrates	−0.11	−0.3 to −0.03
Biomass burning	+0.0	−0.2 to +0.2
Desert dust	−0.10	−0.3 to +0.1
All aerosols	−0.35	−0.85 to +0.15

8.3 Rapid Adjustments to Aerosol–Radiation Interactions

Aerosol–radiation interactions are responsible for a heating rate in the troposphere that modifies the vertical temperature profile and hence atmospheric stability and relative humidity with a potential impact on cloud lifecycle. This effect, previously called the aerosol semi-direct effect, is better thought of as a rapid adjustment to aerosol–radiation interactions. A change in irradiance can be associated with such rapid adjustments although it is more difficult to estimate. In practice it is easier to estimate the sum of the radiative forcing and associated rapid adjustments (known as the effective radiative forcing due to aerosol–radiation interactions or ERFari) as the change in top of atmosphere irradiance between two fixed-sea surface temperature experiments (Sherwood et al. 2015).

It was initially argued by Hansen et al. (1997) that rapid adjustments due to black carbon aerosols could dissipate or "burn of" low-level clouds by heating the atmosphere. This would cause a positive contribution to the effective radiative forcing as low-level clouds tend to have a negative cloud radiative effect. Further, research has shown that several processes were at play with different possible impacts on cloudiness.

Johnson et al. (2004) used a limited area model to simulate a stratocumulus field and introduced a layer of absorbing aerosols at different heights. They showed that aerosol absorption could either strengthen or weaken the cloud layer depending on whether the aerosol is located above or within the cloud field. When the aerosol is located above the stratocumulus cloud, the local heating caused by the aerosol reinforces the temperature inversion at the cloud top, which decreases the entrainment of dry tropospheric air in boundary layer clouds and slows down or prevents their dissipation. Inversely, if the aerosol is located in the stratocumulus, it heats the cloud locally, which for a fixed water vapour content, contributes to decrease relative humidity, and cause the cloud to evaporate. The heating also decreases the temperature inversion with an opposite effect to that induced by an aerosol located

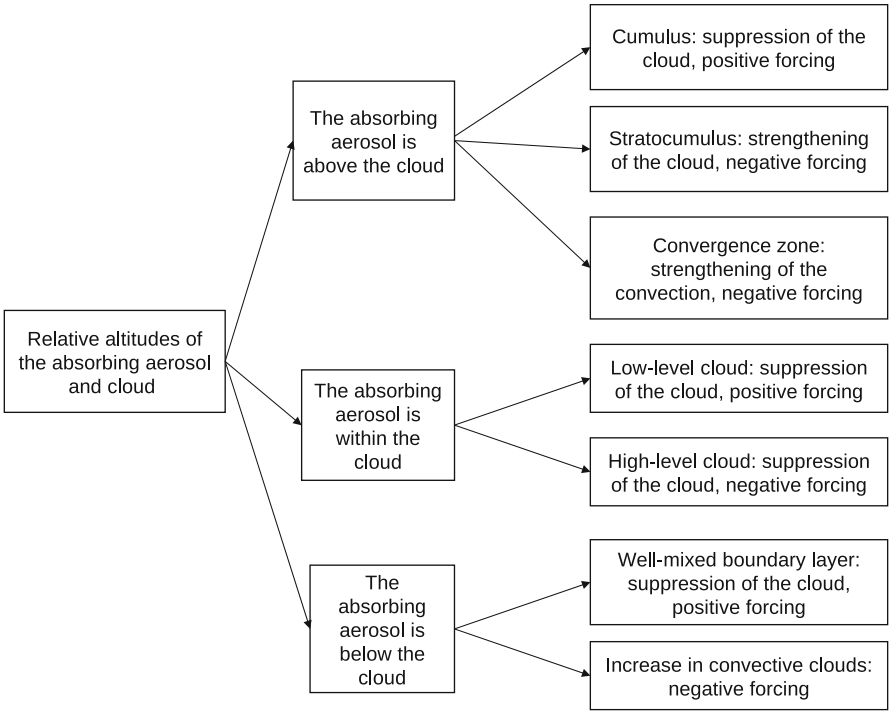


Fig. 8.6 Mechanisms operating in response to aerosol absorption in the atmosphere. (Adapted from Koch and Del Genio (2010). See text for a discussion of the processes involved)

above the cloud. When the aerosol is located under the cloud, the effects are similar to that of an aerosol located within the cloud, as in both cases the heating takes place in the boundary layer which is relatively well-mixed.

Koch and Del Genio (2010) have conceptualized all the different rapid adjustments that follow the absorption of solar radiation by aerosols (Fig. 8.6). The most important factor is the relative position of the aerosol and cloud layers. When the aerosol is located above the cloud, the effect also depends on the cloud types and meteorological situation. In the case of a stratocumulus cloud characterized by a large radiative cooling and temperature inversion at the cloud top, aerosol absorption reinforces the cloud stability, which brings a negative contribution to the effective radiative forcing (ERFari). This situation can occur off the coast of Namibia, where stratocumulus can be seasonally influenced by biomass burning aerosols that are exported from the African continent. In the case of a less stable atmosphere, the aerosol absorption can increase the large-scale upward atmospheric motion and cloud formation, which also contributes a negative contribution to ERFari. When the aerosol is located within the cloud, one would usually expect a reduction of the cloud cover with an ERFari contribution that can vary depending on the aerosol height. When the aerosol is located below the cloud, two effects are possible depending on the degree of stratification of

the boundary layer. If the boundary layer is well-mixed, the cloud dissipation effect prevails, which leads to a positive contribution to the ERFari. If the boundary layer is less well-mixed, aerosol absorption can help to reinforce the convection and bring moisture higherup, thus increasing low-level clouds, which leads to a negative contribution to ERFari. In any case these processes are uncertain and remain poorly represented in models, especially in climate models where a lot of these processes occur at the subgrid scale and need to be parametrized. Large eddy simulations and cloud-resolving models are useful to constrain large-scale models on particular case studies.

8.4 Radiative Impact of Aerosols on Surface Snow and Ice

Aerosols get deposited on land surfaces, oceans and sea ice and their concentration in the snowpack can be measured (Doherty et al. 2010). Absorbing aerosols that are deposited on a snow- or ice-covered surface can modify the albedo of that surface as shown by laboratory measurements (Hadley and Kirchstetter 2012). Snow and ice are very reflective in the ultraviolet and visible part of the solar spectrum. Aerosol absorption at the surface and in the topmost few centimetres of the ice and snowpack decreases their albedo. This effect, called aerosol effect on snow and ice, contributes to heat the climate system and therefore a positive radiative forcing. The reader can refer to the original article by Warren and Wiscombe (1980) and to Appendix D for the description of a simplified radiative transfer model that can quantify the albedo change as a function of physical and optical properties of the snowpack and aerosols. In this chapter, we only describe the process in a more qualitative manner.

Many parameters are needed to estimate this effect. From a purely radiative point of view, the reduction in surface albedo due to the presence of absorbing aerosols depends on the vertical profile of the aerosol that gets deposited and mixed in the snowpack, the amount of aerosol absorption, the size of the snow or ice crystals, the presence of other aerosols and impurities in the snowpack, and the characteristics of the incident solar radiation at the surface (solar zenith angle and distribution between direct and diffuse radiation). Scattering by the aerosols does not matter much because snow and ice are very scattering media themselves. The aerosol parameter that is the most relevant to quantify is the aerosol absorption coefficient (or aerosol absorption optical depth if integrated over a given snow depth). In contrast to aerosol–radiation interactions in the atmosphere, there is no critical single scattering albedo above which the sign of the effect becomes negative (say differently the critical single scattering albedo is large enough to be of no practical use). Some aerosol components, which absorb solar radiation only slightly, can have a cooling effect when they are in the atmosphere, but a warming effect once they are deposited in the snowpack. The reduction in surface albedo is less when snow crystals are larger or in the presence of impurities. For instance, the effect of black carbon will be less if it gets deposited on a snowpack that already contains another absorbing aerosols, such as mineral dust or biomass burning aerosols. Finally, if the snowpack is relatively thin, the albedo reduction also depends on the thickness of the snow layer and the albedo of the

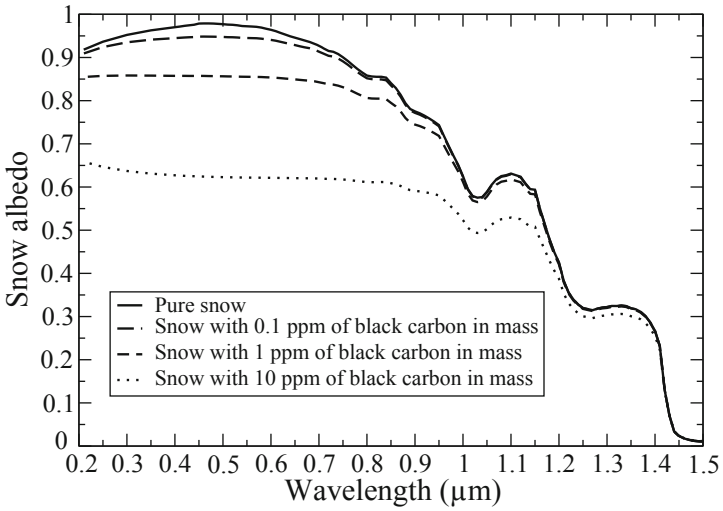


Fig. 8.7 Snow albedo as a function of the wavelength for four different mixing ratios of black carbon in snow. We have used here the model presented in Appendix D. The parameters used for these calculations are: snow depth $L = 10$ mm equivalent water, spherical crystal with a radius $R = 100$ μm, solar zenith angle $\theta_0 = 60°$, an albedo of the underlying surface set to 0

underlying surface. Figure 8.7 illustrates the spectral variations of the albedo of a snow layer in the presence of variable quantities of black carbon. There are a number of processes that can amplify the initial radiative effect due to the darkening of the snowpack induced by black carbon and/or mineral dust. Figure 8.8 illustrates the processes at stake and highlights three feedback loops denoted a, b, and c.

a) First, the darkening of the snow increases absorption of solar radiation in the upper part of the snowpack, which heats up the snow and fastens its aging. In particular snow and ice crystals become larger, which decreases the reflectance of the snow and amplifies the initial radiative forcing.

b) Second, the heating within and above the snowpack increases the melting and sublimation of the snow. This has two distinct effects. On the one hand, the absorbing aerosols (e.g. black carbon) can accumulate at the top of the snowpack, where their radiative effect is maximum. On the other hand, the heating decreases the thickness of the snow layer, and the surface can become clear of snow in places, which uncovers the underlying surface and decreases the surface albedo. This second effect is particularly important when the snow layer depth is small. This forms a second feedback that amplifies the initial radiative forcing.

c) Third, the aerosol effect on snow can modify the climate locally or regionally. A warming can modify circulation and precipitation patterns, cause an increase in liquid precipitation (rain) at the expense of solid precipitation (snow), which once again can contribute to melt the snowpack and decrease the surface albedo (Flanner et al. 2009; Bellouin et al. 2010).

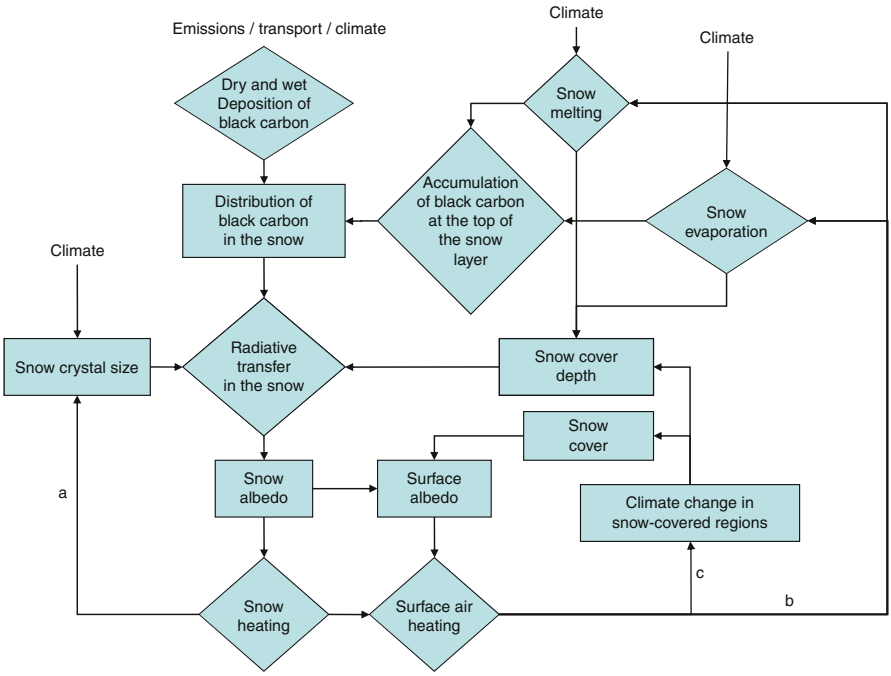


Fig. 8.8 Processes controlling feedbacks of absorbing aerosols on the snowpack and the surface albedo. Physical processes are indicated with *diamonds*, while physical parameters are indicated with *rectangles*. The letters *a*, *b*, and *c* show the three main feedback loops which are described in the text

Once the albedo reduction is known, the radiative forcing at the top of atmosphere due to effect of aerosols on the snow and ice layer can be estimated by doing an atmospheric radiative transfer calculation. The seasonal variation of this radiative effect is important because the regions of snow and ice are found preferentially at high latitudes where there is a strong seasonal cycle in insolation with little or no solar radiation during the winter. Flanner et al. (2007) estimated that the radiative forcing due to the effect of black carbon on snow surfaces is of the order of 0.05 W m^{-2}. Boucher et al. (2013) reviewed existing studies to assess an effective radiative forcing (i.e. including rapid adjustments) of 0.04 W m^{-2} with a 90 % uncertainty range of 0.02–0.09 W m^{-2}. It should be noted however that this forcing mechanism does not lend itself very well to the forcing–adjustment–feedback paradigm as there is a continuum of timescales involved (typically from 1 h to one season). Among the processes discussed above, most act through a local rather than a global change in surface temperature, which makes it difficult to distinguish adjustments from feedbacks in this case. As a consequence it is not easy to differentiate the radiative forcing from the effective radiative forcing for this mechanism and a climate efficacy might still be needed (see Chap. 10).

In conclusion, there are several very strong feedback loops that amplify the initial perturbation induced by the deposition of absorbing aerosols on snow and ice surfaces. It can be estimated that the climate response by unit (effective) radiative forcing is 2–4 times more efficient for the aerosols on snow and ice forcing than for other forcing mechanisms (see Chap. 10).

Exercises

1. Use Eq. 8.10 to estimate the instantaneous radiative effect due to aerosol–radiation interactions at the top of atmosphere for an aerosol layer with optical depth 0.10, single scattering albedo 0.95, upscatter fraction 0.15. The surface albedo and atmospheric transmittance are 0.10 and 0.80, respectively. What is the impact of an uncertainty of ± 0.01 on each of these parameters?

2. We now consider the fact that the upscatter fraction varies with the solar zenith angle and that it is different for downward and upward solar radiation. This implies that the radiative effect due to aerosol–radiation interactions depends on the solar zenith angle. Estimate the aerosol radiative effect for the cases where $\mu_0 = 0.5$, 0.75, and 1. Use $\beta_a(0.5) = 0.25$, $\beta_a(0.75) = 0.15$, and $\beta_a(1) = 0.10$ in rough agreement with Fig. 3.8. The other parameters are $\overline{\beta_a} = 0.20$, $\varpi_a = 0.95$, $T_a(\mu_0) = \overline{T_a} = 0.80$, and $R_s(\mu_0) = \overline{R_s} = 0.10$. Use Eq. 8.13. Why is this model not appropriate for larger solar zenith angles?

3. Use Appendix D and write a computer programme that computes the albedo of a snow layer, $a_s(\mu_0)$, as a function of the snow optical depth τ (ranging from 0.1 to 1000) for the following values of the single scattering co-albedo, $1 - \varpi = 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}$, and 10^{-1}. Use an asymmetry parameter g equals to 0.9 and a solar zenith angle of 60°. For which value of τ does the albedo of the snow layer become within 0.01 of its asymptotic value?

Solutions

1. $\Delta F = -9.2\ \mathrm{W\ m^{-2}}$

$\delta\varpi_a = \pm 0.01 \longrightarrow \delta\Delta F = \pm 0.3\ \mathrm{W\ m^{-2}}$

$\delta\beta_a = \pm 0.01 \longrightarrow \delta\Delta F = \pm 0.7\ \mathrm{W\ m^{-2}}$

$\delta\tau_a = \pm 0.01 \longrightarrow \delta\Delta F = \pm 0.9\ \mathrm{W\ m^{-2}}$

$\delta R_s = \pm 0.01 \longrightarrow \delta\Delta F = \pm 0.3\ \mathrm{W\ m^{-2}}$

$\delta T = \pm 0.01 \longrightarrow \delta\Delta F = \pm 0.2\ \mathrm{W\ m^{-2}}$

2. $R_a = \beta_a(\mu_0)\,\varpi_a\,\tau_a/\cos\theta_0$, $\overline{R_a} = 2\,\overline{\beta_a}\,\varpi_a\,\tau_a$ given the fact that the average of the cosine of solar zenith angle for isotropic diffuse radiation is equal to 0.5.

$\Delta F(\mu_0 = 0.5) = -16\ \mathrm{W\ m^{-2}}$

$\Delta F(\mu_0 = 0.75) = -8\ \mathrm{W\ m^{-2}}$

$\Delta F(\mu_0 = 1) = -3\ \mathrm{W\ m^{-2}}$

This model does not take into account multiple scattering that becomes important and reduces the radiative effect due to aerosol–radiation interactions when the solar zenith angle increases.

References

Bellouin N, Boucher O (March 2010) Climate response and efficacy of snow albedo forcing in the HadGEM2-AML climate model, Hadley Centre Technical Note, HCTN82

Bellouin N, Boucher O, Haywood J, Reddy MS (2005) Estimate of aerosol direct effects over land and oceans from MODIS. Nature 438:1138–1141

Boucher O, Randall D, Artaxo P, Bretherton C, Feingold G, Forster P, Kerminen V-M, Kondo Y, Liao H, Lohmann U, Rasch P, Satheesh SK, Sherwood S, Stevens B, Zhang XY (2013) Clouds and aerosols. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, pp 571–657

Carslaw KS, Lee LA, Reddington CL, Pringle KJ, Rap A, Forster PM, Mann GW, Spracklen DV, Woodhouse MT, Regayre LA, Pierce JR (2013) Large contribution of natural aerosols to uncertainty in indirect forcing. Nature 503:67–71

Charlson RJ, Langner J, Rodhe H, Leovy CB, Warren SG (1991) Perturbation of the northern hemisphere radiative balance by backscattering from anthropogenic sulfate aerosols. Tellus 43B:152–163

Doherty SJ, Warren SG, Grenfell TC, Clarke AD, Brandt RE (2010) Light-absorbing impurities in Arctic snow. Atmos Chem Phys 10:11647–11680

Flanner MG, Zender CS, Randerson JT, Rasch PJ (2007) Present-day climate forcing and response from black carbon in snow. J Geophys Res 112:D11202. doi:10.1029/2006JD008003

Flanner MG, Zender CS, Hess PG, Mahowald NM, Painter TH, Ramanathan V, Rasch PJ (2009) Springtime warming and reduced snow cover from carbonaceous particles. Atmos Chem Phys 9:2481–2497

Ginoux P, Prospero JM, Gill TE, Hsu NC, Zhao M (2012) Global-scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS Deep Blue aerosol products. Rev Geophys 50:RG3005. doi:10.1029/2012RG000388

Hadley OL, Kirchstetter TW (2012) Black-carbon reduction of snow albedo. Nat Clim Change 2:437–440

Hansen JE, Sato M, Ruedy R (1997) Radiative forcing and climate response. J Geophys Res 102:6831–6864

Haywood JM, Shine KP (1995) The effect of anthropogenic sulfate and soot aerosol on the clear sky planetary radiation budget. Geophys Res Lett 22:603–606

Johnson BT, Shine KP, Forster PM (2004) The semi-direct aerosol effect: impact of absorbing aerosols on marine stratocumulus. Q J Royal Meteorol Soc 130:1407–1422

Koch D, Del Genio AD (2010) Black carbon semi-direct effects on cloud cover: review and synthesis. Atmos Chem Phys 10:7685–7696

Sherwood S, Bony S, Boucher O, Bretherton C, Forster P, Gregory J, Stevens B (2015) Adjustments to the forcing-feedback framework for understanding climate change. Bull Am Meteorol Soc 96. doi:10.1175/BAns-D-13-00167.1

Stier P, Schutgens NAJ, Bellouin N, Bian H, Boucher O, Chin M, Ghan S, Huneeus N, Kinne S, Lin G, Ma X, Myhre G, Penner JE, Randles CA, Samset B, Schulz M, Takemura T, Yu F, Yu H, Zhou C (2013) Host model uncertainties in aerosol forcing estimates: results from the AeroCom Prescribed Intercomparison Study. Atmos Chem Phys 13:3245–3270

Warren SG, Wiscombe WJ (1980) A model for the spectral albedo of snow. II: snow containing atmospheric aerosols. J Atmos Sci 37:2734–2745

Further Reading (Textbooks and Articles)

Bellouin N, Boucher O, Vesperini M, Tanré D (2004) Estimating the aerosol direct radiative perturbation: impact of the ocean surface representation and aerosol non-sphericity. Q J Royal Meteorol Soc 130:2217–2232

Haywood JM, Boucher O (2000) Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: a review. Rev Geophys 38:513–543

Kahn RA (2012) Reducing the uncertainties in direct aerosol radiative forcing. Surv Geophys 33:701–721

Chapter 9
Aerosol–Cloud Interactions

Abstract This chapter introduces the various concepts required to understand and assess the radiative forcing due to aerosol–cloud interactions starting from a short description of cloud distribution and properties. A short presentation of the Kelvin effect and Raoult's law leads into the Köhler theory, which provides the basis for the concepts of cloud condensation nucleus and aerosol activation into cloud droplet. A simple air parcel model is used to illustrate the sensitivity of cloud optical properties to the aerosol concentration. The role of rapid adjustments is then discussed as an integral part of the effective radiative forcing due to aerosol–cloud interactions. The discussion is then extended to cover mixed-phased and ice clouds, with an introduction to the concepts of ice nucleus, homogeneous and heterogeneous freezing, the Bergeron process and the possible aerosol invigoration of convective mixed-phased clouds. An assessment of the radiative forcing due to aerosol–cloud interaction is then presented. The chapter ends with a description of the physical mechanisms responsible for contrail formation and aviation-induced cirrus, and an estimate of the associated radiative forcing.

Keywords Cloud · Indirect effect · Radiative forcing · Cloud condensation nucleus · Ice nucleus · Supersaturation

9.1 Introduction

9.1.1 Cloud Formation

Clouds usually form in rising air, which cools when expanding, thus permitting the activation of aerosol particles into cloud droplets and ice crystals in supersaturated air. The evolution of a cloud is then governed by the balance between a number of dynamical, microphysical and radiative processes. Cloud particles of sufficient size become falling hydrometeors, which can be categorized as drizzle drops, raindrops, snow crystals, graupel and hailstones, according to their phase, size and shape. Precipitation formation (or lack of) plays an important role in the lifecycle of a cloud: it is primarily influenced by the distribution of moisture and cloudiness, and to a lesser extent by the concentrations and properties of aerosol particles. Cloud processes relevant to aerosol–cloud interactions are discussed in the relevant sections on liquid and ice clouds.

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_9

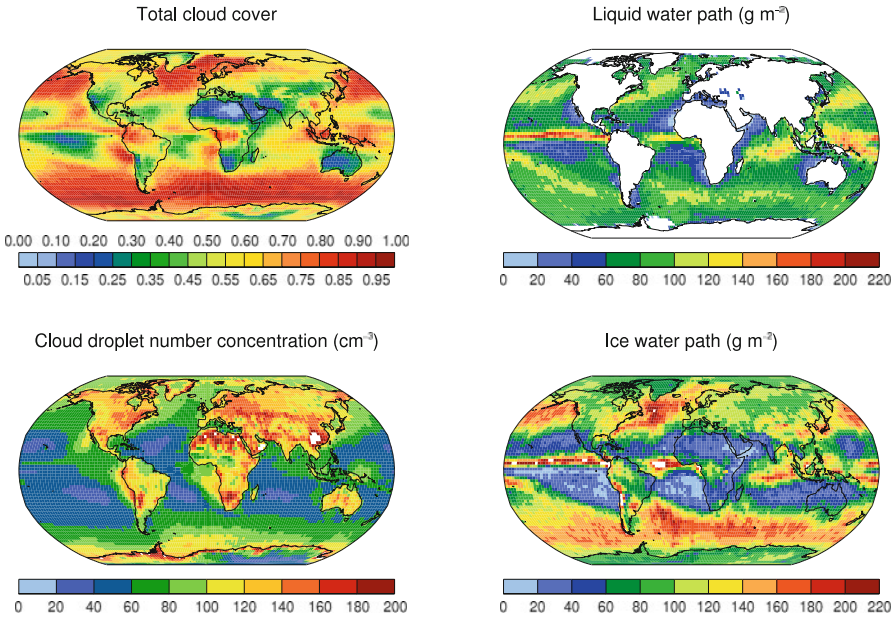


Fig. 9.1 Cloud cover, liquid water path ($g\,m^{-2}$), cloud droplet number concentration (cm^{-3}), and ice water path ($g\,m^{-2}$) averaged over the year 2003. The liquid water path field is from the AMSU instrument while the other fields are from MODIS

9.1.2 Cloud Distribution

Clouds cover approximately two thirds of the globe on average but a more precise value depends on what optical depth threshold is used to define a cloud (Stubenrauch et al. 2013). Cloud cover shows large regional variations that can be explained by the global atmospheric circulation (Fig. 9.1). The mid-latitude oceanic storm tracks and the tropical precipitation belts are particularly cloudy, while continental desert regions and the subsidence regions in the subtropics are relatively cloud-free. While cloud cover is relatively invariant at the global scale, it is much more variable at smaller spatial scales as it strongly depends on the meteorological conditions.

Clouds are composed of liquid water at atmospheric temperatures warmer than 0 °C, ice below about −35 °C, and either or both phases in between. The distributions of cloud liquid and ice water paths are shown on Fig. 9.1.

The effect of clouds on the Earth's radiative budget is measured by the *cloud radiative effect* (CRE) as defined in Eq. 2.2. The cloud radiative effect can be derived from satellite Earth's radiative budget missions by comparing upwelling radiation in cloudy and cloud-free conditions (Ramanathan et al. 1989; Loeb et al. 2009; Fig. 9.2). Clouds reflect solar radiation and enhance the planetary albedo. As a result they exert a global and annual mean shortwave cloud radiative effect, CRE_{SW} of about $-50\ W\,m^{-2}$. Clouds also contribute to the greenhouse effect, with a global

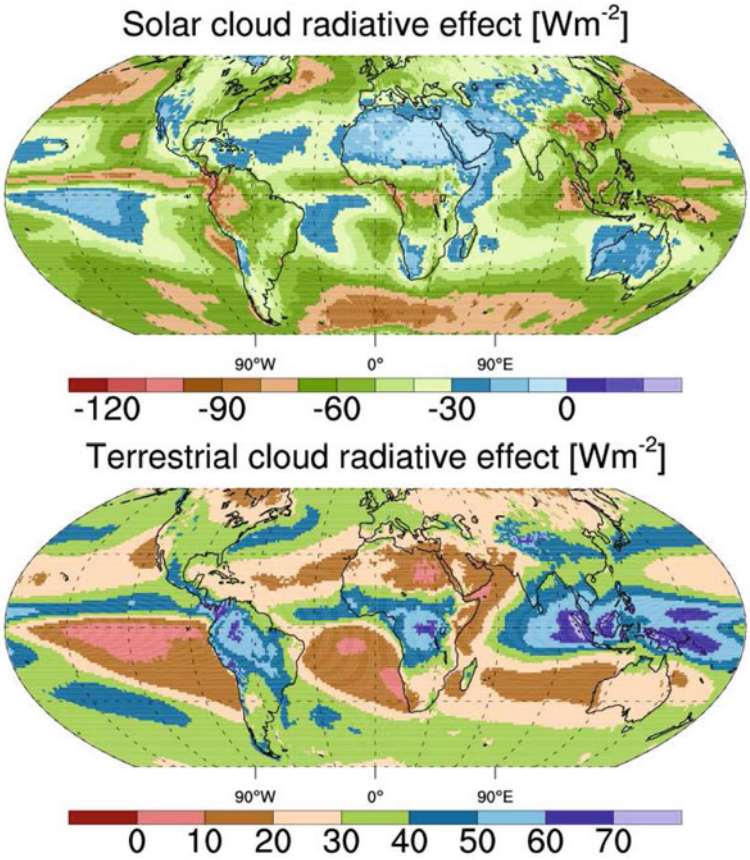


Fig. 9.2 Cloud radiative effects ($\mathrm{W\,m^{-2}}$) in the solar (shortwave) and terrestrial (longwave) part of the electromagnetic spectrum. (The data are from the CERES Terra data. Courtesy of Johannes Quaas)

and annual mean longwave cloud radiative effect, $\mathrm{CRE_{LW}}$, of about $+30\,\mathrm{W\,m^{-2}}$. The net cloud radiative effect is therefore negative, which implies that clouds contribute to cool the planet. Low-level clouds exert more of shortwave and much less of a longwave effect at the top of atmosphere. High-level clouds exert both a shortwave and a longwave effect, the balance between the two depending on the cloud optical thickness and altitude. The patterns in Fig. 9.2 are here to remind us that cloud formation and cloud properties are primarily governed by dynamical processes and that we are looking here at a second-order effect of aerosols on cloud properties.

9.1.3 *Aerosol–Cloud Interactions*

Aerosols influence cloud microphysical properties in a number of ways. These influences are largely known in the scientific literature as *aerosol indirect effects,* but for the sake of clarity are referred here to as aerosol–cloud interactions. Figure 9.3 illustrates the new paradigm discussed in Chap. 2 on naming and categorizing aerosol–cloud interactions. It is useful, at least conceptually, to distinguish some of the processes responsible for aerosol–cloud interactions, although it may be somewhat pointless to quantify each process independently from the others (Boucher et al. 2013).

- Aerosols act as cloud condensation nuclei (CCN) so that an increase in aerosol concentrations generally leads to an increase in the concentration of CCN and therefore, in the concentration of cloud droplets. This process is often referred to as *the aerosol first indirect effect*, *cloud albedo* or *Twomey effect* named after Sean Twomey who first observed the influence of aerosol pollution on clouds (Twomey 1959). This effect can be associated to a radiative forcing (RFaci) calculation, at least in principle if the preindustrial aerosol concentration is known. For a fixed cloud cover and liquid water content, an increase in cloud droplet concentration results in smaller cloud droplets but an increase in the total scattering cross section, and thus an increase in cloud reflectivity. Although only the change in cloud droplet concentration is considered in the original concept, a change in the shape (or dispersion) of the droplet size distribution that is directly induced by the aerosols may also play a role (Liu and Daum 2002).
- The Twomey effect does not occur in isolation. It can be seen as the trigger of a number of subsequent rapid adjustments involving the microphysics of liquid and mixed-phase (i.e. liquid and ice) clouds. These adjustments are multiple and can interact with each other. In particular it has been hypothesized that smaller cloud droplets would further impact on the cloud microphysical and dynamical evolution through a larger vertical and horizontal development of the cloud (Albrecht 1989), delayed precipitation (Givati and Rosenfeld 2004), enhanced entrainment and cloud evaporation (Small et al. 2009) and/or delayed freezing of the cloud droplets into ice crystals (Koren et al. 2005). The delayed freezing would result in the latent heat of freezing being deposited higher up in the cloud and additional condensation as the cloud develops to colder temperatures, with possible impacts on the strength of the convection. As these rapid adjustments take some time to occur, a strict calculation of the radiative forcing is not possible, and one has to revert to the concept of effective radiative forcing (ERFaci) to encompass all these effects into a radiative calculation.
- Aerosols also serve as ice nuclei (IN) so that a change in their concentration could also lead to a change in ice cloud amount and properties. However, there is also an impact of CCN concentrations on ice clouds when these clouds form through the liquid phase. As for liquid clouds, one can distinguish the immediate effects of changes in CNN and IN leading to a Radiative forcing (RF) and the subsequent effects on the cloud microphysics and dynamics leading to an Effective

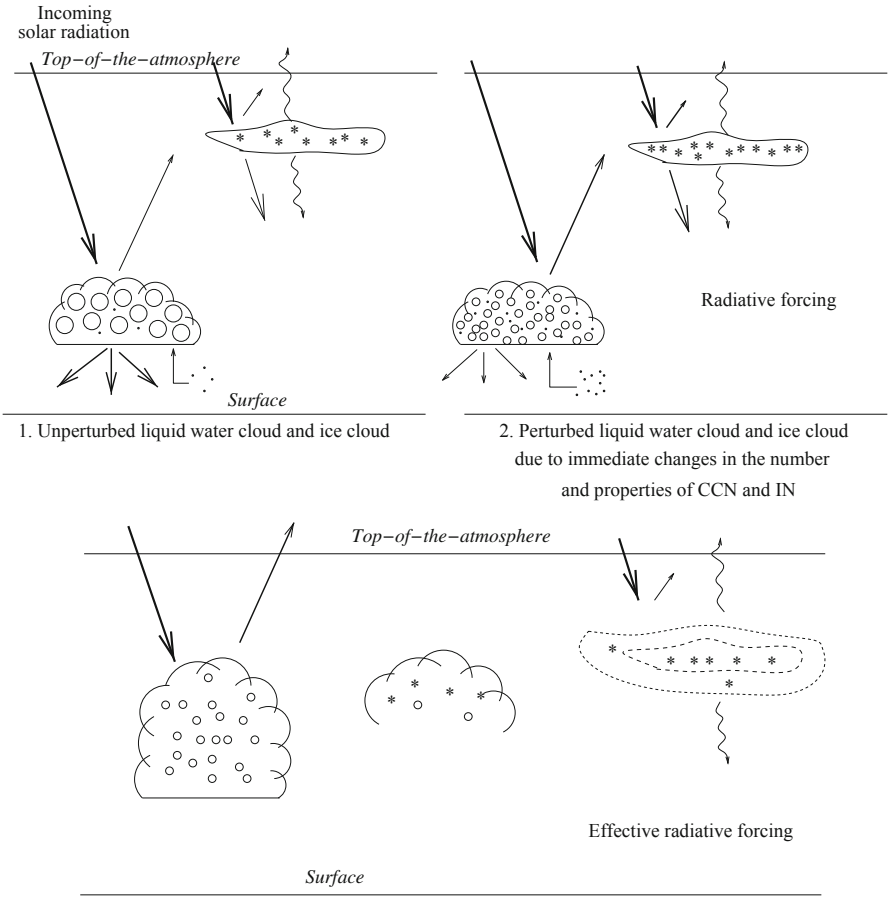


Fig. 9.3 Schematic representing aerosol–cloud interactions. We distinguish immediate changes induced by changes in the number and properties of cloud condensation nuclei (CCN) and ice nuclei (IN), which give rise to a radiative forcing, from further rapid adjustments due to changes in the precipitation development, latent heat release due to the freezing of liquid water into ice, interactions with radiation and convection, which together with the immediate changes give rise to an effective radiative forcing

radiative forcing (ERF). However, ice nuclei acting in different ways, and freezing being sometimes a slower process than condensation itself, it may be easier to encompass all these *aerosol glaciation effects* on mixed-phase and ice clouds as rapid adjustments.

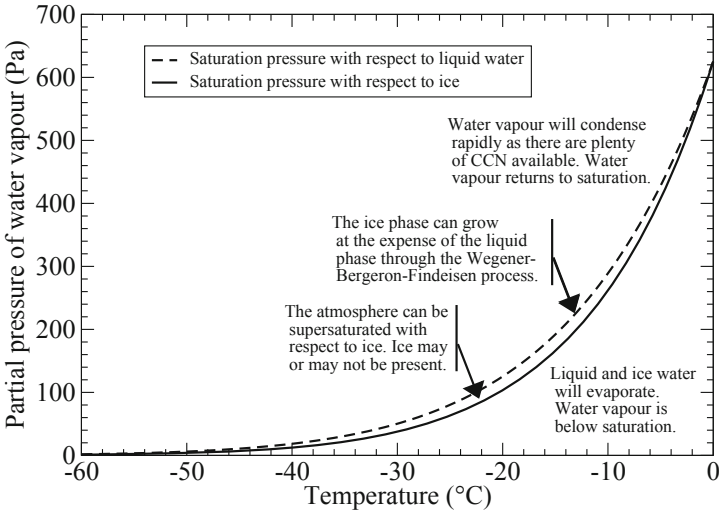


Fig. 9.4 Saturation pressure of water vapour with respect to liquid and ice water in the temperature range of −60 to 0 °C. The text on the figure indicates what can occur in the atmosphere for (P_{H_2O},T) conditions situated above and on the saturation pressure with respect to liquid water, above and on the saturation pressure with respect to ice water

Aerosol–cloud interactions remain poorly understood despite a large amount of research that has been dedicated to this topic. This chapter reviews the various aerosol–cloud interactions based on observations and elements of cloud microphysics.

9.2 Aerosol Effects on Liquid Clouds

9.2.1 Saturation Pressure of Water Vapour

The saturation pressure of water vapour, $P^{sat}_{H_2O}$, is defined as the equilibrium pressure between water vapour and liquid water. From a thermodynamical viewpoint, if the partial pressure of water vapour (or water vapour pressure), P_{H_2O}, exceeds the saturation pressure of water vapour, then there is condensation of water vapour until $P_{H_2O} = P^{sat}_{H_2O}$, which corresponds to the equilibrium point. If instead the partial pressure of water vapour is less than the saturation pressure of water vapour, then there is evaporation of liquid water until $P_{H_2O} = P^{sat}_{H_2O}$. The saturation pressure of water vapour increases with temperature: the warmer it is, the more energetic the water molecules are, the easier it is for them to evaporate, and the more difficult it is for them to condense. This effect is expressed by the *Clausius–Clapeyron relation* that gives the coexistence curve between two phases (here vapour and liquid) as a function of temperature (see Fig. 9.4).

Relative humidity (noted RH and often expressed in %) is defined as the ratio between the partial pressure of water vapour and the saturation pressure of water vapour:

$$\mathrm{RH} = 100\% \times \frac{P_{\mathrm{H_2O}}}{P^{\mathrm{sat}}_{\mathrm{H_2O}}}. \tag{9.1}$$

As relative humidity may exceed 100 % in nonequilibrium conditions, it is convenient to introduce the concept of supersaturation (also expressed in %) equal to

$$S = 100\% \times \left(\frac{P_{\mathrm{H_2O}}}{P^{\mathrm{sat}}_{\mathrm{H_2O}}} - 1 \right) = \mathrm{RH} - 100\%. \tag{9.2}$$

The saturation pressure of water vapour is larger with respect to liquid water than ice (see also Fig. 9.10). Hence it is possible for a cold atmosphere to be supersaturated with respect to ice but not saturated or just saturated with respect to liquid water. This has implications for processes in mixed-phase clouds as discussed in Sect. 9.3.

It should also be noted that the saturation pressure of water vapour, relative humidity and supersaturation, are defined with respect to a plane surface of pure water. These quantities are different when expressed with respect to a curved surface, which is known as the *Kelvin effect* as discussed below.

9.2.2 Kelvin Effect

The saturation pressure for water vapour over a convex surface, such as an aerosol or a cloud droplet, is larger than the saturation pressure over a plane surface. This effect, known as the Kelvin effect, reflects the fact that there is a larger energy barrier for the water vapour to condense onto a rounded surface due to the larger surface tension. The saturation pressure above a spherical droplet can be expressed as:

$$P^{\mathrm{sat,\,curve}}_{\mathrm{H_2O}} = P^{\mathrm{sat}}_{\mathrm{H_2O}} \exp\left(\frac{2\,\sigma\,M_w}{\varrho_w\,R\,T\,r} \right) \tag{9.3}$$

where r is the cloud droplet radius (in m), σ is the surface tension at the water/air interface (about 0.07 J m^{-2}), M_w is the water molar mass (equal to 0.018 kg mol^{-1}), ϱ_w is the density of liquid water (equal to 1000 kg m^{-3}), T is the temperature (in K) and $R = 8.314$ J mol^{-1} K^{-1} is the ideal gas constant ($R = k_B\,A_v$ where $k_B = 1.38\,10^{-23}$ J K^{-1} is the Boltzmann constant and $A_v = 6.02\,10^{23}$ molecules mol^{-1} is the Avogadro number). The saturation pressure over the droplet increases strongly when the cloud droplet radius approaches zero because of the exponential dependence in Eq. 9.3. The supersaturation at equilibrium is about 1 and 10 % for droplets of pure water with radii 100 and 10 nm, respectively. The supersaturation for homogeneous nucleation of water vapour without the presence of an aerosol particle is typically around 300 %.

9.2.3 Raoult's Law

The saturation pressure of water vapour over a solution composed of water and a soluble material is less than over pure water. This means that condensation of water vapour is facilitated in the presence of a soluble aerosol. *Raoult's law* for an ideal, dilute mixture can be expressed as:

$$P_{H_2O}^{\text{sat, solution}} = x_w \; P_{H_2O}^{\text{sat}} \tag{9.4}$$

where x_w is the mole fraction of liquid water (i.e. the amount of water molecules divided by the total amount of molecules from all constituents). In other words, Raoult's law stipulates that the relative reduction in the saturation pressure of water vapour over an ideal solution is proportional to the mole fraction of the solute.

The reduction in the saturation pressure of water vapour is even larger when the solution is not ideal, that is in the presence of interactions between the solvent (the water) and the solute (the dry aerosol). In this case, the water activity, $a_w = \gamma_w \, x_w$, should be substituted to the mole fraction of liquid water, x_w, in Eq. 9.4, with γ_w being the activity coefficient of water.

