Research Topics in Aerospace

Ulrich Schumann *Editor*

Atmospheric Physics

Background—Methods—Trends

DLR

Springer

Research Topics in Aerospace

For further volumes:
http://www.springer.com/series/8625

Ulrich Schumann
Editor

Atmospheric Physics

Background—Methods—Trends

Springer

Editor
Ulrich Schumann
Deutsches Zentrum für Luft- und Raumfahrt
DLR-Institut für Physik der Atmosphäre
82230 Oberpfaffenhofen
Germany

ISBN 978-3-642-30182-7 ISBN 978-3-642-30183-4 (eBook)
DOI 10.1007/978-3-642-30183-4
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012939480

Printed on acid-free paper

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

Preface

The Institute of Atmospheric Physics (Institut für Physik der Atmosphäre, IPA) of the German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt, DLR) in Oberpfaffenhofen near Munich, Germany was founded on 1 July 1962. This book appears on the occasion of the 50th anniversary of this institute. The book is part of the DLR-Springer Series, Research Topics in Aerospace, and as such addresses topics of relevance to aviation, Earth observation from Space, and further research and development fields of DLR including mobility and renewable energy. The book covers topics of special expertise in this institute.

The book is structured into three main sections dealing with general aspects of the atmosphere, methods and instruments used for research in this field, and topics and trends of major importance for present and ongoing research. Each section contains several chapters, in total 51. They cover specific topics, written by experts of this institute and partner institutes. Most of the external authors are former staff members of this institute. Several external authors are now members of the meteorological institute (MIM) of the Ludwig-Maximilians-University (LMU) with which IPA has a long tradition of cooperation and now an institutional link. Each chapter briefly describes the general knowledge in the field, and then explains in detail results obtained by the various teams represented by the authors. A few remarks in the chapters illuminate the historical development of the research topics in this field. The basic *keywords* (printed in italic) of importance in this field are defined in chapters identified by a list of keywords at the end of the book. The references cited refer to important publications of relevance for the field of research discussed, but most of the references refer to work performed in this institute.

The authors intended to describe their field in a manner understandable to a wide community of readers with little reference to background literature. We hope we succeeded and that the book is of interest and help to all who are related to aspects of atmospheric physics—scientific, institutional, educational or in cooperation.

The Institute of Atmospheric Physics has a long tradition in research on meteorology and atmospheric physics in relation to aerospace. It was formed from institutional roots going back to the Rhön-Rossitten society in the year 1924. It has kept its dedication and original name without changes over the last 50 years. Further details can be found in the final chapter.

The staff members of the Institute of Atmospheric Physics of DLR in Oberpfaffenhofen in January 2012 (Photo by Karl-Heinz Koos and Hans Volkert)

At present, the institute performs research on the physical and chemical processes of the troposphere and stratosphere. The knowledge of dynamical, cloud physical, and chemical processes constitutes the basis for many applications of aviation, Earth observation, mobility and renewable energy. On regional and global scales the relevant processes and changes of the state of the atmosphere are quantified and systematically investigated using remote sensing, research aircraft, and computational models. Example research topics include the impact of aviation and other modes of transport on climate, and the development of new measurement techniques for future satellite missions. The institute has access to the major research facilities of the DLR, including the research aircraft Falcon and HALO.

The institute is grateful to all the support it received over the decades from DLR management and cooperating institutes and facilities. We acknowledge the fruitful cooperation with many external partners, and thank for the support provided by Bavarian and Federal ministries, European and international funding partners, scientific societies, and for cooperation with industry.

It is my pleasure to take this occasion to thank all authors for their contributions. About 50 internal and several external reviewers helped essentially to improve the various contributions. Dietrich Heimann coordinated the reviews and helped in finding solutions of all kinds. Nadja Reimann and Sonja Mandelartz kept track of all contributions and formatted many of the lists of references. Reinhold Busen completed the list of abbreviations. Winfried Beer managed the data processing tasks. In particular I thank the lector, Susan Giegerich, for her professional work on the manuscripts in short time. Finally we thank the Springer Verlag for final preparation and publishing this book.

The institute gratefully acknowledges the pioneering work of the former directors of predecessors, Walter Georgii and Harald Koschmieder. Moreover we thank the founder of this institute, Hans Gerhard Müller (director from 1962 to 1972) and the former directors of IPA, Heinz Fortak (1973–1976) and Manfred

Reinhardt (1974–1992), for their pioneering work. I myself had responsibility for this institute since 1982, first jointly with Manfred Reinhardt, and alone since 1992. The institute is presently structured into five departments. The departments and their leaders are as follows: Atmospheric Dynamics (Robert Sausen), Atmospheric Trace Species (Hans Schlager), Atmospheric Remote Sensing (Ralf Meerkötter and Bernhard Mayer, LMU), Cloud Physics and Traffic Meteorology (Thomas Gerz in cooperation with George Craig, LMU) and Lidar (Gerhard Ehret). In spring 2012, the institute had about 125 staff members, most of them can be seen on the photo. I am grateful to all who supported this institute over the years. Finally I wish the readers of this book enjoyable and enlightening reading.

Oberpfaffenhofen, Germany, 1 April 2012 Ulrich Schumann

Contents

Abbreviations

3MI	Multi-viewing multi-channel multi-polarization imager
A2D	ALADIN airborne demonstrator
AATSR	Advanced along-track scanning radiometer
ABI	Advanced baseline imager
ABL	Atmospheric boundary layer
ACARE	Advisory council for aeronautics research in Europe
ACCESS	Arctic climate change economy and society
ACE	Atmospheric chemistry experiment
ACTA	Automatic contrail tracking algorithm
ADM-Aeolus	Atmospheric dynamics mission-Aeolus
ADWICE	Advanced diagnosis and warning system for aircraft icing environments
AerCARE	Impacts of aerosol layers on atmosphere and climate
AERO2k	Global aviation emissions inventory
AERONOX	Impact of NO_x emissions from aircraft on the atmosphere
AEROTROP	Impact of aircraft emissions on the heterogeneous chemistry of the tropopause region
AIC	Aviation induced cloudiness
AIMES	Analysis, integration and modeling of the Earth system
AIMS	Atmospheric chemical ionization mass spectrometer
AIRS	Atmospheric infrared sounder
ALADIN	Atmospheric laser Doppler instrument
ALEX-F	Airborne aerosol Lidar experiment
ALEXIS	Mie backscatter lidar (Phase A study)
AMAN	Arrival manager
AMDAR	Aircraft meteorological data relay
AMMA	African monsoon multidisciplinary analysis
AMS	Aerodyne mass spectrometer
Amu	Atomic mass units
AOCS	Attitude and orbit control system
AOD	Aerosol optical depth

AOGCM	Atmosphere-ocean general circulation models
APICS	Algorithm for the physical investigation of clouds with SEVIRI
APOLLO	AVHRR processing scheme over clouds, land and ocean
A-SCOPE	Advanced space carbon and climate observation of planet Earth
ASTAR	Spring-time transport of pollutants into the Arctic
ATC	Air traffic control
ATLID	Atmospheric Lidar
ATM	Air traffic management
ATR	Average temperature response
A-Train	Afternoon train (a NASA satellite constellation)
A-TReC	Atlantic THORPEX regional campaigns
ATSR-2	Along-track scanning radiometer
ATTAS	Advanced technologies testing aircraft system
ATTICA	European assessment of transport impacts on climate change and ozone depletion
AVA	Aerodynamische Versuchsanstalt
AVAPS	Airborne vertical atmospheric profiling system
AVHRR	Advanced very high resolution radiometer
AVOSS	Aircraft vortex spacing system
BAHAMAS	Basic HALO measurement and sensor system
BBR	Broad band radiometer
BC	Black carbon
BMBF	(German) Federal ministry of education and research
BMVBS	Federal ministry for traffic, construction and urban development
BTD	Brightness temperature difference
C4MIP	Coupled climate-carbon cycle model intercomparison project
CALIOP	Cloud-aerosol lidar with orthogonal polarization
CALIPSO	Cloud-aerosol lidar and infrared pathfinder satellite observations
CAPE	Convective available potential energy
CARIBIC	Civil aircraft for the regular investigation of the atmosphere based on an instrument container
CAS-DPOL	Cloud and aerosol spectrometer
CASI	Compact airborne spectrographic imager
CATS	Climate-compatible air transportation system
CBL	Convective boundary layer
Cb-TRAM	Cumulonimbus tracking and monitoring
CCD	Charge-coupled device
CCM	Chemistry climate model
CCMVal	Chemistry-climate model validation
CCN	Cloud condensation nuclei
CDA	Contrail detection algorithm