9.2.4 Köhler Theory

Köhler theory, which owns its name to Hilding Köhler, results from the combination of the Kelvin effect and Raoult's law (Köhler 1921). On the one hand, the saturation pressure of water vapour above a droplet is higher than over a plane surface. On the other hand, the saturation pressure of water vapour over a solution is lower than over pure water. Altogether the equilibrium saturation pressure of water vapour over a soluble aerosol particle can be expressed as:

$$P_{H_2O}^{\text{sat, aerosol}} = a_w \, \exp\left(\frac{2\,\sigma\,M_w}{\varrho_w\,R\,T\,r}\right) \; P_{H_2O}^{\text{sat}}. \tag{9.5}$$

The first term (a_w) expresses Raoult's law, whereas the second term with the exponential describes the Kelvin effect from curvature. Figure 9.5 illustrates these two effects for a range of dry radii (r_d from 0.01 to 1 μm) for spherical ammonium sulphate particles. The dashed line shows supersaturation over a pure water droplet (Kelvin effect). It increases exponentially when size decreases, which explains why the smallest aerosols cannot serve as cloud condensation nuclei in ambient atmospheric conditions. When the cloud droplet reaches a radius of a few micrometres, then the saturation pressure of water vapour above its surface gets close to that over a plane surface: the droplet can then be considered as being plane.

In the case of a soluble aerosol and when the soluble matter that composes the aerosol is not too diluted, the reduction in the saturation pressure of the water vapour due to the presence of soluble material is more important than the increase due

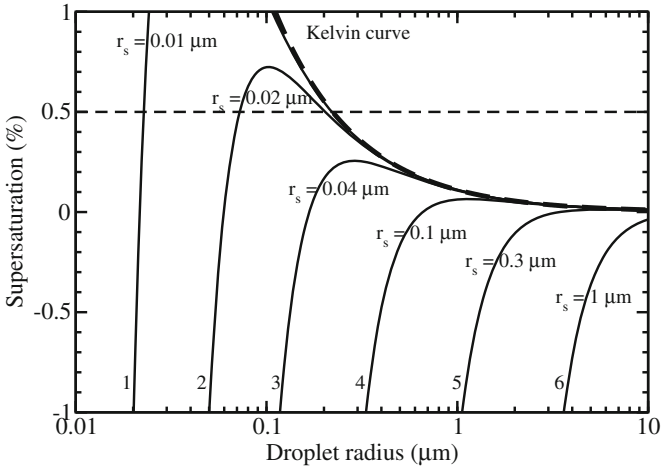


Fig. 9.5 Supersaturation S (in %) at equilibrium over the surface of a pure water droplet (*dashed line*) and the surface of water solution droplets containing increasing amounts of soluble matter (*solid lines*) as a function of the hydrated water droplet radius. The curves are for different values of the radius of the dry aerosol, r_d, which is taken to be ammonium sulphate

to the curvature effect (solid lines on Fig. 9.5). An aerosol can therefore grow by hygroscopicity at negative supersaturations. For instance, a dry aerosol with radius 0.01 µm has a wet equilibrium radius of 0.02 µm for a supersaturation of −1 % (i.e. a relative humidity of 99 %).

It should also be noted that the saturation pressure of water vapour with respect to an aerosol goes through a maximum that corresponds to a critical radius. The aerosol is said to be *activated* when its radius has exceeded this critical radius, which is a function of the dry aerosol size or mass and chemical composition. Let us consider a supersaturation of 0.5 % (dashed horizontal line on Fig. 9.5). A dry aerosol that experiences this supersaturation gets hydrated and its size increases. If the aerosol is too small (curves 1 and 2 on the figure), it cannot pass its critical radius and its size will reach the equilibrium size that corresponds to this supersaturation on the left side of the curve. The aerosol is hydrated but not activated. If instead, the dry aerosol is large enough (curves 3 to 6 on the figure), it can pass its critical radius and continue to grow indefinitely. In this case the aerosol is activated and becomes a cloud droplet.

To each saturation and chemical composition of the dry aerosol corresponds an activation size above which aerosols get activated (assuming there is no kinetic limitations to the activation). Aerosols that get activated at a supersaturation S are called *cloud condensation nuclei* or CCN at supersaturation S. Conversely, a critical supersaturation can be defined for a given chemical composition and dry aerosol size. As all aerosols are not activated at the same supersaturation, it is customary to talk about a spectrum of cloud condensation nuclei as a function of the supersaturation, CCN(S). The CCN concentrations are typically measured and provided for supersaturations of 0.2, 0.4 and 1 %.

John Aitken was the first one in 1880 to show the need for cloud condensation nuclei to trigger the condensation of water vapour given ambient atmospheric conditions (Aitken 1880–1881, 1891, 1900, 1910–1911). This is why aerosols capable of nucleating a cloud water droplet are also called *Aitken nuclei*.

9.2.5 *Extensions to the Köhler Theory*

The Köhler theory only applies to aerosols of homogeneous chemical composition. The presence of insoluble or low-solubility material in the aerosol, such as some organics, and the presence of some volatile species, such as nitric acid, modify the Köhler equation. Various authors have developed extensions to the Köhler theory (e.g. Laaksonen et al. 1998) or have performed accurate calculations of aerosol activation in the presence of multiple chemical species (Abdul-Razzak and Ghan 2000; Nenes et al. 2002). Such calculations are however limited by the complexity of the atmospheric aerosol and the lack of knowledge on the aerosol chemical composition.

Petters and Kreidenweis (2007) have proposed a semiempirical extension of the Köhler theory that relies on the aerosol hygroscopicity parameter, κ. For a pure aerosol species, κ is expressed as:

$$\kappa = \frac{i\,\varrho\,M_w}{\varrho_w\,M} \tag{9.6}$$

where i, ϱ and M are the van't Hoff factor, the density, and the molar mass of the aerosol species, respectively, and ϱ_w and M_w are the water density and molar mass. The κ parameter verifies:

$$a_w = \frac{V_w}{V_w + \kappa\ V_d} \tag{9.7}$$

and expresses how the activity of water deviates from the ratio of the volume of water, V_w, and the total volume of water and dry aerosol particle, $V_w + V_d$.

For a mixture of several soluble species, Petters and Kreidenweis assume that the volume of water can be expressed as the sum of the volumes of water associated with each one of the aerosol species ($V_w = \sum_i V_{wi}$) so that

$$V_w = \frac{a_w}{1-a_w}\sum_i \kappa_i\ V_{si} = \frac{a_w}{1-a_w}\ V_d \sum_i \epsilon_i\ \kappa_i = \frac{a_w}{1-a_w}\ V_d\ \kappa \tag{9.8}$$

where ϵ_i is the volume fraction of species i in the dry aerosol and $\kappa = \sum_i \epsilon_i\ \kappa_i$. Considering the total volume of the particle, V_T, approximated by $V_T = V_w + V_d$, and the radii of the dry aerosol, r_d, and of the wet aerosol, r, a variant of the Köhler equation can be written as:

$$P_{\mathrm{H_2O}}^{\text{sat, aerosol}} = \frac{r^3 - r_d^3}{r^3 - r_d^3\,(1-\kappa)}\ \exp\left(\frac{2\,\sigma\,M_w}{\varrho_w\,R\,T\,r}\right)\ P_{\mathrm{H_2O}}^{\text{sat}}. \tag{9.9}$$

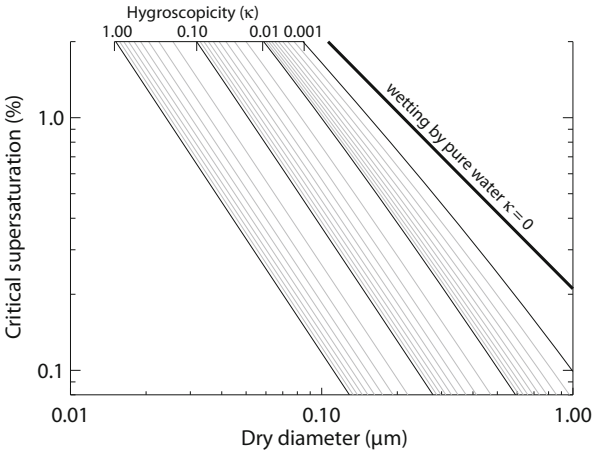


Fig. 9.6 Critical supersaturation as a function of the dry aerosol diameter for κ parameters ranging from 0.001 to 1. (From Petters and Kreidenweis (2007). Creative Commons License)

The advantage of this formulation is that it describes the hygroscopic properties of the aerosol with a single parameter, κ, that combines the aerosol chemical properties (molar mass, density and van't Hoff factor) and can be relatively easily estimated from measurements of the aerosol hygroscopicity and its CCN capability. The parameter κ is 0 for pure water. It takes typical values of 1.4 for a very hygroscopic salt such as sodium chloride, 0.5 for ammonium sulphate and other inorganic salts present in the atmospheric aerosol, and lower values, typically in the range 0 to 0.2, for less soluble organic species (Pringle et al. 2010).

Figure 9.6 shows how the critical supersaturation varies as a function of the dry aerosol size and the hygroscopicity parameter κ. The chemical composition of the aerosol mixture and the aerosol size are key parameters that determine the fraction of aerosols that can be activated at a given supersaturation. Both primary and secondary aerosols contribute to the population of cloud condensation nuclei (Merikanto et al. 2009). Aerosols from the accumulation mode usually provide the majority of the cloud condensation nuclei because their size is sufficient to be activated and their concentration is high. Among aerosols of anthropogenic origin, sulphate aerosols and other inorganic salts are particularly good cloud condensation nuclei, but organic aerosols can also be activated. Aerosols from the coarse mode, even though their number concentration is less, can serve as cloud condensation nuclei when they are soluble, such as is the case for sea spray aerosols, or coated with soluble material, such as can be the case for dust aerosols. Finally, aerosols of the fine mode can also contribute to the CCN concentrations if the supersaturation is large enough. Finally, it has been observed at some locations that given the variability in aerosol properties, aerosol size is generally more important than chemical composition to control the CCN number (Dusek et al. 2006). However, the chemical composition may be important in other locations such as the marine environment, where some primary organic particles have been shown to be very good CCN (Orellana et al. 2011). Finally, it is not clear how important the presence of surfactants is (Prisle et al. 2010).

9.2.6 CCN and Supersaturation in the Cloud

On the basis of observations, Twomey (1959) has introduced the following empirical formulation to express the concentration of active CCN at supersaturation S (expressed again in %):

$$\mathrm{CCN}(S) = c\ S^k \tag{9.10}$$

where c is the concentration of CCN active at 1 % supersaturation and k is a parameter. For atmospheric aerosols, the k parameter depends on the aerosol chemical composition and size distribution; it typically takes values in the range of 0.4 to 0.8. The c parameter depends essentially on the aerosol concentration.

Twomey (1959) has proposed a semi-analytical model to describe the activation of aerosols in a rising air parcel. Such an air parcel experiences an adiabatic expansion, that is accompanied by a cooling and an increase in the relative humidity. There exists a height, called the lifted condensation level or convective condensation level, at which the relative humidity reaches 100 %. If the air rises slightly further up, water vapour in the air parcel will condense and a cloud will form. The supersaturation inside the cloud will result from a competition between two opposing processes: the decrease in the saturation pressure of water vapour as the air temperature decreases (according to the Clausius–Clapeyron relation) and the condensation of water vapour on aerosols and cloud droplets that removes water from the vapour phase. The condensation of water vapour releases latent heat ($L_l = 2.5\ 10^6\ \mathrm{J\,kg^{-1}}$ at 0 °C), which fuels a further ascent of the air parcel. By resolving the conservation equation for water in the air parcel, Twomey shows that the maximum supersaturation that is reached in the cloud can be written as:

$$S_{\max} \propto \left[\frac{w^{3/2}}{c\ k\ B(3/2, k/2)}\right]^{\frac{1}{k+2}} \tag{9.11}$$

where w is the vertical updraft velocity of the parcel and $B(x, y)$ is the mathematical beta function. The cloud droplet concentration corresponds to the concentration of aerosols that has been activated at this supersaturation and can be written as:

$$N = c\ S_{\max}^k \propto c^{\frac{2}{k+2}} \left[\frac{w^{3/2}}{k\ B(3/2, k/2)}\right]^{\frac{k}{k+2}}. \tag{9.12}$$

Given that $k > 0$, it can be observed that the number of activated aerosols N increases less rapidly than the number of cloud condensation nuclei c. It is also very sensitive to the updraft velocity in the cloud which controls the cooling rate in the cloud. A major difficulty in modelling aerosol–cloud interactions in coarse-resolution models is to parameterize the sub-grid scale distribution in cloud updraft velocity.

This simplified model, due to Twomey, relies on the assumption that aerosols are constantly in equilibrium with their environment (i.e. the supersaturation in the cloud). More complex calculations show that even though an aerosol has the potential

to be activated at the maximum supersaturation reached in a cloud, it is not necessarily exposed to this supersaturation for long enough to be activated. This kinetic effect is accentuated in the presence of organic species in the aerosol phase or semi-volatile species in the gas phase such as nitric acid (Nenes et al. 2002).

9.2.7 Dynamical and Radiative Effects in Clouds

The condensation of water vapour on aerosols that accompanies adiabatic cooling is the main mechanism leading to the formation of liquid cloud. There are, however, many other processes that govern the lifecycle of a cloud. The previous section uses the conceptual model of an air parcel that experiences an adiabatic expansion without any exchange of energy and matter with its environment. However, in the real world, the cloud parcel exchanges energy and matter with the surrounding air. There can be *entrainment* of air from the environment into the cloud, that would lead to a dilution of the cloud, and a decrease in the supersaturation inside the cloud. Inversely, cloudy air can be *detrained* into the environment and evaporate. The cloud also exchanges radiative energy with its environment. Although it is very reflective in the shortwave spectrum, it can absorb a small fraction of that radiation, which leads to a heating rate in the cloud. Similarly, absorption of infrared radiation contributes to warm the cloud, while emission of infrared radiation cools it. These radiative processes modify the temperature structure in the cloud, and therefore the relative humidity (through a change in the saturation pressure of water vapour), which impacts on the cloud microphysics. As a result, how a cloud develops and looks like is very sensitive to meteorological conditions, as shown by large variations in cloud types and textures (stratus, stratocumulus, cumulus, etc.).

9.2.8 Principle of the Cloud Albedo Effect

It is well established that the cloud optical depth increases with the concentration of cloud droplets, everything else being equal. Assuming a homogeneous cloud on the horizontal, the cloud optical depth is expressed as:

$$\tau_{\text{cloud}}(\lambda) = \int_{z_b}^{z_t} \int \pi\, r^2\, Q^{\text{ext}}(r,\lambda)\, n(r,z)\, \mathrm{d}r\, \mathrm{d}z \tag{9.13}$$

where z_b and z_t are the altitudes of the base and top of the cloud, and $n(r,z)$ is the cloud droplet size distribution at altitude z. The extinction factor for a cloud droplet is approximately equal to 2 over the whole of the solar spectrum as the cloud droplet size range corresponds to the geometric optics approximation. For a cloud layer not too thick, where $n(r)$ are assumed to be constant over the vertical, the previous equation simplifies to:

$$\tau_{\text{cloud}} = 2\pi\, r_{\text{surface}}^2\, N_c\, \Delta z \tag{9.14}$$

where r_{surface} is the mean surface radius of the cloud droplet size distribution, N_c the cloud droplet number concentration and Δz the cloud geometric thickness. The *liquid water path* is the cloud liquid water content integrated over the vertical. Assuming again the cloud droplet size distribution is constant over the vertical, it can be expressed in kg m^{-2} as:

$$L = \int_{z_b}^{z_t} \int \frac{4}{3} \pi \, r^3 \, \varrho_w \, n(r) \, \mathrm{d}r \, \mathrm{d}z = \frac{4}{3} \pi \, r_{\text{volume}}^3 \, \varrho_w \, N_c \, \Delta z \tag{9.15}$$

where r_{volume} is the mean volume radius of the cloud droplet size distribution. The two equations can be combined to rewrite the cloud optical depth as a function of the cloud properties:

$$\tau_{\text{cloud}} = \frac{3 \, L}{2 \, \varrho_w \, r_e} \tag{9.16}$$

where r_e is the effective radius of the cloud droplet size distribution introduced in Chap. 3. For a fixed liquid water path, the cloud optical depth is inversely proportional to the cloud droplet effective radius. Observations show that the effective radius is proportional to the mean volume radius. It is thus possible to eliminate the effective radius from the equation above and establish that:

$$\tau_{\text{cloud}} \propto L^{\frac{2}{3}} \, N^{\frac{1}{3}} \, \Delta z^{\frac{1}{3}}. \tag{9.17}$$

For a fixed cloud water content, the optical depth is proportional to the cube root of the cloud droplet number concentration. However, this dependence is masked by the fact that the most polluted clouds also tend to be thinner because they originate from more continental air masses that are generally less humid (Brenguier et al. 2003).

Finally, because clouds do not absorb much in the shortwave spectrum, an increase in their optical depth leads automatically to an increase in their albedo. Hansen and Lacis (1974) have shown that the cloud albedo can be approximated by:

$$A \simeq \frac{\tau_{\text{cloud}}}{7.7 + \tau_{\text{cloud}}}. \tag{9.18}$$

An increase in CCN concentration directly translates into an increase in cloud albedo, which is the principle behind the RF due to aerosol–cloud interactions. But as discussed in the introduction, this conceptual model is too simple and there are multiple rapid adjustments at play that give rise to an ERF due to aerosol–cloud interactions.

9.2.9 Observations of the Cloud Albedo Effect

There are multiple observations that indicate an aerosol influence on the cloud droplet number concentration in many regions of the world. First observations show that the CCN concentration increases with the concentration of the dominant aerosol type,

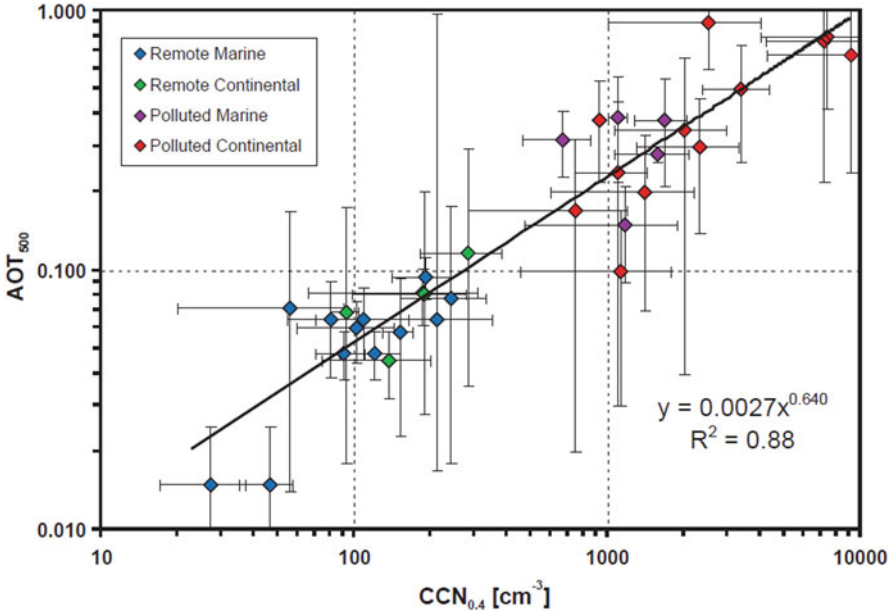


Fig. 9.7 Relationship between the concentration in cloud condensation nuclei at 0.4 % supersaturation ($CCN_{0.4}$ in cm^{-3}) and the aerosol optical depth at 500 nm (AOT_{500}) using averaged data for a range of locations in remote and polluted regions. (From Andreae (2009). Creative Commons License)

be it natural or anthropogenic (e.g. Ayers and Gras 1991; Hegg 1994). Figure 9.7 summarizes a large number of measurements which together display a relationship between the aerosol optical depth (an integrated measure of the aerosol on the vertical that is sensitive to accumulation and coarse-mode aerosols) and the surface concentration of CCN at a supersaturation of 0.4 %. It is remarkable that this relation is valid over several orders of magnitude of CCN concentrations and aerosol optical depths, even though it is not universal (e.g. in the presence of desert dust particles that can contribute a lot to the AOD but less to CCN concentrations).

Observations, further show that the cloud droplet number concentration increases with the aerosol concentration and the CCN concentration (e.g. Werner et al. 2014), but there exists a saturation effect at large concentrations. This saturation is due to the fact that CCN are in competition with each other for the available water vapour as discussed earlier. If more CCN are activated, then the supersaturation will be lowered, which in turn limits the number of CCN that can be activated. Figure 9.8 shows how the cloud droplet number concentration varies with the concentration of sulphate that is dissolved in the cloud water according to aircraft measurements by Leaitch et al. (1992).

The cloud albedo effect is also evident in so-called *ship tracks*, which are bright lines of clouds that materialize behind ships. Ship tracks are generally characterized by an increase in the cloud droplet number concentration resulting from the aerosols

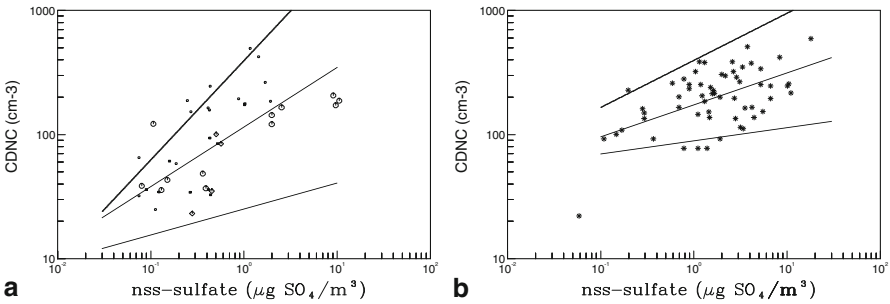


Fig. 9.8 Cloud droplet number concentration (cm^{-3}) as a function of the concentration in dissolved non-sea-salt (*nss*) sulphate in the cloud liquid water for **a** maritime stratiform clouds and **b** continental clouds. (From Leaitch et al. (1992) and Boucher (1995))

emitted by the ship, leading to a decrease in the cloud droplet size and an increase in the cloud albedo (Ackerman et al. 2000), all else being equal. However, it is difficult, if possible at all, to isolate the cloud albedo effect from rapid adjustments also occurring in the cloud field. For instance Coakley and Walsh (2002) showed that cloud water responses can be either positive or negative in ship tracks.

9.2.10 Adjustments in Liquid Water Clouds

Understanding rapid adjustments in liquid water clouds necessitates to consider the time evolution of a cloud system in general, and its ability to develop precipitation in particular. The main process for initiating precipitation in liquid water clouds is called autoconversion.

The continuity equation for the cloud liquid water, q_l, expressed in mixing ratio, can be written as:

$$\frac{\partial q_l}{\partial t} + \frac{1}{\varrho_{\text{air}}} \operatorname{div}(\varrho_{\text{air}} q_l \mathbf{V}) = C - E - R_{ll} - R_{lr} \tag{9.19}$$

where C is the condensation rate, E the evaporation rate at the cloud boundaries, $\mathbf{V}$ the wind speed, R_{ll} the autoconversion rate and R_{lr} the collection (or accretion) rate of liquid water by the flux of rainwater. Here, we deliberately ignore cloud processes that relate to the ice phase. Kessler (1969) has proposed a parameterization of the autoconversion rate that can be generalized under the following form:

$$R_{ll} = c\, q_l\, H(y - y_c) \tag{9.20}$$

where c is an empirical coefficient and H is the Heaviside step function that represents the existence of a threshold below which autoconversion is negligible. The y variable depends on the parameterization and can denote the cloud liquid water content, q_l, as in the original parameterization of Kessler or an average radius of the size distribution

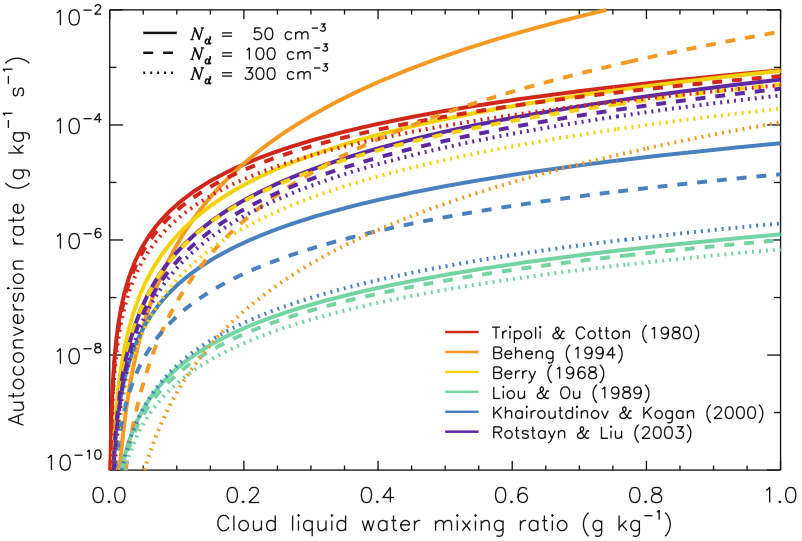


Fig. 9.9 Autoconversion rate (in g precipitating water (kg air)$^{-1}$ s^{-1}) as a function of the cloud liquid water content (in g liquid water (kg air)$^{-1}$) for three different values of the cloud droplet number concentration (N_d) and a range of parameterizations. (From a compilation of data by Andy Jones)

as in other parameterizations (Liou and Ou 1989; Boucher et al. 1995). These parameterizations remain very uncertain because (i) there exist many small-scale processes that are not resolved, (ii) cloud inhomogeneities are not well or not at all taken into account, and (iii) adjustable parameters of the parameterization are not well-constrained by observations. Figure 9.9 shows a compilation of available parameterizations for the autoconversion rate as a function of the cloud liquid water content for different cloud droplet concentrations (and without the threshold included in Eq. 9.20). It can be observed that the autoconversion rate is very weak for small cloud liquid water content, but increases rapidly at higher liquid water content. Furthermore, the sensitivity to the cloud droplet number concentration for a given liquid water content varies a lot among the parameterizations. This highlights the uncertainties related to the autoconversion process but also the lack of a universal parametrization given the variability in cloud regimes. In any case the parametrization of autoconversion induces a large uncertainty in simulations of rapid adjustments in liquid clouds.

The autoconversion rate can also be modelled explicitly if one seeks to represent the cloud droplet size distribution using a sectional or bin approach. The larger the number of size classes to resolve the size distribution, the better the transition between the cloud droplets and raindrops can be represented. Such explicit parameterizations are extremely useful in cloud parcel models, in cloud-resolving models and in large-eddy simulations (see box). They can be used to calibrate simpler parameterizations that rely on bulk or modal approaches discussed above. Rotstayn and Liu (2003) have shown that the autoconversion rate is also very sensitive to the width of the cloud droplet size distribution that cannot be considered as a constant when the cloud droplet number concentration increases.

9.2.11 Rapid Adjustments Occurring in Liquid Clouds

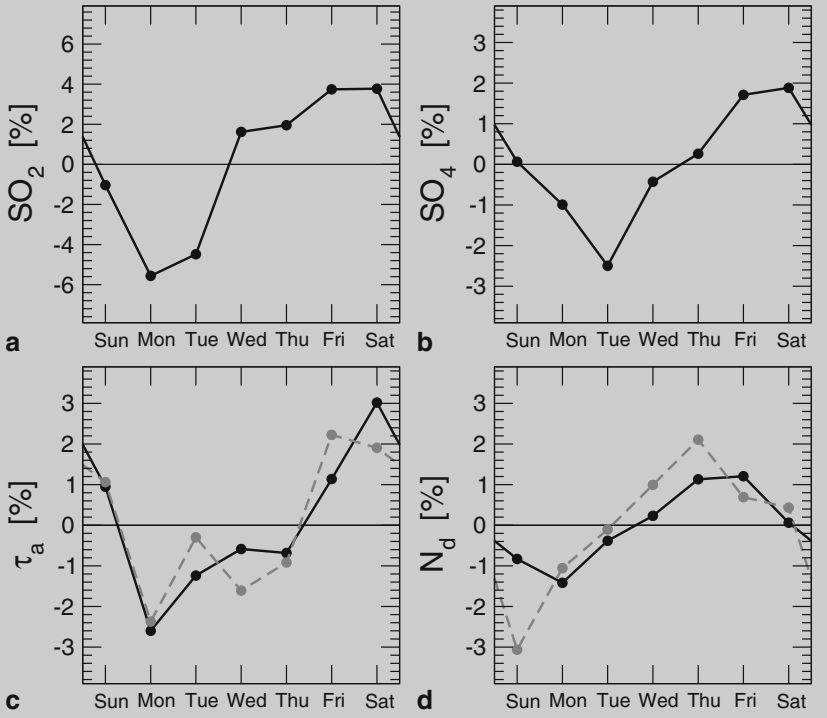
Liquid cloud particles grow by vapour deposition, but also by other microphysical mechanisms such as droplet collision and coalescence. When two cloud droplets collide, they can aggregate and form a new droplet with a larger size. Collisions are mostly due to the fact that cloud droplets of different sizes have different fall velocities in the cloud, but turbulent mixing in the cloud can also play a role. The larger cloud droplets can therefore "catch up" and collect the smaller cloud droplets. This process, called *coalescence*, leads to a shift of the cloud droplet size distribution to larger sizes. Cloud droplets that reach sizes in the range of 100 μm to 1 mm are sufficiently large to have a fall (terminal) velocity that exceeds the cloud updraft velocity. As a consequence they precipitate out of the cloud and can be considered as being raindrops. Coalescence thus, results in the autoconversion of cloud water (composed of droplets with diameter less than about 100 μm) into rain water (composed of drops with diameter larger than about 100 μm). Coalescence is slower for a cloud droplet size distribution that is narrow because the difference in fall velocities is less. As the fall velocity of a cloud droplet depends quadratically on the particle radius in the laminar regime, the relative fall velocities between small and large particles within the spectrum are also smaller when the cloud droplet size distribution is shifted towards smaller radii.

When aerosols lead to a decrease in the cloud droplet size, it is thus expected that coalescence takes longer to produce larger cloud droplets and eventually cloud drops that are large enough to precipitate. It is conceivable also that a cloud can evaporate before it gets a chance to precipitate. In this case, the aerosol is re-emitted in the atmosphere and the water vapour will condense further away on a different aerosol before it eventually precipitates. The initial concept of the cloud lifetime effect is that a less efficient coalescence will delay the development of precipitation and result into an increase in the cloud lifetime. This may be accompanied by an increase in the vertical and horizontal extent of the cloud. This concept is however oversimplistic though in that other adjustments may occur that could mitigate the effect of reduced coalescence. The sensitivity of precipitation formation to the initial cloud droplet size distribution may be weakened when the accretion process dominates. Moreover, precipitation does not always regulate a cloud lifecycle. It has been hypothesized for instance that a larger cloud droplet number concentration may cause an evaporation-entrainment feedback that cancels the initial cloud albedo effect (Wood 2007; Small et al. 2009). Finally, even though reduced or delayed coalescence may have an effect on precipitation at the local scale, this does not necessarily imply a similar change in precipitation at the larger scale because of the microphysical but also the dynamical feedbacks taking place.

Absorbing aerosols whether they are interstitial or embedded in cloud droplets can enhance absorption of solar radiation in clouds, which can affect their lifecycle. There is contradictory evidence regarding the magnitude of this effect.

A Weekly Cycle in Aerosols and Clouds?

Aerosols having a relatively short lifetime in the atmosphere, it is possible to observe a weekly cycle in aerosol concentrations over some industrialised countries that is a direct consequence of a weekly cycle in the emissions. Such a cycle has been observed by several authors from AERONET and surface network aerosol measurements (e.g. Bäumer et al. 2008). The figure below shows a weekly cycle in the concentrations of SO_2 and sulphate over continental Europe as an average over the period 2000–2006. The minimum in SO_2 concentration precedes that in sulphate concentrations by about 1 day, which is consistent with the estimated SO_2 lifetime. The minimum propagates to the aerosol optical depth, and the cloud droplet number concentration. Weekly cycles in cloud cover, convection and precipitation have also been reported (Bell et al. 2008) but remain controversial and disputed on statistical or other methodological grounds (e.g. Sanchez-Lorenzo et al. 2012).



Anomaly of the daily concentrations of (a) SO_2 and (b) sulphate aerosol relative to the weekly average (in %) over the EMEP network for the period 2000–2006. (c) Anomaly of the aerosol optical depth and (d) anomaly of the cloud droplet number concentrations relative to the weekly average (%) for the two MODIS satellite instruments. Adapted from Quaas et al. (2009).

A Range of Cloud Models

There is a range of models that can be used to simulate clouds:

- A cloud parcel model usually has a detailed spectral aerosol–cloud microphysical scheme but often has to parametrize the interaction of the cloud with its environment,
- Large-eddy simulations (LES) resolve explicitly the turbulent eddies that control the cloud dynamics while the effects of the smaller scales of motion are eliminated by filtering and the cloud microphysical processes can be more or less parametrized,
- Cloud-resolving models (CRM) typically have horizontal and vertical grid resolutions that are high enough to be able to explicitly simulate individual clouds, especially those related to deep convection, but low enough to contain cloud systems and to be run long enough to cover several cloud cycles,
- Regional and global models are run with coarser horizontal and vertical grid resolution, which allows to represent more of the feedbacks operating between a cloud field and its environment at the expense of simulating the details of the cloud field, in particular subgrid-scale distributions of water in its gas, liquid and ice phases,
- Global cloud-resolving models (GCRM) and multiscale models, whereby a large-eddy simulation is embedded within each grid-box of a global climate models, are hybrid approaches that help to fill the scale gap between cloud processes and the climate system.

9.3 Aerosols Effects on Mixed-Phased and Ice Clouds

We have focused so far on liquid water clouds, but mixed-phase and ice clouds are also common in the atmosphere and they play an important role in the Earth's radiative budget. As mixed-phase and ice clouds are generally located higher up in the atmosphere than liquid water clouds they have a significant greenhouse effect in addition to their albedo effect. Aerosols may have a significant effect on mixed-phase and ice clouds but the processes governing aerosol interactions with the ice phase are even less well-understood than for the liquid phase. We discuss some of these interactions below.

9.3.1 Elements of Microphysics of Ice Clouds

Liquid water can be found at temperatures below 0 °C, in which case the liquid water is in a metastable state and the water is said to be *supercooled*. Supercooled liquid water is rare in the atmosphere at temperatures below −15 to −20°C, but can exist at temperatures as low as −35°C. Below −35°C to −40°C there is only ice water in the atmosphere.

There are several formation mechanisms for ice crystals in the atmosphere. Firstly, ice crystals can form through *homogeneous freezing* of liquid cloud droplets at temperatures less than about −35°C. Homogeneous freezing occurs spontaneously and does not require the presence of an aerosol other than the aerosol that has initially served as a cloud condensation nucleus. The smaller the liquid cloud droplet, the lower the required temperature for homogeneous freezing, because it is more difficult for a crystal embryo to form inside the droplet given the stronger surface tension.

Secondly, ice crystal can form through *heterogeneous nucleation* or *freezing* involving the presence of an ice nucleus that will trigger the freezing of the water. In this case freezing is not spontaneous but is instead "encouraged" by the aerosol particles. Only some aerosol types, in particular thus having or enabling a crystalline structure, can serve as ice nuclei: this is the case of desert dust particles (Atkinson et al. 2013), primary biogenic aerosol particles such as bacteria and spores (Pratt et al. 2009) and possibly certain types of black carbon particles (Hoose et al. 2012). Heterogeneous freezing can be initiated at the surface of a supercooled liquid cloud droplet upon the collision with an ice nucleus (*contact freezing*), inside a liquid cloud droplet upon immersion of an ice nucleus (*immersion freezing*) or upon deposition of water vapour on the ice nucleus itself (*deposition freezing* or *deposition nucleation*). Because of the relative scarcity of water in a cold atmosphere and of the relative lack of ice nuclei as compared to cloud condensation nuclei, the concentration of crystals in ice clouds is generally much smaller than that of droplets in liquid clouds.

Several other mechanisms can convert the liquid phase into the ice phase (Hallett and Mossop 1974; Cantrell and Heymsfield 2005). Because ice is less dense than water, the freezing of a supercooled water droplet can result in its fragmentation and the expulsion of small ice crystals that can themselves hit other supercooled water droplets and trigger their freezing. This ice multiplication process, known as *ice splintering*, occurs at temperatures warmer than −10°C. When large crystals from aloft fall through a layer of supercooled cloud droplets, they collect these droplets and cause their freezing on the surface of the falling solid precipitation particle. The process is known as *riming* and can also be accompanied by splintering. Ice splintering and riming represent an important mechanism for ice multiplication at temperatures from 0 to −10°C.

Finally, in a mixed-phase cloud, ice crystals can grow through deposition of water vapour at the expense of liquid water droplets. This is because the saturation pressure of water vapour is less over ice than it is over liquid water (Fig. 9.4). This difference is maximal around −13°C. Water vapour can deposit on pre-existing ice crystals—or

suitable ice nuclei–, which depresses the water vapour pressure below the saturation pressure with respect to liquid water, and causes the supercooled liquid droplets to evaporate. This process is particularly efficient to transfer water from the liquid to the ice phase through the vapour phase. It is known as the *Wegener–Bergeron–Findeisen process*, from the names of the three scientists who have contributed to establish the theory (Wegener 1911; Bergeron 1935; Findeisen 1938). Despite the Wegener–Bergeron–Findeisen process, a cloud may persist in the mixed phase if the liquid phase is replenished by condensation induced by the vertical updraft in the cloud (Korolev 2007). As the concentration in ice crystals is generally smaller than that of cloud droplets, the ice crystals can grow to sizes large enough to reach fall velocities larger than the cloud updraft. The Wegener–Bergeron–Findeisen process can therefore lead to precipitation without the coalescence process playing any significant role. It can thus shorten the lifetime of a cloud and suppress cloud cover.

Homogeneous and heterogeneous freezing mechanisms compete with each other and their relative importance depends on atmospheric conditions and aerosol availability (Barahona and Nenes 2009; Lohmann et al. 2004). Given the large variety in formation mechanisms, ice crystals can exhibit very different types and shapes depending essentially on the temperature and updraft velocity of the cloud in which they form and grow. The shapes (or habits) of crystals include hexagons, columns, dendrites and plates.

9.3.2 Impact of Anthropogenic Aerosols on Ice Clouds

Aerosols influence ice clouds directly through their ability to serve as ice nuclei. A Twomey effect may therefore exist for ice clouds, whereby more ice nuclei result in more reflective clouds. There are, however, few observations to corroborate this.

Moreover, aerosols influence ice clouds indirectly through their influence on liquid clouds as a large fraction of the ice in the atmosphere comes from the freezing of liquid droplets. The microphysical properties of mixed-phase and ice clouds therefore depend on the past history of the cloud. It has been suggested in particular that anthropogenic aerosols could impact convection in clouds (Koren et al. 2005; Lebo and Seinfeld 2011). Smaller cloud droplets slow down the coalescence process and can push the cloud development upward to colder temperatures where freezing may occur. Freezing can fuel the convection by releasing additional latent heat (corresponding to the difference between the latent heat for condensation into ice and liquid, i.e. $L_{\text{ice}} - L_{\text{liquid}} \approx 0.35\ 10^6\ \text{J kg}^{-1}$) and by condensing more water as the cloud gets to colder temperatures. Smaller cloud droplets require colder temperatures to freeze, which may also shift upward the level of freezing. The delayed freezing would result in the latent heat of freezing being deposited higher up in the cloud but could also involve additional condensation of water vapour as the cloud develops to colder temperatures. It has been hypothesized that this will deepen the convection, causing an *invigoration effect* that can in turn increase the precipitation

rate of individual cloud systems. While the invigoration effect relies on a physically-based concept and has received modelling support (Wang et al. 2014), observational evidences remain ambiguous, as both aerosol properties and convection are strongly influenced by meteorological factors that are difficult to control for (e.g. Boucher and Quaas 2013). The impact of invigoration on the large-scale is also less clear because in contrast to what may happen at the local scale convective heating and radiative cooling have to balance each other within the atmosphere. Finally, a change in the size distribution of ice crystals could induce some feedback through the sedimentation velocity of the ice crystals and therefore the lifecycle of ice clouds and convective clouds (Fan et al. 2013).

The concentration of ice nuclei in the atmosphere is highly variable (DeMott et al. 2010). A fundamental question remains, which is to know to which extent anthropogenic emissions of aerosol precursors and aerosols have changed the concentration of ice nuclei in the atmosphere, either by making natural aerosols better or worse ice nuclei, or by providing new ice nuclei of anthropogenic origin. It has been hypothesized that black carbon aerosols can serve as ice nuclei, in particular at warmer temperature and/or in the absence of better natural ice nuclei such as desert dust particles and primary biogenic aerosol particles. However, it appears that the capacity of black carbon to serve as ice nucleus diminishes with the degree of mixing with soluble species. For that reason it could be that black carbon directly emitted by aircraft in the upper troposphere might be a more efficient ice nucleus than black carbon emitted at the surface as the latter has time to be coated by soluble species (Hendricks et al. 2005). It is also possible that the ability of desert dust aerosols to serve as ice nuclei is diminished by the condensation of soluble species of anthropogenic origin such as sulphate. In conclusion, it remains very uncertain whether anthropogenic emissions have contributed to increase or decrease the concentrations of ice nuclei in the atmosphere and which one of the process of increase in cloud glaciation or ice nuclei deactivation dominates (Lohmann and Hoose 2009).

9.4 Forcing Due to Aerosol–Cloud Interactions

The radiative forcing due to aerosol–cloud interactions (RFaci) is a theoretical construct for which it is somewhat pointless to provide an estimate. The effective radiative forcing due to aerosol–cloud interactions (ERFaci) can be estimated from observations and from models, but both approaches require to define a reference case.

In the case of observations, calculations generally seek to compare the properties and planetary albedos of cloud fields that have varying degree of influences from anthropogenic aerosols but are otherwise similar. This has traditionally been done by segregating cloud fields by season, region, cloud liquid water path, etc. Because both the aerosol and cloud fields are a strong function of air-mass origin and history, it is difficult to disentangle the effects of aerosols and meteorology on cloud properties. Such issues are increasingly being considered in observationally-based studies of

aerosol effects on clouds and precipitation, e.g. by considering a passive tracer such as CO as a proxy for the aerosols or by considering the effect of aerosols in well-controlled conditions such as, a degasing volcano in the marine boundary layer (Yuan et al. 2011). However, ascribing changes in cloud properties to changes in the aerosol remains a fundamental challenge.

In the case of models, ERFaci cannot be diagnosed through double radiation call with and without the effect of aerosols on clouds because rapid adjustments take some time to develop. Instead ERFaci is diagnosed as the difference in top-of-atmosphere irradiance between two simulations with and without the effect of anthropogenic aerosols. This quantity, also called *quasi-forcing* (Rotstayn and Penner 2001) or *radiative flux perturbation* (Lohmann et al. 2010), includes rapid adjustments but necessitates long model integrations to remove the effects of meteorological and climate variability.

For reasons explained previously related to the complexity of the relevant cloud microphysical processes, the sign and magnitude of the radiative effects of anthropogenic aerosols on mixed-phased and ice clouds remain very uncertain. Some global model studies indicate however that the aerosol glaciation effect partly compensates for the radiative effects on the liquid phase.

Overall available estimates of ERFaci in the literature remain very uncertain. They rely on analysis of satellite data (Quaas et al. 2008), climate model simulations (Lohmann et al. 2010; Shindell et al. 2013), simulations constrained by satellite data (Quaas et al. 2006), or an aerosol reanalysis combined with cloud satellite observations (Bellouin et al. 2013). Boucher et al. (2013) estimated a ERFaci of $-0.45\ \mathrm{W\,m^{-2}}$ with a 90 % uncertainty range of -1.2 to $0\ \mathrm{W\,m^{-2}}$ and a very low confidence.

9.5 Aerosols, Contrails and Aviation-Induced Cloudiness

Contrails and aviation-induced cirrus represent the most visible, and probably the largest of the effects of aircraft emissions on the atmosphere. Although contrails are a particular type of ice clouds, the processes governing their formation is somewhat different to those of other clouds. This justifies examining them separately.

9.5.1 Formation of Condensation Trails

The combustion of kerosene in aircraft engines produces carbon dioxide (CO_2), water vapour (H_2O), nitrogen oxides, aerosol precursors (in particular SO_2), and aerosols (such as black carbon aerosols that could serve as ice nuclei). Condensation trails (or contrails) are not directly linked to the emissions of aerosols although these can influence the formation process. It is rather the water vapour emitted by the aircraft that is the main responsible for the contrail formation.

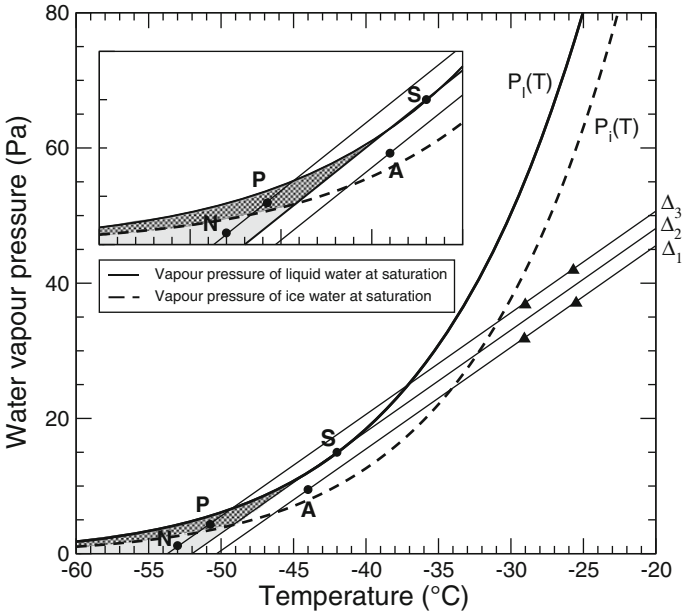


Fig. 9.10 Pressure-temperature diagram for water. The water vapour pressure at saturation is different above a liquid and an ice surface. The A, N and P points indicate different ambient conditions. The Δ lines show two trajectories followed by a volume of air at the exit of the plane engine as it gets diluted with ambient air

As the water vapour pressure above a flat surface of liquid water, $P_l(T)$, is larger than above a flat surface of ice, $P_i(T)$, it is not unusual for the upper atmosphere (at cruising altitudes between 9 to 13 km) to be supersaturated with respect to ice but not with respect to liquid water. These regions can be cloud free because of a too small concentration in suitable ice nuclei. It is the passage of an aircraft that triggers the formation of ice crystals through the liquid phase giving rise to a contrail.

The contrail formation criterion has been established independently by Schmidt (1941) and Appleman (1953). The theory has been further modified by Schumann (1996) who took into account the effect of the aircraft engine efficiency. The combustion of kerosene heats the fuel–air mixture in the engine but also produces water vapour. The hot and humid air that comes out of the engine gets progressively diluted with the cold and dry air of the upper atmosphere. This mixture follows a straight line in a (P_{H_2O}, T) diagram as indicated on Fig. 9.10. This dilution line ends at a point that corresponds to the atmospheric ambient conditions of temperature and humidity (e.g. the point A for the line Δ_1). The slope of the line depends only on the fuel type, the engine efficiency and the ambient pressure. We can define a point S that belongs to the saturation curve with respect to liquid water, $P_l(T)$, where the tangent Δ_2 to this curve has the same slope as the dilution line. The Δ_2 line, the water vapour saturation pressure curves, $P_l(T)$ and $P_g(T)$, and the axes delineate a grey-shaded region

on the figure. For points (P_{H_2O}, T) located in the grey-shaded region, the dilution line cross the saturation curve with respect to liquid water $P_l(T)$; the mixture then becomes supersaturated with respect to liquid water. Unlike ice nucleation which can be kinetically limited, condensation in the liquid phase is a rather fast process, so that in the absence of clouds, crossing the saturation curve leads automatically to the apparition of liquid water droplets at some distance behind the plane. As the mixture continues to dilute and conditions reach that of the ambient atmosphere, condensation is followed by freezing. For ambient conditions corresponding to a point N situated in the dark grey part of the diagram, and after dilution has taken place, the contrail evaporates as there is no longer supersaturation with respect to liquid water (solid line) and ice (dashed line). The contrail is short-lived and said to be nonpersistent. For ambient conditions corresponding to a point P situated in the light grey part of the diagram, the ambient air is supersaturated with respect to ice, and the contrail can grow and persist.

Observations show that going through the liquid phase is a necessary condition for contrail formation. It is not sufficient for the dilution line to cross the saturation curve with respect to ice for the contrail to form because ice formation in these conditions requires adequate ice nuclei and is a relatively slow process. At ordinary cruising altitudes and for current aircraft engines, contrails form generally at temperatures in the range $-40°C$ to $-57°C$ according to the ambient humidity. The engine efficiency is defined as the ratio between the energy used to propel the plane and the total energy released by the combustion in the engine. Of the order of 0.2 at the beginning of commercial aviation, this efficiency has increased to about 0.3 for current engines. It is expected to continue to increase because of technological progress in aircraft and engine design and construction. A higher efficiency implies a lower temperature for the exhaust gases, and therefore a larger slope for the dilution line in Fig. 9.10. The consequence of this is to broaden the grey-shaded areas in the figure, and hence the regions prone to contrail formation.

When ambient conditions allow it, contrails can persist and grow for several hours; in some cases they evolve into a field of cirrus clouds and become indistinguishable from a natural cirrus (Minnis et al. 1998; Haywood et al. 2009). A persistent contrail grows from the deposition of ambient water vapour upon the ice crystals initiated by the passing of the plane; and most of the condensed ice comes from ambient water vapour rather than water vapour emitted by the combustion. As air traffic occurs throughout the day, fresh contrails can coexist with older contrails and contrail-induced cirrus.

9.5.2 Estimate of the Climate Impact of Contrails

Quantifying the impact of contrails on the atmosphere and the climate system requires to know their spatial coverage, thickness, altitude and microphysical properties (size and shape of the ice crystals, ice water content), but also their day-night distribution. The latter parameter is important to determine the relative role of the (generally

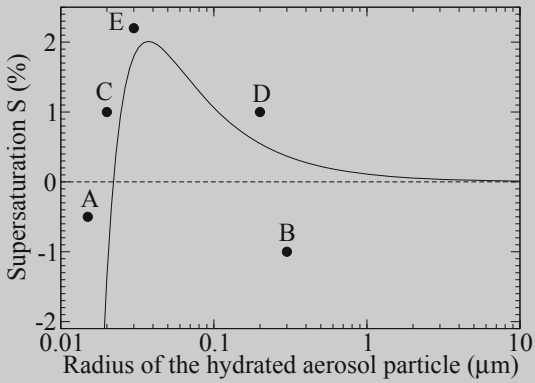
negative) RF contribution in the shortwave and the (positive) RF contribution in the longwave. Satellite observations, which rely on pattern recognition of the linear shape of contrails, indicate that contrail coverage is three times larger during the day than during the night, at least over continental Europe (Mannstein et al. 1999), in agreement with known variations of the air traffic. Estimates of the radiative forcing due to line-shaped contrails (i.e. before any further transformation into cirrus clouds) have been revised downwards since the first studies (Fahey et al. 1999) with a radiative forcing of the order of 0.01 W m^{-2} (Rap et al. 2010a; Boucher et al. 2013) and a fairly small regional impact on surface temperature and precipitation (Rap et al. 2010b).

The radiative forcing due to aviation-induced cirrus is probably larger as indicated by a few studies (Boucher 1999; Stubenrauch and Schumann 2005) that observe an increase in cirrus clouds in air traffic corridors. Burkhardt and Kärcher (2011) estimate the associated RF in the range 0.01 to 0.08 W m^{-2} from a model study representative of the year 2002. Reviewing this and other available studies, Boucher et al. (2013) assessed the combined contrail and contrail-induced cirrus ERF for the year 2011 to be +0.05 (+0.02 to +0.15) W m^{-2} to take into account uncertainties on contrail spreading rate, optical depth, ice particle shape and radiative transfer and the ongoing increase in air traffic.

These estimates do not account for a potential impact of BC ice nuclei emitted by aircraft (Jensen and Toon 1997; Hendricks et al. 2005), which cannot be presently quantified.

Exercises

1. The figure on the right shows the dependence of the supersaturation at equilibrium over a hydrated aerosol with radius r that has evolved from a dry aerosol whose size and composition are given. It is assumed that the supersaturation is hold constant in time. Predict the evolution of aerosols represented by points A, B, C, D and E.



2a. Using a Taylor development, show that the Köhler equation as modified by Petters and Kreidenweis (see Eq. 9.9) can be written in the form:

$$P_{H_2O}^{sat,\,aerosol} \approx \left(1 + \frac{a}{r} - \frac{b}{r^3}\right) P_{H_2O}^{sat}.$$

Explicit the a and b constants.

2b. Deduce an expression for the critical activation radius, r_c, and the critical supersaturation, S_c. How do these two parameters vary as a function of the dry aerosol mass?

3a. We consider a homogeneous cloud layer with liquid water content $q = 0.2$ g m^{-3}, cloud droplet number concentration $N = 100$ cm^{-3} and a geometric thickness $H = 300$ m. Compute the cloud droplet radius, r, assuming a monodisperse distribution. Estimate the cloud optical depth, τ_c, and the cloud albedo, A_c, in the visible part of the electromagnetic spectrum.

3b. Estimate the decrease in cloud droplet radius, Δr, increase in cloud optical depth, $\Delta\tau_c$, and increase in cloud albedo, ΔA_c, for an increase in the cloud droplet number concentration $\Delta N = 50\,\text{cm}^{-3}$ and a fixed liquid water content.

3c. The surface albedo is $A_s = 0.1$. Compute the albedo of the surface-cloud system, A, and its increase, ΔA, for the above-mentioned perturbation ΔN. It is suggested here to follow Chap. 8 to express the albedo of the surface-cloud system as a function of the surface and cloud albedos by neglecting atmospheric radiative effects.

3d. Same question as in 1c but for a surface albedo $A_s = 0.4$.

3e. Repeat the calculations of questions 1a, 1b, 1c and 1d, but for a baseline cloud droplet number concentration of $N = 400$ cm^{-3}. Conclude.

Solutions

1. Aerosols A and C grow until they reach their equilibrium radii of 0.021 and 0.025 μm, respectively. Aerosol B shrinks until it reaches its equilibrium radius of 0.02 μm. Aerosols D and E have gone over the activation barrier and can grow indefinitely.

2a. On the one hand,

$$\exp\left(\frac{2\,\sigma\,M_w}{\rho_w\,R\,T\,r}\right) \approx 1 + \frac{a}{r}$$

with $a = \dfrac{2\,\sigma\,M_w}{\rho_w\,R\,T} \ll 1$. On the other hand,

$$\frac{r^3 - r_s^3}{r^3 - r_s^3(1-\kappa)} \approx 1 - \frac{\kappa\,r_s^3}{r^3} = 1 - \frac{b}{r^3}$$

with $b = \kappa\, r_s^3 = \frac{3\, i\, m_s\, M_w}{4\,\pi\,\rho_w\, M_s} \ll 1$. One can then derive the requested approximation using a Taylor development.

2b. One can differentiate the previous equation with respect to the aerosol radius, which leads to $r_c = (3\,b/a)^{1/2}$ and $S_c = 100\% \times (4\,a^3/27\,b)^{1/2}$. The critical activation radius varies like the square root of the volume or mass of the dry aerosol, m_s. The critical supersaturation varies as the inverse of the square root of the volume or mass of the dry aerosol.

3a. $r = 7.8$ μm, $\tau_c = 11.5$, $A_c = 0.6$.

3b. $\Delta r = -1$ μm, $\Delta\tau_c = 1.7$, $\Delta A_c = 0.032$.

3c. $A = A_c + \frac{(1 - A_c)^2\, A_s}{1 - A_c\, A_s}$ $A = 0.616$, $\Delta A = 0.029$.

3d. $A = 0.684$, $\Delta A = 0.020$.

3e. $R = 4.9$ μm, $\tau_c = 18$, $A_c = 0.7$, $\Delta\tau_c = 0.7$, $\Delta A_c = 0.008$, $A = 0.713$, $\Delta A = 0.008$, $A = 0.752$, $\Delta A = 0.006$. Conclusion: the radiative forcing due to aerosol–cloud interactions (RFaci) saturates when the cloud droplets are small, when the cloud becomes optically thick, and when the surface albedo is high.

References

Abdul-Razzak H, Ghan S (2000) A parameterization of aerosol activation 2. Multiple aerosol types. J Geophys Res 105:6837–6844

Ackerman AS, Toon OB, Taylor JP, Johnson DW, Hobbs PV, Ferek RJ (2000) Effects of aerosols on cloud albedo: evaluation of Twomey's parametrization of cloud susceptibility using measurements of shiptracks. J Atmos Sci 57:2684–2695

Aitken J (1880–1881) On dust, fogs, and clouds. Trans Royal Soc Edinb XXX:337–368

Aitken J (1891) On the solid and liquid particles in clouds. Trans Royal Soc Edinb XXXVI:313–319

Aitken J (1900) On some nuclei of cloudy condensation. Trans Royal Soc Edinb XXXIX:15–25

Aitken J (1910–1911) On some nuclei of cloudy condensation. Proc Royal Soc Edinb XXXII: 478–498

Albrecht BA (1989) Aerosols, cloud microphysics, and fractional cloudiness. Science 245: 1227–1230

Andreae MO (2009) Correlation between cloud condensation nuclei concentration and aerosol optical thickness in remote and polluted regions. Atmos Chem Phys 9:543–556. doi:10.5194/acp-9-543-2009

Appleman H (1953) The formation of exhaust condensation trails by jet aircraft. Bull Am Meteorol Soc 34:14–20

Atkinson JD, Murray BJ, Woodhouse MT, Whale TF, Baustian KJ, Carslaw KS, Dobbie S, O'Sullivan D, Malkin TS (2013) The importance of feldspar for ice nucleation by mineral dust in mixed-phase clouds. Nature 498:355–358

Ayers GP, Gras JL (1991) Seasonal relationship between cloud condensation nuclei and aerosol methanesulphonate in marine air. Nature 353:834–835

Barahona D, Nenes A (2009) Parameterizing the competition between homogeneous and heterogeneous freezing in ice cloud formation—polydisperse ice nuclei. Atmos Chem Phys 9:5933–5948

Bäumer D, Rinke R, Vogel B (2008) Weekly periodicities of aerosol optical thickness over Central Europe—evidence of an anthropogenic direct aerosol effect. Atmos Chem Phys 8:83–90

Beheng KD (1994) A parameterization of warm cloud microphysical conversion processes. Atmos Res 33:193–206

Bell TL, Rosenfeld D, Kim K-M, Yoo J-M, Lee M-I, Hahnenberger M (2008) Midweek increase in US summer rain and storm heights suggests air pollution invigorates rainstorms. J Geophys Res 113:D02209. doi:10.1029/2007JD008623

Bellouin N, Quaas J, Morcrette J-J, Boucher O (2013) Estimates of aerosol radiative forcing from the MACC re-analysis. Atmos Chem Phys 13:2045–2062

Bergeron T (1935) On the physics of clouds and precipitation. Procès Verbaux de l'Association de Météorologie, International Union of Geodesy and Geophysics, pp 156–178

Berry EX (1968) Modification of the warm rain process. In: American Meteorological Society (ed) Proceedings of the First National Conference on Weather modification. State University of New York, Albany, pp 81–88

Boucher O (1995) Etude des interactions aérosol-rayonnement-nuage, Thèse de doctorat, Université Pierre et Marie Curie, 225 pp

Boucher O (1999) Aircraft can increase cirrus cloudiness. Nature 397:30–31

Boucher O, Quaas J (2013) Water vapour affects both rain and aerosol optical depth. Nature Geosci 6:4–5

Boucher O, Le Treut H, Baker MB (1995) Precipitation and radiation modelling in a GCM: Introduction of cloud microphysics. J Geophys Res 100:16395–16414

Boucher O, Randall D, Artaxo P, Bretherton C, Feingold G, Forster P, Kerminen V-M, Kondo Y, Liao H, Lohmann U, Rasch P, Satheesh SK, Sherwood S, Stevens B, Zhang XY (2013) Clouds and aerosols. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, pp 571–657

Brenguier J-L, Pawlowska H, Schüller L (2003) Cloud microphysical and radiative properties for parameterization and satellite monitoring of the indirect effect of aerosol on climate. J Geophys Res 108:8632. doi:10.1029/2002JD002682

Burkhardt U, Kärcher B (2011) Global radiative forcing from contrail cirrus. Nature Clim Change 1:54–59

Cantrell W, Heymsfield A (2005) Production of ice in tropospheric clouds: a review. Bull Am Meteorol Soc 86:795–807

Coakley JA, Walsh CD (2002) Limits to the aerosol indirect radiative effect derived from observations of ship tracks. J Atmos Sci 59:668–680

DeMott PJ, Prenni AJ, Liu X, Kreidenweis SM, Petters MD, Twohy CH, Richardson MS, Eidhammer T, Rogers DC (2010) Predicting global atmospheric ice nuclei distributions and their impacts on climate. Proc Natl Acad Sci U S A 107:11217–11222

Dusek U, Frank GP, Hildebrandt L, Curtius J, Schneider J, Walter S, Chand D, Drewnick F, Hings S, Jung D, Borrmann S, Andreae MO (2006) Size matters more than chemistry for cloud-nucleating ability of aerosol particles. Science 312:1375–1378

Fahey DW, Schumann U, Ackerman S, Artaxo P, Boucher O, Danilin MY, Kärcher B, Minnis P, Nakajima T, Toon OB (1999) Aviation-produced aerosols and cloudiness, Chapter 3. In: Penner

JE, Lister DH, Griggs DJ, Dokken DJ, McFarland M (eds) IPCC special report on aviation and the global atmosphere. Cambridge University Press, Cambridge, pp 65–120
Fan J, Leung LR, Rosenfelf D, Chen Q, Li Z, Zhang J, Yan H (2013) Microphysical effects determine macrophysical response for aerosol impacts on deep convective clouds. Proc Natl Acad Sci U S A doi:10.1073/pnas.1316830110
Findeisen W (1938) Kolloid-meteorologische Vorgänge bei Neiderschlags-bildung. Meteorol Z 55:121–133
Givati A, Rosenfeld D (2004) Quantifying precipitation suppression due to air pollution. J Appl Meteorol 43:1038–1056
Hallett J, Mossop SC (1974) Production of secondary ice particles during the riming process. Nature 249:26–28
Hansen J, Lacis AA (1974) A parameterization for the absorption of solar radiation in the Earth's atmosphere. J Atmos Sci 31:118–133
Haywood JM et al (2009) A case study of the radiative forcing of persistent contrails evolving into contrail-induced cirrus. J Geophys Res 114:D24201. doi:10.1029/2009JD012650
Hegg DA (1994) The cloud condensation nucleus-sulfate mass relationship and cloud albedo. J Geophys Res 99:25903–25907
Hendricks J, Kärcher B, Lohmann U, Ponater M (2005) Do aircraft black carbon emissions affect cirrus clouds on the global scale? Geophys Res Lett 32:L12814. doi:10.1029/2005GL022740
Hoose C, Möhler O (2012) Heterogeneous ice nucleation on atmospheric aerosols: a review of results from laboratory experiments. Atmos Chem Phys 12:9817–9854
Hoose C, Kristjánsson JE, Burrows SM (2010a) How important is biological ice nucleation in clouds on a global scale? Environ Res Lett 5:024009
Hoose C, Kristjánsson JE, Chen JP, Hazra A (2010b) A classical-theory-based parameterization of heterogeneous ice nucleation by mineral dust, soot, and biological particles in a global climate model. J Atmos Sci 67:2483–2503
Jensen EJ, Toon OB (1997) The potential impact of soot particles from aircraft exhaust on cirrus clouds. Geophys Res Lett 24:249–252. doi:10.1029/96GL03235
Kessler E (1969) On the distribution and continuity of water substance in atmospheric circulation, Meteorological Monographs No 32, American Meteorological Society, 84 pp
Khairoutdinov M, Kogan Y (2000) A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus. Mon Weather Rev 128:229–243
Köhler H (1921) Zur Kondensation des Wasserdampfes in der Atmosphäre. Geofysik Publikation 2, No. 3 and No. 6
Koren I, Kaufman YJ, Rosenfeld D, Remer LA, Rudich Y (2005) Aerosol invigoration and restructuring of Atlantic convective clouds. Geophys Res Lett 32:L14828. doi:10.1029/2005GL023187
Korolev A (2007) Limitations of the Wegener–Bergeron–Findeisen mechanism in the evolution of mixed-phase clouds. J Atmos Sci 64:3372–3375
Laaksonen A, Korhonen P, Kulmala M, Charlson RJ (1998) Modification of the Köhler equation to include soluble trace gases and slightly soluble substances. J Atmos Sci 55:853–862
Leaitch WR, Isaac GA, Strapp JW, Banic CM, Wiebe HA (1992) The relationship between cloud droplet number concentrations and anthropogenic pollution: Observations and climatic implications. J Geophys Res 97:2463–2474
Lebo ZJ, Seinfeld JH (2011) Theoretical basis for convective invigoration due to increased aerosol concentration. Atmos Chem Phys 11:5407–5429
Liou KN, Ou SC (1989) The role of cloud microphysical processes in climate: an assessment from a one-dimensional perspective. J Geophys Res 94D:8599–8607
Liu Y, Daum PH (2002) Anthropogenic aerosols: indirect warming effect from dispersion forcing. Nature 419:580–581
Loeb NG, Wielicki BA, Doelling DR, Smith GL, Keyes DF, Kato S, Manalo-Smith N, Wong T (2009) Toward optimal closure of the Earth's top-of-atmosphere radiation budget. J Clim 22:748–766

Lohmann U, Hoose C (2009) Sensitivity studies of different aerosol indirect effects in mixed-phase clouds. Atmos Chem Phys 9:8917–8934

Lohmann U, Kärcher B, Hendricks J (2004) Sensitivity studies of cirrus clouds formed by heterogeneous freezing in the ECHAM GCM. J Geophys Res 109:D16204. doi:10.1029/2003JD004443

Lohmann U, Rotstayn L, Storelvmo T, Jones A, Menon S, Quaas J, Ekman AML, Koch D, Ruedy R (2010) Total aerosol effect: radiative forcing or radiative flux perturbation? Atmos Chem Phys 10:3235–3246

Mannstein H, Meyer R, Wendling P (1999) Operational detection of contrails from NOAA-AVHRR data. Int J Remote Sens 20:1641–1660

Merikanto J, Spracklen DV, Mann GW, Pickering SJ, Carslaw KS (2009) Impact of nucleation on global CCN. Atmos Chem Phys 9:8601–8616

Minnis P, Young DF, Garber DP, Nguyen L, Smith WL Jr, Palikonda R (1998) Transformation of contrails into cirrus during SUCCESS. Geophys Res Lett 25:1157–1160. doi:10.1029/97GL03314

Nenes A, Charlson RJ, Facchini MC, Kulmala M, Laaksonen A, Seinfeld JH (2002) Can chemical effects on cloud droplet number rival the first indirect effect? Geophy Res Lett 29(17):1848. doi:10.1029/2002GL015295

Orellana MV, Matrai PA, Leck C, Rauschenberg CD, Lee AM, Coz E (2011) Marine microgels as a source of cloud condensation nuclei in the high Arctic. Proc Natl Acad Sci U S A 108:13612–13617

Petters MD, Kreidenweis SM (2007) A single parameter representation of hygroscopic growth and cloud condensation nucleus activity. Atmos Chem Phys 7:1961–1971

Pratt KA, DeMott PJ, French JR, Wang Z, Westphal DL, Heymsfield AJ, Ywohy CH, Prenni AJ, Prather KA (2009) In situ detection of biological particles in cloud ice crystals. Nature Geosci 2:398–401

Pringle KJ, Tost H, Pozzer A, Pöschl U, Lelieveld J (2010) Global distribution of the effective aerosol hygroscopicity parameter for CCN activation. Atmos Chem Phys 10:5241–5255

Prisle NL, Raatikainen T, Laaksonen A, Bilde M (2010) Surfactants in cloud droplet activation: mixed organic-inorganic particles. Atmos Chem Phys 10:5663–5683

Quaas J, Boucher O, Lohmann U (2006) Constraining the total aerosol indirect effect in the LMDZ and ECHAM4 GCMs using MODIS satellite data. Atmos Chem Phys 6:947–955

Quaas J, Boucher O, Bellouin N, Kinne S (2008) Satellite-based estimate of the combined direct and indirect aerosol climate forcing. J Geophys Res 113:D05204. doi:10.1029/2007JD008962

Quaas J, Boucher O, Jones A, Weedon G, Kieser J, Joos H (2009) Exploiting the weekly cycle as observed over Europe to analyse the aerosol indirect aerosol effects in two climate models. Atmos Chem Phys 9:8493–8501

Ramanathan V, Cess RD, Harrison EF, Minnis P, Barkstrom BR, Ahmad E, Hartmann D (1989) Cloud-radiative forcing and climate: results from the Earth Radiation Budget Experiment. Science 243:57–63

Rap A, Forster PM, Jones A, Boucher O, Haywood JM, Bellouin N, De Leon RR (2010a) Parameterization of contrails in the UK Met Office climate model. J Geophys Res 115:D10205. doi:10.1029/ 2009JD012443

Rap A, Forster PM, Haywood JM, Jones A, Boucher O (2010b) Estimating the climate impact of linear contrails using the UK Met Office climate model. Geophys Res Lett 37:L20703. doi:10.1029/2010GL045161

Rotstayn LD, Penner JE (2001) Indirect aerosol forcing, quasi forcing, and climate response. J Clim 14:2960–2975

Rotstayn LD, Liu Y (2003) Sensitivity of the first indirect aerosol effect to an increase of cloud droplet spectral dispersion with droplet number concentration. J Clim 16:3476–3481

Sanchez-Lorenzo A, Laux P, Hendricks Franssen HJ, Calbó J, Vogl S, Georgoulias AK, Quaas J (2012) Assessing large-scale weekly cycles in meteorological variables: a review. Atmos Chem Phys 12:5755–5771

Schmidt E (1941) Die Entstehung von Eisnebel aus den Auspuffgasen von Flugmotoren. Schriften der Deutschen Akademie der Luftfahrtforschung, Heft 44. Verlag R. Oldenbourg, München, pp 1–15

Shindell DT, Lamarque J-F, Schulz M, Flanner M, Jiao C, Chin M, Young PJ, Lee YH, Rotstayn L, Mahowald N, Milly G, Faluvegi G, Balkanski Y, Collins WJ, Conley AJ, Dalsoren S, Easter R, Ghan S, Horowitz L, Liu X, Myhre G, Nagashima T, Naik V, Rumbold ST, Skeie R, Sudo K, Szopa S, Takemura T, Voulgarakis A, Yoon J-H, Lo F (2013) Radiative forcing in the ACCMIP historical and future climate simulations. Atmos Chem Phys 13:2939–2974

Small JD, Chuang PY, Feingold G, Jiang H (2009) Can aerosol decrease cloud lifetime? Geophys Res Lett 36:L16806. doi:10.1029/2009GL038888

Stubenrauch CJ, Schumann U (2005) Impact of air traffic on cirrus coverage. Geophys Res Lett 32:L14813. doi:10.1029/2005GL022707

Stubenrauch CJ, Rossow WB, Kinne S, Ackerman S, Cesana G, Chepfer H, Di Girolamo L, Getzewich B, Guignard A, Heidinger A, Maddux BC, Menzel WP, Minnis P, Pearl C, Platnick S, Poulsen C, Riedi J, Sun-Mack S, Walther A, Winker D, Zeng S, Zhao G (2013) Assessment of global cloud datasets from satellites: project and database initiated by the GEWEX Radiation Panel. Bull Am Meteorol Soc 94:1031–1049

Tripoli GJ, Cotton WR (1980) A numerical investigation of several factors contributing to the observed variable intensity of deep convection over South-Florida. J Appl Meteorol 19:1037–1063

Twomey S (1959) The nuclei of natural cloud formation. Part II: the supersaturation in natural clouds and the variation of cloud droplet concentration. Pure Appl Geophys 43:243–249

Twomey S (1974) Pollution and the planetary albedo. Atmos Environ 8:1251–1256

Wang Y, Wang M, Zhang R, Ghan SJ, Lin Y, Hu J, Pan B, Levy M, Jiang JH, Molina MJ (2014) Assessing the effects of anthropogenic aerosols on Pacific storm track using a multiscale global climate model. Proc Natl Acad Sci U S A doi:10.1073/pnas.1403364111

Wegener A (1911) Thermodynamik der Atmosphäre. Barth, Leipzig, 331 pp

Werner F, Ditas F, Siebert H, Simmel M, Wehner B, Pilewskie P, Schmeissner T, Shaw RA, Hartmann S, Wex H, Roberts GC, Wendisch M (2014) Twomey effect observed from collocated microphysical and remote sensing measurements over shallow cumulus. J Geophys Res Atmos 119:1534–1545

Wood R (2007) Cancellation of aerosol indirect effects in marine stratocumulus through cloud thinning. J Atmos Sci 64:2657–2669

Yuan T, Remer LA, Yu H (2011) Microphysical, macrophysical and radiative signatures of volcanic aerosols in trade wind cumulus observed by the A-Train. Atmos Chem Phys 11:7119–7132

Further Reading (Textbooks and Articles)

Bohren CF, Albrecht BA (1998) Atmospheric thermodynamics. Oxford University Press, Oxford, 402 pp

Boucher O, Lohmann U (1995) The sulfate-CCN-cloud albedo effect: a sensitivity study using two general circulation models. Tellus 47B:281–300

Gettelman A, Liu X, Barahona D, Lohmann U, Chen C-C (2012) Climate impacts of ice nucleation. J Geophys Res 117:D20201. doi:10.1029/2012JD017950

Heintzenberg J, Charlson RJ (eds) (2009) Clouds in the perturbed climate system: their relationship to energy balance, atmospheric dynamics, and precipitation. Massachusetts Institute of Technology Cambridge, Massachusetts

Levin Z, Cotton WR (eds) (2009) Aerosol pollution impact on precipitation, a scientific review. Springer, Berlin and Heidelberg, 386 pp

Lohmann U, Feichter J (2005) Global indirect aerosol effects: a review. Atmos Chem Phys 5:715–737

Pruppacher HR, Klett JD (1997) Microphysics of clouds and precipitation. Kluwer, Dordrecht 976 pp (Revised and enlarged edition)
Rogers RR, Yau MK (1989) A short course in cloud physics, 3rd edn. Butterworth-Heinemann Oxford.
Stevens B, Bony S (2013) Water in the atmosphere. Phys Today 6:29–34
Storelvmo T, Kristjánsson JE, Lohmann U, Iversen T, Kirkevåg A, Seland ⊘ (2008) Modeling of the Wegener–Bergeron–Findeisen process: implications for aerosol indirect effects. Environ Res Lett 3:045001
Vali G (1985) Atmospheric ice nucleation—a review. J Rech Atmos 19:105–115

Chapter 10
Climate Response to Aerosol Forcings

Abstract This chapter examines how the changes in radiative fluxes, in particular those induced by changes in aerosols, impact the climate system. It elaborates on the concepts of radiative forcing, adjustments, feedbacks, and climate sensitivity before discussing the climate response to greenhouse gases and aerosols. The chapter also includes a discussion of past and future emissions.

Keywords Radiative forcing · Feedback · Response · Rapid adjustment · Emissions

10.1 Introduction

The previous chapters have focused on aerosol properties and aerosol processes responsible for aerosol–radiation and aerosol–cloud interactions. We have shown in particular how such interactions translate into changes in radiative fluxes. This chapter is concerned with the next stage, which is to understand how changes in the radiative fluxes impact the climate system. We first elaborate on the concepts of radiative forcing, adjustments and feedbacks introduced in Chap. 2, before discussing the climate response to greenhouse gases and aerosols. As climate modelling is a constantly evolving field, this chapter aims to focus on the concepts and the climate responses to past and future aerosol forcings *per se* are essentially for illustrative purpose.

10.2 Radiative Forcing, Feedbacks and Climate Response

10.2.1 Radiative Forcing

It is useful to introduce first the usual paradigm of *radiative forcing*, *climate feedbacks*, and *climate response*. To this effect we consider a climate system that is in an equilibrium state, with a balanced radiative budget, for which the amount of outgoing (infrared) radiative energy is equal to the amount of incoming (solar) radiative energy, as illustrated on Fig. 5.17. This equilibrium is only realized on average over a long enough (typically multidecadal) period because the climate system generates

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_10

naturally an internal variability that can take the system out of radiative balance on shorter timescales. In the presence of a radiative perturbation, the climate system evolves towards a new equilibrium for which the radiative budget, but also other parameters, differ from those of the initial equilibrium state.

In reality, the climate system is never strictly speaking at equilibrium because there are a number of external perturbations that act upon it on various time scales ranging from a year and less to a hundred thousand years and more. Some perturbations are natural and can be more or less random, such as the volcanic activity, or exhibit some cycles, such as changes in solar activity and changes in parameters of the Earth orbit around the Sun. All these perturbations act simultaneously so that the climate is in perpetual evolution, as can be inferred from observed climate change during the geological period and the more recent past. Other perturbations have an anthropogenic origin such as the emissions of long-lived greenhouse gases, aerosols, aerosol precursors and land use change. As discussed in the introductory chapter, anthropogenic perturbations are now more important and more rapid than natural perturbations and have dominated the recent evolution of the climate system (say over the last 50 to 100 years).

These perturbations of the climate system, also called *climate forcings*, are often of radiative origin and can be characterized through a quantity called *radiative forcing*. A radiative forcing, denoted as ΔF, corresponds to the radiative imbalance (measured in W m^{-2}) caused by a climate perturbation. It is often estimated at the tropopause, before the Earth's surface and the troposphere respond to the radiative perturbation (that is with fixed surface properties and vertical profiles of temperature, humidity and clouds), but after stratospheric adjustment of the temperature profile. Stratospheric adjustment was recognized early on as being important because it is essentially in radiative equilibrium and decoupled from the troposphere. Accounting for the stratospheric adjustment makes the radiative forcing a better predictor of the eventual surface temperature change (Hansen et al. 1995; Forster et al. 2007). In practice a radiative forcing is diagnosed with a double call to the radiative transfer scheme in an atmospheric or climate model with and without the radiative perturbation. Accounting for the stratospheric adjustment requires to equilibrate the stratosphere before recomputing the radiative perturbation. The difference in the net radiative flux at the tropopause (sometimes approximated at the top of the atmosphere) corresponds to the traditional definition of the radiative forcing.

10.2.2 Climate Feedbacks

The climate system responds to a perturbation through a number of mechanisms referred to as *climate feedbacks*. A feedback occurs through modifications of the Earth's surface, atmosphere, ocean and cryosphere that can amplify (positive feedback) or dampen (negative feedback) the initial perturbation induced by the radiative forcing (Fig. 10.1).

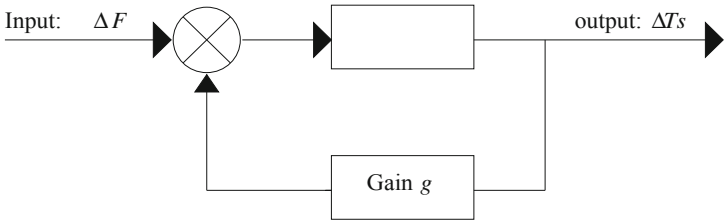


Fig. 10.1 Schematic representation of the feedback gain

The climate is stable to a small radiative perturbation because there is an increase (a decrease) in outgoing longwave radiation when the Earth's surface and atmosphere warms up (cools down). This is referred to as the black body feedback because it follows the Stefan–Boltzmann law, whereby a black body at temperature T emits radiation as $\sigma\, T^4$ (see Chap. 5). This process is a fundamental feedback that acts to stabilize the climate.

Assuming the Earth's system can be approximated by a black body, the increase in temperature, ΔT_0, required to compensate for a radiative forcing, ΔF, is obtained by differentiating the Stefan–Boltzmann law:

$$\Delta F = 4\,\sigma\, T_e^3\, \Delta T_0 \tag{10.1}$$

where T_e is the average temperature of emission of the planet (which corresponds to the average temperature of the atmosphere at an altitude of about 5 km given the actual concentrations of greenhouse gases) which can be taken equal to 255 K. In the absence of any other climate feedback, the ratio between the change in surface temperature and the radiative forcing, which is called *climate sensitivity parameter*, would be equal to:

$$\lambda_0 = \Delta T_0/\Delta F = (4\,\sigma\, T_e^3)^{-1} = 0.26\ \mathrm{K\ (W\ m^{-2})^{-1}}. \tag{10.2}$$

Other climate feedbacks are approximately proportional to the global mean surface temperature change and modify the Earth's radiative budget so that their impact on the global mean surface temperature change can be expressed in the following way:

$$\Delta T = \lambda_0 \left(\Delta F + \sum_i c_i\, \Delta T \right) \tag{10.3}$$

where c_i are the *feedback parameters* in unit W $\mathrm{m^{-2}\,K^{-1}}$. This equation can be rewritten as:

$$\Delta F - \frac{1}{\lambda_0}\, \Delta T + \sum_i c_i\, \Delta T = 0 \tag{10.4}$$

where it becomes evident that the black body feedback parameter is equal to $-1/\lambda_0$ and is, therefore, roughly equal to -4 W m^{-2} K^{-1}. Alternatively, the other climate feedbacks can be expressed relative to the black body feedback, as is done by rewriting the previous equation into:

$$\Delta T = \frac{\lambda_0 \, \Delta F}{1 - \lambda_0 \, \sum_i c_i} = \frac{\Delta T_0}{1 - \sum_i g_i} \tag{10.5}$$

where the $g_i = \lambda_0 \, c_i$ designate the normalized feedback parameters (unitless). The climate system can only be stable if $\sum_i g_i$ is smaller than 1. The gain factor associated to feedback i is equal to:

$$G_i = \frac{\Delta T_i}{\Delta T_0} = \frac{1}{1 - \lambda_0 \, c_i} = \frac{1}{1 - g_i} \tag{10.6}$$

where ΔT_i is the global mean change in surface temperature in the presence of feedback i and the fundamental black body feedback. It should be noted that the feedback parameters, c_i and g_i, are additive but the G_i gains are not[1]. If we have two feedbacks working together, the individual gain factors G_1 and G_2 combine into a net gain:

$$G_{12} = \frac{G_1 \, G_2}{G_1 + G_2 - G_1 \, G_2} \tag{10.7}$$

A large number of climate feedbacks exist, among which we can cite:

- **The water vapour feedback**. The water vapour saturation pressure increases with temperature, and so does the water vapour concentration in the atmosphere because its main sink through condensation and precipitation requires saturation to occur. Water vapour being a greenhouse gas, the increase in water vapour concentration reinforces the greenhouse effect and amplifies the initial warming. The water vapour feedback is a powerful positive climate feedback that is only partly compensated by a change in the atmospheric lapse rate. For this reason, the water vapour feedback is often considered in combination with the lapse rate feedback to form the water vapour–lapse rate feedback. Its feedback parameter is estimated to be $+1.1$ ($+0.9$ to 1.3) W m^{-2} K^{-1} (corresponding to a normalized feedback parameter $g_{H_2O} \approx 0.29$ and a gain factor $G_{H_2O} \approx 1.40$);
- **The cloud feedback**. Climate change induced by a radiative forcing modifies the distribution, quantity and properties of high-level, mid-level and low-level clouds. There is not one but many cloud feedbacks that are differentiated according to the regions, the circulation regimes and the cloud types. Amongst the most robust cloud feedbacks are the fixed anvil temperature (or FAT) feedback, whereby the top of tropical clouds rises to maintain the same cloud-top temperature, and the

[1] The reader should be cautious here as different conventions and definitions can be found for feedback parameters and gains in the scientific literature.

feedbacks associated with a broadening of the Hadley cell and poleward shift of the storm track regions. The sign of the total cloud feedback remains uncertain but is likely positive according to the latest IPCC assessment (Boucher et al. 2013). The feedback parameter has been assessed to be +0.6 (−0.2 to 2) W m^{-2} K^{-1} (corresponding to a normalized feedback parameter $g_{\text{cloud}} \approx 0.16$ and a gain factor $G_{\text{cloud}} \approx 1.18$);

- **The surface albedo feedback**. Warming of the climate system induces a melting of sea ice and the snowpack, which decreases the surface albedo and further enhances the warming at high latitudes. The feedback parameter is assessed to be + 0.3 ± 0.1 W m^{-2} K^{-1} (corresponding to a normalized feedback parameter $g_{\text{surface}} \approx 0.08$ and a gain factor $G_{\text{surface}} \approx 1.08$).