CDC	Collision dissociation chamber
CERES	Clouds and the Earth radiant energy system
CFC	Chlorofluorocarbons
CFH	Cryogenic frost point hygrometer
CFL	Courant-Friedrichs-Lewy criterion
CG	Cloud-to-ground
CHARM®	CH_4 airborne remote monitoring
CI	Chemical ionization
CIMS	Chemical ionization mass spectrometry
CIN	Convective inhibition energy
CIP	Cloud imaging probe
CiTTyCAT	Cambridge tropospheric trajectory model of chemistry and transport
CIWSIR	Cloud ice water submillimeter imaging radiometer
CLEOPATRA	Cloud experiment Oberpfaffenhofen and transports
CMD	Count median diameter
CMIP5	Coupled model intercomparison project
CN	Condensation nuclei
CNES	Centre national d'Etudes Spatiales
CNRS	National center for scientific research, France
CoCiP	Contrail cirrus prediction model
COCS	Cirrus optical properties derived from CALIOP and SEVIRI
CONCERT	Contrail and cirrus experiment
CONTRACE	Convective transport of trace gases into the middle and upper troposphere over Europe: budget and impact on chemistry
COPS	Convective and orographically-induced precipitation study
COSMO	Consortium for small-scale modeling, NWP model of DWD
COST	European cooperation in science and technology
CPC	Condensation particle counter
CPI	Cloud particle imager
CPR	Cloud profiling radar
CPU	Central processing units
CRDS	Cavity ring-down spectroscopy
CRPSS	Conditional ranked probability skill score
CRS	Cloud resolving simulation
CS	Calibration source
CTBTO	Comprehensive nuclear-test-ban treaty organization
CTH	Cloud-top-height
CW	Continuous-wave
D2P	Deterministic two-phase wake vortex model
DAOD	Differential atmospheric optical depth
DE	Detection efficiency
DE	Deutschland (Germany)
DERA	Defence evaluation and research agency, UK

DFD	Deutsches Fernerkundungsdatenzentrum, DLR Oberpfaffenhofen
DFG	Deutsche Forschungsgemeinschaft
DFS	Deutsche Forschungsanstalt für Segelflug
DFS	Deutsche Flugsicherung (German Air Traffic Control)
DFVLR	Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt
DGVM	Dynamic global vegetation models
DIAL	Differential absorption lidar
DLH	Deutsche Lufthansa AG
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German aerospace center)
DMA	Differential mobility analyzer
DMPS	Differential mobility particle sizer
DMS	Dimethylsulfide
DMT	Droplet measurement technologies, Boulder CO, USA
DNS	Direct numerical simulation
DOA	Approved design organization
DOAS	Differential optical absorption spectroscopy
DOC	Direct operating costs
DU	Dobson units
DVL	Deutsche Versuchsanstalt für Luft- und Raumfahrt
DWD	Deutscher Wetterdienst / German meteorological service
DWL	Doppler wind lidar
EADS	European aeronautic defense and space company
EARLINET	European aerosol research lidar network
EarthCARE	Earth clouds, aerosols and radiation explorer
EC	European commission
ECHAM5	ECMWF and Hamburg version 5 of an atmospheric GCM
ECMWF	European centre for medium-range weather forecasts
EC-TRAM	Electrically-charged cells tracking and monitoring
ECV	Essential climate variables
EF	Energy forcing
ELITE	European contribution to LITE
EMAC	ECHAM/MESSy atmospheric chemistry model
ENIAC	Electronic numerical integrator and computer
ENSO	El-Niño southern oscillation
ENVISAT	Environmental multi instrument satellite (launched 2002)
EOS	Earth observing system
EPS	Ensemble prediction system
ERBE	Earth radiation budget experiment
ERBS	Earth radiation budget satellite
ESA	European space agency
ESM	Earth system model
ESMVal	Earth system model validation

ESRIN	European space research institute
ETH	Swiss federal institute of technology Zurich
EU	European union
EU15	Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, United Kingdom
EUCAARI	European integrated project on aerosol, cloud, climate, and air quality interactions
EUFAR	European facility for airborne research
EULAG	Eulerian/semi-lagrangian numerical model for fluids
EULINOX	European lightning nitrogen oxides experiment
EUMETSAT	European organisation for the exploitation of meteorological satellites
EUROCONTROL	European air traffic management centre
EUPHORE	European photoreactor
EUPLEX	European polar stratospheric cloud and lee-wave experiment
EURASER	European research aircraft and sensors for environmental research
EXPORT	European export of precursors and ozone by long-range transport
FAA	Federal aviation administration
FAAM	Facility for airborne atmospheric measurements
FACE	Free atmospheric CO_2 enrichment
FAR	False alarm ratio, false alarm rate
FBI	Frequency bias index
FCI	Flexible combined imager
FFM	Flugwissenschaftliche Forschungsanstalt München
FIONA	Formal intercomparisons of observations of nitrous acid
FMG	Flughafen München GmbH
FSC	Fuel sulfur content
FSSP	Forward scattering spectrometer probe
FU Berlin	Freie Universität Berlin
FWHM	Full-width at half-maximum
FZJ	Forschungszentrum Jülich (Jülich Research Center)
GAC	Gulfstream aerospace corporation
GALION	GAW aerosol lidar observation network
GAW	Global atmosphere watch
GCM	General circulation model
GCOS	Global climate observing system
GEOSAR	Geostationary search and rescue
GERB	Geostationary Earth radiation budget
GFS	Global forecast system
GHG	Greenhouse gas
GOES-R	Geostationary operational environmental satellite, R-series
GOME	Global Ozone monitoring experiment on the ERS-2 satellite

GOSAT	Greenhouse gases observing satellite
GPS	Global positioning system
GRACE	German contribution to POLARCAT
GSD	Geometric standard deviation
GTP	Global temperature potential
GWP	Global warming potential
HALO	High altitude and long range research aircraft
HALOE	Halogen occultation experiment
HAMP	HALO/Hamburg microwave package
HCC	Hub control center
HGF	Helmholtz association of German research centers
HIAPER	High-performance instrumented airborne platform for environmental research, operated by NCAR
HIPPO	HIAPER pole-to-pole observations
HLOS	Horizontal line-of-sight direction
HPC	High performance computing
HRDL	High resolution Doppler lidar
HRV	High resolution visible channel
HSR	High spectral resolution
HSRL	High spectral resolution lidar
HYSPLIT	Hybrid single particle lagrangian integrated trajectory model
IAMAS	International association of meteorology and atmospheric sciences
IC	Intracloud
ICA	Independent column approximation
ICAO	International civil aviation organization
ICAROHS	Inter-comparison of aerosol retrievals and observational requirements for multi-wavelength HSRL systems
ICARTT	International consortium for atmospheric research on transport and transformation
IfT	Leibniz-Institut für Troposphärenforschung e.V
IGBP	International geosphere biosphere programme
IHOP	International H_2O project
ILAMB	International land model benchmarking
IN	Ice nuclei
INCA	Interhemispheric differences in cirrus properties from anthropogenic emissions
INSU	Institut national de sciences de l'universe (CNRS)
INTEX	Intercontinental pollutant transport in the northern hemisphere
IOP	Intensive operational period
IPA	Institut für Physik der Atmosphäre, DLR
IPCC	Intergovernmental panel on climate change
IPDA	Integrated path differential absorption
IS	Ion source

ISA	International standard atmosphere
ISSR	Ice-supersaturated region
ITMS	Ion trap mass spectrometer
ITWS	Integrated terminal weather system
J(NO2)	Photolysis rate of NO_2
J-process	Junction process
KIT	Karlsruhe institute of technology
KTP	Potassium titanyl phosphate
LACE	Lindenberg aerosol characterization experiment
LBA	Luftfahrtbundesamt (the German national aviation authority)
LCL	Lifting condensation level
LCM	Lagrangian cirrus model
LDR	Linear depolarization ratio
LEEA	Low emissions effect aircraft
LES	Large eddy simulation
LF	Low frequency
LFC	Level of free convection
Lidar	Light detection and ranging
LIDO-OC	An operational flight planning tool of Lufthansa systems
LINET	Lightning detection network
LINOX	Lightning-produced NO_x
LIS	Lightning imaging sensor
LITE	Lidar in space technology experiment
LMU	Ludwig-Maximilians-Universität München
LNB	Level of neutral buoyancy
LNOx	Lightning-induced nitrogen oxides
LOPAP	Long path absorption photometer
LOS	Line of sight
LQMS	Linear quadrupole mass spectrometer
LS	Lower stratosphere
LW	Longwave
LWC	Liquid water content
MADE-IN	Modal aerosol dynamics model including insoluble modes
MAG	Mission advisory group
MAP	Marine aerosol production from natural sources
MAP	Mesoscale alpine program
MCS	Mesoscale convective system
MeCiDA	Meteosat cirrus detection algorithm
MEMOSA	Measurement and modeling of air pollution transport in the alpine region
MERIS	Medium spectral resolution imaging spectrometer
MERLIN	Methane remote sensing lidar mission
MESOSCOP	Mesoscale flow and cloud model Oberpfaffenhofen
MESSy	Modular Earth submodel system
MET	Aviation-meteorological information

METEOSAT	Geostationary meteorological satellite operated by EUMETSAT
MIM	Meteorologisches Institut München (LMU)
MINOS	Pollutant transport into the mediterranean boundary layer
MIPAS	Michelson interferometer for passive atmospheric sounding
MISR	Multiangle imaging spectroradiometer
ML-Cirrus	Mission on mid-latitude cirrus clouds, HALO experiment
MLS	Microwave limb sounder
MODIS	Moderate resolution imaging spectroradiometer
MOL-RAO	Meteorologisches Observatorium Lindenberg/Richard Assmann Observatorium
MOPITT	Measurements of pollutants in the troposphere
MOZAIC	Measurement of ozone and water vapor by Airbus in-service aircraft
MPG	Max Planck society
MPI	Message passing interface
MPI-K	Max-Planck-Institut für Kernphysik, Heidelberg
MPI-K HD	Atmospheric physics research group at the MPI-K
MSA	Methane sulfonic acid
MSG	METEOSAT second generation
MSI	Multi-spectral imager
MSP	Mie spectrometer
MTG	METEOSAT third generation
MTOW	Maximum take-off weight
MUC	Munich international airport
MULIS	Multiwavelength lidar system
MVIRI	METEOSAT visible and infrared imager
MYSTIC	Monte carlo code for the physically correct tracing of photons in cloudy atmospheres
NAIS	Neutral cluster / air ion spectrometer
NASA	National aeronautics and space administration
NARVAL	Next-generation aircraft remote-sensing for validation studies
NAT	Nitric acid trihydrate
NCAR	National center for atmospheric research
Nd:YAG	Neodymium-doped yttrium-aluminum garnet
NDIR	Non-dispersing infrared spectroscopy
NILU	Norwegian institute for air research (NILU), Kjeller, Norway
NLO	Nonlinear optical
NMHC	Non-methane hydrocarbon
NOAA	National oceanographic and atmospheric administration
NOy	Total reactive nitrogen; odd nitrogen

NPOESS	National polar-orbiting operational environmental satellite system
NWP	Numerical weather prediction
OC	Organic carbon
OCS	Carbonyl sulfide
OFP	Ozonforschungsprogramm, German national ozone research programme
OLEX	Ozone lidar experiment
OLR	Outgoing longwave radiation
OMO	Oxidation mechanisms observations HALO experiment
ONERA	Office national d'etudes et recherches aérospatiales, french national aerospace research center
OPA	Optical parametric amplifier
OPC	Optical particle counter
OPO	Optical parametric oscillator
OSSE	Observing system simulation experiment
OTD	Optical transient detector
P2P	Probabilistic two-phase wake vortex model
PAN	Peroxyacetyl nitrate ($CH_3C(O)O_2NO_2$)
PARTEMIS	Measurement and predictions of the emission of aerosols and gaseous precursors from gas turbine engines
PAZI	Partikel und Zirrus, a DLR project
PBL	Planetary boundary layer
PCASP-0X	Passive cavity aerosol spectrometer probe
PDE	Partial differential equation
PDF	Probability density function
PFA	Perfluoroalkoxy
PFC	Perfluorocarbon
PM	Particulate matter
PM	Phase-matching
PMC	Polar mesospheric clouds
PMCP	Perfluromethylcyclopentane (C_6F_{12})
PMS	Particle measuring systems, Boulder CO, USA
PN	Polar nephelometer
PNJ	Polar night jet
POD	Probability of detection
POLARCAT	Polar study of climate, chemistry, aerosols, and transport
POLDIRAD	Polarization diversity Doppler radar
POLINAT	Pollution from aircraft emissions in the North Atlantic flight corridor
POLIS	Portable lidar system
POLSTAR	Polar stratospheric aerosol experiment
POM	Particulate organic matter
PPA	Plane parallel assumption
PPI	Plan position indicator

PRF	Pulse repetition frequency
PRT	Pulse repetition time
PSAP	Particle soot absorption photometer
PSC	Polar stratospheric clouds
PSL	Polysterene latex particles
RSP	Rayleigh spectrometer
RSR	Reflected solar radiation
PTH	Pressure, temperature, humidity
PV	Potential vorticity
PVU	Potential vorticity unit
QBO	Quasi-biennial oscillation
QPM	Quasi-phase matching
QUANTIFY	Quantifying the climate impact of global and european transport systems
radar	Radio detection and ranging
Rad-TRAM	Radar tracking and monitoring
REACT4C	Reducing emissions from aviation by changing trajectories for the benefit of climate
RECONCILE	Reconciliation of essential process parameters for an enhanced predictability of arctic stratospheric ozone loss and its climate interaction
RF	Radiative forcing
RF	Radio frequency
RH	Relative humidity
RHI	Range height indicator
RHi, RHI	Relative humidity with respect to ice
RUAG	International technology company for aerospace and defense
RVSM	Reduced vertical separation minimum
SAC	Schmidt-Appleman criterion
SAF	Satellite application facility
SALTRACE	Saharan aerosol long-range transport and aerosol-cloud-interaction experiment
SAMUM	Saharan mineral dust experiment
SANA	Origin and distribution of air pollutants over eastern Germany
SAR	Specific air range
SCIAMACHY	Scanning imaging absorption spectrometer for atmospheric cartography
SCOUT-O3	Stratospheric-climate links with emphasis on the upper troposphere and lower stratosphere
SDR	Solar direct radiation
SeaKLIM	Impact of ship emissions on atmosphere and climate
SEEPS	Stable equitable error in probability space
SESAME	Second european stratospheric arctic and mid-latituide experiment

SESAR	Single European sky ATM research programme
SEVIRI	Spinning enhanced visible and infra-red imager
SLD	Supercooled liquid droplet
SFM	Sum frequency mixing
SGS	Sub-grid scale
SHA	Simplified hazard areas
SHIVA	Stratospheric ozone: halogen impacts in a varying atmosphere
SID	Small ice detector
SIP	Simulation ice particles
SMPS	Scanning mobility particle sizer
SOA	Secondary organic aerosol
Sodar	Sonic detection and ranging
SP2	Single particle soot photometer
SPARC	Stratospheric processes and their role in climate
SPEC	Stratton park engineering company, Boulder CO, USA
SPDB	Secondary power distribution box
SPL	Sound pressure level
SRES	IPCC emission scenarios (Special report emissions scenario)
STAAARTE	Scientific training and access to aircraft to atmospheric research throughout Europe
STD	Standard deviation
STISS	Second THORPEX international science symposium
STP	Standard pressure and temperature (0 °C, 1013.25 hPa)
SUCCESS	Subsonic aircraft: contrail and cloud effects special study
SULFUR	Sulfur experiment of DLR-IPA
SW	Shortwave
SYNOP	Numerical reporting code for surface synoptic observations
SYNRAD	Synthetic radar
SYNSAT	Synthetic satellite
TAC	Transport, atmosphere and climate conference
TACTS	Transport and composition in the upper troposphere and lowermost stratosphere HALO mission
TAF	Terminal aerodrome forecast
TANSO	Thermal and near-infrared sensor for carbon observation
TCAS	Traffic alert and collision avoidance system
TDL	Tunable diode laser
TGI	Trace gas inlet for HALO
THESEO	Third European stratospheric experiment on ozone
THORPEX-IPY	The observing system research and predictability experiment-international polar year
TIL	Tropopause inversion layer
TIROS	Television infrared observation satellite
TKE	Turbulent kinetic energy
TLE	Transient luminous events