The quantity

$$\lambda = \frac{\lambda_0}{1 - \lambda_0 \sum_i c_i} \tag{10.8}$$

is the climate sensitivity parameter in the presence of climate feedbacks and is measured in K (W $m^{-2})^{-1}$. Models show that the relation between ΔT and ΔF is approximately linear, that is to say that the climate sensitivity parameter, λ, is relatively constant for a range of radiative forcing that is not too large. Climate sensitivity is also relatively constant for different radiative forcing mechanisms: with a few exceptions, ΔF is a good predictor of the global mean surface temperature change whatever the cause of climate change. The effective radiative forcing, that folds rapid adjustments into the radiative forcing, is a better predictor than the radiative forcing, because it isolates better the temperature-dependent feedbacks that are common to all forcing mechanisms. The relationship between ΔF and ΔT_{eq} is only verified at the global scale, and is not valid at the regional scale because the climate system can redistribute energy spatially. Thus a positive (negative) radiative forcing at the regional scale does not necessarily implies a corresponding regional warming (cooling).

The climate sensitivity parameter is estimated to be in the range 0.4–1.2 K (W $m^{-2})^{-1}$. Climate sensitivity often refers to the equilibrium change in the annual global mean surface temperature following a doubling of the CO_2 atmospheric concentration rather than a unit change in radiative forcing. As a CO_2 doubling corresponds to a radiative forcing of approximately 3.7 W m^{-2}, the climate sensitivity is roughly in the range from 1.5 to 4.5°C.

10.2.3 Rapid Adjustments and Effective Radiative Forcing

It is useful at this point to introduce further *rapid adjustments* and differentiate them from climate feedbacks. Rapid adjustments (sometimes called rapid responses) arise when the forcing mechanisms affect the cloud cover or other components of the climate system in a short loop, thus altering the global radiative budget, without

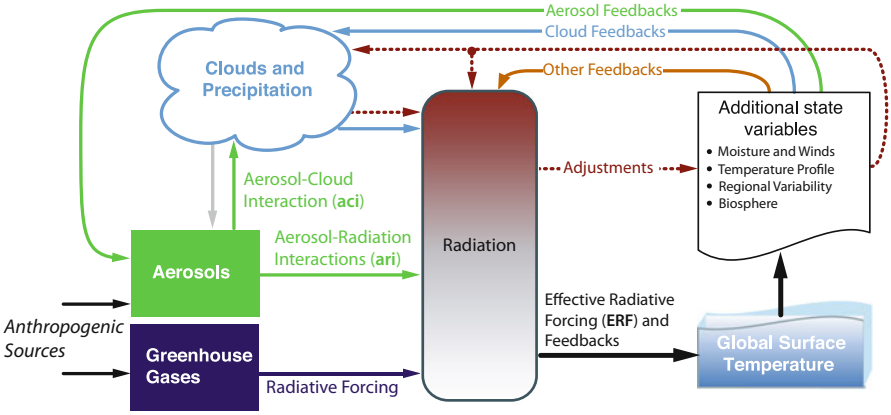


Fig. 10.2 Schematic representation of forcing and feedback pathways involving greenhouse gases, aerosols and clouds. Forcing agents are in the *green* and *dark blue boxes*, with forcing mechanisms indicated by the *solid green* and *dark blue arrows*. The forcing is modified by rapid adjustments whose pathways are independent of changes in the globally averaged surface temperature and are denoted by *brown dashed arrows*. Feedback loops, which are ultimately rooted in changes ensuing from changes in the surface temperature, are represented by curving arrows (*blue* denotes cloud feedbacks; *green* denotes aerosol feedbacks; and *orange* denotes other feedback loops such as those involving the lapse rate, water vapour and surface albedo). The final temperature response depends on the effective radiative forcing (ERF) that is felt by the system, that is, after accounting for rapid adjustments, and the feedbacks. (From Boucher et al. (2013). © IPCC)

operating through changes in the global mean surface temperature (Fig. 10.2). Rapid adjustments are therefore related to the re-equilibration of fast components of the climate system, such as the atmospheric water cycle and the energetics of the atmosphere. Given the small thermal capacity of the atmosphere and the rapidity of the atmospheric water cycle, the timescale associated with these rapid adjustments are of the order of hours to weeks. In contrast climate feedbacks are associated with changes in climate variables that are mediated by a change in global mean surface temperature and materialize themselves as the climate response develops (Gregory et al. 2004). As different forcing mechanisms operate in different ways, the rapid adjustments associated to them are different, but the climate feedbacks are much more similar (Andrews et al. 2010; Huneeus et al. 2014).

The radiative forcing due to CO_2 has an atmospheric component that is quickly –on timescales of days to weeks–compensated by a reduction in the latent heat flux in the atmosphere, so that the energy budget of the atmosphere is maintained (Andrews and Forster 2008). Water vapour and liquid water having a short lifetime in the troposphere, this reduction in the condensation rate of atmospheric water vapour is rapidly compensated by an equivalent reduction in the evaporation rate at the surface, and precipitation rate in the atmosphere. Rapid adjustments associated with the CO_2 radiative forcing lead to a modification of clouds and a reduction in the global precipitation flux, whereas climate feedbacks result in an increase in the evaporation and precipitation rates as global warming proceeds. For a radiative forcing that

increases slowly year after year, the effects of climate feedbacks dominate, but for a radiative forcing that increases very quickly (such as a step change in the CO_2 concentration as is considered in idealized climate model experiments), the rapid adjustments dominate initially until the effects of climate feedbacks prevail.

Scattering aerosols only exert a very small radiative forcing in the atmosphere and as a consequence the rapid adjustment on precipitation is small. In contrast absorbing aerosols heat the atmosphere, and rapid adjustments contribute to decrease the global mean precipitation. It can be generalized that rapid adjustments respond to the atmospheric radiative forcing defined as the difference between the top-of-atmosphere and surface radiative forcing:

$$\Delta F^{\text{atm}} = \Delta F^{\text{TOA}} - \Delta F^{\text{surface}} \tag{10.9}$$

with the convention that a positive ΔF corresponds to a gain to the system.

The concept of rapid adjustments offers the opportunity to redefine the radiative forcing as the change in top-of-atmosphere irradiance after rapid adjustments have taken place but before the average surface temperature has changed. This can typically be computed in climate models in two different ways. The first method consists in diagnosing the change in top-of-atmosphere irradiance in an experiment where sea surface temperature are fixed and the forcing mechanism is included (Hansen et al. 2005). In the second method, the change in top-of-atmosphere irradiance is regressed against the globally-averaged surface temperature change in a coupled ocean-atmosphere simulation where the forcing mechanism is imposed as a step change. The radiative forcing estimate is obtained as the interpolated top-of-atmosphere irradiance for globally-averaged surface temperature change $\Delta T = 0$ (Gregory et al. 2004). In practice, the difference in radiative fluxes between a simulation that includes the forcing mechanism and a control simulation is used in order to eliminate the effect of a potential residual radiative imbalance in the model.

This new definition of the radiative forcing is called *effective radiative forcing* in Boucher et al. (2013) and subsequent publications (Sherwood et al. 2015) to avoid confusion with the traditional definition of radiative forcing. It is however not a completely new concept and has been known before as quasi-forcing or radiative forcing perturbation. As effective radiative forcing includes rapid adjustments from the atmosphere and land surface into the radiative forcing concept, it is generally a better predictor of the eventual surface temperature change induced by a forcing mechanism (Hansen et al. 2005). This relatively new concept is also attractive for aerosols as already discussed in Chap. 2. The semidirect effect is better seen as a rapid adjustment associated to aerosol–radiation interactions. The effective radiative forcing due to aerosol–radiation interactions (ERFari) is therefore the sum of what have formally been called the direct and semidirect effects. The various components of the forcing due to aerosol–cloud interactions, often referred to as indirect aerosol effects, can more easily be lumped together into an effective radiative forcing due to aerosol–cloud interactions (ERFaci). These indirect effects can be interpreted as rapid adjustments associated with the initial modification of the concentration of cloud condensation and ice nuclei. Lohmann et al. (2010) showed that the radiative forcing

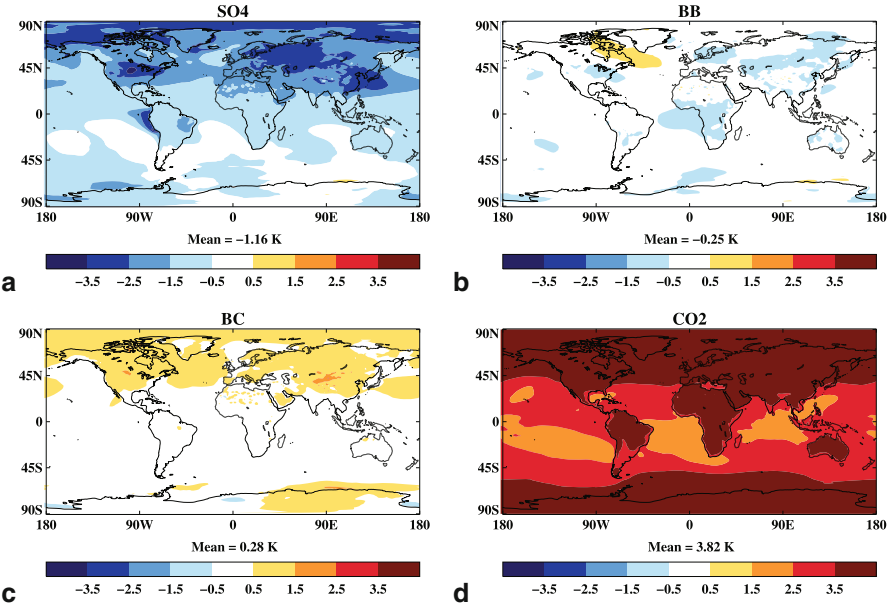


Fig. 10.3 Response of the surface air temperature (°C) to the radiative forcing due to sulphate aerosols (SO4), biomass burning aerosols (BB), black carbon aerosols (BC) and a doubling of the CO_2 atmospheric concentration (CO2). (From Jones et al. (2007). © British Crown Copyright 2007, Met Office)

and effective radiative forcing estimates are equivalent when only the cloud albedo effect is considered, thus giving some legitimacy to the concept. One inconvenient of the effective radiative forcing is that available methods to compute it result in noisy estimates, especially for small forcing values.

10.2.4 Climate Response and Climate Efficacy

Figure 10.3d represents the change in global mean surface temperature in response to a doubling of the CO_2 atmospheric concentration. The warming is not uniform: it is more pronounced over the continents than over the oceans and it is amplified in polar regions of the northern hemisphere. The amplification of the warming at high latitudes is caused by the snow and sea ice surface albedo, but also by the fact that the Stefan–Boltzmann law requires a larger temperature change at lower temperature to evacuate a given amount of excess energy.

Climate feedbacks scale well to the global mean of the surface temperature change but may depend on its spatial and temporal distribution which itself depends partly on the type and distribution of the radiative forcing. It is usual to compare the climate sensitivity parameter to a forcing mechanism, λ_i, to that of the CO_2 forcing, λ_{CO_2}. The

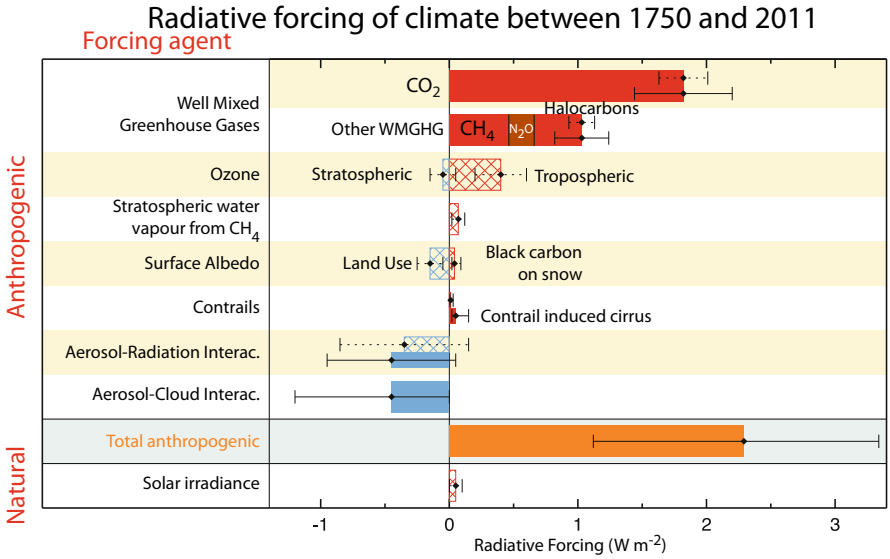


Fig. 10.4 Natural and anthropogenic radiative forcings over the period 1750–2011. *Hatched bars* and *solid bars* are for RF and ERF estimates, respectively. Uncertainty ranges correspond to a 90 % confidence range. (From Myhre et al. (2013). © IPCC)

ratio λ_i/λ_{CO_2} is the *climate efficacy* for forcing mechanism i. Moving to a framework where rapid adjustments are folded into the effective radiative forcing and climate sensitivity parameter is computed relative to the effective radiative forcing instead of the radiative forcing lessens but does not eliminate the spread in climate efficacies across a range of forcing mechanisms. There remain a few pathological forcing mechanisms, such as the radiative forcing due to aerosol–snow interactions, for which we have to accommodate a climate efficacy significantly different from unity. We provide climate efficacies for the aerosol radiative forcings in the next section.

10.3 Climate Response to Aerosol Forcings

10.3.1 Equilibrium Response

Figure 10.4 summarizes the latest estimates of natural and anthropogenic radiative forcings and effective radiative forcings in year 2011 relative to the preindustrial period set in 1750. It appears that aerosol–radiation and aerosol–cloud interactions cause a negative effective radiative forcing that is large enough to compensate for a significant fraction of the warming expected from the anthropogenic greenhouse effect. The effective radiative forcing due to aerosols is also much less uniform than that due to well-mixed greenhouse gases. This is due to the heterogeneity of

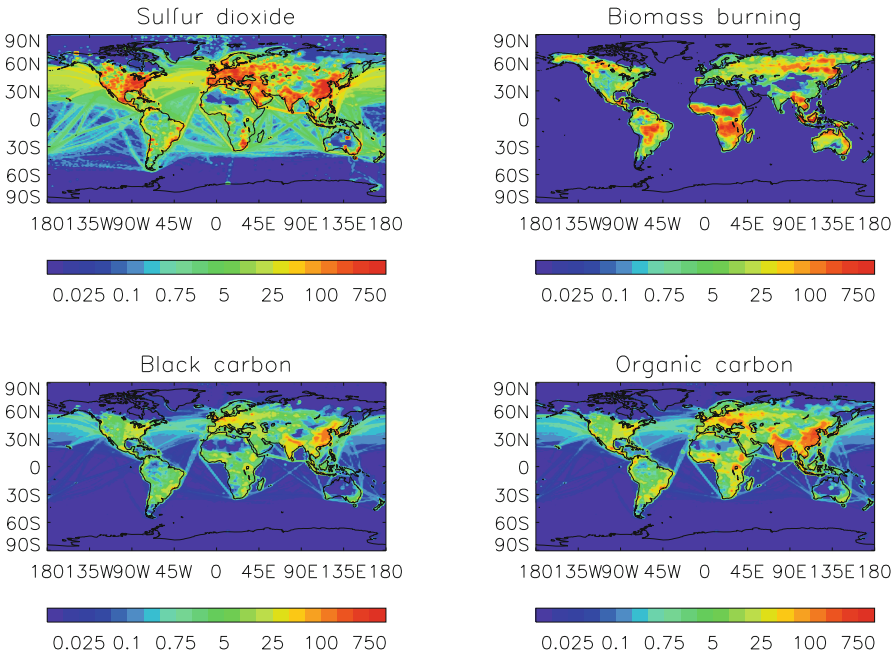


Fig. 10.5 Spatial distribution of the emissions of sulphur dioxide (10^{-13} kg S $m^{-2}\,s^{-1}$), biomass burning aerosols (10^{-13} kg C $m^{-2}\,s^{-1}$), black carbon and primary organic matter (10^{-13} kg C $m^{-2}\,s^{-1}$) from fossil fuel combustion according to the emission inventory prepared for the CMIP5 exercise. Note that the colour scales are not linear. (Redrawn from data in Lamarque et al. (2010))

the emissions, and hence the concentrations, of the main anthropogenic primary aerosols and aerosol precursors. Present-day emissions of some aerosols and aerosol precursors are shown on Fig. 10.5.

Table 10.1 shows some elements of the climate response to a range of aerosol radiative forcings. There is a smaller climate efficacy associated with the atmospheric forcing of black carbon aerosols–which is probably due to the fact that heating of the troposphere does not penetrate all the way to the surface–and a much larger climate efficacy associated with the snow forcing of black carbon aerosols. The global mean forcing and temperature change values hide large regional spatial variability. The pattern of the surface temperature response shows both a global component and a more regional component that is related to the radiative forcing pattern but is not necessarily identical to it. For instance, the radiative forcing by sulphate aerosols is essentially concentrated over and downwind of industrial regions of the northern hemisphere. The equilibrium response encompasses the northern hemisphere and presents a polar amplification (Fig. 10.3a) that is also observed for the greenhouse gas forcing. For biomass burning aerosols, whose forcing is largely confined to tropical regions, the climate response is weaker and limited to the same regions (Fig. 10.3b).

Table 10.1 Effective radiative forcing (ERF, W m^{-2}), surface temperature change (K), climate sensitivity (K $(W\ m^{-2})^{-1}$) and climate efficacy (unitless) for a doubling of the CO_2 atmospheric concentration, and for an increase in various types of anthropogenic aerosols between 1860 and 2000. *ARI* aerosol–radiation interactions, *ACI* aerosol–cloud interactions, *S* Snow. Uncertainties refer to the evaluation method of the radiative forcing and climate response in the model and not to the true uncertainty on these quantities. (Adapted from Jones et al. (2007) and Flanner et al. (2009))

Forcing	ERF ($W\ m^{-2}$)	ΔT (K)	λ ($K\,(W\,m^{-2})^{-1}$)	Climate efficacy (%)
$2\times CO_2$	$+3.83 \pm 0.06$	$+3.82 \pm 0.04$	1.01	100
Sulphates (ARI + ACI)	-1.15 ± 0.06	-1.16 ± 0.02	1.01	101
Biomass burning (ARI + ACI)	-0.29 ± 0.07	-0.25 ± 0.03	0.86	87
Black carbon (ARI)	$+0.39 \pm 0.06$	$+0.28 \pm 0.03$	0.71	71
Black carbon (S)	+0.06	+0.08	1.4	230

For the black carbon aerosol, the climate response is globally similar to that to sulphate aerosols but of opposite sign, and the polar amplification is less than for the response to sulphate aerosols (Fig. 10.3c). Figure 10.6 shows how the patterns of emissions, concentrations, radiative forcing (RFari) and climate response differ from each other for the black carbon aerosol.

10.3.2 Past Emissions

The climate responses described above are for illustration only because they come from a particular climate model when in fact it would be necessary to consider an ensemble of climate models to estimate the robustness of the simulated response. Moreover these are equilibrium climate responses that have to be convoluted with past variations in emissions in different regions of the world.

Schematically emissions of aerosols and aerosol precursors of industrial origin have increased steadily during the twentieth century (Fig. 10.7), at least until the beginning of the 1980s in "western" industrialized countries. Environmental concerns surrounding the impact of *acid rain* on ecosystems (see Chap. 11), and an increased awareness of the impact of air pollution on human health, have led to air quality policies that have targeted emission reductions of initially sulphur dioxide followed by other aerosols and aerosol precursors. As an example Fig. 10.8 shows the evolution of sulphate aerosol concentrations in Europe and sulphate deposition over United States over the period 1980–2000. Emissions reductions have been substantial and have continued into the twenty-first century. While the general trend in industrialized countries is now for a decrease in emissions there are particular aerosol types, such as nitrates, for which emissions may increase. Unlike what is observed in industrialized countries, emissions of sulphur dioxide and most other anthropogenic

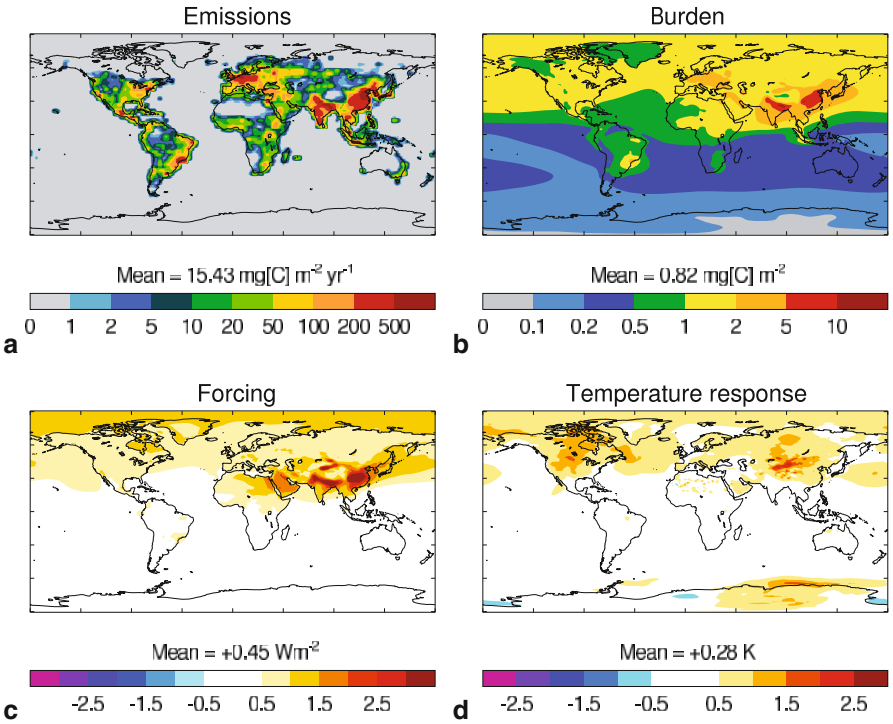


Fig. 10.6 **a** Emission (mg C m^{-2} yr^{-1}), **b** atmospheric burden (mg C m^{-2}), **c** effective radiative forcing (W m^{-2}) and **d** change in surface air temperature (°C) from anthropogenic sources of black carbon from fossil fuel and biofuel combustion. (From Jones et al. (2007). © British Crown Copyright 2007, Met Office)

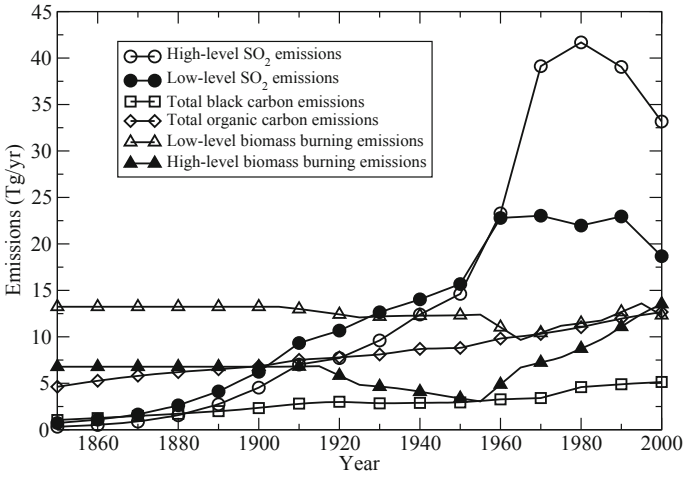


Fig. 10.7 Past evolution of anthropogenic emissions of sulphur dioxide (as S), black carbon, primary organic carbon and biomass burning aerosols over the period 1850–2000. (From the historical inventory developed by Lamarque et al. (2010))

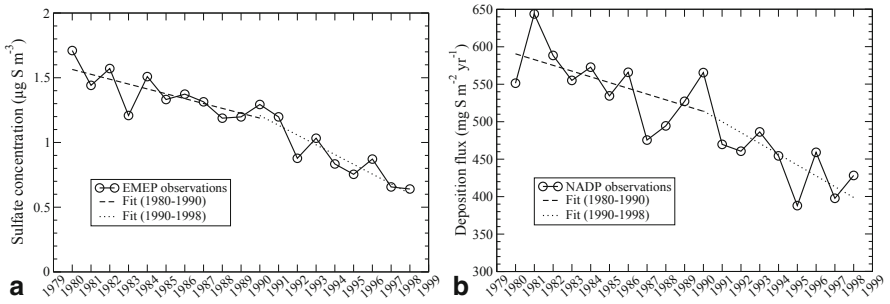


Fig. 10.8 **a** Temporal evolution of the concentrations of sulphate aerosols as measured by the EMEP European network. **b** Temporal evolution of the sulphate deposition as measured by the American NADP network. (Adapted from Boucher and Pham (2002))

aerosol precursors and aerosols have continued to increase in developing countries especially in Asia. These countries may have reached a point where environmental awareness and concerns on human health impact are expected to drive a decrease in emissions in the near future.

It is more problematic to know accurately the past history of biomass burning aerosol emissions because it can be difficult to disentangle natural and anthropogenic sources. Emission of biomass burning aerosols have increased strongly in regions where deforestation has been pursued actively, such as Amazonia and other tropical regions. It is however difficult to estimate how the number of fires, their intensity and the associated emissions have evolved in ecosystems, such as savannas and some coniferous forests, where wild fires form a natural and essential process that maintains the vitality and renewal of the ecosystem. On the one hand anthropogenic pressures have increased ignition sources, but on the other hand anthropogenic management of the ecosystems and fire suppression policies have contributed to reduce the number and duration of fires and their aerosol emissions in some regions. It is possible that fire suppression policies may have led to an accumulation of dry biomass in some ecosystems that make fires stronger and more difficult to control, thus leading to fewer but more intense fires. Overall, the time evolution of this aerosol source over the last century is not well known even though the large interannual variability in emissions due to climate variability and anthropogenic activities are now well documented.

10.3.3 Detection and Attribution of Aerosol Impacts

It is a legitimate question to ask whether the impacts of anthropogenic aerosols can be detected in the climate record of the last century, especially in terms of surface temperature. Climate models cannot reproduce the observed trend in surface temperature without accounting for the radiative forcing by greenhouse gases. Specifically, *detection and attribution* techniques consist in simulating the climate response to each radiative forcing or group of radiative forcings taken separately (e.g. natural

forcings, anthropogenic greenhouse gases, aerosols) and estimating the optimal combination of radiative forcings needed to reproduce the spatial and temporal variations (i.e. so-called *fingerprints*) observed in the climate records:

$$\Delta X_{\text{obs}} = \sum_i \alpha_i \, \Delta X_i + \epsilon \tag{10.10}$$

where ΔX_{obs} is the observed change in climate variable X, ΔX_i is the simulated change in climate variable X due to forcing mechanism i, α_i is a scaling parameter for the simulated response ΔX_i, and ϵ is the residual that we seek to minimize by adjusting the α_i coefficients. Detection and attribution is essentially a signal to noise problem given the large climate variability on many timescales. A good knowledge and estimate of the amount of natural (unforced) climate variability is therefore key to attributing observed climate change to a given forcing mechanism. Uncertainties in the ΔX_i climate responses and natural variability are both accounted for when minimizing the residual. An observed change ΔX_{obs} is said to be attributed to process i is the scaling parameter α_i is found to be significantly different to zero to some level. The scaling parameters in Eq. 10.10 and their associated uncertainty range provide an additional independent constraint on the radiative forcing estimates.

The spatial and temporal signature of the surface cooling induced by the aerosols, as simulated by different climate models, can be detected in observed time series of surface temperature (Stott et al. 2006; Bindoff et al. 2013). The compared analysis of models and observations suggest a total radiative forcing by the aerosols of the order of -0.4 to -1.4 W m^{-2}, which is consistent with independent estimates of the radiative forcing, but does not really help to reduce the uncertainty. Jones et al. (2011) also detected the warming effect induced by black carbon aerosols in observations, but only over the period 1950–2000 and with more limited confidence. There are few studies that have tried to attribute observed changes on climate variables other than surface temperature to anthropogenic factors. However, Gedney et al. (2014) found the signature of increased and decreased aerosol radiative effect in the record of European river flow. It remains, thus, difficult to quantify accurately the aerosol climate effects from observations of climate change.

10.3.4 Future Emissions Scenarios

One cannot forecast with any certainty what aerosol emissions and concentrations will be in the future, but it is possible to build scenarios of future emissions based on various socio-economic assumptions. Such scenarios rely on integrated assessment models that predict emissions of aerosols and aerosol precursors under the assumption of future economic activity and the constraints of assumed climate and air quality policies. Figure 10.9 shows emissions of CO_2 and aerosols for three representative future scenarios with varying degree of climate change mitigation policies. It is predicted that global aerosol emissions will reach a plateau, and then decrease at

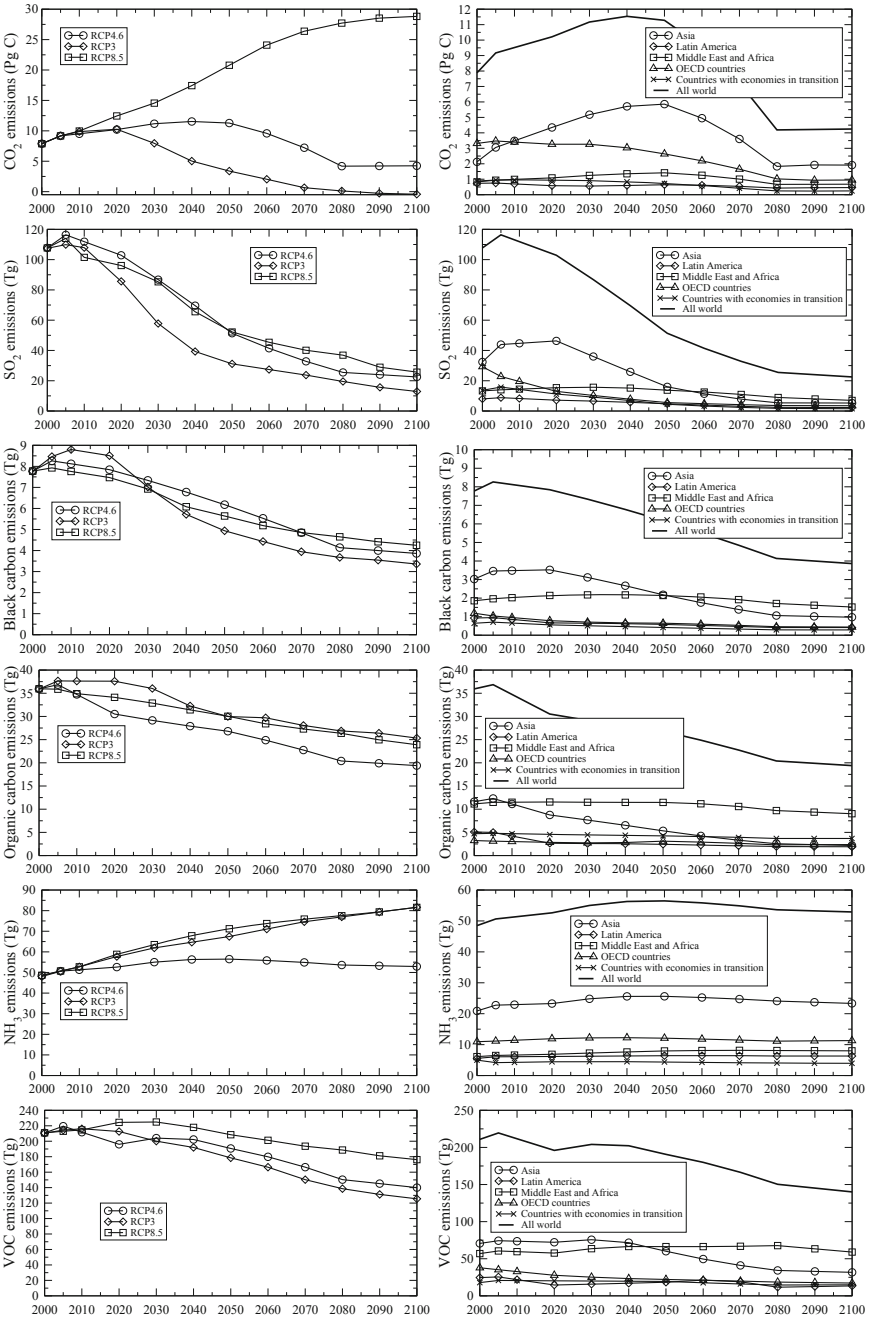


Fig. 10.9 Future emissions scenarios for CO_2, SO_2, black carbon, primary organic carbon, NH_3 and volatile organic compounds. The RCP3-PD (strong and rapid climate mitigation), RCP4.5 (intermediate climate mitigation) and RCP8.5 (no climate mitigation) from Moss et al. (2008) are shown on the left column. The right column shows the emissions by region for the intermediate RCP4.5 scenario

a rate that is faster for the more aggressive climate mitigation policies. These models generally assume an inverse relationship between economical development and emissions of short-lived atmospheric pollutants, which reflects the fact that societies become more environmentally-aware as they get richer. Moreover climate mitigation is also accompanied by a decrease in aerosol emissions because of the increase in combustion-free energy sources. Ammonia (NH_3) forms an exception to this rule because it is mainly emitted from the agricultural sector which intensifies in most future scenarios to feed a growing population and meet an increase demand for biofuels. The increase in NH_3 emissions accompanied by a decrease in SO_2 emissions can induce a significant modification of the chemical composition of the inorganic aerosol in the future and is expected to reinforce the role of nitrate aerosols (Bellouin et al. 2011). A decrease in anthropogenic aerosols in the future would also result in an accelerated warming in the future as the aerosol cooling effect is reduced.

10.4 Nuclear Winter

The climate impact of aerosols is also relevant in the context of a hypothetical, hopefully unlikely, nuclear conflict on the regional to global scale. The explosion of a nuclear bomb on the Earth's surface would result in a huge "mushroom cloud" that would pulverize and inject large amount of smoke and dust into the stratosphere. A series of nuclear explosions would, therefore, lead to large concentrations of aerosols in the stratosphere that would cool the climate system for a period that is long enough to have a significant impact on ecosystems and societies. This effect would be exacerbated if the explosions occur in cities where the available quantity of inflammable material is large. This phenomenon is known as *nuclear winter* (Turco et al. 1983).

The possibility of a nuclear winter has been alluded to during the cold war, at a time when the two superpowers had accumulated an enormous amount of nuclear weapons. Turco et al. (1990) estimated that the absorption aerosol optical depth could reach a value of 2 or more, with a surface cooling of 5–20°C and a dramatic decrease in precipitation during the months that would follow a global nuclear war. In this case, absorbing aerosols cool the climate, because they are emitted at an altitude and in quantities that are sufficient to stabilize the atmosphere and cool the surface by reducing the incoming solar radiation without heating the troposphere much. The plumes of smoke and dust emitted in the northern hemisphere would diffuse to the southern hemisphere and the cooling could last several years. There are other factors that need to be accounted for to estimate the magnitude and length of a nuclear winter, such as coagulation between particles when present in the stratosphere in large concentrations, the possibility that large concentrations of absorbing aerosols may modify the structure of the stratosphere and its circulation, and the effect of deposition of smoke and dust on snow and ice surfaces that can persist for years (Warren and Wiscombe 1985).

The topic has received less attention after the collapse of the Soviet Union and the end of the cold war, but it is still relevant given the increasing number of countries with smaller nuclear arsenals, and the risk for regional conflicts. Toon et al. (2007) estimated that the explosion of a hundred of Hiroshima-size nuclear bombs (which represents a mere 0.03 % of the world's nuclear arsenal) in the Tropics would inject between 1 to 5 Tg of soot in the stratosphere. Robock et al. (2007) simulated the climate impact that would follow using the upper bound of this estimate. They computed a residence time of the soot in the stratosphere of 6 years, which is much longer than for volcanic aerosols because of the absorbing nature of the soot. The stratospheric soot would induce a surface cooling of about 1 °C and a 10 % reduction in precipitation that would take longer than 10 years to disappear. The reduction in solar radiation at the Earth's surface, the reduction of temperature and precipitation would lead to a reduction of primary productivity which would affect the global production of food.

Exercises

1. Derive Eq. 10.7.

2. Compute the climate sensitivity parameter in the presence of the three following climate feedbacks: $c_{\text{water vapour}} = 1.0\,\text{W m}^{-2}\,\text{K}^{-1}$, $c_{\text{surface}} = 0.3\,\text{W m}^{-2}\,\text{K}^{-1}$ and $c_{\text{clouds}} = 0.5\,\text{W m}^{-2}\,\text{K}^{-1}$, in addition to the canonical black body feedback.

3. A number of fixed-SST experiments are performed with a climate model using prescribed aerosol distributions and properties:

- EXP1: a fixed-SST experiment with total (anthropogenic and natural) aerosols used for both aerosol–radiation and aerosol–cloud interactions, a second radiation call made at each timestep with natural aerosols only for aerosol–radiation interactions and a third radiation call made at each timestep with natural aerosols only for aerosol–cloud interactions,
- EXP2: a fixed-SST experiment with total aerosols used for aerosol–radiation interactions and natural aerosols used for aerosol–cloud interactions,
- EXP3: a fixed-SST experiment with natural aerosols used for aerosol–radiation interactions and total aerosols used for aerosol–cloud interactions,
- EXP4: a fixed-SST experiment with natural aerosols used both for aerosol–radiation interactions and aerosol–cloud interactions.

Describe what diagnostics and calculations should be made to diagnose the various components of the aerosol radiative forcing and effective radiative forcing, what difficulties may arise and how to circumvent them. What if the aerosols are not prescribed but interactive in the climate model?

Solutions

2. The climate sensitivity parameter can be written as:

$$\lambda = \frac{\lambda_0}{1 - \lambda_0 \left(c_{\text{water vapour}} + c_{\text{surface}} + c_{\text{clouds}}\right)}$$

which leads to $\lambda = 0.49$ K (W m^{-2})$^{-1}$.

3. Differences in top-of-atmosphere radiative fluxes can be computed from the double radiation calls at each timestep in EXP1 to estimate RFari and RFaci. Differences in top-of-atmosphere radiative fluxes between EXP1 and EXP2 can be computed to estimate ERFaci. Differences in top-of-atmosphere radiative fluxes between EXP1 and EXP3 can be computed to estimate ERFari. Differences between EXP1 and EXP4 can be computed to see if ERFari and ERFaci are additive. Long enough simulations will have to be performed to reduce the level of noise in the ERF estimates. The effect of land cooling on the ERF estimates remains a problem. Double radiation calls are difficult to perform when aerosols are interactive and a different strategy to diagnose the RF should be used.