TMA	Terminal maneuvering area
TMI	TRMM microwave imager
TOA	Time of arrival
TOA	Top of the atmosphere
TOVS	TIROS operational vertical sounder
T-PARC	THORPEX Pacific Asian regional campaign
TQMS	Triple quadrupole mass spectrometer
TREND	Trend-type landing forecasts
TRMM	Tropical rainfall measuring mission
TROCCINOX	Tropical convection, cirrus and nitrogen oxides experiment
TTL	Tropical transition layer
TU	Technical university
UARS	Upper atmosphere research satellite
UFO	Umweltgerechte Flugroutenoptimierung (a BMBF project on environmentally compatible flight route optimization)
UFS	Environmental research station Schneefernerhaus
UHSAS	Ultra-high sensitivity aerosol spectrometer
UNFCCC	United Nations framework convention on climate change
UPD	Updraft velocity
UT	Upper troposphere
UTC	Universal time coordinated
UTLS	Upper troposphere and lower stratosphere
UV	Ultraviolet
VAAC	Volcanic ash advisory center
VEI	Volcanic explosivity index
VERA	Vienna enhanced resolution analysis
VERTICATOR	Vertical transport and orography project
VESA	Vortex encounter severity assessment
VEU	Verkehrsentwicklung und Umwelt
VHF	Very high frequency
VHRR	Very high resolution radiometer
VII	Visible/infrared imager
VIIRS	Visible/infrared imager/radiometer suite
VISSR	Visible infrared spin scan radiometer
VLF	Very low frequency
VOC	Volatile organic carbons
VolcATS	Volcanic ash impact on the air transport system, a DLR project
WakeScene	Wake vortex scenarios simulation
WALES	Water vapor lidar experiment in space
WARAN	Water vapor atmospheric analyzer
WCB	Warm conveyor belt
WCRP	World climate research programme
WEPS	Wake encounter prevention system
WEZARD	Weather hazards of aviation

WGCM	Working group on coupled modeling
WGNE	Working group on numerical experimentation
WIND	Wind infrared doppler lidar
WMO	World meteorological organization
WSVBS	Wake vortex prediction and monitoring system
WWO	Winter weather object
WWRP	WMO world weather research programm
WxFUSION	Weather forecast user-oriented system including object nowcasting
XGHG	Column-integrated dry-air greenhouse gas mixing ratio
ZDR	Differential reflectivity

Part I

Background

The Atmosphere: Vast, Shallow, and Full of Subtleties

1

Andreas Dörnbrack

Abstract

This chapter introduces relevant atmospheric terms used throughout the book. It is also an attempt to describe the hydrostatic nature of the atmosphere and to touch on the fundamental tropospheric and stratospheric circulations. Besides these large-scale circulations, two selected atmospheric subtleties investigated at DLR's Institute of Atmospheric Physics (IPA) are presented: the extra-tropical tropopause inversion layer and mesoscale stratospheric mountain waves.

1.1 Introduction

The primary properties of the *atmosphere* are its pressure, density, temperature and composition. Motion is certainly a further fundamental property, and as Tuck (2008) opened his book, "The atmosphere consists of molecules in motion, …." A classical way to experience changes in atmospheric properties is to travel. Hiking up a high mountain from sea level to the peak gives an impression of the dramatic variations of pressure, temperature, and constituents such as oxygen or water vapor over a few kilometers in altitude. Usually, even the vegetation changes from tropical wilderness to arctic sparseness, a transition which otherwise only occurs over a horizontal distance of thousands of kilometers. What does this

A. Dörnbrack (✉)
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Andreas.Dörnbrack@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_1,

spatial difference in the rate of change, however, mean for scientists? It means that the atmosphere is an extremely thin, shallow spherical shell, reflecting the Greek roots of the word: ατμός, vapor and σφαῖρα, sphere. Geometrically, the great circle length is 40 000 km, the depth of the atmosphere about 40 km; this gives an extremely large aspect ratio of 1 000:1!

The tallest mountains on Earth do not reach the altitudes which were investigated by IPA scientists. For this purpose, different aircraft were employed to observe the physical and chemical properties of the atmosphere over the last fifty years, with changing research foci. Figure 1.1 exemplifies selected vertical temperature profiles measured on board the Geophysica and Falcon research aircraft at three different sites. These observations show that vertical changes over short distances are much more dramatic at one location than horizontal ones over thousands of kilometers. Nevertheless, scientists on long-range research flights to Australia, East Asia, Africa, South and North America were faced with quite diverse atmospheric phenomena.

It is the goal of this chapter to briefly overview the vertical structure of the atmosphere and to sketch the main circulation patterns for later reference. Obviously, a complete overview, or even a thorough introduction to atmospheric circulations and their subtleties cannot be given here. Therefore, the description will be intentionally kept short and fragmentary. For more detailed information the reader is referred to textbooks, such as Vallis (2006).

1.2 The Vertical Structure of the Atmosphere

The changes in atmospheric properties with height are the result of specific physical conditions and processes on Earth and in its atmosphere. The actual temperature of the air depends on the energy its molecules receive from *radiation*. There are two main sources of radiation, incoming solar radiation and the outgoing longwave radiation from Earth's surface. The sun's radiation is mostly visible (44 %), near-infrared (37 %), and ultraviolet (7 %), while Earth's radiation is mostly far-infrared. Absorption of the sun's radiation heats the earth's surface, which then warms the air above it.

Research aircraft employed by IPA record numerous vertical temperature profiles mostly during climb and descent. In Fig. 1.1, a few soundings are shown to reflect the variability researchers are faced with at their campaign bases in different parts of the world. The observed vertical temperature distribution has a characteristic structure with a couple of principal layers. The outstanding characteristic of the lower atmosphere in general, above the ground surface, both summers as winter, is the tendency for the temperature to decrease with height. The layer above the earth's surface about 8–16 km deep is called the *troposphere*. In a standard atmosphere representing average conditions the temperature decreases from the surface upward nearly evenly by $\gamma \approx 6.5$ K km^{-1}, as sketched by the thick gray line in Fig. 1.1. This rate of temperature change observed in

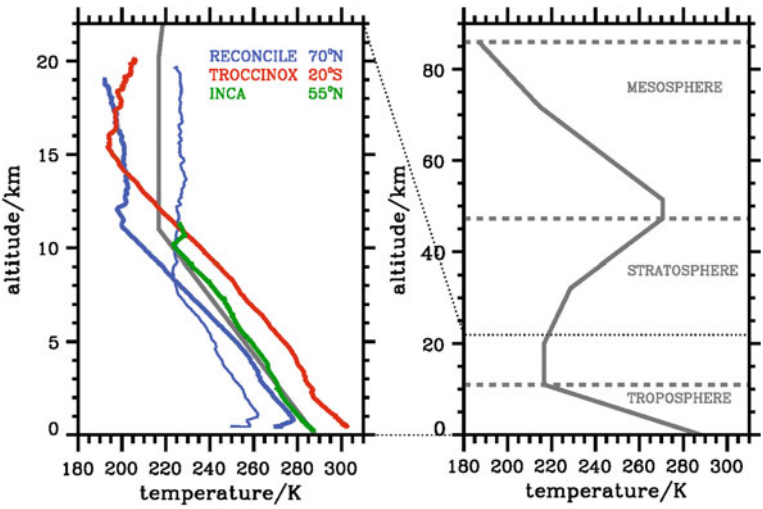

Fig. 1.1 *Left panel* vertical temperature profiles recorded during descents of the Geophysica research aircraft at Kiruna (Sweden) in February 2010 (*blue*) and Aracatuba (Brazil) in February 2006 (*red*), and of the Falcon aircraft in Prestwick (Scotland) in September 2000 (*green*). The two arctic profiles illustrate the remarkable stratospheric temperature change during a sudden stratospheric warming event, thick line before, thin line after the event. *Right panel* sketch of the principal atmospheric layers. The temperature profile of the *standard atmosphere* is plotted in *gray*

passing upward from one layer of air to the next is called the lapse rate $\gamma = -\partial T/\partial z$. The actual *lapse rate* γ observed in the atmosphere varies with location, time of day, weather conditions and season, among other factors.

The density and pressure decrease with height as consequence of gravity, which sorts the masses in the atmosphere in such a way that horizontal gradients tend to vanish. In other words, gravity restores a *hydrostatic* equilibrium effectively. A horizontal layering of clouds and aerosol particles is the visible manifestation of the restoring forces. As a consequence, in most cases it is possible to calculate the surface pressure as the weight of a 1.01972 kg column of dry air above a 1 cm^2 ground surface. The pressure at a certain elevation depends on the weight of air above that height. Under standard conditions (288 K, $\rho = 1.225$ kg m^{-3}) the pressure decreases $\Delta p \approx 1$ hPa per $\Delta h \approx 8$ m according to $-p = \rho g h$, where g = 9.81 m s^{-2} is the gravitational acceleration. When warm air near the ground rises due to buoyancy, the air expands and cools. According to the thermal energy equation $p = \rho c_P T$, the temperature drops by $\Delta T \approx 0.082$ K per 8 m, which is nearly 1 K/100 m. This value is close to the dry adiabatic lapse rate of $-\Delta T/\Delta h = g/c_P = 9.8$ Kkm^{-1}. Diabatic processes, such as phase changes in moist air or radiation, lead to deviations from this value.

The troposphere contains approximately 80 % of the atmosphere's mass and 99 % of its water vapor. The word "tropos" (Greek: $\tau\rho\acute{o}\pi o\varsigma$) means "turn" and hints at the ceaseless mixing character of motion in this portion of the atmosphere (Dutton 1986). Because warm air tends to rise and cool air tends to sink, the troposphere is a location of turbulence. Hence, the troposphere as the home of all significant weather events is described by meteorologists as being well-mixed. Besides vertical exchange processes caused by thermal convection, horizontal temperature gradients drive the winds in the troposphere.

The lowest part of the troposphere, where direct contact with the earth's surface influences air flow by momentum and heat exchange, is called the *atmospheric boundary layer* (ABL). This narrow layer is typically a few tens of meters to several kilometers deep depending on the geographical location, season, and time of day. Inside the ABL, the temperature can increase with height. When this happens, it is called an *inversion*, see, e.g., the pronounced inversions near the surface in both arctic profiles in Fig. 1.1. Most often, the ABL is capped by such an inversion layer, which limits the vertical exchange of heat, momentum and constituents. In the early years of IPA, instrumented aircraft such as motor gliders were extensively used to investigate dynamical and chemical processes in this layer. The portion above the ABL is often called the *free atmosphere* as it is only weakly impacted by exchange processes with the earth's surface, and the dynamical processes can be treated as for an ideal fluid. At the top of the troposphere the rapid temperature decrease ceases. This boundary where the temperature attains its minimum (Fig. 1.1) is called the *tropopause*, see (Hoinka 1997) for a historical review of the discovery of the tropopause.

In contrast to the troposphere, radiative processes determine significantly the vertical temperature distribution in the *stratosphere*. In the standard atmosphere, the stratosphere is marked by an inversion up to 50 km altitude, where the temperature increases to about 270 K, just slightly below the freezing point of water. The *ozone layer*, located near 25 km above the Earth's surface, absorbs ultraviolet radiation from the sun which, in turn, increases the motion of the ozone molecules. The ozone molecules then collide with other molecules in the air, increasing its temperature. As in the troposphere, the stratospheric air is stably stratified; however, as the lapse rate $\gamma \approx 0$, the *Brunt-Väisälä frequency* N (see Box 1.1) is about twice as large. This means there are few overturning air currents and, thus, the stratosphere is a region of little vertical mixing, in accordance with the original meaning of the syllable strato (lat.: *strātum* = *stratus)*. Tropospheric particles injected into the stratosphere can stay aloft for many years without returning to the ground. This was observed after large volcanic eruptions such as the Mount Pinatubo eruption in June 1991 that emitted ash into the stratosphere, where it remained for years and impacted the stratospheric chemistry.

On various flight missions to investigate convective storms at low latitudes, the *tropical tropopause layer* was observed by the Geophysica aircraft. Visually, the bottom of the stratosphere becomes apparent by the tops of these storms, which rarely overshoot into the lower stratosphere. Because of the apparent sharp temperature inversion—compare the tropical profile with the arctic profiles in

Fig. 1.1—air rising in these thunderstorms eventually becomes cooler than its environment in the stratosphere and stops rising.

So far, the principal layers higher up in the atmosphere were not the focus of research at IPA. Embedded between the stratosphere and the *thermosphere* is the *mesosphere*, named in accordance with the meaning of the Greek word $\mu\acute{\varepsilon}\sigma o\varsigma$ = middle. The temperature decreases with increasing height in the mesosphere, and the *mesopause* as its upper boundary can be the coldest naturally occurring place in the near-Earth atmosphere at $T < 130$ K.

In addition to the classical division of the atmosphere into its principal layers, there is a remarkable variation in the depth of the layers from place to place and from time to time. The fine structure in each profile exemplified in Fig. 1.1 is the result of various thermodynamic processes which cannot be explained here. However, the transition from the troposphere to the stratosphere is one of the most fascinating subtleties in each sounding.

Box 1: Potential Temperature and Brunt-Väisäla Frequency

In addition to the absolute temperature T, the potential temperature Θ is usually used in atmospheric sciences. The absolute temperature T, measured in Kelvin (K), characterizes the conditions at the location where it is measured. In contrast, the potential temperature is not given for the location it is measured but at another, vertically displaced one. Θ is the temperature a parcel of dry air located at pressure p attains if it would be brought adiabatically and reversibly down to standard pressure $p_0 = 1\,000$ hPa. So the potential temperature is defined as

$$\Theta = T \left(p_0/p \right)^{\kappa},$$

where the exponent κ is assumed to be 2/7, the ratio of the gas constant R to the specific heat capacity c_P at constant pressure for an ideal diatomic gas. Normally, the potential temperature Θ increases with altitude. This condition $\partial\Theta/\partial z > 0$ defines stably stratified conditions where a vertically displaced air parcel will be forced to return to its origin. The natural frequency N at which such a displaced air parcel will oscillate around its origin is called the Brunt-Väisälä frequency and is given as

$$N = \sqrt{\frac{g}{\Theta}\frac{\partial\Theta}{\partial z}}.$$

This frequency is not defined in statically unstable air, $\partial\Theta/\partial z < 0$, and it is zero in statically neutral air, $\partial\Theta/\partial z = 0$. The frequency of internal gravity waves propagating through the atmosphere cannot exceed the local Brunt-Väisälä frequency.

1.3 The Tropopause Inversion Layer

As mentioned above, the troposphere and the stratosphere can be distinguished by their different thermal stratifications where the static stability parameters are $N_{TROPO} \approx 0.01\ s^{-1}$ and $N_{STRATO} \approx 0.02\ s^{-1}$, respectively. The *thermal tropopause* is defined as the lowest level at which the lapse rate γ falls below $\gamma_{TP} = 2\ K\ km^{-1}$, if the average of γ between this level and all higher levels within 2 km remains below γ_{TP}. The *dynamical tropopause* utilizes the *Ertel's potential vorticity* (PV) (see Box 2) by determining the lowest level at which PV exceeds a certain threshold, usually ranging between 1.5 and 4 PVU. Another definition of the boundary between troposphere and stratosphere uses thresholds of specific tracer concentrations (*chemical tropopause*), as for example ozone (Gettelman et al. 2011).