References

Andrews T, Forster PM (2008) CO_2 forcing induces semi-direct effects with consequences for climate feedback interpretations. Geophys Res Lett 35:L04802. doi:10.1029/2007GL032273

Andrews T, Forster PM, Boucher O, Bellouin N, Jones A (2010) Precipitation, radiative forcing and global temperature change. Geophys Res Lett 37:L14701. doi:10.1029/2010GL043991

Bellouin N, Rae J, Johnson C, Haywood J, Jones A, Boucher O (2011) Aerosol forcing in Hadley Centre CMIP5 simulations and the role of nitrate. J Geophys Res 116:D20206. doi:10.1029/2011JD016074

Bindoff NL, Stott PA, AchutaRao KM, Allen MR, Gillett N, Gutzler D, Hansingo K, Hegerl G, Hu Y, Jain S, Mokhov II, Overland J, Perlwitz J, Sebbari R, Zhang X (2013) Detection and attribution of climate change: from global to regional. In:Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Boucher O, Pham M (2002) History of the sulfate aerosol radiative forcings. Geophys Res Lett 29:1308. doi:10.1029/2001GL014048

Boucher O, Randall D, Artaxo P, Bretherton C, Feingold G, Forster P, Kerminen V-M, Kondo Y, Liao H, Lohmann U, Rasch P, Satheesh SK, Sherwood S, Stevens B, Zhang XY (2013) Clouds and aerosols. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, pp 571–657

Flanner MG, Zender CS, Randerson JT, Rasch PJ (2007) Present-day climate forcing and response from black carbon in snow. J Geophys Res 112:D11202. doi:10.1029/2006JD008003

Flanner MG, Zender CS, Hess PG, Mahowald NM, Painter TH, Ramanathan V, Rasch P J (2009) Springtime warming and reduced snow cover from carbonaceous particles. Atmos Chem Phys 9:2481–2497

Forster PM de F, Ramaswamy V, Artaxo P, Berntsen T, Betts RA, Fahey DW, Haywood JA, Lean J, Lowe DC, Myhre G, Nganga J, Prinn R, Raga G, Schulz M, Van Dorland R (2007) Changes in atmospheric constituents and in radiative forcing. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Gedney N, Huntingford C, Weedon GP, Bellouin N, Boucher O, Cox PM (2014) Detection of aerosol effects on historical Northern Hemisphere river flow. Nat Geosci 7:796–800

Gregory JM, Ingram WJ, Palmer MA, Jones GS, Stott PA, Thorpe RB, Lowe JA, Johns TC, Williams KD (2004) A new method for diagnosing radiative forcing and climate sensitivity. Geophys Res Lett 31:L03205. doi:10.1029/2003GL018747

Hansen J et al (2005) Efficacy of climate forcings. J Geophys Res 110:D18104. doi:10.1029/2005JD005776

Huneeus N, Boucher O, Alterskjaer K, Cole JNS, Curry CL, Ji D, Jones A, Kravitz B, Kristjánsson JE, Moore J, Muri H, Niemeier U, Robock A, Schmidt H, Schulz M, Tilmes S, Watanabe S (2014) Solar irradiance reduction and increased CO_2 radiative forcings and feedbacks in the GeoMIP ensemble. J Geophys Res Atmos 119 5226–5239. doi:10.1002/2013JD021110

Jones A, Haywood JM, Boucher O (2007) Aerosol forcing, climate response and climate sensitivity in the Hadley Centre climate model HadGEM2-AML. J Geophys Res 112:D20211. doi:10.1029/2007JD008688

Jones GS, Christidis N, Stott PA (2011) Detecting the influence of fossil fuel and bio-fuel black carbon aerosols on near surface temperature changes. Atmos Chem Phys 11:799–816

Lamarque J-F, Bond TC, Eyring V, Granier C, Heil A, Klimont Z, Lee D, Liousse C, Mieville A, Owen B, Schultz MG, Shindell D, Smith SJ, Stehfest E, Van Aardenne J, Cooper OR, Kainuma M, Mahowald N, McConnell JR, Naik V, Riahi K, van Vuuren DP (2010) Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. Atmos Chem Phys 10:7017–7039

Lohmann U, Rotstayn L, Storelvmo T, Jones A, Menon S, Quaas J, Ekman AML, Koch D, Ruedy R (2010) Total aerosol effect: radiative forcing or radiative flux perturbation? Atmos Chem Phys 10:3235–3246

Moss R, Babiker M, Brinkman S, Calvo E, Carter T, Edmonds J, Elgizouli I, Emori S, Erda L, Hibbard K, Jones R, Kainuma M, Kelleher J, Lamarque J-F, Manning M, Matthews B, Meehl J, Meyer L, Mitchell J, Nakicenovic N, O'Neill B, Pichs R, Riahi K, Rose S, Runci P, Stouffer R, van Vuuren D, Weyant J, Wilbanks T, van Ypersele J-P, Zurek M (2008) Towards new scenarios for analysis of emissions, climate change, impacts, and response strategies. Intergovernmental Panel on Climate Change. Geneva, 132 pp

Myhre G, Shindell D, Bréon F-M, Collins W, Fuglestvedt J, Huang J, Koch D, Lamarque J-F, Lee D, Mendoza B, Nakajima T, Robock A, Stephens G, Takemura T, Zhang H (2013) Anthropogenic and natural radiative forcing. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate Change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Robock A, Oman L, Stenchikov GL, Toon OB, Bardeen C, Turco RP (2007) Climatic consequences of regional nuclear conflicts. Atmos Chem Phys 7:2003–2012

Sherwood S, Boucher O, Bony S, Bretherton C, Forster P, Gregory J, Stevens B (2015) Adjustments to the forcing-feedback framework for understanding climate change. Bull Am Meteorol Soc doi:10.1175/BAMS-D-13-00167.1

Stott PA, Mitchell JFB, Allen MR, Delworth TL, Gregory JM, Meehl GA, Santer BD (2006) Observational constraints on past attributable warming and predictions of future global warming. J Clim 19:3055–3069. doi:10.1175/JCLI3802.1

Toon O, Robock A, Turco RP, Bardeen C, Oman L, G. Stenchikov L (2007) Consequences of regional-scale nuclear conflicts. Science 315:1224–1225

Turco RP, Toon OB, Ackerman TP, Pollack JB, Sagan C (1983) Nuclear winter: global consequences of multiple nuclear explosions. Science 222:1283–1292

Turco RP, Toon OB, Ackerman TP, Pollack JB, Sagan C (1990) Climate and smoke: an appraisal of nuclear winter. Science, 247:166–176

Warren SG, Wiscombe WJ (1985) Dirty snow after nuclear war. Nature 313:467–470

Further Reading (Textbooks and Articles)

Ghan SJ (2013) Technical note: estimating aerosol effects on cloud radiative forcing. Atmos Chem Phys 13:9971–9974

Haywood JM, Bellouin N, Jones A, Boucher O, Wild M, Shine KP (2011) The roles of aerosol, water vapour and cloud in future global dimming/brightening. J Geophys Res 116:D20203. doi:10.1029/2011JD016000

McGuffie K, Henderson-Sellers A (2005) A climate modelling primer, 3rd edn. Wiley, 296 pp

Peixoto JP, Ort AH (2007) Physics of climate. Springer, 564 pp

Stevens B, Schwartz SE (2012) Observing and modeling Earth's energy flows. Surv Geophys 33:779–816

Chapter 11
Biogeochemical Effects and Climate Feedbacks of Aerosols

Abstract Aerosols have a climate impact through aerosol–radiation interactions and aerosol–cloud interactions, but can also interact with a number of biogeochemical cycles and have an impact on ecosystems. Physical and biological processes are sources of aerosols that can be modified by climate change, allowing for the possibility for climate feedback loops that can play a role in natural climate variability. This chapter provides a short description of some of the processes involved in these biogeochemical effects, such as the impact of aerosols on diffuse radiation at the surface and photosynthesis by terrestrial ecosystems, deposition of nutrients, and acidification. It is followed by a discussion of possible feedback loops involving aerosols, looking separately at the different aerosol species such as sulphate aerosols from dimethylsulphide, sea spray, secondary organic aerosols, biomass burning aerosols, and desert dust.

Keywords Feedback · Biogeochemical cycle · Nutrient · Vegetation · Plankton

11.1 Introduction

As discussed in earlier chapters, the natural environment is a major source of natural aerosols, be it dust and soil particles, sea spray, primary biogenic aerosol particles (PBAP), secondary organic aerosols of biogenic origin, biomass burning aerosols, sulphate aerosols coming from dimethysulphide (DMS) emissions, or other types of inorganic aerosols. These aerosols have a climate impact through aerosol–radiation interactions (Chap. 8) and aerosol–cloud interactions (Chap. 9). They can influence other components of the climate system, like the cryosphere, but also marine and terrestrial ecosystems, and interact with a number of biogeochemical cycles. Moreover, physical and biological processes that are a source of aerosols can be modified by climate change, allowing for the possibility for climate feedback loops that can play a role in natural climate variability.

There are multiple evidence that aerosols have experienced large variations during past climates. In particular, measurement in ice cores and loess deposits show that dust deposition is relatively small during warmer periods (i.e. interglacial periods and interstadials) and relatively large during colder periods (i.e. glacial periods and

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_11

stadials). Variations of natural aerosols on long timescales, such as those of glacial-interglacial cycles of the last million years, can bring insight of what may happen in the context of the current climate change.

However we are only starting to grasp the complex mechanisms by which aerosols can influence biogeochemical cycles, including the most important one which is the carbon cycle. The objective of this chapter is to provide a short description of some of the processes involved in these biogeochemical effects, as well as a discussion of possible feedback loops involving aerosols. More complete discussions can be found in Carslaw et al. (2010) and Mahowald et al. (2011).

11.2 Impact of Aerosols on Terrestrial Ecosystems

11.2.1 Diffuse Radiation and Primary Productivity

Natural and anthropogenic aerosols interact with solar radiation and modify the incoming solar radiation at the Earth's surface. This is the case in particular of radiation in the wavelength range of 0.45–0.75 μm, also called photosynthetically active radiation (or PAR), which is used by vegetation for photosynthesis. For this reason aerosols can have a direct impact on ecosystem productivity and hence on the carbon cycle.

Looking at this in more details, aerosols exert antagonist effects on terrestrial ecosystems. First of all they can decrease the total amount of incoming solar radiation at the surface, which can decrease photosynthesis where light is the limiting factor. However, as we have seen earlier (Fig. 8.4), aerosols can also contribute to increase diffuse radiation at the surface. It has been shown that increased diffuse radiation can increase the primary productivity of terrestrial ecosystems because plants can make a better use of radiation across the different stages of the canopy. Diffuse radiation can penetrate deeper and more homogeneously into the canopy than direct radiation. A larger fraction of the surface of leaves, which is where photosynthesis takes place, is then illuminated. For moderate aerosol optical depths, the positive effect due to the increase in diffuse radiation wins over the negative effect due to the decrease in direct radiation. This has been verified experimentally for various types of ecosystems (Rocha et al. 2004; Mercado et al. 2009) and is well known to agronomists. At large aerosol optical depth, the opposite will be true and primary productivity can decrease. It has been estimated that the increase in diffuse radiation at the surface that has followed Mount Pinatubo eruption in 1991 is partly the cause for the vegetation greening and the increase in the terrestrial carbon sink that has then been observed (Gu et al. 2003). It should be noted, however, that the cooling associated with the volcanic eruption can also explain part of the observed increase in the carbon sink because of a decrease in soil respiration that essentially depends on soil temperature and humidity.

If stratospheric aerosols can have an impact on the primary productivity of terrestrial ecosystems, then it is conceivable, if not likely, that anthropogenic aerosols

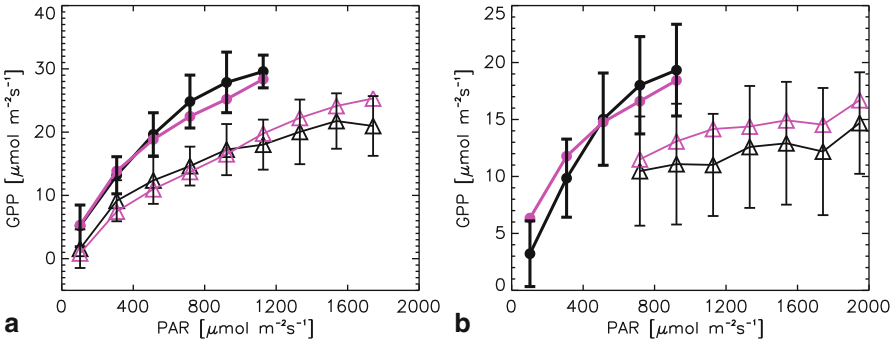


Fig. 11.1 Primary productivity (μmol m^{-2} s^{-1}) as a function of photosynthetically active radiation (PAR) for direct radiation (*triangles*) and diffuse radiation (*circles*) for **a** a temperate deciduous forest and **b** a temperate evergreen forest. (Adapted from Mercado et al. (2009))

have a similar impact. Mercado et al. (2009) have shown that the steady increase in aerosol concentrations since the middle of the twentieth century is responsible for an increase in the carbon sink by vegetation and soils. When aerosol concentrations stabilize, the effect on primary productivity continues, but the impact on the carbon budget fades away because the carbon cycle returns to a new equilibrium. The amount of carbon stored in the vegetation and soils stops increasing, and carbon sources to the atmosphere through vegetation and soil respiration can catch up. When aerosol concentrations decrease, which is what has started to occur over Europe and North America, then primary productivity decreases to return to its initial level and the carbon sink can turn into a carbon source. This is what could happen in the twenty-first century as air quality control policies will decrease even further aerosol concentrations (Fig. 11.1).

11.2.2 Aerosols as a Source of Nutrients

There exist other interactions between the aerosol cycle and the biogeochemical cycles of carbon and methane. Dry and wet deposition of aerosol particles at the surface is a source of nutrients that are necessary for plant growth. Nitrogen and phosphorus are two important macronutrients to which atmospheric aerosols contribute to the budget both in the marine and terrestrial environment. Anthropogenic inputs of nitrogen to terrestrial ecosystems come from various sources and result from the deposition of NO_y and NH_x compounds, including those of ammonium and nitrate aerosols. Nitrogen inputs have contributed to fertilise ecosystems and have contributed to the increase storage of carbon in vegetation and soils during the last century (Thornton et al. 2007). Amazonia is a region where the nutrient input from atmospheric aerosols is thought to be beneficial, especially in terms of phosphorus contained in mineral dust aerosols (Okin et al. 2004; Mahowald et al.

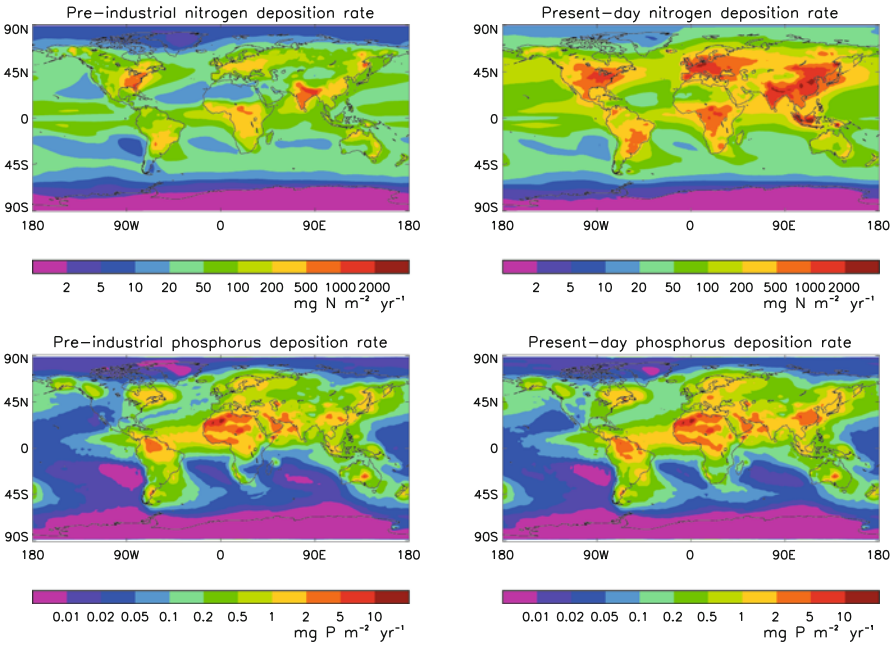


Fig. 11.2 Preindustrial (*left panels*) and present-day (*right panels*) deposition flux of nitrogen (N, *top panels*) and phosphorus (P, *bottom panels*). The nitrogen flux includes all gas-phase and aerosol nitrogen contributions. The phosphorus flux comes from desert dust, biomass burning and industrial aerosol. (Data are from Mahowald et al. (2008) and Lamarque et al. (2010). Adapted from Peñuelas et al. (2013))

2008). It has been shown that the Sahara region is an important source of dust for the Amazonian forest (Reichholf 1986; Koren et al. 2006) where ancient soils have been depleted in phosphorous content. Additionally, biomass burning aerosols also contain mineral substances that were in plants. Natural and anthropogenic vegetation fires contribute to redistribute macro and micronutrients spatially within terrestrial ecosystems. Likewise it is possible that PBAP redistribute some P and N nutrients on short distances, especially from adjacent unmanaged to managed ecosystems.

It should be noted that dry deposition of aerosols on leaves can also alter the humidity level of the leaves, gas exchanges through the stomata, photosynthesis and the plant resistance to pathogens. So, clearly the effects of aerosols on plants are multiple (Fig. 11.2).

11.2.3 Acidification of Precipitation

Raindrops contain a number of trace gases, some of which are acidic, some of which are alkaline, which come from aerosols scavenged by cloud- and rainwater. Aerosols

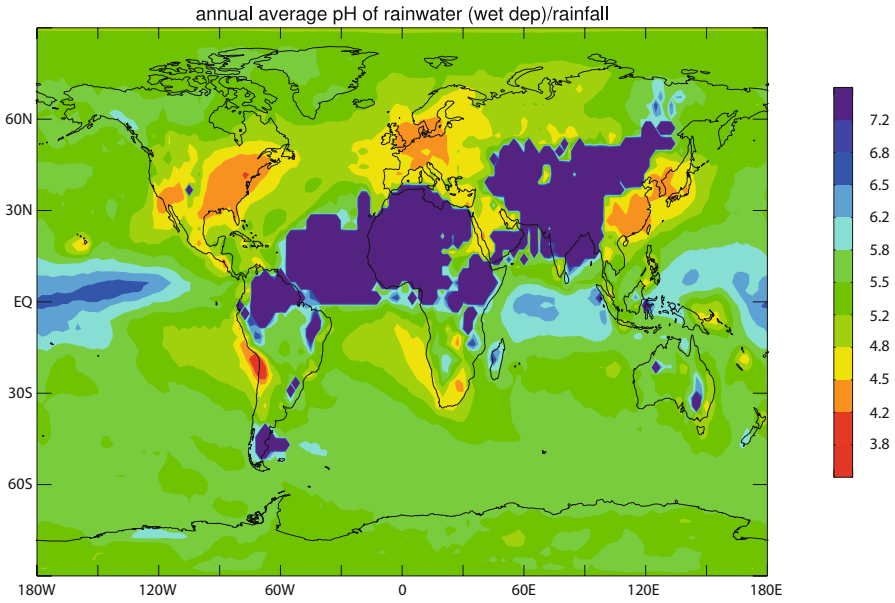


Fig. 11.3 Simulated distribution of rainwater pH. (From Rodhe et al. (2002))

therefore contribute to determine the pH of rainwater. It can be considered that in the absence of pollution the rain pH is generally about 5 to 6, although lower or larger pH can be found in some places. Rainwater becomes acidic in the presence of sulphuric acid (H_2SO_4) and nitric acid (HNO_3), both of which are strong acids. Organic acids can also contribute to decrease the rainwater pH. On the contrary ammonia (NH_3) and calcium carbonate ($CaCO_3$) are weak bases and can increase the pH of precipitations. The pH of rainwater can indeed be in the alkaline range in semiarid regions such as North Africa and India because of the alkaline property of soil aerosols that are scavenged by the rain. Overall, there are fairly large spatial variations in rainwater pH across the globe with both natural and anthropogenic variations (Rodhe et al. 2002; Fig. 11.3).

Anthropogenic emissions of SO_2 have led to rain acidification with typical pH values of 4 over North America and Europe, and values as low as 2 over the most polluted regions. Acid rain and, more generally acid deposition under all its forms, contribute to the acidification of soils and freshwater bodies such as lakes and streams. Acid rain has deleterious effects on some ecosystems and started to receive public attention in the 1960s (Likens and Bormann 1974). The extent to which soils can neutralize acid rain depends on many factors including the soil type and its chemical composition, its thickness, and how water runs off in the soil. Acid deposition adds hydrogen ions which swap places with some metal cations such as calcium, potassium, and magnesium ions which are then leached (i.e. washed out) from the top soil layer and are no longer available as micronutrients to plants. The severity

of the effects of acid rain on the vegetation therefore, depends on the acidity of the rainwater and the sensitivity of the soils to the acid deposition (Rodhe et al. 2002). Acidic deposition have damaged many forests in North America, and Central and Northern Europe, with associated effects on the carbon balance of these forests. Developed countries have then engaged into active emission controls of SO_2 and NO_x to reduce acid deposition and improve the health of their forests. This resulted in the "Gothenburg Protocol to Abate Acidification, Eutrophication and Ground-level Ozone" which was signed in 1999.

Acid deposition also has an impact on wetlands where it is suspected to decrease methane emissions. In anaerobic conditions (i.e. in the absence of oxygen), sulphur species favour sulphate-reducing bacteria and archaea at the expense of methanogenic archaea which produce methane. Gauci et al. (2004) estimated that acidic deposition have decreased the methane source from wetlands by 8 % since the preindustrial era. If acidic deposition decreases in the future, then wetland methane emissions could start increasing again. The same processes and changes in methane emissions also apply to managed wetlands such as paddy fields (Gauci et al. 2008).

11.3 Impact of Aerosols on Marine Ecosystems

Aerosols also have an impact on marine ecosystems, especially through an input of nutrients. In the coastal ocean, the nutrient budget is dominated by the riverine input, but in the open ocean the input from aerosol deposition can be an essential complement to the local recycling of nutrients. The role of aerosols is particularly important in so-called *high-nutrient low-chlorophyll* regions. It is considered that there is one micronutrient, most often iron, that is the limiting factor for photosynthesis by the phytoplankton. Such regions have been observed in the Southern Ocean and the equatorial and subarctic part of the Pacific Ocean. Desert (mineral) dust contain some iron and dust deposition that can turn out to be an essential source of iron for the phytoplankton (Jickells et al. 2005; Mahowald et al. 2005), including in some oceanic regions which are fairly remote from desertic areas. The iron that is contained in desert dust is relatively insoluble. However, the chemical composition of the dust can be altered as the aerosol ages, and the iron can become more soluble. These aging processes, which are essential to make the iron usable by the phytoplankton, is not well understood, but seem to depend on chemical reactions on the particle surface in the presence of a coating of sulphuric acid or organic acids. The bioavailability of iron for the phytoplankton depends on the iron solubility, which itself varies with particle size and aging. The input of iron by aerosols can modify the primary productivity in the ocean and therefore the air–sea gas exchange of carbon dioxide. Iron deposition is likely to have played a role in the Southern Ocean during glacial periods when the dust cycle was more active than in the present-day climate. The deposition of mineral aerosols would have stimulated photosynthesis and the biological pump. Watson et al. (2000) estimated that this process could explain a 40 ppm decrease in the atmospheric concentration of CO_2 between glacial and interglacial periods.

Aerosols can also modify the quantity and quality of solar radiation which penetrates the ocean and is used by the phytoplankton for photosynthesis. The ocean being a very diffusive medium, the repartition between direct and diffuse radiation at the ocean surface is not so important for marine ecosystems unlike for terrestrial ecosystems. However, the total amount of photosynthetically active radiation matters, as well as the amount of ultraviolet radiation reaching the surface (see Box). Ultraviolet radiation can stress on phytoplankton and induces some changes in the speciation and/or the amount of trace gases emitted by the phytoplankton such as DMS (Larsen 2005).

11.4 Aerosols–Atmospheric Chemistry Interactions

11.4.1 Interactions with Tropospheric Chemistry

Aerosols interact with atmospheric chemistry in a number of ways. Some aerosol precursors, such as DMS, SO_2, and volatile organic compounds, have to be oxidized in less volatile compounds before they can condense in the particulate phase. The concentration of oxidants control in part the atmospheric lifetime of these gaseous precursors and contribute to limit the formation of secondary aerosols. For instance, the availability of OH and H_2O_2 can be a limiting factor for the production of sulphate aerosols when the SO_2 concentrations are large and/or the concentrations of oxidants are low, as it is the case in mid- and high-latitudes of the winter hemisphere. Other oxidants such as O_3 or catalytic metal oxidation cycles in the aqueous phase may become important (Mauldin III et al. 2012). It is more unusual, however that the formation of secondary organic aerosols consume so much oxidant as to feedback on their concentrations. This is nevertheless the case for hydrogen peroxide (H_2O_2) which contributes to oxidizes SO_2 in aqueous phase and on which production is limited in polluted regions during wintertime.

Finally, it should be noted that the emissions of some gaseous precursors can modify the formation of secondary aerosols through a modification of the oxidizing capacity of the atmosphere. Unger et al. (2006) have suggested for instance that NO_x emissions are responsible for an increase in the concentration of sulphate aerosols.

Ultraviolet Radiation

A key characteristic of electromagnetic radiation is its wavelength (see Chap. 5). Given the temperature of its external layers, the Sun mostly emits in the ultraviolet, visible and near-infrared parts of the electromagnetic spectrum. Ultraviolet radiation (UV) only represents a small fraction of the energy radiated by the Sun, however this radiation plays an important role in the atmosphere. One usually distinguishes:

- UV-C radiation, with wavelengths 100–280 nm, is very energetic but is absorbed high up in the atmosphere and does not reach the Earth's surface;
- UV-B radiation, with wavelengths 280–315 nm, is energetic and the quantity that reaches the surface depends on the stratospheric ozone column but can be modulated by cloudiness, tropospheric ozone, stratospheric, and tropospheric aerosols;
- UV-A radiation, with wavelengths 315–400 nm, is less energetic and is not absorbed much by atmospheric ozone; the quantity received at the surface depends essentially on cloudiness and the amount of stratospheric and tropospheric aerosols.

The amount of UV-A and UV-B received at the surface also depends on the altitude (there is more ultraviolet radiation at the mountain level than at the sea level), the position of the Sun in the sky (which itself depends on the latitude, season, and hour in the day) and the surface reflectivity (reflective surfaces such as snow, white sand or a rough sea increases UV radiation at the surface). Finally, the amount of UV received by a surface depends on its orientation relative to the Sun: A horizontal surface can receive more or less UV radiation than a vertical surface. The shorter the wavelength, the more energetic is the UV radiation, and therefore the more harmful it is to human health and to ecosystems. The efficacy of UV radiation is maximum between 280 and 300 nm and decreases exponentially between 300 and 340 nm. A one-unit dose of UV at 300 nm wavelength has the same effect as a 5000-unit dose at 340 nm wavelength. It is therefore usual to weigh the spectrum of UV radiation at the surface with an action spectrum that takes into account the harmfulness of the different wavelengths in a differentiated way. A typical action spectrum for the skin, called erythemal action spectrum, is often used. When such a spectrum is considered, UV-B have the largest impact but UV-A cannot be neglected. The erythemal function can be defined as the integral over the ultraviolet spectrum of the product of the flux of ultraviolet radiation at the surface with the erythemal action spectrum. When multiplied by an arbitrary factor of 40, this erythemal function provides a UV index, which is simple to use as it varies on a scale ranging typically from 0 to 16, although in principle larger values are possible. One can then estimate a maximum exposure time for each UV index and skin type before there is a risk of sunburn.

Aerosol surfaces are also the locus for chemical reactions that can modify the concentrations of a number of tropospheric chemical species. Dust particles have been identified as important for heterogeneous chemistry early on (Dentener et al. 1996). This represents a sink for tropospheric ozone and other atmospheric reactive species (Bauer et al. 2004). The magnitude of the ozone sink depends however on the aerosol surface concentration, the ozone sticking coefficient on the aerosol particles, and the potential deactivation of active sites when a particle ages and has its surface properties altered. Other aerosol types, such as carbonaceous aerosols, may be important for heterogeneous chemistry, but their role is even less understood (Nienow and Roberts 2006).

Finally, aerosols modify the distribution of solar radiation in the atmosphere which affect actinic fluxes (see Sect. 5.6.6 and Fig. 5.18), photolysis rates and hence a number of photochemical species such as ozone.

11.4.2 Impact of Stratospheric Aerosols on the Ozone Layer and Ultraviolet Radiation

Stratospheric ozone offers an indispensable protection against ultraviolet radiation originating from the Sun (see Box). Stratospheric aerosols contribute to the formation of polar stratospheric clouds (PSC). Chemical reactions occur on the surface of these cloud particles, which plays a central role in the destruction of the ozone layer by chlorinated species stemming from chlorofluorocarbons.

In the absence of a polar stratospheric cloud, stratospheric aerosols can also trigger heterogeneous chemical reactions which take place on their surface and lead to the activation of chlorinated species and to the destruction of stratospheric ozone. A number of measurements confirm that large concentrations in stratospheric aerosols, such as those encountered after the eruption of the Pinatubo volcano, were responsible for an enhanced destruction of stratospheric ozone at the beginning of the 1990s (Solomon 1999).

11.5 Climate Feedbacks Involving Marine Aerosols

11.5.1 Sulphate Aerosols from DMS Emissions

Dimethylsulphide originates from the decomposition of dimethylsulphoniopropionate (DMSP) produced by some marine organisms, in particular by some phytoplankton species. It is found in very variable concentrations in surface waters of the ocean (Kettle and Andreae 2000; Lana et al. 2011) where it can be emitted to the atmosphere and contribute to the formation of aerosols. The oxidation of dimethyl sulphide (DMS) starts with the reaction of DMS with the OH and NO_3

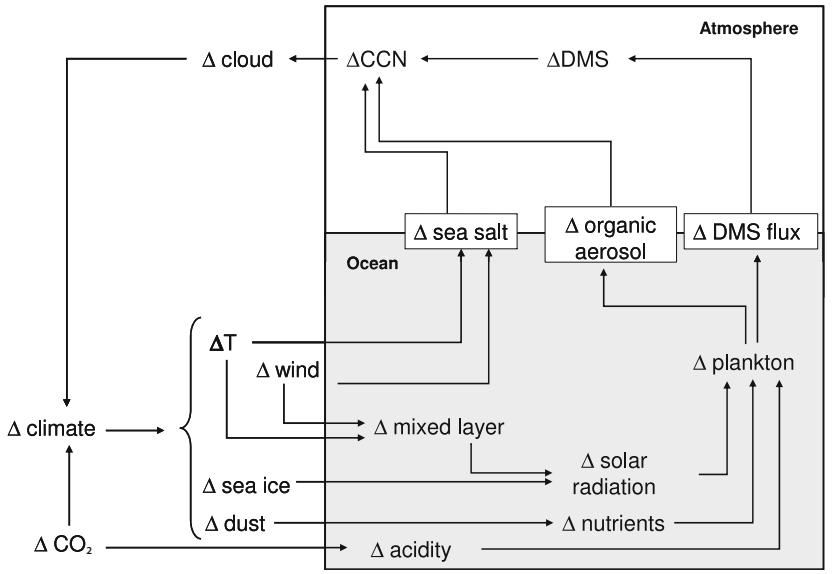


Fig. 11.4 Earth system feedbacks involving marine aerosols. (Adapted from Carslaw et al. (2010))

radicals but the oxidation chain also involves ozone (O_3), the HO_2 radicals and hydrogen peroxide (H_2O_2) and leads to the production of sulphur dioxide (SO_2), which can get deposited to the surface, or can be further oxidized to methanesulphonic acid (MSA) or sulphuric acid (H_2SO_4) which can then condense on pre-existing particles or nucleate to form new particles. The climate impact of DMS depends essentially on the capacity of its oxidation products to serve as cloud condensation nuclei and to modify the planetary albedo through aerosol–cloud interactions (e.g. Charlson et al. 1987). The idea has been put forward that DMS emissions modify the concentrations of cloud condensation nuclei, cloud microphysical and radiative properties and hence climate, to which phytoplankton and DMS emissions will then respond. This hypothesis is known as the CLAW hypothesis (from the initials of the surname of its authors). It has often been presented as a negative feedback loop by which phytoplankton can help to stabilize the climate system: a cooler climate system would have less DMS emissions, less cloud condensation nuclei (CCN), and brighter clouds, which would warm the climate. However, the sign of the feedback is uncertain and has been presented as such by Charlson et al. (1987). It is unclear if this feedback loop has been important in past climates, but if it were the case, it is certainly less important in the present climate, in particular in the Northern Hemisphere and polluted regions of the Southern Hemisphere where the sulphur cycle is completely dominated by anthropogenic emissions of SO_2. Moreover, the more this feedback loop is studied, the more scientists realize how complex it is (Fig. 11.4; Ayers and Cainey 2007; Cameron-Smith et al. 2011). For instance, the climate change associated with a modification of DMS emissions would not only change sea surface

temperatures, but also surface winds which would modify the air–sea transfer of DMS and the depth of the oceanic mixed layer where the phytoplankton resides. If the ocean stratification increases and the mixed layer depth is reduced, as could be the case with global warming, this would concentrate the phytoplankton in surface waters, where available solar radiation is larger, which would contribute to increase primary productivity. It has been shown indeed that the production of DMSP increases with the amount of solar radiation available to the phytoplankton. Inversely, an increase in ocean stratification could lead to a reduction of the flux of nutrients from depth, with an opposite effect on phytoplankton, primary productivity and DMS production. Climate change may also lead to a shift in the equilibrium between different phytoplanktonic species. A modification of surface winds could play in one direction or the other on the mixed layer depth and the DMS flux. Other processes could also play a role:

- The retreat of the sea-ice cover can expose a larger fraction of the ocean to solar radiation, which would increase the DMS production (Gabric et al. 2005);
- Ocean acidification could also disturb the ecosystem equilibrium and DMS emissions. However, the response of the various species of phytoplankton to the seawater pH is not well understood;
- A modification in the iron deposition flux from airborne dust could have an impact on the primary productivity in certain regions of the ocean which are iron-limited;
- The amount of ultraviolet radiation can modify both the primary productivity and the rate of DMS destruction in the seawater;
- A modification of the transport and mixing in the atmosphere could modify the efficacy with which the DMS is converted into cloud condensation nuclei.
- It is increasingly realized that additional DMS emissions may not enhance CCN concentrations that much because DMS tend to condense onto pre-existing particles rather than to forming new particles (Woodhouse et al. 2010; Quinn and Bates 2011).

A number of simulations have been realised with earth system models which include some but in any case not all of the above-mentioned processes. Overall, these simulations predict a modest increase in the DMS flux to the atmosphere in response to global warming (Bopp et al. 2004; Vallina et al. 2007; Kloster et al. 2007). Regional variations are important and it is conceivable that the CLAW hypothesis can modify the atmospheric aerosol if not the climate at the regional scale.

11.5.2 Marine Aerosols

Sea spray aerosols interact with radiation, they are also good cloud condensation nuclei and can be present in sufficient quantities to modify the cloud microphysical properties under suitable conditions. Their production rate depends principally on the surface windspeed, the sea surface conditions and to a lesser extent the sea surface

temperature. A change in the intensity of surface winds in response to global warming could therefore modify the emissions of sea spray.

The surface wind speed has increased significantly in mid- and high-latitudes of the Southern Hemisphere, with an impact on air–sea exchange of carbon dioxide (Le Quéré et al. 2007). This could be a specific response to the depletion of the stratospheric ozone layer in the Antarctic vortex, which is expected to be reversible when the stratospheric ozone layer will have recovered towards the second half of the twenty-first century. It is possible that emissions of DMS and sea spray have also increased and in response to this surface windspeeds increses (Korhonen et al. 2010). Satellite observations show that the increase in surface windspeeds is not restricted to the high latitudes of the Southern Hemisphere (Young et al. 2011). Processes responsible for this increase are not known.

Penner et al. (2001) have shown that some models predict an increase in the intensity of surface winds while this is not the case of other models such as HadGEM2 (Jones et al. 2007). However, these authors have shown that sea-ice melting will mechanically increase emissions of sea spray as the surface of open ocean increases. Moreover the roughness length of the open ocean is shorter than that of sea ice, which further increases the windspeed at the surface.

11.5.3 Other Aerosols of Maritime Origin

The ocean is a source of primary and also secondary organic aerosols. The correlation between marine organic aerosols and chlorophyll, and the pronounced seasonal cycle in the concentrations of organic aerosols suggest a biogenic source for this organic matter (O'Dowd et al. 2004). The source of marine organic aerosols is uncertain and not well understood so it is only possible to speculate on their radiative effects and their potential role to climate feedback loops. Spracklen et al. (20008b) have estimated a source of marine organic carbon of 8 Tg yr^{-1}. It is conceivable that biogenic organic aerosols modify the number, concentration and properties of the cloud condensation nuclei (sulphates and sea spray) over marine regions.

11.6 Climate Feedbacks Involving Continental Aerosols

The terrestrial biosphere emits primary biological aerosol particles and trace gases that get oxidized into organic compounds, some of which can condense and form secondary aerosols. The terrestrial biosphere can also emit large amounts of aerosols during biomass burning episodes. We now turn our attention to how these aerosol species may respond to climate change.

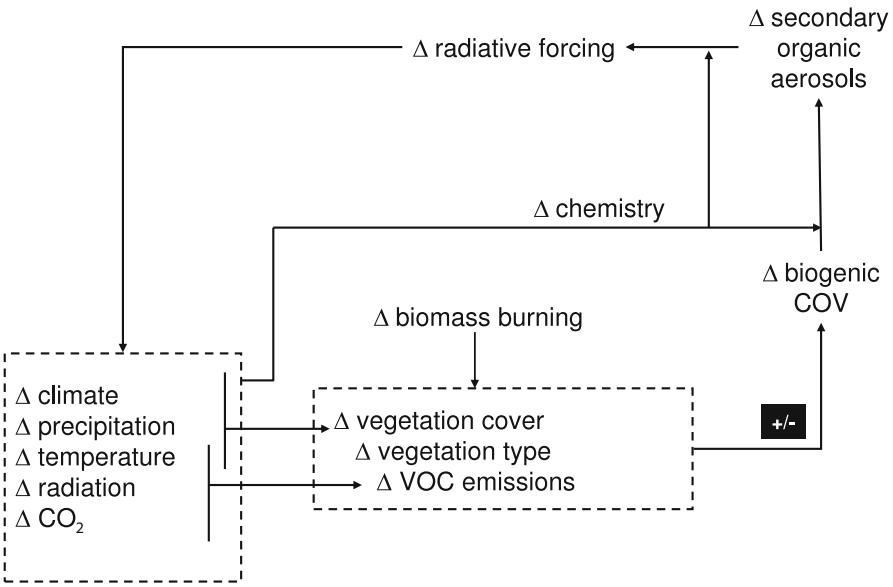


Fig. 11.5 Earth system feedbacks involving terrestrial biogenic aerosols. (Adapted from Carslaw et al. (2010))

11.6.1 Secondary Organic Aerosols

The biosphere emits biogenic volatile organic compounds (BVOC) such as isoprene, terpenes, acetone and methanol with a global source estimated at 1100 Tg C yr^{-1} (Guenther et al. 1995, 2012). The global isoprene flux is in the range of 400–600 Tg C per year and the global monoterpene flux is in the range of 30–130 Tg C per year. These species react in the atmosphere to form secondary organic aerosols with a source term typically estimated to be 12–70 Tg C per year, but even that range may not span all uncertainties of the aerosol formation mechanism. It has been shown that biogenic secondary aerosols can be the dominant aerosol type in forested continental regions where the influence of anthropogenic emissions is less. Moreover, secondary organic aerosols are generally of submicronic size and can serve as cloud condensation nuclei.

Figure 11.5 shows possible feedback loops associated with secondary organic aerosols. Meteorology and climate control strongly BVOC emissions (Peñuelas and Staudt 2010). It is thought that an increase in surface temperature would lead to an increase in BVOC emissions and cooling effect from secondary organic aerosols (Makkonen et al. 2012). Everything else being equal, this would imply a negative feedback loop. However, the atmospheric concentration in organic aerosols is not only controlled by BVOC emissions. For instance, the condensation rate of volatile organic compounds decreases with increasing temperature. Moreover, BVOC emissions do not depend only on air temperature. Laboratory measurements and field

experiments have shown that emissions are controlled by a complex interplay of parameters such as leaf temperature, solar radiation, nutrient availability, soil humidity, vegetation type, biomass density or atmospheric concentration of VOCs (Peñuelas and Staudt 2010). Some measurements suggest that the CO_2 atmospheric concentration has an inhibition effect on isoprene emissions (Arneth et al. 2007), which would offset the strength of the above-mentioned feedback loop (Fig. 4.4). Palmer et al. (2006) have detected large interannual variations in isoprene emissions of the order of 20–30 % over the United States, which they attribute essentially to temperature variations. Several studies suggest an increase in isoprene emissions of 20 % to more than 50 % by 2100 for fixed vegetation type and amount. There are fewer studies looking at future emissions of monoterpenes but these could increase for the same reasons. The response of vegetation to climate change could increase BVOC emissions beyond what would be expected from a temperature change only. However, the replacement of natural vegetation by crops could decrease the emissions of isoprene as compared to the present-day situation (Lathière et al. 2010), although there are some specific species (e.g. used for biofuel production) that could lead to an increase in emissions locally. It is unclear whether higher CO_2 concentrations can inhibit monoterpene emissions as it seems to be the case for isoprene. In conclusion, uncertainties are large but rough estimates indicate that climate change could increase the concentrations of secondary organic aerosols of biogenic origin by 25–150 % by 2100. This would lead to a radiative forcing due to aerosol–radiation interactions of a few tenths of Wm^{-2} and ultimately a fairly small feedback factor (Paasonen et al. 2013).

11.6.2 Primary Aerosols of Biogenic Origin

Continental ecosystems emit different types of primary biological aerosol particles (PBAP) among which viruses (diameters less than 0.3 μm), bacteria (diameters between 0.3 and 10 μm), spores (diameters between 1 and 30 μm), pollen (diameters between 10 and 100 μm) and plant debris can be found. These aerosols sources are poorly quantified and factors controlling the emissions remain not well understood and partly unknown. It is therefore difficult to estimate how climate change could modify the concentration of PBAP. The feedback loop associated with the interaction of these aerosols with radiation is probably very weak because their radiative effects are weak in the first place. However, the ability of these particles to serve as ice nuclei may give them a disproportionate role in the water cycle and we cannot exclude a feedback loop that would operate through aerosol–cloud interactions.

The role of surface winds in the emission process of primary biogenic aerosol particles is not known but it should be kept in mind that, just like other than the ocean, climate change may induce a change in surface windspeed. Vautard et al. (2010) have observed a slowdown of winds over continental surfaces over the last decades which they have partly attributed to an increase in the surface roughness because of vegetation changes.

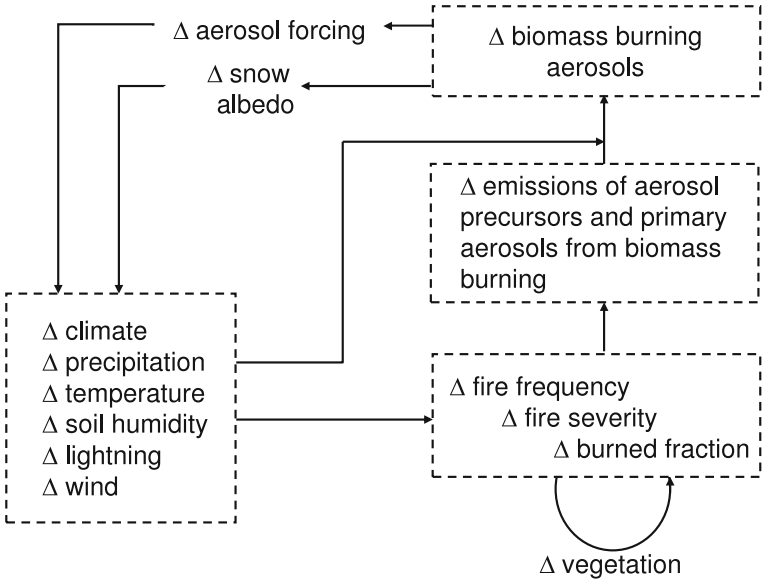


Fig. 11.6 Earth system feedbacks involving biomass burning aerosols. (Adapted from Carslaw et al. (2010))

11.6.3 Aerosols from Vegetation Fires

Aerosols from vegetation fires are also important because of their light-absorbing properties. There is an uncertainty in the sign of the radiative forcing associated with biomass burning aerosols but this radiative forcing is thought to be weak. However, their atmospheric radiative forcing is larger and there is likely to be an impact on precipitation at least on the regional scale. Aerosol–cloud interactions, aerosol–vegetation interactions and aerosol–surface interactions may also play a role in such a feedback.

There is no doubt that climate can control the frequency and severity of natural vegetation fires. Most published studies suggest an increase in natural vegetation fires in the future. Moreover, it is not known that how feedback loops could operate (Fig. 11.6). In the Arctic region, it is possible that vegetation fires in neighbouring regions induce further snowmelt in the northernmost regions because of the deposition of absorbing aerosols on the snowpack.