Box 2: Vorticity and Potential Vorticity

The vorticity of a fluid particle is determined with respect to an absolute coordinate system. The absolute vorticity vector is defined by $2\Omega + \nabla \times u$, where Ω is the Earth's angular velocity vector and u is the three-dimensional relative velocity vector. The vertical component η of the absolute vorticity vector is given by the sum of the vertical component of the vorticity with respect to Earth (the relative vorticity) ζ and the vorticity of Earth (equal to the Coriolis parameter) f:

$$\eta = \zeta + \mathrm{f}$$

The specific volume $\alpha = 1/\rho$ times the scalar product of the absolute vorticity vector and the gradient of potential temperature defines the potential vorticity:

$$P = \alpha(2\Omega + \nabla \times u)\nabla\Theta.$$

In the absence of friction and heat sources, the potential vorticity P is a materially conservative property, i.e., DP/Dt = 0. The potential vorticity P has the SI unit $m^2\ s^{-1}\ K\ kg^{-1}$. It has become accepted to define $10^{-6}\ m^2\ s^{-1}\ K\ kg^{-1}$ as one potential vorticity unit (1 PVU). For a review on the potential vorticity see Thorpe and Volkert (1997).

In contrast to previous tropopause climatologies, e.g., Hoinka (1998), Birner et al. (2002), and Birner (2006) investigated the stratification and the structure of the wind field of the extratropical tropopause by analyzing multiyear sets of high-vertical-resolution radiosondes.

It was known that on any particular day the tropopause detected in this data set is usually identifiable as a rather distinct discontinuity marking the abrupt transition from γ_{TP} to stratospheric values of $\gamma \approx 0$ or even $\gamma > 0$; see, e.g., the green

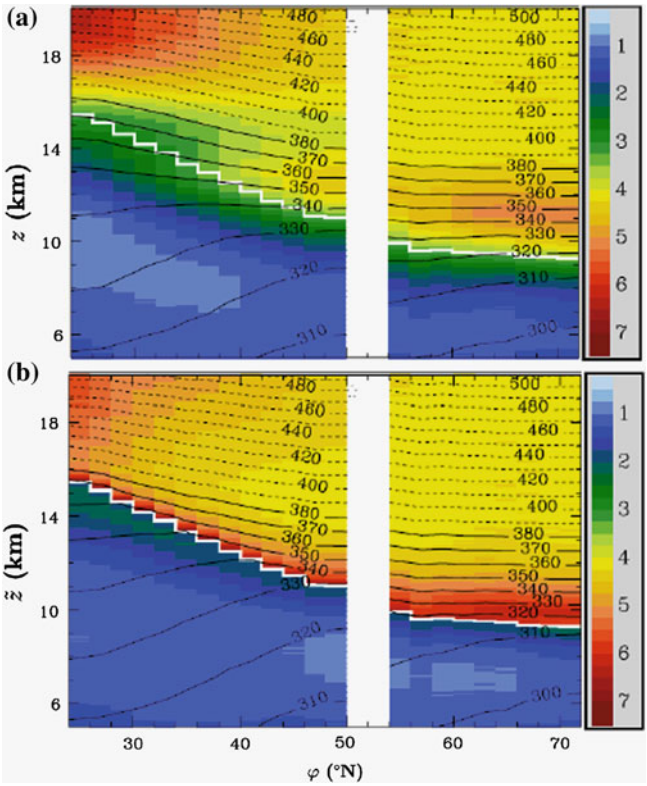

Fig. 1.2 Zonally averaged climatological annual mean values calculated from high-vertical-resolution radiosondes: N^2 (10^{-4} s^{-2}, color shading) and Θ (K, contours) as a function of latitude and altitude. **a** Conventional average with respect to the geometrical altitude z. **b** Tropopause-based average with a modified altitude $\tilde{z} = z - z_{TP} + \bar{z}_{TP}$, where $\bar{z}_{TP}$ is the mean tropopause height over all latitudes. Thick white lines denote the thermal tropopause. White areas represent data gaps. Figure is taken from Birner (2006)

line in Fig. 1.1. However, as the tropopause height rises and sinks discontinuously as a function of time, weather system and season, the mean tropopause is usually located within a comparatively thick transition zone in the lapse rate γ (Fig. 1.2a). For the first time, Birner et al. (2002) calculated temporal averages using the actual tropopause elevation as a common reference level; see Fig. 1.2b. Despite averaging over many years, a surprisingly strong temperature inversion at the tropopause was found. That is, the temperature increases with altitude at the base of the stratosphere and forms a distinct inversion in the mean. Consequently, the static stability parameter N^2 attains maximum values at the tropopause and decays nearly exponentially toward the characteristic stratospheric value N^2_{STRATO}. This *tropopause inversion layer*, nowadays abbreviated TIL, exists on average throughout the extratropics from about 30 to 70°N.

1.4 Atmospheric Circulation Patterns

A description of the vertical structure of the static atmosphere might be written in a couple of paragraphs, as attempted above. However, the atmospheric circulations are too complex to permit more than a brief descriptive overview. The annual mean temperature distribution observed throughout the atmosphere remains nearly unchanged from year to year. At a certain time, however, different regions of the atmosphere receive more heat than others. This heat supply to particular portions of the earth differs on a daily and seasonal cycle. Apparently, the most important difference in net heating and cooling of the earth's surface and atmosphere exists from equator to pole.

A simple tank experiment, here conducted in a virtual laboratory by means of a high-resolution numerical simulation, reflects the latitudinal distribution of heating and cooling and the attending fundamental circulation patterns (Ghil et al. 2010). In this differentially-heated rotating annulus, the polar region is imitated by the inner, cold cylinder whereas the outer cylinder substitutes the effects of the warm equatorial regions at lower latitudes (Fig. 1.3a). The magnitude of the temperature contrast, the aspect ratio of the tank, viscosity, thermal diffusivity, and the angular velocity determine the resulting flow patterns. Assume that the tank is sitting on a turntable at rest. Air on the inner, cold side of the annulus sinks, whereas on the outer side the air is ascending as long as both cylinders are kept at the respective temperatures. Due to the hydrostatic pressure difference between north and south, a *thermally driven*, *direct circulation* develops where warm air flows poleward aloft and cold air with reverse direction at the ground moves in an equal amount equatorward. Except near the lateral boundaries, isothermal surfaces remain horizontal (Fig. 1.3a), a state which is called *barotropic*.

If the turntable is switched on, centrifugal force inhibits the poleward flow of warmer air aloft, whereas the colder air stream near the ground is enhanced. As a result, isothermal surfaces become tilted (Fig. 1.3b), i.e., baroclinicity is amplified. A vertical shear of the zonal wind is generated in response to the *meridional* temperature gradient. The evolving *baroclinic instability* leads to nonaxisymmetric perturbations visible as wave-like structures in the horizontal cross-sections. Figure 1.3c–d show snapshots of selected flow patterns at two successive stages of their evolution. In the atmosphere, these secondary circulations are called *baroclinic Rossby waves* and constitute the familiar low and high pressure systems, often called *cyclones* and *anti-cyclones* when the circulation patterns are closed. Baroclinic waves resulting from this instability have a complicated vertical and horizontal structure and life cycle precluding a detailed description here. Briefly, the job of these large-scale vortices embedded in the wave propagation is to exchange heat and momentum horizontally across the planetary atmosphere and to eliminate the meridional gradients due to the differential radiative heating between the poles and the equator.

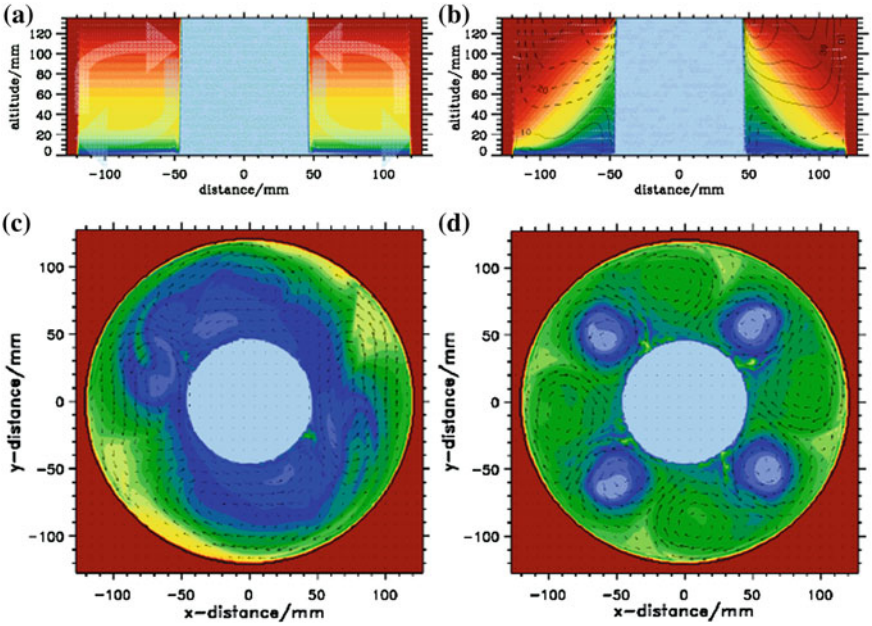

Fig. 1.3 Baroclinic instability in a numerically simulated rotating tank experiment. The inner cold (*light blue*) and outer warm (*dark red*) boundaries of the annulus are represented by immersed boundaries. The temperature contrast between cold and warm boundaries amounts to 6 K; the color shading is linear with increments of 0.2 K. The geometry of the computational domain is adapted from a laboratory experiment (Harlander et al. 2011). The vertical sections in (**a** and **b**) are along the y = 0 line. Panels (**c** and **d**): Formation of trough-ridge systems depicted at two selected stages of their evolution by the temperature and wind fields in the lower part of the rotating annulus

Turning our attention to the real atmosphere, Fig. 1.4 exemplifies the horizontal variability of the atmosphere as expressed by a south–north cross section showing the vertical distribution of the horizontal wind and the potential temperature at winter solstice 2010. This snapshot compliments the same cross section depicting the water vapor concentration; see Fig. 11.2 by Kiemle et al. Both figures, based on operational analyses of the ECMWF,[1] could be regarded as virtual observations as the variables (except PV) are measurable with diverse sensors on the research aircraft deployed by IPA. Currently, such long distances are not accessible by available research aircraft. Yet, with the new HALO aircraft such flights along the meridian from the equatorial regions to the pole are expected to become possible.

It is astonishing to find essential signatures of the direct and secondary circulations in both snapshots shown in Fig. 1.4 below and in Fig. 11.2 by Kiemle et al. The elevated tropopause at low latitudes marks the ascending branch of the *Hadley cell*, a tropical atmospheric circulation driven by the heating near the equator.

[1] European Centre for Medium-Range Weather Forecasts; www.ecmwf.int.

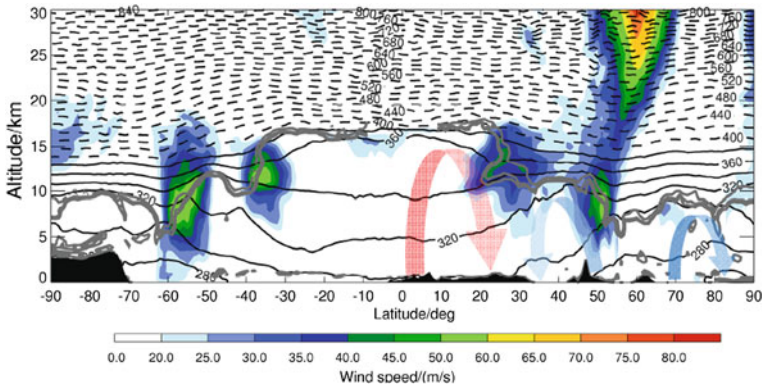

Fig. 1.4 Vertical section of potential temperature Θ (K, *black contour lines*) and the magnitude of horizontal wind V_H (m s^{-1}, color shades) along 11°E from south to north at 12 UTC 21 December 2010. The *thick light gray* bands denote the 1.5 and 2.5 PVU contours marking the dynamical tropopause. The *dark gray* shaded orography is from south to north: Antarctica, Africa, the Alps, and Norway. Locations of the northern hemispheric Hadley, Ferrel and polar circulation cells are indicated by *thick arrows*, respectively, from *left* to *right*

Strong vertical transport of heat, momentum, and constituents occurs in the equatorial low pressure belt of tropical convection (*inter tropical convergence zone*), which explains the observed elevations of the tropical tropopause layer. As a consequence of the upward motion, air must be replenished and streams equatorward at lower levels. Deflection by the Coriolis force forms the well-known *trade winds*. At upper levels, the reverse, poleward flow is found, which subsides over the subtropics.

Indeed, in contrast to the one single, direct thermal circulation cell in the non-rotating annulus, two further cells are assigned for atmospheric circulations. The *polar cell* is also a direct thermal circulation where the upper and low-level flows are reversed. The *Ferrel cell* has a thermally indirect circulation which is induced by the action of the two others. The actual air flow is characterized by that of mid-latitudes baroclinic Rossby waves. In Fig. 1.4, the corresponding latitude belts are reflected by the stepwise equatorward progression of the tropopause. At each of these steps wind maxima mark the *polar* and the *subtropical jet stream*. These locations are also characterized by a sharp horizontal PV gradient on isentropic surfaces ∇_ΘPV: For example, following the 360 K isentropic surface in Fig. 1.4 from the equator poleward, an abrupt PV gradient is seen near ±30°. The above-mentioned baroclinic waves need a restoring force which is—based on theoretical considerations—just proportional to this PV-gradient ∇_ΘPV. Regions of sharp horizontal PV gradients constitute effective guides for *Rossby wave trains* (Dritschel and McIntyre 2008). It is known that upper-level Rossby wave trains are an important component of weather systems in the Atlantic and European sector (Madden 2007) as they are responsible for the sequential initiation of high impact weather events (Glatt et al. 2011).

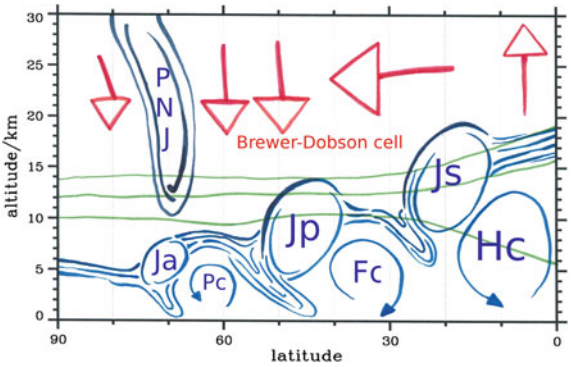

Fig. 1.5 Schematic overview of the wintertime circulation patterns in the troposphere and stratosphere. In the stratosphere, the *residual circulation* of the *Brewer-Dobson cell* is indicated by red arrows. The tropospheric circulation cells are denoted by Hc (Hadley cell), Fc (Ferrel cell), and Pc (polar cell). The locations of the jet streams are marked by PNJ (*polar night jet*). Js (subtropical jet), Jp (polar jet), and Ja (arctic jet). Green lines indicate selected isentropic surfaces

At mid-latitudes, the selected meridional cross-section (Fig. 1.4) shows characteristic dips of the dynamical tropopause. These filamentary tropopause deflections develop in the course of amplifying and breaking upper-level Rossby waves. Usually, these so-called *tropopause folds* are associated with intrusions of dry stratospheric air, see Fig. 11.2 by Kiemle et al., but they are also locations of intense exchange processes between the tropospheric and stratospheric air. Typically, the air in mid-latitudes shows these and other filamentary structures on a broad range of meridional scales from a few hundred meters to several hundred kilometers (Flentje et al. 2005). Large-scale circulations are also assumed to be responsible for forming and maintaining the extratropical TIL, as the detected well-mixed PV suggests (Birner 2006). However, as the involved processes are not fully understood at present, this is still an active area of research (Birner 2010).