It is important to underline that the distinction between natural and human-induced vegetation fires is difficult to make. Schematically one can say that (i) emissions due to deforestation in the tropics have increased with population during the last two centuries and are largely anthropogenic; (ii) forest fires in the northern hemisphere midlatitudes are largely natural but have decreased during the last century because of fire control policies; (iii) some ecosystems such as savannas are essentially regulated

by fires and would burn regularly even in the absence of anthropogenic pressures, although they would probably do so less frequently.

Vegetation fires are extremely variable from one year to the next, in particular, in response to drought conditions such as that induced regionally by the El Niño phenomenon (van der Werf et al. 2004). Decadal variations in vegetation fires have also been observed in some regions (Westerling et al. 2006). Variations on longer timescales relate to both climate change and anthropogenic activities (Marlon et al. 2008; Archibald et al. 2011).

Vegetation fires require the presence of fuel (organic matter), suitable conditions for the combustion to occur and an ignition source. Humidity is the most important climate variable in relation to vegetation fires, which is why drought severity indices are often used as an indicator of fire risk. It is also possible that climate change leads to an increase in lightning and cloud-to-ground strikes and therefore in potential for fire ignition (Price and Rind 1994). Many studies predict an increase in vegetation fires in the future, be it in United States or Canada (Flannigan et al. 2005), Russia (Stocks et al. 1998) or in Europe (Moriondo et al. 2006). Even though a majority of studies predict an increase in vegetation fires, some studies suggest a decrease in burned areas in some regions, such as Eastern Canada because of the increase in precipitation predicted in boreal regions (Flannigan et al. 2001). Moreover, all these studies do not take into account a potential modification of vegetation.

Aerosol emissions are expected to increase with the number of vegetation fires. However, aerosol emissions may not increase as rapidly as burned surfaces because there could be a negative feedback loop by which an increase in the number of vegetation fires may result in a decrease in the quantity of biomass susceptible to burn.

The interaction between vegetation fires and climate is not restricted to aerosols. A more comprehensive approach would require to simultaneously take into account greenhouse gas emissions, aerosol emissions, modification of surface albedo, and possible feedbacks between vegetation and production of biogenic volatile organic compounds (Randerson et al. 2006; Spracklen et al. 2008a).

11.6.4 Desert Dust

Desert dust particles interact with solar radiation, which contributes to cool the climate, and with terrestrial radiation, which contributes to increase the greenhouse effect and warm the climate. The relative impact of these two effects depend on the particle size, their mineral composition, their altitude and the surface albedo. Aerosol–cloud interactions specific to desert dust are much less well-understood. Soluble compounds get deposited on desert dust particles, which can become giant cloud condensation nuclei, and play a role in rain initiation. Desert dust can also diminish the concentration of cloud condensation nuclei by condensing soluble species on their surface rather than on a larger number of smaller particles. Desert dust particles are also known to be good ice nuclei, although this again may depend on their

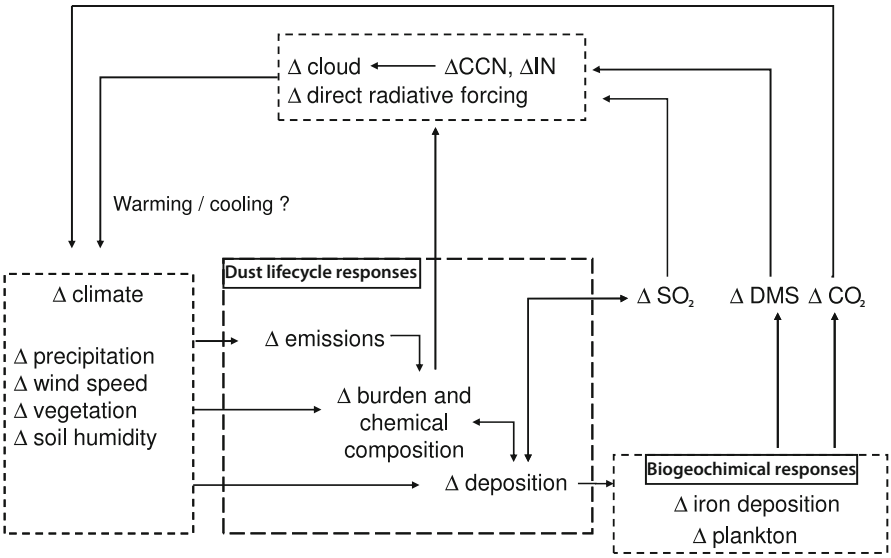


Fig. 11.7 Earth system feedbacks involving mineral dust aerosols. (Adapted from Carslaw et al. (2010))

mineralogical composition (Hoose and Möhler 2012; Atkinson et al. 2013). Dry and wet deposition of dust on snow- and ice-covered surfaces contribute to decrease the surface albedo and warm the surface (Krinner et al. 2006). Finally, desert dust can be an important source of nutrients for some ecosystems, such as the Amazon forest (Okin et al. 2004).

Surface wind speed, soil humidity, and vegetation cover are three important factors that govern desert dust emissions. These factors vary over multiple timescales because of both natural and anthropogenic drivers. Attributing recent variations in desert dust to a natural and/or anthropogenic drivers is therefore difficult, and choosing a reference period is essentially arbitrary. Some studies have attributed a significant part of the increase in desert dust to a change in land use in the Sahel region (Moulin and Chiapello 2006; Mulitza et al. 2010). However, this increase could also be due to natural decadal modes of variability in the climate system. Long-term measurements made in the Barbados in the Western Tropical Atlantic Ocean show a fourfold increase in concentrations since 1965 with interannual and interdecadal variations in dust concentrations related to drought index in the Sahel (Middleton 1985) and the North Atlantic Oscillation (Chiapello and Moulin 2002). Ginoux et al. (2012) attribute 75 % of the global dust emissions to natural sources and the remaining 25 % to anthropogenic sources. Over the Sahel, the fraction of dust emissions of anthropogenic origin could be less.

It is unclear whether future climate change will contribute to increase or decrease the emissions of desert dust (Fig. 11.7). Woodward et al. (2005) simulated a threefold increase by 2100 because of a decrease in the vegetation cover simulated by their

model in response to shifts in precipitation patterns associated with climate change. In contrast Mahowald et al. (2006) predicted a 60 % decrease in emissions for a doubling CO_2. The CO_2 fertilization effect on vegetation could instead increase the vegetation cover in semiarid regions and decrease the fraction of bare soil.

A change in the cycle of desert dust could have an impact on marine ecosystems. Case studies have shown correlations between the deposition flux of desert dust and primary productivity in regions limited by nutrients. It is conceivable that future variations in the dust cycle could modulate DMS emissions and their impact on aerosol concentrations and properties (Kloster et al. 2007).

11.7 Climate Feedbacks Involving Stratospheric Aerosols

We have seen in Chap. 12 that stratospheric aerosols form a layer between the tropopause and an altitude of about 25 km. In background conditions (i.e. in the absence of volcanic eruptions in the stratosphere), this layer is very thin with an optical depth of 0.005 at 550 nm. It is therefore very unlikely that a feedback loop that would increase or decrease the background stratospheric aerosol layer could have a significant climate impact.

Climate change is also very unlikely to modify the frequency and magnitude of explosive volcanic eruptions in any significant manner. It is conceivable however that climate change modifies the atmospheric cycle of volcanic aerosols when they are injected and formed in the stratosphere, which would modulate their climate impact. Climate change could indeed modify the transport and mixing of aerosols in the stratosphere. Butchart and Scaife (2001) showed that tropospheric–stratospheric exchanges will increase with global warming. This increase is observed in many climate models at an average rate of 2 % per decade. The acceleration of the Brewer–Dobson circulation could result in a small reduction in the residence time of stratospheric aerosols.

Exercise

1. A coupled vegetation–atmosphere-chemistry model predicts that a 2°C global-mean warming induces an increase in natural BVOC emissions of 10 %. Assuming a radiative effect due to aerosol–radiation interactions of -1 Wm^{-2} for natural secondary organic aerosols, estimate the feedback parameter of this feedback. How does it compare to the strength of some known physical feedbacks?

Solution

1. The feedback parameter is $\alpha = -0.05$ $Wm^{-2}{}^{\circ}C^{-1}$.

References

Archibald S, Staver AC, Levin SA (2011) Evolution of human-driven fire regimes in Africa. Proc Natl Acad Sci U S A 109:847–852

Arneth A, Miller PA, Scholze M, Hickler T, Schurges G, Smith B, Prentice IC (2007) CO_2 inhibition of global terrestrial isoprene emissions: Potential implications for atmospheric chemistry. Geophys Res Lett 34:L18813. doi:10.1029/2007GL030615

Atkinson JD, Murray BJ, Woodhouse MT, Whale TF, Baustian KJ, Carslaw KS, Dobbie S, O'Sullivan D, Malkin TS (2013) The importance of feldspar for ice nucleation by mineral dust in mixed-phase clouds. Nature 498:355–358

Ayers GP, Cainey JM (2007) The CLAW hypothesis: a review of the major developments. Environ Chem 4:366–374

Bauer SE, Balkanski Y, Schulz M, Hauglustaine DA, Dentener F (2004) Global modeling of heterogeneous chemistry on mineral aerosol surfaces: influence on tropospheric ozone chemistry and comparison to observations. J Geophys Res 109:D02304. doi:10.1029/2003JD003868

Bopp L, Boucher O, Aumont O, Belviso S, Dufresne J-L, Pham M, Monfray P (2004) Will marine dimethylsulphide emissions alleviate global warming? A model study. Can J Fish Aquat Sci 61(5):826–835

Butchart N, Scaife AA (2001) Removal of chloroflurocarbons by increased mass exchange between the stratosphere and troposphere in a changing climate. Nature 410:799–802

Cameron-Smith P, Elliott S, Maltrud M, Erickson D, Wingenter O (2011) Changes in dimethyl sulfide oceanic distribution due to climate change. Geophys Res Lett 38:L07704. doi:10.1029/2011GL047069

Carslaw K, Boucher O, Spracklen DV, Mann G, Rae JGL, Woodward S, Kulmala M (2010) Aerosol in the earth system: a review of interactions and feedbacks. Atmos Chem Phys 10:1701–1737

Charlson RJ, Lovelock JE, Andreae MO, Warren SG (1987) Oceanic phytoplankton, atmospheric sulfur, cloud albedo and climate. Nature 326:655–661

Chiapello I, Moulin C (2002) TOMS and METEOSAT satellite records of the variability of Saharan dust transport over the Atlantic during the last two decades (1979–1997). Geophys Res Lett 29:1176. doi:10.1029/ 2001GL13767

Dentener FJ, Carmichael GR, Zhang Y, Lelieveld J, Crutzen PJ (1996) Role of mineral aerosol as a reactive surface in the global troposphere. J Geophys Res 101:22869–22889

Flannigan M, Campbell I, Wotton M, Carcaillet C, Richard P, Bergeron Y (2001) Future fire in Canada's boreal forest: paleoecology results and general circulation model regional climate model simulations. Can J For Res 31:854–864

Flannigan MD, Loganv KA, Amiro BD, Skinner WR, Stocks BJ (2005) Future area burned in Canada. Clim Change 72:1–16

Gabric AJ, Qu B, Matrai P, Hirst AC (2005) The simulated response of dimethylsulfide production in the Arctic Ocean to global warming. Tellus 57B:391–403

Gauci V, Matthews E, Dise N, Walter B, Koch D, Granberg G, Vile M (2004) Sulfur pollution suppression of the wetland methane source in the 20th and 21st centuries. Proc Natl Acad Sci U S A 101:12583–12587

Gauci V, Dise NB, Howell G, Jenkins ME (2008) Suppression of rice methane emission by sulfate deposition in simulated acid rain. J Geophys Res 113:G00A07. doi:10.1029/2007JG000501

Ginoux P, Prospero JM, Gill TE, Hsu NC, Zhao M (2012) Global-scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS Deep Blue aerosol products. Rev Geophys 50:RG3005. doi:10.1029/2012RG000388

Gu L, Baldocchi DD, Wofsy SC, Munger JW, Michalsky JJ, Urbanski SP, Boden TA (2003) Response of a deciduous forest to the Mount Pinatubo eruption: enhanced photosynthesis. Science 299:2035–2038

Guenther A, Hewitt CN, Erickson D, Fall R, Geron C, Graedel T, Harley P, Klinger L, Lerdau M, Mckay WA, Pierce T, Scholes B, Steinbrecher R, Tallamraju R, Taylor J, Zimmerman P (1995) A global model of natural volatile organic compound emissions. J Geophys Res 100:8873–8892

Guenther AB, Jiang X, Heald CL, Sakulyanontvittaya T, Duhl T, Emmons LK, Wang X (2012) The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions. Geosci Model Dev 5:1471–1492

Hoose C, Möhler O (2012) Heterogeneous ice nucleation on atmospheric aerosols: a review of results from laboratory experiments. Atmos Chem Phys 12:9817–9854

Jickells TD, An ZS, Andersen KK, Baker AR, Bergametti G, Brooks N, Cao JJ, Boyd PW, Duce RA, Hunter KA, Kawahata H, Kubilay N, laRoche J, Liss PS, Mahowald N, Prospero JM, Ridgwell AJ, Tegen I, Torres R (2005) Global iron connections between desert dust, ocean biogeochemistry, and climate. Science 308:67–71

Jones A, Haywood JM, Boucher O (2007) Aerosol forcing, climate response and climate sensitivity in the Hadley Centre climate model HadGEM2-AML. J Geophys Res 112:D20211. doi:10.1029/2007JD008688

Kettle AJ, Andreae MO (2000) Flux of dimethylsulfide from the oceans: a comparison of updated data sets and flux models. J Geophys Res 105:26973–26808

Kloster S, Six KD, Feichter J, Maier-Reimer E, Roeckner E, Wetzel P, Stier P, Esch M (2007) Response of dimethylsulfide (DMS) in the ocean and atmosphere to global warming. J Geophys Res 112:G03005. doi:10.1029/2006JG000224

Koren I, Kaufman YJ, Washington R, Todd MC, Rudich Y, Martins JV, Rosenfeld D (2006) The Bodélé depression: a single spot in the Sahara that provides most of the mineral dust to the Amazon forest. Environ Res Lett 1:014005. doi:10.1088/1748-9326/1/1/014005

Korhonen H, Carslaw KS, Forster PM, Mikkonen S, Gordon ND, Kokkola H (2010) Aerosol climate feedback due to decadal increases in Southern Hemisphere wind speeds. Geophys Res Lett 37:L02805. doi:10.1029/2009GL041320

Krinner G, Boucher O, Balkanski Y (2006) Reduced glacial ice sheet extent in Northern Asia and Alaska owing to deposition of mineral dust on snow. Clim Dyn 27:613–625. doi:10.1007/s00382-006-0159-z

Lamarque J-F, Bond TC, Eyring V, Granier C, Heil A, Klimont Z, Lee D, Liousse C, Mieville A, Owen B, Schultz MG, Shindell D, Smith SJ, Stehfest E, Van Aardenne J, Cooper OR, Kainuma M, Mahowald N, McConnell JR, Naik V, Riahi K, van Vuuren DP (2010) Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. Atmos Chem Phys 10:7017–7039

Lana A, Bell TG, Simó R, Vallina SM, Ballabrera-Poy J, Kettle AJ, Dachs J, Bopp L, Saltzman ES, Stefels J, Johnson JE, Liss PS (2011) An updated climatology of surface dimethlysulfide concentrations and emission fluxes in the global ocean. Global Biogeochem Cycles 25:GB1004. doi:10.1029/2010GB003850

Larsen SH (2005) Solar variability, dimethyl sulphide, clouds and climate. Global Biogeochem Cycles 19:GB1014. doi:10.1029/2004GB002333

Lathière J, Hewitt CN, Beerling DJ (2010) Sensitivity of isoprene emissions from the terrestrial biosphere to 20th century changes in atmospheric CO_2 concentration, climate, and land use. Global Biogeochem Cycles 24:GB1004. doi:10.1029/2009GB003548

Le Quéré C, Rödenbeck C, Buitenhuis ET, Conway TJ, Langenfelds R, Gomez A, Labuschagne C, Ramonet M, Nakazawa T, Metzl N, Gillett N, Heimann M (2007) Saturation of the Southern Ocean CO_2 sink due to recent climate change. Science 316:1735–1738

Likens GE, Bormann FH (1974) Acid rain: a serious regional environmental problem. Science 184:1176–1179

Mahowald NM, Baker AR, Bergametti G, Brooks N, Duce RA, Jickells TD, Kubilay N, Prospero JM, Tegen I (2005) Atmospheric global dust cycle and iron inputs to the ocean. Global Biogeochem Cycles 19:GB4025. doi:10.1029/2004GB002402

Mahowald NM, Muhs DR, Levis S, Rasch PJ, Yoshioka M, Zender CS, Luo C (2006) Change in atmospheric mineral aerosols in response to climate: last glacial period, preindustrial, modern, and doubled carbon dioxide climates. J Geophys Res 111:D10202. doi:10.1029/2005JD006653

Mahowald N, Jickells TD, Baker AR, Artaxo P, Benitez-Nelson CR, Bergametti G, Bond TC, Chen Y, Cohen DD, Herut B, Kubilay N, Losno R, Luo C, Maenhaut W, McGee KA, Okin

GS, Siefert RL, Tsukuda S (2008) Global distribution of atmospheric phosphorus sources, concentrations and deposition rates, and anthropogenic impacts. Global Biogeochem Cycles 22:GB4026. doi:10.1029/2008GB003240

Mahowald N, Ward DS, Kloster S, Flanner MG, Heald CL, Heavens NG, Hess PG, Lamarque J-F, Chuang PY (2011) Aerosol impacts on climate and biogeochemistry. Annu Rev Environ Res 36:45–74

Makkonen R, Asmi A, Kerminen VM, Boy M, Arneth A, Guenther A, Kulmala M (2012) BVOC–aerosol–climate interactions in the global aerosol–climate model ECHAM5.5-HAM2. Atmos Chem Phys 12:10077–10096

Marlon JR, Bartlein PJ, Carcaillet C, Gavin DG, Harrison SP, Higuera PE, Joos F, Power MJ, Prentice IC (2008) Climate and human influences on global biomass burning over the past two millennia. Nat Geosci 1:697–702

Mauldin III RL, Berndt T, Sipilä M, Paasonen P, Petäjä T, Kim S, Kurtén T, Stratmann F, Kerminen V-M, Kulmala M (2012) A new atmospherically relevant oxidant of sulphur dioxide. Nature 488:193–196

Mercado LM, Bellouin N, Sitch S, Boucher O, Huntingford C, Cox P (2009) Impact of changes in diffuse radiation on the global land carbon sink. Nature 458:1014–1017

Middleton NJ (1985) Effect of drought on dust production in the Sahel. Nature 316:431–434

Moriondo M, Good P, Durao R, Bindi M, Giannakopoulos C, Corte-Real J (2006) Potential impact of climate change on fire risk in the Mediterranean area. Clim Res 31:85–95

Moulin C, Chiapello I (2006) Impact of human-induced desertification on the intensification of Sahel dust emission and export over the last decades. Geophys Res Lett 33:L18808. doi:10.1029/2006GL025923

Multiza S, Heslop D, Pittauerova D, Fischer HW, Meyer I, Stuut J-B, Zabel M, Mollenhauer G, Collins JA, Kuhnert H, Schulz M (2010) Increase in African dust flux at the onset of commercial agriculture in the Sahel region. Nature 466:226–228

Nienow AM, Roberts JT (2006) Heterogeneous chemistry of carbon aerosols. Annu Rev Phys Chem 57:105–128

O'Dowd CD, Facchini MC, Cavalli F, Ceburnis D, Mircea M, Decesari S, Fuzzi S, Yoon YJ, Putaud J-P (2004) Biogenically driven organic contribution to marine aerosol. Nature 431:676–680

Okin GS, Mahowald N, Chadwick OA, Artaxo P (2004) Impact of desert dust on the biogeochemistry of phosphorus in terrestrial ecosystems. Global Biogeochem Cycles 18:GB2005. doi:10.1029/2003GB002145

Paasonen P et al (2013) Warming-induced increase in aerosol number concentration likely to moderate climate change. Nat Geosci 6:438–443

Palmer PI, Abbot DS, Fu TM, Jacob DJ (2006) Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column. J Geophys Res 111:D12315. doi10.1029/2005JD006689

Penner JE, Andreae M, Annegarn H, Barrie L, Feichter J, Hegg D, Jayaraman A, Leaitch R, Murphy D, Nganga J, Pitari G (2001) Aerosols, their direct and indirect effects, in: climate change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, edited by: J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden and D. Xiaosu, Cambridge University Press, UK, Chap 5, pp 289–348

Peñuelas J, Staudt M (2010) BVOCs and global change. Trends Plant Sci 15:133–144

Peñuelas J, Poulter B, Sardans J, Ciais P, van der Velde M, Bopp L, Boucher O, Godderis Y, Llusia L, Nardin E, Vicca S, Obersteiner M, Janssens IA (2013) Human-induced nitrogen-phosphorus imbalances alter ecosystems across the globe. Nat Commun 4:2934. doi:10.1038/ncomms3934

Price C, Rind D (1994) The impact of a 2 $\times$ CO_2 climate on lighting-caused fires. J Clim 7:1484–1494

Quinn PK, Bates TS (2011) The case against climate regulation via oceanic phytoplankton sulphur emissions. Nature 480:51–56

Randerson JT, Liu H, Flanner MG, Chambers SD, Jin Y, Hess PG, Pfister G, Mack MC, Treseder KK, Welp LR, Chapin FS, Harden JW, Goulden ML, Lyons E, Neff JC, Schuur EAG, Zender CS (2006) The impact of boreal forest fire on climate warming. Science 314:1130–1132

Reichholf JH (1986) Is Saharan dust a major source of nutrients for the Amazonian rain forest? Stud Neotrop Fauna Environ 21:251–255

Rocha AV, Su H-B, Vogel CS, Schmid HP, Curtis PS (2004) Photosynthetic and water use efficiency responses to diffuse radiation by an aspen-dominated northern hardwood forest. For Sci 50:793–801

Rodhe H, Dentener F, Schulz M (2002) The global distribution of acidifying wet deposition. Environ Sci Technol 36:4382–4388

Solomon S (1999) Stratospheric ozone depletion: a review of concepts and history. Rev Geophys 37:275–316

Spracklen DV, Bonn B, Carslaw KS (2008a) Boreal forests, aerosols and the impacts on clouds and climate. Philos Trans R Soc A 366:4613–4626 doi:10.1098/rsta.2008.0201

Spracklen DV, Arnold SR, Sciare J, Carslaw KS, Pio C (2008b) Globally significant oceanic source of organic carbon aerosol. Geophys Res Lett 35:L12811. doi:10.1029/2008GL033359

Stocks B, Fosberg M, Lynham T, Mearns L, Wotton B, Yang Q, Jin J, Lawrence K, Hartley G, Mason J, McKenney D (1998) Climate change and forest fire potential in Russian and Canadian boreal forests. Clim Change 38:1–13

Thornton P, Lamarque J-F, Rosenbloom NA, Mahowald NM (2007) Influence of carbon-nitrogen cycle coupling on land model response to CO_2 fertilization and climate variability. Global Biogeochem Cycles 21: GB4028. doi:10.1029/2006GB002868

Unger N, Shindell DT, Koch DM, Streets DG (2006) Cross influences of ozone and sulfate precursor emissions changes on air quality and climate. Proc Natl Acad Sci U S A 103:4377–4380

Vallina SM, Simó R, Gassó S, de Boyer-Montégut C, del Rio E, Jurado E, Dachs J (2007) Analysis of a potential solar radiation dose-dimethylsulfide-cloud condensation nuclei link from globally mapped seasonal correlations. Global Biogeochem Cycles 21:GB2004. doi:10.1029/2006GB002787

van der Werf GR, Randerson JT, Collatz GJ, Giglio L, Kasibhatla PS, Arellano AF, Olsen SC, Kasischke ES (2004) Continental-scale partitioning of fire emissions during the 1997 to 2001 El Niño/La Niña period. Science 303:73–76

Vautard R, Cattiaux J, Yiou P, Thépaud J-N, Ciais P (2010) Northern hemisphere atmospheric stilling partly attributed to an increase in surface roughness. Nat Geosci 3:756–761

Watson AJ, Bakker DCE, Ridgwell AJ, Boyd PW, Law CS (2000) Effect of iron supply on Southern Ocean CO_2 uptake and implications for glacial atmospheric CO_2. Nature 407:730–733

Westerling AL, Hidalgo HG, Cayan DR, Swetnam TW (2006) Warming and earlier spring increase Western U.S. forest wildfire activity. Science 313:940–943

Woodhouse MT, Mann GW, Carslaw KS, Boucher O (2008) The impact of oceanic iron fertilization on cloud condensation nuclei. Atmos Environ 42:5728–5730

Woodhouse MT, Carslaw KS, Mann GW, Vallina SM, Vogt M, Halloran PR, Boucher O (2010) Low sensitivity of cloud condensation nuclei to changes in the sea-air flux of dimethyl-sulphide. Atmos Chem Phys 10:7545–7559

Woodward S, Roberts DL, Betts RA (2005) A simulation of the effect of climate change-induced desertification on mineral dust aerosol. Geophys Res Lett 32:L18810. doi:10.1029/2005GL023482

Young IR, Zieger S, Babanin AV (2011) Global trends in wind speed and wave height. Science 333:451–455

Further Reading (Textbooks and Articles)

Fischer H, Siggaard-Andersen M-L, Ruth U, Röthlisberger R, Wolff E (2007) Glacial/interglacial changes in mineral dust and sea-salt records in polar ice cores: sources, transport, and deposition. Rev Geophys 45:RG1002. doi:10.1029/2005RG000192

Mahowald NM, Lamarque JF, Tie XX, Wolff E (2006) Sea-salt aerosol response to climate change: last glacial maximum, preindustrial, and doubled carbon dioxide climates. J Geophys Res 111:D05303. doi:10.1029/2005JD006459

Masson-Delmotte V, Schulz M, Abe-Ouchi A, Beer J, Ganopolski A, González Rouco JF, Jansen E, Lambeck K, Luterbacher J, Naish T, Osborn T, Otto-Bliesner B, Quinn T, Ramesh R, Rojas M, Shao X, Timmermann A (2013) Information from paleoclimate archives. In: Climate Change 2013: the Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge

Chapter 12
Stratospheric Aerosols

Abstract This chapter discusses the specificities of aerosol processes and aerosol properties in the stratosphere. This includes background material on stratospheric circulation, a discussion of the sulphur budget in the stratosphere, and a description of aerosol size distribution in volcanic and nonvolcanic (background) conditions. The chapter further provides a description of past major volcanic eruptions and their impact on the climate system.

Keywords Stratosphere · Junge aerosol layer · Ozone · Volcano · Eruption · Brewer–Dobson · Pinatubo

12.1 Introduction

The processes that govern stratospheric aerosols and their interactions with the climate system are essentially the same as for tropospheric aerosols. However, the relative importance of the various processes can be quite different from those in the troposphere because of the particularities of the stratosphere and its meteorological conditions. First of all, the absence of wet deposition in the stratosphere and the small rate of mixing between the troposphere and the stratosphere are such that the lifetime for stratospheric aerosols is typically 6 months to 2 years instead of a few days to 2 weeks in the troposphere. This much longer lifetime justifies a distinct chapter for stratospheric aerosols. We will describe here the physical and chemical properties of stratospheric aerosols, the equilibrium between sources and sinks, what is known about the past evolution of stratospheric aerosols, and their impact on the climate system.

12.2 Atmospheric Cycle of Stratospheric Aerosols

12.2.1 Sources of Stratospheric Aerosols

The transport of air from the troposphere to the stratosphere takes place almost exclusively in the Tropics. Such air is generally very cold and dry because it has gone through the tropical tropopause region where it is dehydrated. As a result, it

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_12

also generally contains relatively few aerosols of tropospheric origin, because such aerosols will have been scavenged in the process. However, the quantity of tropospheric aerosols entering the stratosphere is not known with precision. There are observations of convection overshooting into the stratosphere, which also has the potential to inject aerosols coming from the troposphere (e.g. Nielsen et al. (2007)). Overall the contribution of tropospheric aerosols to the stratospheric aerosol budget is nevertheless thought to be small. Some aerosol precursor gases are not scavenged effectively in the troposphere because they are not sufficiently soluble. As a result, they can reach the upper troposphere and from there into the stratosphere. Such gases include dimethylsulphide (DMS) emitted by the oceans, and to a small extent by the vegetation, hydrogen sulphide (H_2S) and sulphur dioxide (SO_2). However, these three species have a relative short lifetime in the troposphere of about 1 day because they are oxidized by the OH radical. Moreover, these gases are not completely insoluble. Their concentrations in the upper troposphere and their flux into the stratosphere are therefore relatively small.

The main gaseous precursor for stratospheric aerosols during nonvolcanic periods is thought to be the carbonyl sulphide, whose chemical formula is COS (Chin and Davis 1995). COS is relatively stable in the troposphere. It is therefore relatively well-mixed in the troposphere and can be transported into the stratosphere, where its mixing ratio decreases with altitude (Barkley et al. 2008). Sources of COS in the atmosphere are mostly natural; there is 1.3 Tg S per year coming from oceans, soils, wetlands and from the oxidation of DMS and CS_2, but there is also a small anthropogenic source of the order of 0.1 Tg S per year (Watts 2000). Carbonyl sulphide is photodissociated into SO_2 in the stratosphere, where it represents the dominant nonvolcanic source of sulphur gases in this region of the atmosphere. The end product of the oxidation of sulphur species is sulphuric acid (H_2SO_4), which has a relatively low saturation vapour pressure at low temperatures that are typical of the lower stratosphere. For this reason, gaseous H_2SO_4 condenses with water vapour to form sulphuric acid aerosols in the lower stratosphere. Figure 12.1 summarizes our understanding of how the sulphur budget in the stratosphere maintains a background sulphate aerosol layer in the absence of major volcanic eruption.

There are other sources of nonvolcanic aerosols to the stratosphere but these are poorly known because of the lack of appropriate measurements. Deep convection and pyroconvection have been shown to occasionally inject biomass burning aerosols into the stratosphere (Fromm et al. 2000). Carbonaceous aerosols have been detected in the lower stratosphere (Murphy et al. 2007). There is also some evidence for the presence of some aerosol material of meteoritic origin in the stratosphere (Cziczo et al. 2001; Renard et al. 2010), as suggested since the 1970s (e.g. Turco et al. 1981).

Volcanic eruptions are a sporadic but important source of gases and aerosols in the atmosphere. Small eruptions emit in the troposphere and have a limited impact on the climate in light of the large anthropogenic sulphur sources at the Earth's surface (Andres and Kasgnoc 1998). However, sulphur emitted during the largest eruptions can reach the stratosphere when the volcanic plume has enough buoyancy to go through the tropopause. The amount of material emitted by volcanoes, as well as the ratio between emitted volcanic ash and sulphur species (H_2S and SO_2

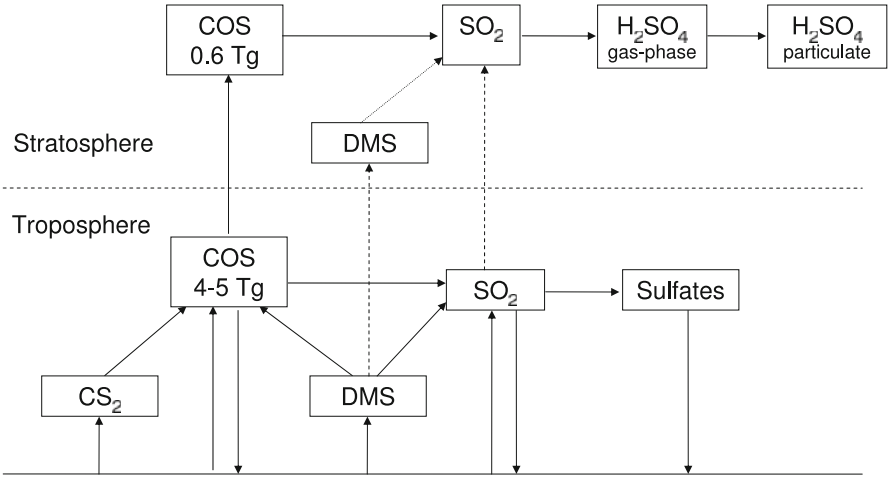


Fig. 12.1 Budget of sulphur chemical species that play a role in maintaining a background layer of stratospheric sulphate aerosols

taken together), vary enormously from one eruption to another. The residence time of volcanic ash aerosols in the atmosphere (whether they are injected in the troposphere or the stratosphere) is relatively short (i.e. typically hours to days for the larger particles and weeks for smaller particles), so that it is the emissions of sulphur gases that control the climate response to a volcanic eruption. The Volcanic Explosivity Index (VEI), ranging from 0 to 8 and used by geologists to provide a relative measure of the explosive energy of volcanic eruptions (Newhall and Self 1982), is generally not indicative of a volcano's impact on climate. A "small" but sulphur-rich eruption in the stratosphere can therefore have more impact on the stratospheric aerosol layer and the climate system than a "large" but sulphur-poor eruption. For instance the 1980 Mount Saint Helen's eruption had a large VEI of 5, but was sulphur-poor and thus had no impact on climate (Robock 1981). When they inject material in the stratosphere, tropical eruptions are more efficient than extra-tropical ones at perturbing the climate because gases and aerosols have longer residence times in the stratosphere when they enter it through the tropics. We have a particularly interesting and spectacular case study with the eruption of Mount Pinatubo on June 15, 1991 (Fig. 12.2), but other, more modest eruptions also contribute to influence stratospheric aerosols (Vernier et al. 2011).

12.2.2 Transport in the Stratosphere

Unlike the troposphere where vertical motions are important, the stratosphere is very stratified because of the inversion in the vertical temperature gradient. The stratosphere is separated from the troposphere by the tropopause that acts as a barrier

Fig. 12.2 Photography of the Pinatubo eruption on June 12, 1991, preceding the major eruption of June 15, 1991. (© U.S. Geological Survey)

between these two parts of the atmosphere. The circulation in the stratosphere tends to be zonal (i.e. it exhibits little change with longitude) and winds can be very strong. The meridional (i.e. latitudinal) component of the stratospheric circulation is called the Brewer–Dobson circulation. Schematically it originates in the tropical tropopause where air masses associated with deep convection penetrate into the stratosphere. The Brewer–Dobson circulation consists in a slow ascending motion of air in the Tropics. At the same time, it is advected towards mid-latitudes and polar regions of both hemispheres where there is a slow descending motion. Air returns to the Tropics in the lower stratosphere and also reenters the troposphere at high latitudes. The transit time of an air mass in the stratosphere is typically of the order of several years. One way to track the stratospheric circulation is through the age of air, which represents the time elapsed since an air mass has entered the stratosphere (Fig. 12.3; Diallo et al. 2012). The driver of the Brewer–Dobson circulation is not deep convection for the ascending branch or intrusions of stratospheric air in the troposphere for the descending branch, but is rather driven by the breakdown and dissipation of gravity and planetary waves in the stratosphere and mesosphere.

The mixing in the stratosphere is much faster in the longitudinal than in the latitudinal direction. Volcanic material injected at a given latitude tends to be rapidly transported in the zonal direction resulting in fairly homogeneous concentrations along a latitudinal band. A volcanic eruption at mid- or high latitudes increases aerosol concentrations in these regions and is followed by poleward transport but little transport towards the Equator (Kravitz and Robock 2011; Jégou et al. 2013). In contrast, a tropical eruption tends to increase concentrations at all latitudes (Trepte et al. 1993).

Another aspect of the stratospheric circulation that needs to be considered is the *quasi-biennal oscillation* (QBO) by which the zonal wind in the equatorial stratosphere switches from easterlies to westerlies with a period of about two years (Fig. 12.4; Baldwin et al. 2001). This oscillation in the wind regime originates in the upper stratosphere and propagates downwards at an approximative speed of 1 km per month until it reaches the tropopause. Wind speeds are stronger in the easterly phase than in the westerly phase. The phase of the quasi-biennal oscillation has a

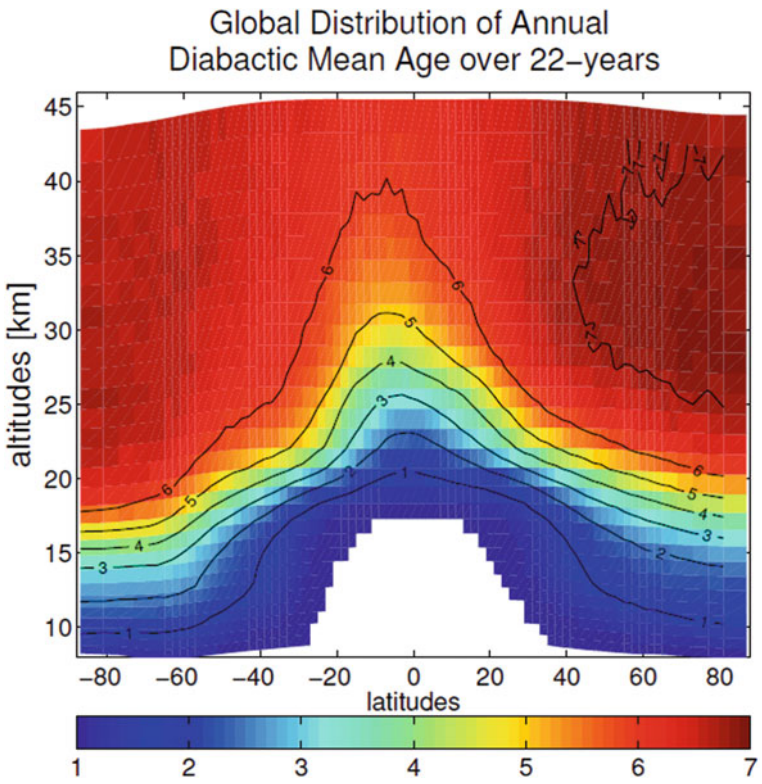


Fig. 12.3 Latitude-altitude distribution of the age of air (years) in the stratosphere averaged over the period 1989–2010 according to the ERA-interim meteorological reanalysis. (Reproduced from Diallo et al. (2012). Creative Commons License)

direct impact on the way and the speed with which chemical species and aerosols injected in the tropical lower stratosphere are transported vertically and dispersed towards mid-latitudes. Generally speaking poleward longitudinal transport is more efficient during the westerly shear phase of the QBO.

12.3 Physics and Chemistry of Stratospheric Aerosols

There exist some particularities regarding processes that govern the physical and chemical properties of aerosols in the stratosphere. Wet scavenging does not play a role as it does not rain in the stratosphere. Dry deposition is also absent as a sink because there is no surface upon which the aerosols can stick. Therefore, the only way to remove aerosols from the stratosphere is through transport through its lower boundary. In this context, coagulation and sedimentation are both important processes, even for accumulation-mode aerosols because of their longer residence time. Several processes contribute to aerosols sinks in the stratosphere: (i) the slow but

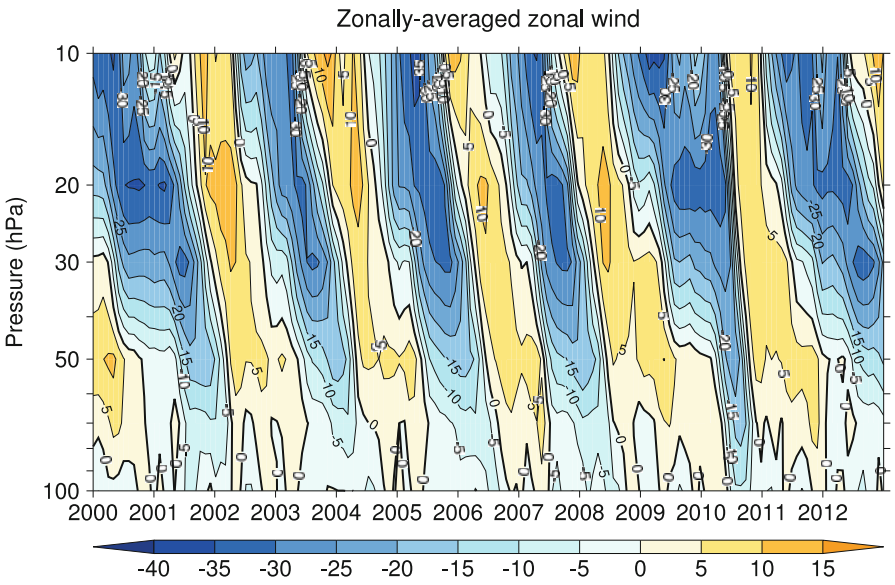


Fig. 12.4 Zonally-averaged zonal wind (ms^{-1}) in the equatorial stratosphere showing the downward propagation of the quasi-biennal oscillation. Redrawn from ERA-interim meteorological data

permanent sedimentation of aerosols in the stratosphere down to the troposphere, (ii) downward advective transport through the tropopause (for instance during tropopause folding episodes), and marginally (iii) turbulent vertical diffusion in the lower stratosphere. The characteristic timescale associated with vertical advective transport is of the order of several years, and is significantly longer than that associated with sedimentation which is about one year.

As the main precursors to stratospheric aerosols are sulphur gases, stratospheric aerosols are mostly composed of sulphuric acid and water. Aerosols are essentially located in the lower stratosphere, between the tropopause and 30–35 km. Although the lower stratosphere is very cold, sulphuric acid aerosols are liquid. Above 30–35 km, where the temperature is higher, aerosols evaporate, which defines the upper limit of the stratospheric aerosol layer. The stratospheric aerosol layer was first observed by C. Junge in the early 1960s (Junge et al. 1961) and is also called the *Junge layer*.

One can note that stratospheric aerosols are efficient condensation nuclei for *polar stratospheric clouds* (PSC), which as indicated by their name can form in the polar lower stratosphere, usually in wintertime. PSCs form at very low temperatures (below 195 K). Some of them are large liquid ternary solution aerosols composed of sulphuric acid (H_2SO_4), water vapour (H_2O) and nitric acid (HNO_3). Thermodynamic equilibrium between the gas and liquid phase occurs at more extreme temperature than in the troposphere, which necessitates an improved database to determine the chemical composition of the ternary solution aerosol as a function of the gas-phase concentrations of H_2SO_4, H_2O and HNO_3.

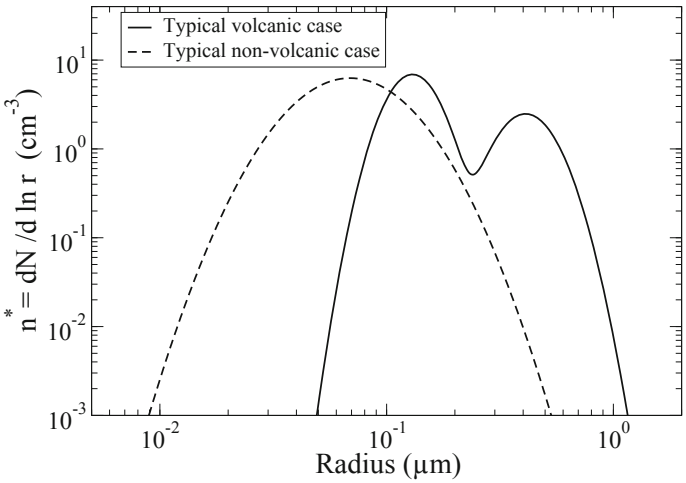


Fig. 12.5 Typical size distributions of stratospheric aerosols during nonvolcanic (*dashed line*) and volcanic (*solid line*) periods. (Drawn from data by Deshler et al. (2003))

The size distribution of stratospheric aerosols varies spatially and temporally according to the seasonal variations in the stratospheric circulation and the relative importance of the microphysical processes discussed above. The distribution of aerosols also fluctuates on an interannual timescale because of variations in sources, in particular those related to major volcanic eruptions. Figure 12.5 shows typical size distributions for stratospheric aerosols during a nonvolcanic (i.e. background) and a volcanic period. Schematically, a major volcanic eruption injects large quantities of sulphur gases (mainly SO_2) into the stratosphere which are then oxidized progressively to sulphuric acid on a timescale of a couple of months. The sulphuric acid in the gas phase can nucleate and form new particles or instead condense on pre-existing aerosols. Stratospheric aerosols also grow progressively in size under the effect of coagulation. As aerosols grow and the size distribution shifts to larger sizes, the importance of sedimentation increases. As a result of these effects, the size distribution which is monomodal in nonvolcanic periods, becomes bimodal in the months following a volcanic period (Fig. 12.5). While the bimodal distribution has been clearly observed after large eruptions, it may not necessarily appear after smaller eruptions.

12.4 Volcanic Aerosol Records

12.4.1 Volcanic Aerosols During the 1750–2010 Period

On the monthly timescale, stratospheric aerosols are well-mixed in longitude (i.e. at a given latitude). It is therefore enough to have measurements which are well

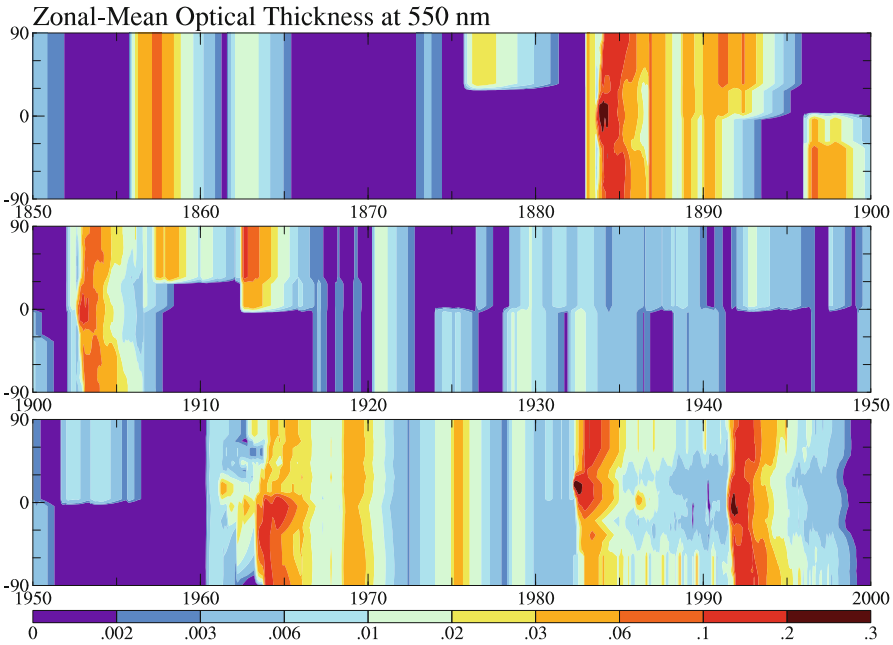


Fig. 12.6 Temporal evolution of the stratospheric aerosol optical depth at 550 nm by latitudinal bands for the period 1850–2000. One can notice the signature of large volcanic eruptions. (Updated from Sato et al. (1993). © NASA GISS)

distributed in latitude and altitude to construct a climatology. While the recent period is well-covered by satellite measurements and ground-based lidar measurements, aerosol concentrations from the more distant past have to be inferred from indirect measurements of the volcanic activity. Figure 12.6 shows how the stratospheric aerosol optical depth has varied over the period 1750–2000. These variations are estimated from temporal time series of proxies of volcanic eruption such as the concentration of sulphate in ice cores. The stratosphere tends to relax towards a fairly thin aerosol layer which thickens every time that a major volcanic eruption fills up the stratospheric reservoir. The relaxation towards a background aerosol layer typically follows an exponential decrease. The e-folding time associated with the relaxation from a volcanic state to the nonvolcanic (i.e. background) state is about one year for a tropical injection such as that of Mount Pinatubo. This means that it takes at least several years to recover to a background stratosphere after a major volcanic eruption. In the absence of any significant volcanic eruption for more than a decade after Mount Pinatubo eruption in June 1991, the stratospheric aerosol layer probably reached a minimum in the early 2000s which had not been experienced for a long time. For smaller eruptions, and for eruptions that occur outside the tropical region, the e-folding time can be much shorter (e.g. Kravitz and Robock 2011; Jégou et al. 2013).

Figure 12.7 extends the evolution of the average stratospheric aerosol optical depth over the recent period. A few volcanic eruptions in 2009 and subsequent years have terminated a long nonvolcanic period. However these eruptions were either located in the mid-latitudes, such as that of the Sarychev in 2009, where they are less effective than tropical eruptions at perturbing durably the stratospheric aerosol layer, and/or had a limited magnitude, such as that of the Nabro in 2011.

12.4.2 A Few Recent Climate-Relevant Eruptions

A few eruptions have had a marked impact on the recent climate of the Earth. It is useful to mention them here, not only to understand the past climate of the Earth, but also possible risks to the future climate. Each eruption is different in terms of the quantity of emitted aerosol precursors, the latitude of the volcano, and the altitude range of the emissions.

Ice cores show the signature of a very large eruption that took place in year 1257 or 1258 (Oppenheimer 2003). The location of the volcano is still being debated but the Salamas volcano of the Rinjani volcanic complex in Indonesia is thought to be the culprit (Lavigne et al. 2013). The stratospheric load of sulphate aerosols could have been 10 times larger than that of the 1991 Mount Pinatubo eruption. Yet proxy data show that the surface cooling was rather limited, which suggests that the radiative forcing associated to volcanic SO_2 emissions does not scale linearly with emissions and sulphate loading because of changes in the size distribution (Timmreck et al. 2009).

Laki is an Icelandic volcano at 64°N, which became active in June 1783 and experienced an eruption that lasted until February 1784. It has been estimated that the volcano injected about 120 Mt of SO_2 and 8 Mt of hydrogen fluoride (HF) in the atmosphere. This amount of SO_2 represents a little less than the present-day worldwide anthropogenic emissions in a year. Several explosive eruptions emitting 10 Mt of SO_2 each were superimposed on continuous emissions in the troposphere (Oman et al. 2006). Emissions of HF and SO_2 caused an acidic fog which has been advected over Europe. A significant increase in mortality and morbidity has followed, which can be partly attributable to the decrease in air quality. It has been estimated that Laki was also responsible for an increase in the concentrations of stratospheric aerosols in the northern hemisphere, but also for a substantial increase in cloud condensation nuclei which may not have been limited to the northern hemisphere (Schmidt et al. 2010). The aerosol perturbation is thought to have caused some circulation changes and monsoon failures in some regions, which triggered some famines. Another potential consequence is the very cold 1783–1784 winter over Europe and North America.

Indonesia is a very active volcanic region, with three major eruptions in the last 200 years. Mount Tambora is a stratovolcano located at the latitude of 8°S and experienced a major explosive eruption in April 1815. Large quantities of SO_2 reached the stratosphere, which caused a cooling of about 1°C in the northern hemisphere

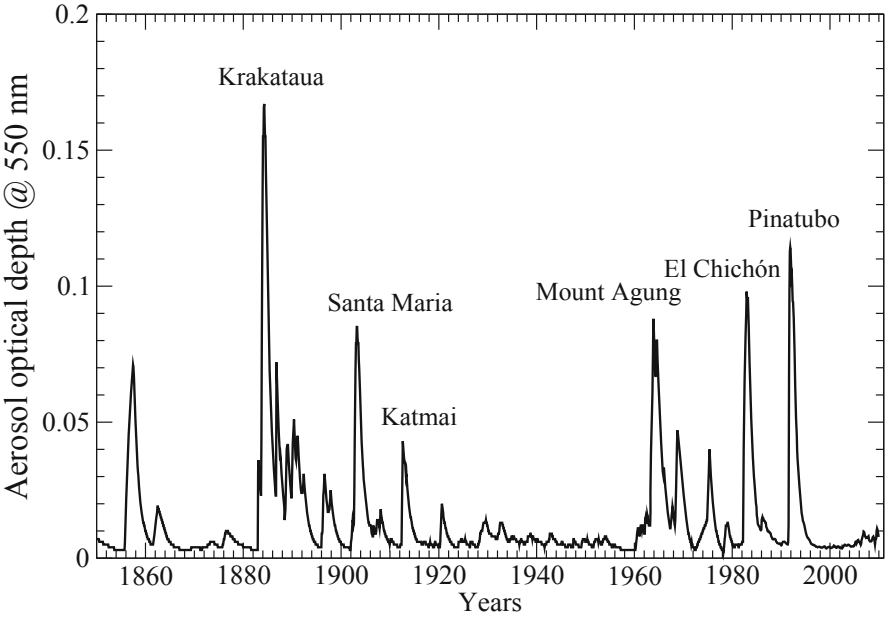


Fig. 12.7 Optical depth of stratospheric aerosols (at 550 nm) for the period 1850–2009. We have combined the climatology due to M. Sato from 1850 to 1985 with the climatology from Vernier et al. (2011) from 1985 to 2009. The Sato climatology was corrected before 1960 to include a small background stratospheric aerosol. (Updated from Sato et al. (1993))

in the year 1816, also known as the "year without a summer" in Europe (Cole-Dai et al. 2009). The Krakataua is another Indonesian volcano located at the latitude of 6°S whose last eruption was in August 1883. As for Mount Tambora, the eruption was explosive in nature and injected a major quantity of SO_2 in the stratosphere. Its importance in Fig. 12.7 may be exaggerated. Indonesia is definitely a very active volcanic region since Mount Agung also influenced the stratospheric aerosol layer after its eruption in 1963.

El Chichón is a Mexican volcano located at the latitude of 17°N that emitted an estimated 7 Mt of SO_2 during its eruption in April 1982. This eruption was not followed by a cooling of the climate system, because it coincided with an El Niño episode that itself tends to warm the climate. The winter 1982–1983 was warmer than usual over North America, Europe and Siberia.

Mount Pinatubo is a volcano in the Philippines (15°N) which erupted on June 15, 1991. It is not only the largest volcanic eruption of the twentieth century but also the eruption which has been the most observed and studied in terms of its impact on the atmosphere and the climate system. It has been estimated that Mount Pinatubo eruption released 20 Mt of SO_2 in the stratosphere. A decrease in incident solar radiation at the Earth's surface and a cooling of the order of 0.6°C in the northern hemisphere and 0.4°C in the global mean have been observed.

12.4.3 Mega-Eruptions

The Earth has certainly experienced volcanic eruptions that are much larger than those described in the previous section. It has been suggested that the Toba eruption in Indonesia, which dates back about 70,000 years, could be the largest explosive eruption of the last few millions years. The amount of sulphur species which has been emitted could be 100 times larger than for Mount Pinatubo eruption. The possibility that the Toba eruption has modified the climate for an extended period of time is still being debated (Rampino et al. 1992; Robock et al. 2009).

12.5 Role of Stratospheric Aerosols on the Climate System

Stratospheric aerosols have a climatic impact which occurs primarily through its aerosol–radiation interactions (ARI effect as defined in Chap. 2) as there are no water clouds in the stratosphere (except for PSCs). Aerosol–radiation interactions from stratospheric aerosols result in a negative shortwave forcing because the aerosols are primarily scattering. For Mount Pinatubo eruption an anomaly in the top-of-atmosphere shortwave radiation of the order of -4 to $-5\ \mathrm{Wm}^{-2}$ (Fig. 12.8) has been measured. Stratospheric aerosols also exert a greenhouse effect because they absorb and emit longwave radiation at a temperature that is generally colder than the emission temperature of longwave radiation they absorb. This additional greenhouse effect offsets radiatively about one third of the aerosol effect in the shortwave. However, the degree of compensation between the greenhouse and solar albedo effects depends on the aerosol size distribution and is therefore a function of time after a major volcanic eruption. Altogether the net radiative forcing per unit aerosol optical depth is of the order of $-20\ \mathrm{W\,m}^{-2}$. It has been suggested that stratospheric aerosols sedimenting in the upper troposphere can influence the formation of cirrus (Wang et al. 1995); however this remains largely an unproved hypothesis, which is also inconsistent with our understanding of what consists a good ice nucleus. However, changes in cirrus clouds may also be due to changes in the vertical profile of heating rates caused by stratospheric aerosols.

The transient radiative forcing caused by a volcanic eruption such as that of Mount Pinatubo is responsible for a temporary cooling and a decrease in the water vapour content of the troposphere, as illustrated on Fig. 12.8. Although the radiative forcing is relatively homogeneous spatially, the associated cooling is not, first because there is no reason for the cooling to be as homogeneous as the forcing itself, and second because it is difficult to disentangle the climate response to the volcanic eruption from natural variability. The Mount Pinatubo eruption was followed by a warmer than usual winter over Eurasia and North America. Because this has been observed for a number of past volcanic eruptions, and simulated in some climate models, winter warming after volcanic eruptions is now well-established. It is understood as a dynamical response to the stratospheric aerosol forcing (Robock and Mao 1992). The

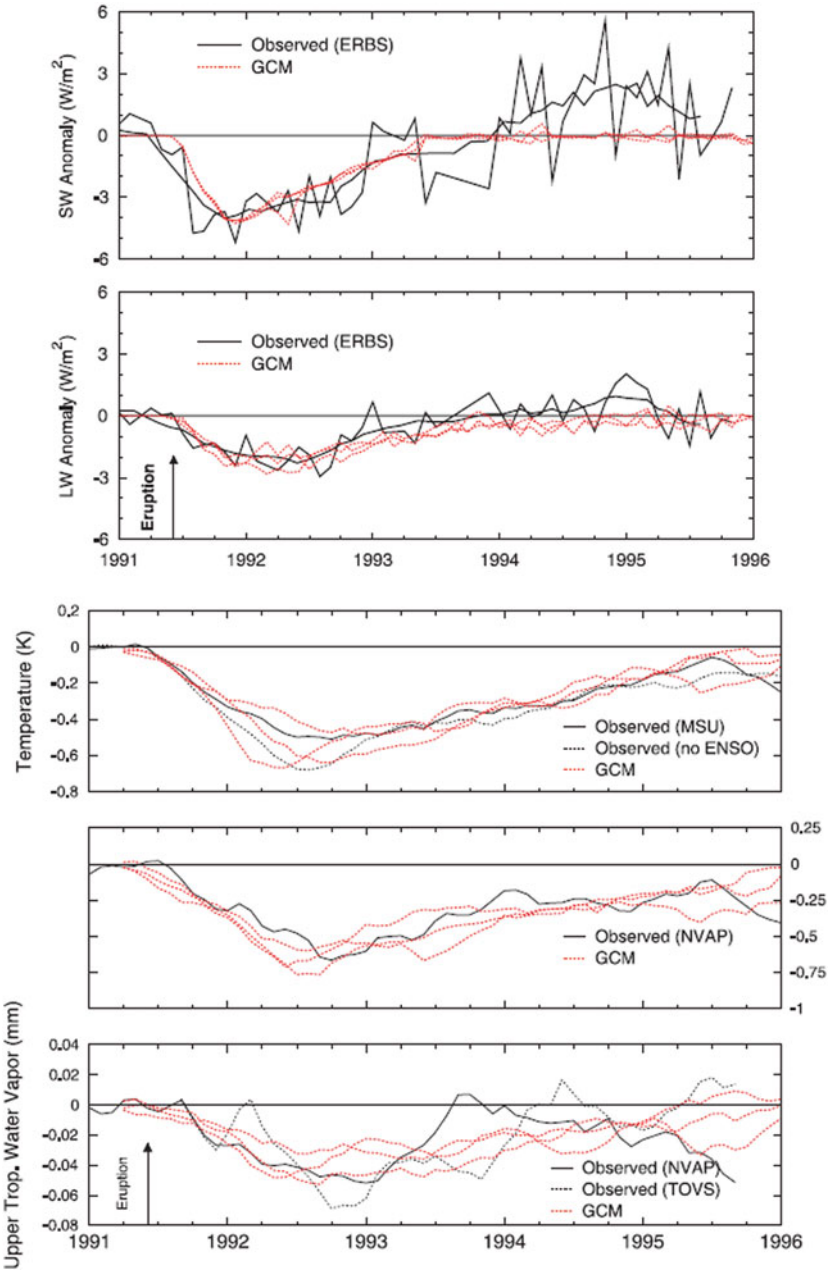


Fig. 12.8 Climate response to the radiative forcing due to volcanic aerosols from the 1991 Pinatubo eruption. From *top* to *bottom*: change in shortwave radiative flux at the *top* of the atmosphere (Wm^{-2}), change in longwave radiative flux at the *top* of the atmosphere (Wm^{-2}), change in surface air temperature (K), change in water vapour column (mm or kg water m^{-2}) in the upper troposphere. The result from a climate model (in *red*) is superposed to observations (in *black*). (From Soden et al. (2002). © American Geophysical Union)

aerosol layer initially heats the tropical stratosphere, before the temperature anomaly propagates to middle latitudes. This increases the temperature gradient between the midlatitudes and the North Pole, resulting in a strengthening of the zonal circulation in the stratosphere. This anomaly of westerlies then propagates in the troposphere and induces changes in the tropospheric circulation, leading to warmer winters in some regions. The Mount Pinatubo eruption was also followed by a decrease in the atmospheric water vapour content and a global decrease in precipitation. Moreover, enhanced stratospheric aerosols levels increased the amount of diffuse solar radiation at the Earth's surface, which is likely to have increased carbon sinks in terrestrial ecosystems (Gu et al. 2003).

Exercises

1a. The IASI instrument has shown that the explosion of the Sarychev volcano in June 2009 has been followed by an increase in the stratospheric SO_2 burden of $\Delta B_0 = 0.17$ Dobson unit (DU) on average in the Northern Hemisphere (Haywood et al. 2010). Compute the quantity of SO_2 that was emitted during the eruption. It is reminded that a Dobson unit correspond to a thickness of 10 μm of gas at standard conditions for temperature and pressure and that the ideal gas law is $P\,V = n\,R\,T$ in its common form.

1b. The SO_2 burden decreases to $\Delta B_1 = 0.05$ Dobson unit at a time $t_1 = 14$ days after the eruption. Estimate the e-folding time of the SO_2 in the stratosphere (i.e. the time when the SO_2 burden has decreased by a factor e).

1c. Estimate the maximum perturbation induced on the stratospheric aerosol optical depth at 670 nm by assuming that aerosols remain confined to the Northern Hemisphere, the conversion of SO_2 into sulphuric acid is complete, and the extinction efficiency of the hydrated aerosol per unit mass of sulphuric acid is $\alpha^{\mathrm{ext}}(670\ \mathrm{nm}) = 3\ \mathrm{m^2 g^{-1}}$.

Solutions

1a. The total volume of SO_2 emitted by the volcano at standard conditions for temperature and pressure is $V_{SO_2} = 2\,10^{-5}\,\pi\,R^2_{\mathrm{Earth}}\,\Delta B$ where ΔB is expressed in Dobson unit. The emitted quantity of SO_2 is

$$E = n_{SO_2}\,M_{SO_2} = \frac{2\,10^{-5}\,\pi\,R^2_{\mathrm{Earth}}\,\Delta B\,P_0\,M_{SO_2}}{R\,T_0} = 1.1\ \mathrm{Tg\ SO_2}$$

which underestimates somewhat the true value.

1b. The e-folding time of SO_2 in the stratosphere following the Sarychev

eruption is $\tau_{SO_2} = -t_1 / \ln(\Delta B_1 / \Delta B_0) = 11$ days.

1c. The maximum additional aerosol optical depth due to the eruption is:

$$\tau = \frac{P_0\, B_0\, M_{H_2SO_4} \alpha^{ext} \times 10^{-5}}{R\, T_0} = 0.02$$

which is about twice the value observed by satellite.

References

Andres RJ, Kasgnoc AD (1998) A time-averaged inventory of subaerial volcanic sulfur emissions. J Geophys Res 103(25):25251–25261. doi:10.1029/98JD02091

Baldwin MP, Gray LJ, Dunkerton TJ, Hamilton K, Haynes PH, Randel WJ, Holton JR, Alexander MJ, Hirota I, Horinouchi T, Jones DBA, Kinnersley JS, Marquardt C, Sato K, Takahashi M (2001) The quasi-biennial oscillation. Rev Geophys 39(2):179–229. doi:10.1029/1999RG000073

Barkley MP, Palmer PI, Boone CD, Bernath PF, Suntharalingam P (2008) Global distributions of carbonyl sulfide in the upper troposphere and stratosphere. Geophys Res Lett 35:L14810. doi:10.1029/2008GL034270

Chin M, Davis DD (1995) A reanalysis of carbonyl sulfide as a source of stratospheric background sulfur aerosol. J Geophys Res 100:8993–9005. doi:10.1029/95JD00275

Cole-Dai J, Ferris D, Lanciki A, Savarino J, Baroni M, Thiemens MH (2009) Cold decade (AD 1810–1819) caused by Tambora (1815) and another (1809) stratospheric volcanic eruption. Geophys Res Lett 36:L22703. doi:10.1029/2009GL040882

Cziczo DJ, Thomson DS, Murphy DM (2001) Ablation, flux, and atmospheric implications of meteors inferred from stratospheric aerosol. Science 291:1772–1775

Deshler T, Hervig ME, Hofmann DJ, Rosen JM, Liley JB (2003) Thirty years of in situ stratospheric aerosol size distribution measurements from Laramie, Wyoming (41°N), using balloon-borne instruments. J Geophys Res 108:4167. doi:10.1029/2002JD2514

Diallo M, Legras B, Chédin A (2012) Age of stratospheric air in the ERA-Interim. Atmos Chem Phys 12:12133–12154

Fromm M, Alfred J, Hoppel K, Hornstein J, Bevilacqua R, Shettle E, Servranckx R, Li Z, Stocks B (2000) Observations of boreal forest fire smoke in the stratosphere by POAM III, SAGE II, and lidar in 1998. Geophys Res Lett 27:1407–1410

Gu L, Baldocchi DD, Wofsy SC, Munger JW, Michalsky JJ, Urbanski SP, Boden TA (2003) Response of a deciduous forest to the Mount Pinatubo eruption: enhanced photosynthesis. Science 299:2035–2038

Haywood J, Jones A, Clarisse L, Bourassa A, Barnes J, Telford P, Bellouin N, Boucher O, Agnew P, Clerbaux C, Carslaw K (2010) Observations and simulations of the eruption of the Sarychev volcano using HadGEM2. J Geophys Res 115:D21212. doi:10.1029/2010JD014447

Jégou F, Berthet G, Brogniez C, Renard J-B, François P, Haywood JM, Jones A, Bourgeois Q, Lurton T, Auriol F, Godin-Beekmann S, Guimbaud C, Krysztofiak G, Gaubicher B, Chartier M, Clarisse L, Clerbaux C, Balois J-Y, Verwaerde C, Daugeron D (2013) Stratospheric aerosols from the Sarychev volcano eruption in the 2009 Arctic summer. Atmos Chem Phys 13:6533–6552

Junge CE, Chagnon CW, Manson JE (1961) Stratospheric aerosols. J Meteorol 18:81–108

Kravitz B, Robock A (2011) The climate effects of high latitude volcanic eruptions: the role of the time of year. J Geophys Res 116:D01105. doi:10.1029/2010JD014448

Lavigne F, Degeai J-P, Komorowski J-C, Guillet S, Robert V, Lahitte P, Oppenheimer C, Stoffeld M, Vidal CM, Surono, Pratomo I, Wassmer P, Hajdas I, Sri Hadmoko D, de Belizal E (2013) Source of the great A.D. 1257 mystery eruption unveiled, Samalas volcano, Rinjani Volcanic Complex, Indonesia. Proc Nat Acad Sci U S A 110:16742–16747

Murphy DM, Cziczo DJ, Hudson PK, Thomson DS (2007) Carbonaceous material in aerosol particles in the lower stratosphere and tropopause region. J Geophys Res 112:D04203. doi:10.1029/2006JD007297

Newhall CG, Self S (1982) The volcanic explosivity index (VEI) an estimate of explosive magnitude for historical volcanism. J Geophys Res 87(C2):1231–1238

Nielsen JK, Larsen N, Cairo F, Di Donfrancesco G, Rosen JM, Durry G, Held G, Pommereau JP (2007) Solid particles in the tropical lowest stratosphere. Atmos Chem Phys 7:685–695

Oman L, Robock A, Stenchikov GL, Thordarson T, Koch D, Shindell DT, Gao C (2006) Modeling the distribution of the volcanic aerosol cloud from the 1783–1784 Laki eruption. J Geophys Res 111:D12209. doi:10.1029/2005JD006899

Oppenheimer C (2003) Ice core and palaeoclimatic evidence for the timing and nature of the great mid-13th century volcanic eruption. Int J Climatol 23:417–426

Rampino MR, Self S (1992) Volcanic winter and accelerated glaciation following the Toba super-eruption. Nature 359:50–52

Renard J-B, Berthet G, Salazar V, Catoire V, Tagger M, Gaubicher B, Robert C (2010) In situ detection of aerosol layers in the middle stratosphere. Geophys Res Lett 37:L20803. doi:10.1029/2010GL044307

Robock A (1981) The Mount St. Helens volcanic eruption of 18 May 1980: minimal climatic effect. Science 212:1383–1384

Robock A, Mao J (1992) Winter warming from large volcanic eruptions. Geophys Res Lett 19:2405–2408. doi:10.1029/92GL02627

Robock A, Ammann CM, Oman L, Shindell D, Levis S, Stenchikov G (2009) Did the Toba volcanic eruption of ~74 ka B.P. produce widespread glaciation? J Geophys Res 114:D10107. doi:10.1029/2008JD011652

Sato M, Hansen JE, McCormick MP, Pollack JB (1993) Stratospheric aerosol optical depths, 1850–1990. J Geophys Res 98:22987–22994

Schmidt A, Carslaw KS, Mann GW, Wilson M, Breider TJ, Pickering SJ, Thordarson T (2010) The impact of the 1783–1784 AD Laki eruption on global aerosol formation processes and cloud condensation nuclei. Atmos Chem Phys 10:6025–6041

Soden BJ, Wetherald RT, Stenchikov GL, Robock A (2002) Global cooling after the eruption of Mount Pinatubo: a test of climate feedback by water vapor. Science 296:727–730

Timmreck C, Lorenz SJ, Crowley TJ, Kinne S, Raddatz TJ, Thomas MA, Jungclaus JH (2009) Limited temperature response to the very large AD 1258 volcanic eruption. Geophys Res Lett 36:L21708. doi:10.1029/2009GL040083

Trepte CR, Veiga RE, McCormick MP (1993) The poleward dispersal of Mount Pinatubo volcanic aerosol. J Geophys Res 98:18563–18573

Turco RP, Toon OB, Hamill P, Whitten RC (1981) Effects of meteoric debris on stratospheric aerosols and gases. J Geophys Res 86:1113–1128

Vernier J-P, Thomason L, Pommereau J-P, Bourassa A, Pelon J, Garnier A, Hauchecorne A, Blanot L, Trepte C, Degenstein D, Vargas F (2011) Major influence of tropical volcanic eruptions on the stratospheric aerosol layer during the last decade. Geophys Res Lett 38:L12807. doi:10.1029/2011GL047563

Wang P-H, Minnis P, Yue GK (1995) Extinction coefficient (1 μm) properties of high-altitude clouds from solar occultation measurements (1985–1990): evidence of volcanic aerosol effect J Geophys Res 100:3181–3199

Watts SF (2000) The mass budgets of carbonyl sulfide, dimethyl sulfide, carbon disulfide and hydrogen sulfide. Atmos Environ 34:761–779

Further Reading (Textbooks and Articles)

Myhre, G., Shindell D, Bréon F-M, Collins W, Fuglestvedt, J, Huang J, Koch D, Lamarque J-F, Lee D, Mendoza B, Nakajima T, Robock A, Stephens G, Takemura T, Zhang H (2013) Anthropogenic and natural radiative forcing. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge

Robock A (2000) Volcanic eruptions and climate. Rev Geophys 38:191–219. doi:10.1029/1998RG000054

Solomon S (1999) Stratospheric ozone depletion: a review of concepts and history. Rev Geophys 37:275–316

Chapter 13
Aerosol-Based Climate Engineering

Abstract This chapter introduces climate engineering as a range of techniques aiming to alleviate the impacts of climate change. Two aerosol-based techniques, known as stratospheric aerosol injection and marine cloud brightening, are discussed in terms of their potential, side effects, and shortcomings. The climate's response to aerosol-based climate engineering is assessed using the concepts of climate feedbacks and rapid adjustments. The chapter ends on the issue of the termination and other risks associated to such techniques.

Keywords Geo-engineering · Climate engineering · Marine cloud brightening · Stratospheric aerosol injection

13.1 Introduction

It is proving difficult to curb the world emissions of greenhouse gases and mitigate climate change. Some scientists have suggested to cool the Earth's climate artificially in order to compensate for anthropogenic climate change. Intentional modification of the climate system is often referred to as *geoengineering* or *climate engineering*. We prefer the second term as it is more specific (Boucher et al. 2014). In the absence of an agreed definition, we take here climate engineering to encompass all techniques aiming to modify the Earth's system on the large scale in order to counter anthropogenic climate change and its impacts. Such a definition excludes those techniques which are restricted to the small scale (such as weather modification), whose primary aim is not to cool the climate, or whose aim is to reduce greenhouse gas emissions at the source point (e.g. carbon capture and storage). As illustrated in Fig. 13.1, one can classify climate engineering techniques in two broad categories: techniques which aim to suppress the greenhouse effect by capturing or destroying greenhouse gases in the atmosphere (known as carbon dioxide removal or more generally greenhouse gas removal techniques) and techniques which aim to rebalance the radiative budget of the planet by reducing the amount of solar radiation that is absorbed by the atmosphere and the surface (known as solar radiation management techniques). More sophisticated categorizations have been suggested in the literature (Boucher et al. 2014). In the case of carbon dioxide, the capture and storage can happen through biological processes (using the terrestrial or the oceanic biosphere) or

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1_13

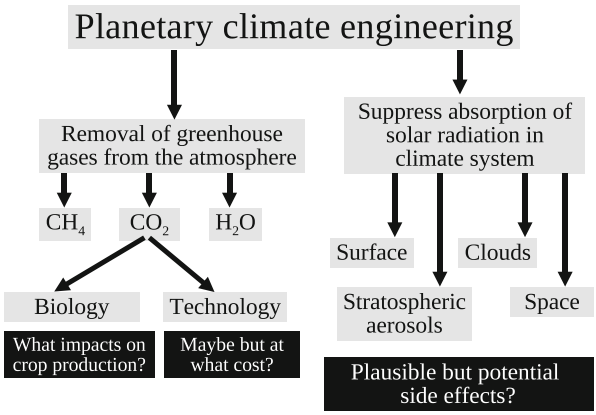


Fig. 13.1 A categorization of different types of climate engineering techniques

chemical processes (using the ocean, rocks or some industrial process). This chapter is only concerned with solar radiation management (SRM) techniques because the most realistic of these techniques are those that rely on the large leverage effect of aerosol–radiation and aerosol–cloud interactions.

Several techniques have been proposed to decrease the amount of solar radiation absorbed by the Earth. The first technique would consist in placing a large number of small mirrors or other reflecting material between the Earth and the Sun so as to decrease the amount of solar radiation reaching the top of the atmosphere (Angel 2006). Such a technique is hardly technologically feasible, would probably be very expensive and for these reasons is not discussed further here. Other techniques aim to increase the albedo (i.e. the fraction of solar radiation that is reflected back to space) of the surface–atmosphere system. This could be achieved by injecting large quantities of sea spray aerosols into the marine boundary layer in order to reflect more solar radiation both in clear sky (through aerosol scattering) and in cloudy sky (through an increase in cloud condensation nuclei). It has also been suggested that aerosols or aerosol precursors could be injected into the stratosphere to reflect more solar radiation back to space, thus mimicking the effects of a natural volcanic eruption. Other techniques would consist in increasing the reflectivity of the Earth's surface, either above the oceans (using bubbles, foam or some other floating material) or the continents (though modifying ecosystems, crops or the built environment). The potential of these techniques remains uncertain, which is why we only discuss here the feasibility, potential, benefits, and disbenefits of the two techniques that rely on the injection of atmospheric aerosols.

13.2 Stratospheric Aerosol Injection

Crutzen (2006) has proposed to inject intentionally gaseous aerosol precursors or aerosols into the stratosphere in order to offset some of the anthropogenic greenhouse effect and mask global warming. The idea was not new, and has already been

proposed in the 1970s by a Russian scientist named Budyko. There exists a natural analogue to this idea since explosive volcanic eruptions can generate a transient global cooling effect when large quantities of sulphur are injected into the stratosphere (see Chap. 12). However, in the case of climate engineering injection into the stratosphere would have to take place continuously because of the limited residence time of aerosols in the stratosphere. Several technological solutions have been envisaged to inject the aerosol precursors into the stratosphere: a fleet of aircraft carriers, a pipe lifted by a giant stratospheric balloon tethered to the ground or the massive use of artillery. The technological feasibility of some of these options is probably debatable, but it is reasonable to think that using a dedicated fleet of aircraft is a rather simple, robust, and scalable possibility. Another option would be to emit at the surface a sulphur species that is stable in the troposphere and can diffuse to the stratosphere where it is photodissociated by ultraviolet radiation. Moreover, large quantities of sulphur compounds are available, only because fossil fuels are massively desulphurized before or after being burned. The amount of sulphur species that would be required to cool the climate through stratospheric injection remains significantly smaller than that emitted at the surface through fossil fuel combustion. This can be explained by the longer residence time of the aerosols in the stratosphere as compared to that in the troposphere.

The optimal aerosol injection strategy is one that seeks to maximize the time-integral of the radiative impact of the aerosols for a given mass of aerosol or aerosol precursor, for a given energy input, or for a given cost. The residence time of stratospheric aerosols is much longer when aerosols are injected above the tropical tropopause, i.e. in the rising branch of the Brewer–Dobson circulation. The altitude of the tropopause decreases from about 20 km at the Equator down to about 6 km over the Poles. However an injection at mid- or high latitudes results in a much shorter aerosol residence time and has much less impact, especially if the altitude of injection is just above the tropopause. The injection system therefore requires altitudes of the order of 20 km to be optimal. It can been envisaged to emit aerosols at different latitudes in order to induce a radiative forcing that has a latitudinal distribution that resembles mostly due to anthropogenic greenhouse gases. However, it is not known to which extent the two radiative forcings can be compensated at the top of atmosphere given that the the two forcing mechanisms have different seasonal behaviours. The maximum radiative forcing that can be achieved through the method is also uncertain. As the rate of emission of aerosol precursors increases, coagulation between particles becomes more important, which increases the average aerosol size, may decrease the aerosol mass scattering efficiency and increase the sedimentation flux out of the stratosphere. Our current understanding suggests that the radiative forcing by stratospheric sulphur aerosols would saturate beyond a few (say 4) Wm^{-2}, which corresponds approximately to a doubling of the CO_2 atmospheric concentration.

The artificial injection of stratospheric aerosols is far from being an ideal solution in response to global warming. Injection of sulphate aerosols in the stratosphere would carry an impact on human health as the aerosols sediment to the surface where they can be breathed but the risk is thought to be small given the rather small emission

rates of aerosols as compared to current emission rates of the same species at the surface. Stratospheric aerosols are likely to reduce the thickness of the stratospheric ozone layer and increase the levels of ultraviolet radiation at the surface (Heckendorn et al. 2009). Moreover, the spectral and directional characteristics of solar radiation would be modified, with less direct solar radiation, more diffuse solar radiation and less total solar radiation reaching the Earth's surface. While an increase in the fraction of diffused solar radiation can be beneficial to some terrestrial ecosystems with an increase in net primary productivity, everything else being equal (Mercado et al. 2009), impacts on oceanic ecosystems and long-term impacts on terrestrial ecosystems are largely unknown. One can also wonder what the impacts of a major volcanic eruption would be if there is already a large artificial stratospheric aerosol layer.

13.3 Marine Cloud Brightening

Latham et al. (2008) suggested to inject sea spray aerosols in the marine boundary layer in order to increase the number concentration of cloud condensation nuclei and brighten low-level clouds. The idea is to use aerosol–cloud interactions as a leverage of the effect of aerosols on climate, although the aerosol effects in clear-sky may be significant as well (Korhonen et al. 2010; Jones and Haywood 2012). Latham et al. (2008) and Salter et al. (2008) have studied some aspects of the method, including the optimal size of the sea spray aerosols that need to be generated, and the technical feasibility of sea spray generating vessels. The manufacturing and dispersion of sea salt aerosols of submicronic size poses a significant technological challenge given the large amount of impurities in the sea water. Generating somewhat larger aerosols that evaporate and shrink in the boundary layer might be easier but would cool down the air mass that contains the aerosol plume and prevent its mixing within the marine boundary layer.

The method presents some advantages from an Earth system point of view. Sea spray aerosols are natural and their deposition over the ocean has no impact on marine ecosystems. The impact on terrestrial ecosystems and soils is unknown but could be limited by not spraying aerosols in the coastal regions. This would also limit but not remove the impact on human health.

As it has been discussed in Chap. 9, processes governing aerosol–cloud interactions are uncertain and as a consequence processes governing the climate engineering method proposed by Latham et al. are comparably uncertain. Trails left by ships in the marine cloudy boundary layer (known as ship tracks) provide a basis for this cloud seeding technique. However, there is no guarantee that cloud seeding would work in all oceanic regions and meteorological conditions. Some observations and modelling studies support the idea that cloud seeding may have an effect that is opposite to the desired one for some cloud types or under some conditions (Alterskjær et al. 2012). It is vain to believe that the cloud seeding technique could result in a forcing that is uniform in space and time. The forcing that is induced by cloud

seeding would depend on the presence or absence of clouds, the cloud susceptibility to additional cloud condensation nuclei (CCN) concentration, and the additional CCN concentration which itself varies with the aerosol lifetime. It is likely that the forcing induced by a fleet of cloud seeding vessels would exhibit some maxima in cloudy regions, especially those with stratocumulus clouds, where the cloud cover is large and clouds can be non-precipitating. If aerosol–cloud interactions produce a forcing on the upper side of the uncertainty range, then it is likely that seeding of the marine boundary layer could produce a negative radiative forcing of several Wm^{-2} and several °C of cooling in the global mean (Jones et al. 2009). Finally it is likely that injection of sea spray aerosols can induce a significant radiative forcing in clear-sky regions through backscattering of solar radiation.

13.4 Role of Rapid Adjustments and Feedbacks

The compensation of the radiative forcing of greenhouse gases by an increase in the planetary albedo does not result in a cancellation of the climate response. Firstly it would be difficult to produce a radiative forcing that matches exactly the spatial and temporal distribution of the radiative forcing by greenhouse gases. At the very least the two forcings would differ in their diurnal cycle. Secondly the two mechanisms do not force the atmosphere in the same way. We have seen in Chap. 10 that the climate response to a radiative perturbation can be decomposed into rapid adjustments and surface temperature-dependent feedbacks. The rapid adjustments are largely independent of the change in surface temperature and represent a direct response of the atmosphere and surface to the atmospheric forcing. For instance, the radiative forcing of greenhouse gases is different at the surface than at the top of atmosphere: it implies a convergence of radiative energy in the atmosphere which is compensated by a reduction in the latent heat flux in the atmosphere, which has to be matched by a reduction in the latent heat flux at the surface and a decrease in the global precipitation rate. In contrast the radiative forcing due to scattering aerosols does not involve much of a change in the vertical heating of the atmosphere and is mostly exerted at the surface. It does not therefore contribute to significant rapid adjustments in terms of latent heat flux. The combination of the two forcing mechanisms cancels the feedback responses of the climate system (as there is no change in surface temperature) but leaves the rapid adjustments, which leads to a reduction in the global precipitation (Schmidt et al. 2012). It is therefore not possible to keep both the global-mean surface temperature and precipitation rate unchanged. Either the global-mean surface temperature is kept unchanged and there is a reduction in the global-mean precipitation, or the global-mean precipitation is kept unchanged and there is an increase in global-mean surface temperature.

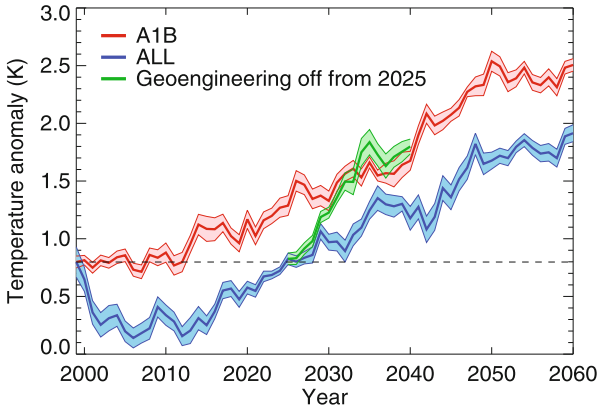


Fig. 13.2 Temporal evolution of the change in the global-mean surface air temperature for the A1B climate change scenario (in *red*) and for a scenario where the cloud reflectivity has been arbitrarily increased in three regions of stratocumulus clouds starting in year 2000 (in *blue*). The *green curve* shows the evolution of the global-mean surface air temperature if the climate engineering scheme is stopped in 2025. (From Jones et al. (2009). © British Crown Copyright 2009, Met Office)

13.5 A Thought on Timescales

While carbon dioxide removal has a long-term effect in terms of avoided radiative forcing, solar radiation management techniques such as those discussed above have an effect that only lasts for the time they are applied. Soon after the technique is stopped, the warming by greenhouse gases that was masked by the solar radiation management technique will materialize and the Earth will experience an accelerated warming for a couple of decades. This so-called 'termination' effect would cause a warming rate that would be much larger than without any climate engineering intervention, as illustrated on Fig. 13.2 which shows the evolution of the global-mean temperature for an A1B scenario with and without solar radiation management and with solar radiation management stopped after 30 years. A larger warming rate would make adaptation to climate change more of a problem (Boucher et al. 2009).

Moreover solar radiation management techniques would not solve the issue of ocean acidification which is induced by the increase in the atmospheric concentration of CO_2. Carbon dioxide is a soluble gas, and gets dissolved in the ocean, where it can react with a molecule of water to form a molecule of carbonic acid (H_2CO_3), which then dissociates in bicarbonate ion (HCO_3^-) and carbonate ion (CO_3^{2-}) with the release of H^+ cations. At current concentrations, an increase in atmospheric CO_2 results in more CO_2 being dissolved in the ocean, an increase in bicarbonate ion and a decrease in pH. It has been estimated that the ocean pH has decreased by about 0.1 unit since the beginning of the industrial period, and could decrease by another 0.3 to 0.5 unit until 2100. A more acidic ocean (or rather a less alkaline ocean) could be problematic for certain marine organisms that rely on calcium carbonate for their shell. The impacts of acidification on marine productivity and biodiversity remain however largely unknown.

In conclusion it can be said that solar radiation management techniques, even though they may effectively contribute to cooling the climate system, have three major drawbacks. First of all, they cannot offset the climate effects of greenhouse gases completely and there would necessarily be a residual climate change and side effects. Second, they imply a long-term commitment if they are used as a surrogate for emissions reductions because the atmospheric concentrations of long-lived greenhouse gases can only decrease fairly slowly once emissions are reduced because of the long timescales associated with the carbon cycle. Third, solar radiation management techniques do not address the problem of ocean acidification caused by increasing CO_2 atmospheric concentration. For all those reasons, solar radiation management techniques can only be seen as a last resort solution and in association with ambitious climate mitigation policies, if and when the impacts of climate change becomes untenable, for instance in the case of a climate sensitivity in the upper part of the uncertainty range. Only in such a scenario could solar radiation management be envisaged to maintain global surface temperatures below a desired target and for a limited time. In any case the social acceptability and governance of solar radiation management are likely to be problematic.

Exercise

1. Estimate the emission rate of SO_2 that it would be necessary to inject into the stratosphere continuously in a steady state regime to produce a radiative forcing at the top of atmosphere of –1 Wm^{-2}. We will assume that all of the SO_2 is converted in sulphate aerosols and that the aerosol has a residence time of 1 year in the stratosphere with a radiative efficiency of –100 W $(g\ SO_4^{2-})^{-1}$. Compare this emission rate to the quantity of SO_2 that was emitted by the Pinatubo eruption of June 1991.

Solution

1. $E = 2.5\ Tg\ SO_2\ yr^{-1}$.

References

Alterskjær K, Kristjánsson JE, Seland Ø (2012) sensitivity to deliberate sea salt seeding of marine clouds—observations and model simulations. Atmos Chem Phys 12:2795–2807

Angel R (2006) Feasibility of cooling the Earth with a cloud of small spacecraft near the inner Lagrange Point (L1). Proc Natl Acad Sci U S A 103:17184–17189

Boucher O, Lowe J Jones CD (2009) Constraints of the carbon cycle on timescales of climate-engineering options. Clim Change 92:261–273

Boucher O, Forster PM, Gruber N, Ha-Duong M, Lawrence M, Lenton TM, Maas A, Vaughan N (2014) Rethinking climate engineering categorization in the context of climate change mitigation and adaptation. WIREs Clim Change 5:23–35. doi:10.1002/wcc.261

Crutzen PJ (2006) Albedo enhancement by stratospheric sulfur injections: a contribution to resolve a policy dilemma? Clim Change 77:211–220. doi:10.1007/s10584-006-9101-y
Heckendorn P, Weisenstein D, Fueglistaler S, Luo BP, Rozanov E, Schraner M, Thomason LW, Peter T (2009) The impact of geoengineering aerosols on stratospheric temperature and ozone. Environ Res Lett 4:045108
Jones A, Haywood JM (2012) Sea-spray geoengineering in the HadGEM2-ES earth-system model: radiative impact and climate response. Atmos Chem Phys 12:10887–10898
Jones A, Haywood J, Boucher O (2009) Climate impacts of geoengineering marine stratocumulus clouds. J Geophys Res 114:D10106. doi:10.1029/2008JD011450
Korhonen H, Carslaw KS, Forster PM, Mikkonen S, Gordon ND, Kokkola H (2010) Aerosol climate feedback due to decadal increases in Southern Hemisphere wind speeds. Geophys Res Lett 37:L02805. doi:10.1029/2009GL041320
Latham, J, Rasch P, Chen C-C, Kettles L, Gadian A, Gettelman A, Morrison H, Bower K, Choularton T (2008) Global temperature stabilization via controlled albedo enhancement of low-level maritime clouds. Philos Trans R Soc A 366:3969–3987
Mercado LM, Bellouin N, Sitch S, Boucher O, Huntingford C, Cox P (2009) Impact of changes in diffuse radiation on the global land carbon sink. Nature 458:1014–1017
Salter S, Sortino G, Latham J (2008) Sea-going hardware for the cloud albedo method of reversing global warming. Philos Trans R Soc A 366:3989–4006
Schmidt H, Alterskjaer K, Bou Karam D, Boucher O, Jones A, Kristjansson JE, Niemeier U, Schulz M, Aaheim A, Benduhn F, Lawrence M Timmreck C (2012) Solar irradiance reduction to counteract radiative forcing from a quadrupling of CO_2: climate responses simulated by four earth system models. Earth Syst Dyn 3:63–78

Further Reading (Textbooks and Articles)

Boucher O, Randall D, Artaxo P, Bretherton C, Feingold G, Forster P, Kerminen V-M, Kondo Y, Liao H, Lohmann U, Rasch P, Satheesh SK, Sherwood S, Stevens B, Zhang XY (2013) Clouds and aerosols. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, pp 571–657
English, JT, Toon OB, Mills MJ (2012) Microphysical simulations of sulfur burdens from stratospheric sulfur geoengineering. Atmos Phys Chem 12:4775–4793

Appendix A
Units and Physical Constants

Table A.1 Physical constants and data

Symbol	Name	Value	Unit
A_v	Avogadro number	$6.022\ 10^{23}$	molecules mol^{-1}
c	Speed of light in vacuum	$2.99792\ 10^{8}$	ms^{-1}
g	Gravitational constant	9.81	ms^{-2}
h	Planck constant	$6.62607\ 10^{-34}$	Js
κ	von Kármán constant	0.35–0.40	
k_B	Boltzmann constant	$1.38065\ 10^{-23}$	JK^{-1}
r	Average Sun–Earth distance	$1.496\ 10^{11}$	m
R	Ideal gas constant ($= k_B\ A_v$)	8.314	$Jmol^{-1}K^{-1}$
R_e	Earth's radius	$6.370\ 10^{6}$	m
ϱ_w	Water density	1000	kgm^{-3}
S	Solar constant	1362	Wm^{-2}
σ	Stefan–Boltzmann constant	$5.67\ 10^{-8}$	$Wm^{-2}K^{-4}$

Table A.2 Prefixes for units

Prefix	Symbol	Factor
Exa	E	10^{18}
Peta	P	10^{15}
Tera	T	10^{12}
Giga	G	10^{9}
Mega	M	10^{6}
Kilo	k	10^{3}
Milli	m	10^{-3}
Micro	μ	10^{-6}
Nano	n	10^{-9}
Pico	p	10^{-12}

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1

Table A.3 Symbols and molar masses (g mol^{-1}) for a selection of elements, molecules and ions

Symbol	Name	Molar mass
H	Hydrogen	1
C	Carbon	12
N	Nitrogen	14
O	Oxygen	16
S	Sulphur	32
CH_3SCH_3	Dimethylsulphide	62
CH_4	Methane	16
CO	Carbon monoxide	28
CO_2	Carbon dioxide	44
COS	Carbonyl sulfide	60
CS_2	Carbon disulfide	76
H_2O	Water	18
H_2S	Hydrogen sulfide	34
HNO_3	Nitric acid	63
H_2SO_4	Sulfuric acid	98
N_2	Nitrogen gas	28
NH_3	Ammonia	17
NH_4^+	Ammonium ion	18
NH_4NO_3	Ammonium nitrate	80
$(NH_4)HSO_4$	Ammonium bisulphate	115
$(NH_4)_2SO_4$	Ammonium sulphate	132
NO_3^-	Nitrate ion	62
O_2	Dioxygen	32
O_3	Ozone	48
SO_2	Sulphur dioxide	32
SO_4^{2-}	Sulphate ion	64

Appendix B
Properties of the Log-Normal Size Distribution

The log-normal size distribution is widely used in aerosol sciences. It is useful here to spell out the definitions and applications surrounding the log-normal size distribution in a clear way. The unnormalized log-normal size distribution $n^*(r)$ can be written as:

$$n^*(r) = \frac{N}{\sqrt{2\pi}\ \sigma_0}\ \exp\left[-\frac{1}{2}\left(\frac{\ln r - \ln r_0}{\sigma_0}\right)^2\right] \tag{B.1}$$

where $n^*(r)\ \mathrm{d}\ln r$ is the number of particles for which the logarithm of the radius is between $\ln r$ and $\ln r + \mathrm{d}\ln r$, N is the total concentration in particles, r_0 is a characteristic radius and σ_0 is a measure of the dispersion of the distribution.

An interesting property of the log-normal distribution is that the distribution of the random variable $z = r^n$ (radius to the power n) also follows a log-normal law:

$$f_z^*(r) = \frac{N\ r_0^n\ \exp\left(\frac{n^2\ \sigma_0^2}{2}\right)}{\sqrt{2\pi}\ \sigma_0}\ \exp\left[-\frac{1}{2}\left(\frac{\ln r - \ln\left(r_0 \exp\left(n\ \sigma_0^2\right)\right)}{\sigma_0}\right)^2\right] \tag{B.2}$$

$f_z^*(r)\ \mathrm{d}\ln r$ is the "quantity of z" for which the logarithm of the radius is between $\ln r$ and $\ln r + \mathrm{d}\ln r$. It can be checked that this distribution law is equivalent to that of $n^*(r)$ when $n = 0$. The total quantity of z integrated over the size distribution is:

$$M_0^z = N\ r_0^n\ \exp\left(\frac{n^2\sigma_0^2}{2}\right). \tag{B.3}$$

The mode and the median of the distribution of the radius to the power n are both equal to $r_0\ \exp\left(n\ \sigma_0^2\right)$.

For a log-normal size distribution, the moments can be calculated through a double integration by substitution:

$$\overline{r^n} = \int r^n\ n^*(r)\ \mathrm{d}\ln r\ /\ \int n^*(r)\ \mathrm{d}\ln r = r_0^n\ \exp\left(\frac{n^2\ \sigma_0^2}{2}\right) \tag{B.4}$$

so that the equivalent radius for the moment of order n is

$$r_n = \left(\overline{r^n}\right)^{\frac{1}{n}} = r_0\ \exp\left(\frac{n\ \sigma_0^2}{2}\right). \tag{B.5}$$

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1

The mean geometric radius (denoted r_g, r_{g0} or sometimes r_{gN}) verifies

$$\ln r_g = \frac{\int \ln r \; n^*(r) \, \mathrm{d} \ln r}{\int n^*(r) \, \mathrm{d} \ln r} \tag{B.6}$$

and is precisely equal to r_0 for a log-normal distribution. In other words the mode, median and mean geometric radius are the same quantity, which is a unique property of the log-normal distribution. The geometric standard deviation verifies:

$$(\ln \sigma_g)^2 = \int (\ln r)^2 \, n^*(r) \, \mathrm{d} \ln r - \left(\int \ln r \; n^*(r) \, \mathrm{d} \ln r \right)^2 . \tag{B.7}$$

For a log-normal distribution, one can show that

$$\ln \sigma_g = \sigma_0 . \tag{B.8}$$

Finally, the geometric radii of higher orders r_{gn} for $n \geq 1$, are defined as

$$\ln r_{gn} = \frac{\int \ln r \; n^*(r) \, r^n \, \mathrm{d} \ln r}{\int n^*(r) \, r^n \, \mathrm{d} \ln r} \tag{B.9}$$

and can again be calculated through a double integration by substitution:

$$r_{gn} = r_0 \; \exp(n \, \sigma_0^2) = r_0 \; \exp(n \, (\ln \sigma_g)^2) . \tag{B.10}$$

The geometric radius of order n is, therefore, equal to the mode and median of the distribution of the radius to the power n.

Appendix C
Mie Theory

C.1 Calculation of the Extinction Factor and Asymmetry Parameter

We follow the notations of van de Hulst (1981) to describe the Mie theory. The size parameter is defined as

$$x = \frac{2\,\pi\,r}{\lambda} \tag{C.1}$$

where r is the particle radius and λ is the wavelength. The extinction, scattering and absorption factors can be written as combinations of sums and products of a_n and b_n coefficients. The usual expression for these coefficients is given by:

$$\begin{cases} a_n = \dfrac{\psi_n'(y)\,\psi_n(x) - m\,\psi_n(y)\,\psi_n'(x)}{\psi_n'(y)\,\xi_n(x) - m\,\psi_n(y)\,\xi_n'(x)} \\ b_n = \dfrac{m\,\psi_n'(y)\,\psi_n(x) - \psi_n(y)\,\psi_n'(x)}{m\,\psi_n'(y)\,\xi_n(x) - \psi_n(y)\,\xi_n'(x)} \end{cases} \tag{C.2}$$

where m is the complex index of refraction and $y = m\,x$. The functions ψ and ξ of a complex variable are the Riccati–Bessel functions defined by:

$$\begin{cases} \psi_n(z) = \sqrt{\dfrac{\pi\,z}{2}}\,J_{n+1/2}(z) \\ \xi_n(z) = \sqrt{\dfrac{\pi\,z}{2}}\,H^{(2)}_{n+1/2}(z) \end{cases} \tag{C.3}$$

where $J_{n+1/2}$ and $H^{(2)}_{n+1/2}$ are the Bessel functions of the first and third types, respectively.

The extinction and scattering factors and the asymmetry parameter of a particle with a size parameter x can be written as:

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1

$$\begin{cases} Q^{\text{ext}} = \dfrac{2}{x^2}\displaystyle\sum_{n=1}^{\infty}(2n+1)\,\text{Re}(a_n + b_n) \\ Q^{\text{sca}} = \dfrac{2}{x^2}\displaystyle\sum_{n=1}^{\infty}(|a_n|^2 + |b_n|^2) \\ g = \dfrac{4}{x^2 Q^{\text{sca}}}\displaystyle\sum_{n=1}^{\infty}(2n+1)\Big[\dfrac{n(n+2)}{n+1}\text{Re}(a_n a_{n+1}^* + b_n b_{n+1}^*) + \dfrac{2n+1}{n(n+1)}\text{Re}(a_n b_n^*)\Big] \end{cases} \tag{C.4}$$

where Re designate the real part of a complex number and z^* is the conjugate of the complex number z.

The a_n and b_n coefficients can be computed through the algorithm of Kattawar and Plass (1967):

$$\begin{cases} a_n = \dfrac{\psi_n(x)}{\xi_n(x)}\left(\dfrac{D_n(y) - m\ D_n(x)}{D_n(y) - m\ G_n(x)}\right) \\ b_n = \dfrac{\psi_n(x)}{\xi_n(x)}\left(\dfrac{m\ D_n(y) - D_n(x)}{m\ D_n(y) - G_n(x)}\right) \end{cases} \tag{C.5}$$

where the functions D_n and G_n are defined as:

$$\begin{cases} D_n(z) = [\ln \psi_n(z)]' \\ G_n(z) = [\ln \xi_n(z)]'. \end{cases} \tag{C.6}$$

One can use an upward recurrence to compute $G_n(x)$:

$$\begin{cases} G_0(x) = -i \\ G_n(x) = -\dfrac{n}{x} + \left[\dfrac{n}{x} - G_{n-1}(x)\right]^{-1}. \end{cases} \tag{C.7}$$

The same recurrence relationship is verified by $D_n(z)$ but an upward recurrence does not converge. A downward recurrence should be used instead to determine $D_n(x)$ and $D_n(y)$:

$$\begin{cases} D_{N_0}(z) = 0 \\ D_n(z) = \dfrac{n+1}{z} - \left[\dfrac{n+1}{z} + D_{n+1}(z)\right]^{-1}. \end{cases} \tag{C.8}$$

The recurrence relationships provided by Toon and Ackerman (1981) can be used to compute $\psi_n(x)$ and $\xi_n(x)$:

$$\begin{cases} \xi_{-1}(x) &= \cos x - i\ \sin x \\ \xi_0(x) &= \sin x + i\ \cos x \\ \xi_n(x) &= \dfrac{2n-1}{x}\ \xi_{n-1}(x) - \xi_{n-2}(x) \\ \psi_n(x) &= \text{Re}[\xi_n(x)]. \end{cases} \tag{C.9}$$

The number of terms required to compute the sums in Eq. C.4 is given by Wiscombe (1980):

$$N = \begin{cases} \mathrm{E}(x + 4\,x^{1/3} + 1) & \text{si } 0.02 \leq x \leq 8 \\ \mathrm{E}(x + 4.05\,x^{1/3} + 2) & \text{si } 8 < x < 4200 \\ \mathrm{E}(x + 4\,x^{1/3} + 2) & \text{si } 4200 \leq x \leq 20000. \end{cases} \tag{C.10}$$

N_0 can be taken equal to $N + 10$ for the downward recurrence.

C.2 Calculation of the Phase Function

The calculation of the phase function requires the calculation of two more series of coefficients, which we note $\pi_n(\mu)$ and $\tau_n(\mu)$, where μ is the cosine of the scattering angle. These coefficients are defined by:

$$\begin{cases} \pi_n(\mu) = P_n'(\mu) \\ \tau_n(\mu) = \mu\,\pi_n(\mu) - (1 - \mu^2)\,\pi_n'(\mu) \end{cases} \tag{C.11}$$

where P_n is the Legendre polynomial of order n.

We use the recurrence relationship from Wiscombe (1980):

$$\begin{cases} \pi_0(\mu) = 0 \\ \pi_1(\mu) = 1 \\ s = \mu\,\pi_n(\mu) \\ t = s - \pi_{n-1}(\mu) \\ \tau_n(\mu) = n\,t - \pi_{n-1}(\mu) \\ \pi_{n+1}(\mu) = s + \dfrac{n+1}{n}\,t. \end{cases} \tag{C.12}$$

The series of coefficients $\pi_n(\mu)$ and $\tau_n(\mu)$ allow the calculation of the S_1 and S_2 complex scattering amplitudes for two orthogonal directions of incident polarization:

$$\begin{cases} S_1(\mu) = \displaystyle\sum_{n=1}^{\infty} \frac{2\,n+1}{n\,(n+1)}\,[a_n\,\pi_n(\mu) + b_n\,\tau_n(\mu)] \\ S_2(\mu) = \displaystyle\sum_{n=1}^{\infty} \frac{2\,n+1}{n\,(n+1)}\,[a_n\,\tau(\mu) + b_n\,\pi_n(\mu)]. \end{cases} \tag{C.13}$$

If we ignore polarization of the radiation, the normalized phase function (i.e. $\int_{-1}^{1} P(\mu)\,\mathrm{d}\mu = 2$) can be written as:

$$P(\mu) = \frac{4}{x^2\,Q^{\mathrm{sca}}}\,\left(|S_1(\mu)|^2 + |S_2(\mu)|^2\right). \tag{C.14}$$

One can use the same number of terms in the computation of the $S_1(\mu)$ and $S_2(\mu)$ as in the computation of Q^{ext} and g.

One can further estimate the monodirectional and isotropic upscatter fraction by appropriate integration of the phase function (Wiscombe and Grams 1976). It is possible to check that the discretization of $P(\mu)$ is sufficiently accurate by computing the asymmetry parameter g by integrating the phase function $P(\mu)$ and compare this value to the value that can be directly computed from the a_n and b_n coefficients.

For an aerosol size distribution, it is necessary to compute the a_n, b_n, $\pi_n(\mu)$ and $\tau_n(\mu)$ coefficients for each radius and the different parameters can be computed by appropriate integrations on the size distribution.

C.3 Polarized Phase Function

The previous section was only concerned with predicting the phase function for the intensity of the electromagnetic radiation. As seen in Sect. 5.6.7, scattering depends on the state of polarization of the radiation, and can also in return polarize unpolarized incident radiation. The elements of the scattering matrix for a homogeneous sphere can be obtained from the $S_1(\mu)$ and $S_2(\mu)$ parameters in the following way:

$$\begin{cases} S_{11} = \frac{1}{2}(|S_1|^2 + |S_2|^2) \\ S_{12} = \frac{1}{2}(|S_2|^2 - |S_1|^2) \\ S_{33} = \frac{1}{2}(S_1^* S_2 + S_1 S_2^*) \\ S_{34} = \frac{i}{2}(S_1 S_2^* - S_1^* S_2) \end{cases} \tag{C.15}$$

where the μ dependence has been omitted here. If the incident light is unpolarized, then applying Eq. 5.77 shows that the Stokes parameters of the scattered radiation are $I_s = S_{11} I_i$, $Q_s = S_{12} I_i$, and $U_s = V_s = 0$, which indicates a partial linear polarization. The degree of linear polarization for scattered radiation resulting from unpolarized radiation is therefore defined as $-S_{12}/S_{11}$, and varies between -1 and 1, with the sign indicative of the direction of linear polarization. The polarized phase function follows:

$$\mathbf{P} = \frac{4}{x^2 \, Q^{\text{sca}}} \mathbf{S}. \tag{C.16}$$

C.4 Extension of the Mie Theory and Other Theories

Toon and Ackerman (1981) extended the Mie theory to concentric stratified spheres. This model can be used to compute the optical properties of a solid aerosol covered with a liquid coating, such as a black carbon core with a sulphate shell. Fuller (1994, 1995a) established the theory for externally-aggregated spheres while Fuller (1995b)

treated the case for spheres containing arbitrarily located spherical inhomogeneities. Mishchenko and Travis (1994) have developed a technique to compute the optical properties of polydispersions of randomly-oriented oblate or prolate spheroids. The so-called T-matrix method can be used to treat any type of geometry. All these developments are relevant and important for atmospheric aerosols but fall outside the scope of this textbook. The bibliography lists a small number of textbooks on absorption and scattering of light by small particles.

References

Fuller KA (1994) Scattering and absorption cross sections of compounded spheres I. Theory for external aggregation. J Opt Soc Am 11:3251–3260

Fuller KA (1995a) Scattering and absorption cross sections of compounded spheres II. Calculations for external aggregation. J Opt Soc Am 12:881–892

Fuller KA (1995b) Scattering and absorption cross sections of compounded spheres III. Spheres containing arbitrarily located spherical inhomogeneities. J Opt Soc Am 12:893–904

Kattawar GW, Plass GN (1967) Electromagnetic scattering from absorbing spheres. Appl Opt 6:1377–1382

Mishchenko MI, Travis LD (1994) Light scattering by polydispersions of randomly oriented spheroids with sizes comparable to wavelength of observation. App Opt 33:7206–7225

Toon A, Ackerman TP (1981) Algorithms for the calculation of scattering by stratified spheres. Appl Opt 20:3657–3660

van de Hulst HC (1981) Light scattering by small particles. Dover Publications, New York, 470 pp (reprinted)

Wiscombe W (1980) Improved Mie scattering algorithms. Appl Opt 19:1505–1509

Wiscombe WJ, Grams GW (1976) The backscattered fraction in two-stream approximations. J Atmos Sci 33:2440–2451

Further Reading

Bohren CF, Huffman DR (1998) Absorption and scattering of light by small particles. Wiley Science Paperback Series, New York, 544 pp

Mishchenko MI, Hovenier JC, Travis LD (eds) (1999) Light scattering by nonspherical particles, theory, measurements and applications. Academic, 690 pp

Mishchenko MI, Travis LD, Lacis AA (2002) Scattering, absorption, and emission of light by small particles. Cambridge University Press, Cambridge, 480 pp

Appendix D
Radiative Impact of Aerosols on Snow and Ice

Wiscombe and Warren (1980) have proposed a parameterization of snow albedo which we have used to produce Fig. 8.7 and is explicited here. First of all, one needs to estimate the absorption and scattering properties of the snow and aerosol mixture. Snow crystals being very strong scatterers, only the absorption by aerosols matters. Generally speaking, one can estimate the properties of the mixture in the following way:

$$\begin{cases} \tau = \tau_{\text{snow}} + \tau_{\text{aerosol}} \\ \omega = \dfrac{\omega_{\text{snow}}\ \tau_{\text{snow}} + \omega_{\text{aerosol}}\ \tau_{\text{aerosol}}}{\tau_{\text{snow}} + \tau_{\text{aerosol}}} \\ g = \dfrac{g_{\text{snow}}\ \omega_{\text{snow}}\ \tau_{\text{snow}} + g_{\text{aerosol}}\ \omega_{\text{aerosol}}\ \tau_{\text{aerosol}}}{\omega_{\text{snow}}\ \tau_{\text{snow}} + \omega_{\text{aerosol}}\ \tau_{\text{aerosol}}}. \end{cases} \tag{D.1}$$

Wiscombe and Warren use a delta-Eddington approximation which consists in removing the forward scattering peak of snow crystals by redimensioning the three radiative parameters, optical depth, τ, single scattering albedo, ω, and asymmetry factor, g, into τ^*, ω^* and g^* according to the equations:

$$\begin{cases} g^* = \dfrac{g}{1+g} \\ \omega^* = \dfrac{(1-g^2)\,\omega}{1-g^2\,\omega} \\ \tau^* = (1-g^2\,\omega)\,\tau. \end{cases} \tag{D.2}$$

One can note that this equivalent layer scatters less in the forward direction ($g^* < g$), is more absorbing ($\omega^* < \omega$) and is optically thinner ($\tau^* < \tau$) but has the same absorption optical depth as the initial layer since $(1-\omega)\,\tau = (1-\omega^*)\tau^*$.

For direct solar radiation with a zenith angle $\theta_0 = \arccos \mu_0$ and an underlying Lambertian surface with albedo A, Wiscombe and Warren provides a solution to estimate the albedo under the delta-Eddington approximation:

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1

$$a_s(\mu_0) = \frac{1}{Q}\left[2\left(P(1-\gamma+\omega^* b^*) + \omega^*(1+b^*)\frac{\gamma\epsilon\mu_0 - P}{1-\epsilon^2\mu_0^2}\right)\exp(-\tau^*/\mu_0)\right.$$
$$\left. -\omega^* b^*(Q^+ - Q^-) + \omega^*(1+b^*)\left(\frac{Q^+}{1+\epsilon\mu_0} - \frac{Q^-}{1-\epsilon\mu_0}\right)\right] \quad \text{(D.3)}$$

where

$$\begin{cases} a^* = 1 - \omega^* g^* \\ b^* = g^*/a^* \\ \epsilon = \sqrt{3a^*(1-\omega^*)} \\ P = 2\epsilon/(3a^*) \\ \gamma = (1-A)/(1+A) \\ Q^+ = (\gamma + P)\,\exp(\epsilon\tau^*) \\ Q^- = (\gamma - P)\,\exp(-\epsilon\tau^*) \\ Q = (1+P)\,Q^+ - (1-P)\,Q^-. \end{cases} \quad \text{(D.4)}$$

When τ tends to infinity, the albedo $a_s(\mu_0)$ tends towards an asymptotic value $a_s^\infty(\mu_0)$ equal to:

$$a_s^\infty(\mu_0) = \frac{\omega^*}{1+P}\frac{1-b^*\epsilon\,\mu_0}{1+\epsilon\,\mu_0}. \quad \text{(D.5)}$$

The albedo of a snow layer for isotropic diffuse solar radiation can be calculated from the following integral:

$$a_d = 2\int_0^1 \mu_0\, a_s(\mu_0)\,\mathrm{d}\mu_0 \quad \text{(D.6)}$$

for which Wiscombe and Warren (1980) provide an analytical solution which we do not reproduce here.

References

Wiscombe WJ, Warren SG (1980) A model for the spectral albedo of snow. I: pure snow. J Atmos Sci 37:2712–2733

Further Reading

Warren SG, Wiscombe WJ (1980) A model for the spectral albedo of snow. II: snow containing atmospheric aerosols. J Atmos Sci 37:2734–2745

Index

A
absorption, 32, 33, 96
  band, 113
  coefficient, 43, 97, 101
  cross section, 36, 43, 97, 100
  factor, 36
  line, 88, 97
  mass efficiency, 37, 101
  volume efficiency, 37, 42, 101
  intensity, 92, 97, 107
  shape, 93
absorptivity, 102, 124, 178
accretion, 208
acid rain, 237, 250
acidification
  ocean, 292
  rain, 250
actinic flux, 119, 255
action spectrum, 254
activation, 193, 201, 202, 204
adding–doubling method, 112
adjoint model, 163, 167
aerodynamic diameter, 150
AERONET, 132, 135, 138, 166
aerosol
  biomass burning, 14, 31, 74, 179, 185, 236, 261
  black carbon, 15, 30, 33, 42, 45, 74, 165, 185, 213, 215, 236
  continental, 11, 258
  definition, 4, 9
  desert dust, 11, 12, 33, 130, 185, 249, 252, 262
  direct effect, 21, 175
  distribution, 15
  emissions, 15
  filter, 152
  indirect effect, 22
  inorganic, 30
  lifetime, 15, 23, 51, 271
  marine, 11, 257
  mixture, 29
  modelling, 73
  nitrate, 30, 185, 237, 249
  non-spherical, 45, 97, 101, 142, 302
  organic, 31
  organic carbon, 45
  primary, 10
  primary biogenic, 14, 260
  residence time, 15, 78
  sea salt, 12, 55, 72, 257
  secondary, 10
  secondary organic, 14, 62, 259
  sink, 15, 275
  sources, 15
  stratospheric, 130, 248, 255, 264, 271, 277, 288, 293
  sulphate, 185, 236, 255
  transport, 73
  urban, 11
  volcanic, 11, 13, 182, 215, 242
aerosol–cloud interactions, 22, 193, 233, 290
aerosol–radiation interactions, 21, 173, 233
aerosol–surface interactions, 22, 173
air mass (factor), 135
Aitken nucleus, 200
albedo
  cloud, 206
  snow, 187, 305
  surface, 120, 175
almucantar, 138
ammonia, 66, 250
ammonium sulphate, 30, 200
Ångström coefficient, 44, 165

O. Boucher, *Atmospheric Aerosols,* DOI 10.1007/978-94-017-9649-1

absorption, 45
Ångström parameter, *see* Ångström coefficient
asymmetry parameter, 38, 99, 100, 299
atmospheric boundary layer, 2
atmospheric window, 54, 113, 130
aureole, 138
autoconversion, 208, 210
aviation, 216

B
backscatter fraction, 38, 98
Beer-Lambert law, 107, 110
bidirectional reflectance function, 141
biogeochemical cycles, 5, 247
black body, 102, 144, 228
  feedback, 228
Boltzmann constant, 199
Bouguer's law, 97
Bouguer-Langley method, 135
boundary layer, 2
Brewer-Dobson circulation, 264, 273, 289
Bruggeman model, 34
burden, 15

C
CALIPSO, 148, 166
carbonyl sulphide, 272
chlorophyll, 61, 252, 258
Clausius–Clapeyron relation, 198
climate change, 6, 227
climate efficacy, 234, 236
climate engineering, 287
climate feedback, 227, 228, 255, 258, 264
  black body, 228
  cloud, 230
  fast, 291
  slow, 291
  surface albedo, 231
  water vapour, 230
climate response, 227, 228, 234
climate sensitivity, 231
  parameter, 228, 243
climate system, 1, 227
cloud, 4, 68, 69, 193
  optical depth, 205, 220
cloud albedo effect, 196
cloud condensation nucleus, 19, 22, 29, 69, 196, 200, 201, 204, 255, 290
cloud invigoration, 214
cloud radiative effect, 21, 185, 194
cloud seeding, 290
cloud-resolving model, 211
  global, 211
coagulation, 67, 75, 275, 289
coefficient
  absorption, 43, 97, 101
  extinction, 43, 101, 147
  scattering, 43, 101
collection, 208
condensation
  semi-volatile compounds, 67
  water, 198
condensation nucleus, 150, 201
conservation equation, 52
continuity equation, 52, 72, 208
contrail, 216
creeping, 56
cross section, 36, 100
  absorption, 36, 43, 97, 100
  extinction, 36, 100
  geometric, 36, 100
  mass absorption, 37, 101
  mass extinction, 37
  mass scattering, 37, 101
  scattering, 36, 43, 99, 100, 122
crystallisation, 36

D
data assimilation, 155, 161
  variational, 163
deforestation, 261
deliquescence, 36
delta-Eddington approximation, 117, 305
deposition
  dry, 71
  nucleation, 213
  velocity, 71
  wet, 69
desert dust, 12, 33, 56, 130, 249, 252, 262
detection and attribution, 6, 239
differential mobility analyzer, 151
dimethylsulphide, 14, 60, 77, 255, 264, 272
dimethylsulphide propionate, 60
direct effect, 21, 175
  exact computation, 179
  simplified equation, 175, 178
discrete-ordinate method, 112

E
ecosystems, 239, 289
  marine, 252, 264
  terrestrial, 54, 248, 259–261
effect
  aerosol direct, 173
  aerosol on snow, 22, 187, 242, 305
  direct, 21
  indirect, 22

semi-direct, 22, 185, 233
effective medium theory, 34
effective radiative forcing, 185, 196, 216, 231, 233, 235
efflorescence, 36
emission, 52
biogenic aerosol, 14
biomass burning, 14
desert dust, 12, 56
dimethylsulphide, 14, 60
fossil fuels, 53
future, 240
past, 237
sea salt, 12, 55, 257
vegetation fires, 54, 239
volatile organic compounds, 62
volcanic, 13
emissivity, 103, 182
entrainment, 205
eruption
1258, 279
El Chichón, 280
Krakataua, 279
Laki, 279
Mount Agung, 279
Mount Pinatubo, 248, 255, 280
Mount Tambora, 279
Toba, 281
volcanic, 248, 272
extinction
coefficient, 43, 101, 147
cross section, 36, 100
factor, 36
mass efficiency, 37, 167
volume efficiency, 37
extinction-to-backscatter ratio, 147, 155

F
factor
absorption, 36, 41, 100, 299
extinction, 36, 41, 100, 136, 299
scattering, 36, 41, 100, 299
feedback, 188, 227, 247
feedback parameter, 229
fire radiative power, 54
flaming, 54
freezing
heterogeneous, 213
homogeneous, 213

G
gain matrix, 163
Gamma function, 28, 79
Gamma law, 28, 46
modified, 28
gaseous precursor, 15, 253, 271, 288
geoengineering, 287
geometric standard deviation, 28, 46, 297
glaciation effect, 22, 196, 212
global warming, 6, 227
gravitational settling, 71
greenhouse effect, 4, 124, 212
aerosol, 182
greenhouse gases, 4, 113, 227
growth factor, 151

H
Henry's law, 68
heterogeneous chemistry, 255
heterogeneous freezing, *see* freezing 213
homogeneous freezing, *see* freezing 213
hydrogen sulphide, 13, 77, 271
hydrometeor, 3, 69
hygroscopicity, 29, 36, 151, 202
hysteresis, 36

I
ice nucleus, 19, 29, 196, 213
ideal gas constant, 67, 199, 295
impactor, 150
in situ measurement, 54, 129, 150, 166
indirect effect, 22
first, 22, 196, 205
glaciation, 22, 196, 212
second, 22, 196
isoprene, 62, 259

J
Jacobian matrix, 163
Junge law, 28
Junge layer, 276

K
Kalman filter, 164
Kelvin effect, 67, 199, 200
Kirchhoff's law, 102
Köhler theory, 219
Köhler theory, 200, 202
Koschmieder equation, 45

L
Lambertian surface, 138, 305
lapse rate, 2, 230
large-eddy simulation, 211
latent heat, 4, 118, 204, 232, 291
leaf area index, 63
lidar, 129, 146, 166, 277

equation, 146, 155, 167
Raman, 148
lifetime, 15, 23, 51, 271
limb, 130, 136
liquid water path, 205
log-normal distribution, 28, 47, 297

M
marine cloud brightening, 290
mass spectrometer, 152
Maxwell-Garnett equation, 34
methanesulphonic acid, 67, 77, 255
Mie
extension of the theory, 302
parameter, 40, 100, 299
theory, 40, 41, 100, 299
mineral dust, 12, 249, 252, 262
mixed-phase cloud, 19, 212
mixing ratio, 52
mode
accumulation, 25, 71, 75, 122, 142, 175, 179, 203
Aitken, 25, 75
coarse, 25, 75, 122, 142, 175, 203
fine, 25, 203
nucleation, 25, 75
supercoarse, 25, 75
ultrafine, 25
MODIS, 132, 166, 211
multiple scattering, 87, 112, 175

N
nephelometer, 153
new particle formation, 65
North Atlantic Oscillation, 263
nuclear winter, 242
nucleation, 25, 65
hetero-molecular, 65
homogeneous, 65
homomolecular, 65
ice crystal, 213
water droplet, 199
nutrient, 22, 249, 259, 262, 264

O
occultation, 136
optical depth, 107, 135
aerosol, 43, 165, 175, 277
cloud, 205, 220
molecular, 99
ozone layer, 2, 255, 258, 289

P
particle counter, 150
condensation, 150
pH
cloud water, 68
ocean, 292
rainwater, 250
phase function, 38, 98, 101, 137, 301
photolysis rate, 255
photometer, 130, 138
PHOTONS, 135
photosynthesis, 248
photosynthetically active radiation, 63, 248
phytoplankton, 60, 252, 255
Pinatubo, 248, 255, 277, 280
Planck constant, 295
Planck's law, 103
polar stratospheric cloud, 255, 276
polarization, 121, 130, 302

Q
quasi-biennal oscillation, 274

R
radiance, 86, 102, 134, 137–139, 145
radiation
absorption, 19, 96
diffuse, 138, 182, 248
scattering, 19, 96, 99, 137, 138
ultraviolet, 113, 254, 255, 289
radiative budget, 4, 117, 227
radiative flux, 86, 115, 228
radiative forcing, 227, 228
direct, 183
effective, 232
indirect, 193
radiative transfer, 105, 132, 179
radius
critical, 201
effective, 27
mean, 26
mean geometric, 297
mean geometric volume, 27
mean surface, 26
mean volume, 26
mean volume geometric, 297
median, 26
Raman effect, 96, 148
Raoult's law, 200
rapid adjustment, 22, 185, 196, 210, 216, 227, 231, 232, 291
reflectance, 112, 138, 145, 176
refractive index, 32, 96, 100, 136
relative humidity, 36, 66, 177, 198
remote sensing, 129

active, 129, 146
limb, 130
nadir, 130
passive, 129, 134
resuspension, 63
riming, 213

S
saltation, 56
sandblasting, 56
saturation pressure of water vapour, 198, 204, 213, 217
scattering
coefficient, 43, 101
cross section, 36, 43, 99, 100, 122
factor, 36
mass efficiency, 37, 101
Mie, 96
Raman, 96
Rayleigh, 96
volume efficiency, 37, 42, 101
scavenging
below-cloud, 19, 69
in-cloud, 19, 69
wet, 69
sea salt, 12, 55, 72, 257, 288
sea spray, 12, 72, 179, 257, 288
sedimentation, 71, 275
semi-direct effect, 22, 185, 233
sensible heat, 4, 118
ship track, 207, 290
single particle soot photometer, 154
single scattering, 87, 111, 137, 139, 175
single scattering albedo, 36, 100, 137, 178
size distribution, 19, 26, 57, 73, 136, 276, 297
size parameter, 100, 299
smoldering, 54
snow, 22, 187, 305
solar constant, 118, 175
solid angle, 85, 145, 146
specular reflection, 140
splintering, 213
Stefan–Boltzmann constant, 117, 295
Stefan–Boltzmann law, 117, 228
Stokes parameters, 121
Stokes vector, 121
Stokes velocity, 72, 78
Stokes' law, 72, 78
stratocumulus, 120, 185, 290
stratosphere, 2, 228, 242, 264, 271, 273, 288
sulphate, 68, 77, 177, 289
sulphur cycle, 77, 177, 256
stratosphere, 271
sulphur dioxide, 13, 77, 165, 237, 255, 271, 272
sulphuric acid, 30, 65–67, 255, 275
sunphotometer, 134
calibration, 135
supercooled water, 213
supersaturation, 65, 150, 198, 204, 217
surface albedo, 141, 175, 179, 231, 262
surface reflectance, 138, 140, 141
surface tension, 67, 199

T
tangent linear model, 163
terminal velocity, 72, 78, 210
transmittance, 107, 112, 113, 176
tropopause, 2, 273, 289
troposphere, 2
Twomey effect, 196

U
upscatter fraction, 38, 98, 175, 302
UV, 254

V
van't Hoff factor, 202
vegetation fires, 54, 239, 261
visibility, 45
Voigt profile, 95
volatile organic compound, 14, 62, 259
volcanic ash, 13, 272
volcanic explosivity index, 272

W
water activity, 200, 202
water cycle, 5, 193
wavelength, 84
wavenumber, 84
weekly cycle, 211
Wegener–Bergeron–Findeisen process, 213