The sharpness of the tropopause itself affects the propagation, amplification and breaking of Rossby waves along the *wave guides*. On the other hand, propagation of gravity waves through the tropopause is strongly modified by a sharp tropopause, mainly through the almost step-like behavior of N^2. Not only the last-mentioned dynamical processes are influenced; the distribution of constituents across the tropopause is also visible at the tropopause level, a *barrier* to the transport of trace gases and aerosol particles; see the recent, comprehensive review by Gettelman et al. (2011).

A remarkable discovery related to the troposphere/stratosphere transition and the distribution of constituents in the stratosphere was the explanation of the very dry conditions observed during a series of research flights in the lower stratosphere over southern England in the late 1940s (Tuck et al. 1997). Originally, the proposed hemispheric stratospheric circulation assumed a poleward advection of tropospheric air entering the stratosphere in the tropics, and downward motion

elsewhere (Dobson et al. 1946; Brewer 1949). This circulation cell was named the *Brewer-Dobson circulation*. Later on, this concept was revised.

One might think of the Brewer-Dobson circulation as being formed by radiative cooling of the polar stratosphere near the winter pole. There, constant pressure surfaces are descending and the associated meridional height gradient would produce a poleward force. However, Earth is rotating and due to the Coriolis force an eastward deflection forms the polar night jet encircling the *polar vortex*, a stratospheric cyclone. The mindful reader might already have noticed the signature of this prominent region of high wind exceeding 70 m s^{-1} in the winter stratosphere in Fig. 1.4. Any poleward movement of tropical air parcels along the meridian in the stratosphere would be hindered by such a wind band, usually called a "*transport barrier*". Meridional exchange is possible in periods of unstable polar vortex. Then, *planetary waves* propagate upward from the troposphere, disturb the polar night jet and break in the so-called "*surf zone*" (McIntyre and Palmer 1984). In this way, momentum and constituents can be transported poleward. It is apparent that the Brewer-Dobson circulation as indicated in Fig. 1.5 does not represent the actual physical motion of the air; it is rather a statistical, long-term mean of the eddy fluxes. It is noteworthy that this residual circulation has a long "overturning" period of years in contrast to the aforementioned tropospheric circulations characterized by advection, see Fig. 1.5.

1.5 Regional Atmospheric Circulations

Starting in the 1990s, IPA participated in a series of international airborne *field campaigns* to study arctic ozone chemistry. Not only the gradually elaborating airborne *lidar* instruments allowed new insights in the structure and dynamics of the arctic stratosphere, also the operational availability of numerical prediction models focusing on regional scales enabled meticulous flight planning (Dörnbrack et al. 1997).

One surprising discovery during the exploration of the arctic stratosphere was how regional, also called mesoscale atmospheric phenomena contribute to the formation of *polar stratospheric clouds* (PSCs), see Carslaw et al. (1998). Under certain circumstances, the polar night jet and the polar jet in the troposphere become aligned and the strong flow across the northern mountain barriers of Greenland, Iceland, Svalbard and Scandinavia excites vertical oscillations which are able to propagate upward to stratospheric heights. At these elevations, the adiabatic expansion in the ascending branches of these mountain waves leads to *mesoscale temperature anomalies*, as shown in Fig. 1.6, disturbing the otherwise smooth and balanced stratosphere. There, complicated nucleation mechanisms form PSCs. In recent years, new algorithms to analyze satellite data from CALIPSO have been developed to allow for a more quantitative investigation of the impact of wave ice PSCs on ozone chemistry (Pitts et al. 2011).

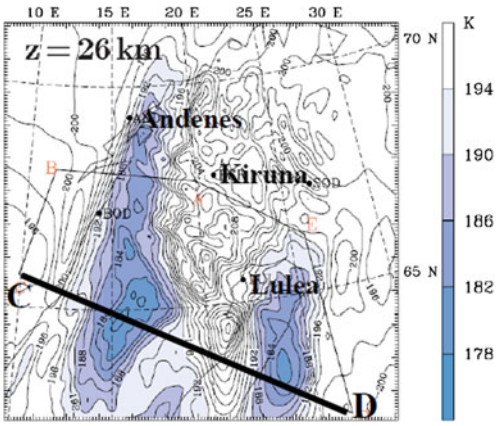

Fig. 1.6 Simulated stratospheric temperature at 26 km altitude at 14 UTC 26 January 2000 using the MM5 (mesoscale model 5) with a horizontal grid size of 8 km. The blue shaded areas highlight adiabatically cooled mesoscale temperature anomalies generated by upward propagating *mountain waves* above the Scandinavian mountains. The flight leg of the DLR Falcon is indicated by a black line, the segment as shown in Fig. 45.2 by Dameris and Loyola is CD. The figure is taken from (Dörnbrack et al. 2002)

The circulations mentioned above occur on a global scale or regionally in the stratosphere. Naturally, meteorological systems occurring in the vicinity of the IPA in southern Germany were also studied comprehensively with ground-based and airborne instruments during recent years. The institute participated in various campaigns, and organized and accomplished a successful series of international field experiments. Due to the proximity of the Alps, a majority of them studied orographically induced flows such as foehn, gravity waves, convection, and the modification of fronts through the Alps.

Certainly, so much more could be written about the subtleties occurring in the atmosphere; the long and the short of it is that it remains a truly fascinating subject of research. During recent years IPA scientists investigated certain phenomena which are embedded in the atmospheric motion spectrum from its vast extension in the horizontal directions to its shallowness. The following chapters of this book give an impression of the broad variety of these studies.

References

Birner, T., Dörnbrack, A., Schumann, U.: How sharp is the tropopause at midlatitudes? Geophys. Res. Lett. **29**(14), 1700 (2002). doi:10.1029/2002GL015142

Birner, T.: Fine-scale structure of the extratropical tropopause region. J. Geophys. Res. **111**, D04104 (2006). doi:10.1029/2005JD006301

Birner, T.: Residual circulation and tropopause structure. J. Atmos. Sci. **67**, 2582–2600 (2010)

Brewer, A.W.: Evidence for a world circulation provided by measurements of the helium and water vapour distribution in the stratosphere. Quart. J. Roy. Meteor. Soc. **75**, 351–363 (1949)

Carslaw, K.S., Wirth, M., Tsias, A., Luo, B.P., Dörnbrack, A., Volkert, H., Leutbecher, M., Renger, W., Bacmeister, J.T., Reimer, E., et al.: Increased stratospheric ozone depletion due to mountain-induced atmospheric waves. Nature **391**, 675–678 (1998)

Dobson, G.M.B., Brewer, A.W., Cwilong, B.M.: Meteorology of the lower stratosphere. Proc. R. Soc. Lond. **A185**, 144–175 (1946)

Dörnbrack, A., Leutbecher, M., Volkert, H., Wirth, M.: Mesoscale forecasts of stratospheric mountain waves. Meteor. Appl. **5**, 117–126 (1997)

Dörnbrack, A., Birner, T., Fix, A., Flentje, H., Meister, A., Schmid, H., Browell, E.V., Mahoney, M.J.: Evidence for inertia gravity waves forming polar stratospheric clouds over Scandinavia. J. Geophys. Res. **107**(20), 8287 (2002). doi:10.1029/2001JD000452

Dritschel, D.G., McIntyre, M.E.: Multiple jets as PV staircases: The Philips effect and the resilience of eddy transport barriers. J. Atmos. Sci. **65**, 855–874 (2008)

Dutton, J.F.: The Ceaseless Wind: An Introduction to the Theory of Atmospheric Motion. Dover Publications, New York (1986)

Flentje, H., Dörnbrack, A., Ehret, G., Fix, A., Kiemle, C., Poberaj, G., Wirth, M.: Water vapor heterogeneity related to tropopause folds over the North Atlantic revealed by airborne water vapor differential absorption lidar. J. Geophys. Res. **110**, D03115 (2005). doi: 10.1029/2004JD004957

Gettelman, A., Hoor, P., Pan, L.L., Randel, W.J., Hegglin, M.I., Birner, T.: The extratropical upper troposphere and lower stratosphere. Rev. Geophys. **49**, RG3003 (2011). doi: 10.1029/2011RG000355

Ghil, M., Read, P., Smith, L.: Geophysical flows as dynamical systems: The influence of Hide's experiments. Astron. Geophys. **51**, 4.28–24.35 (2010). doi:10.1111/j.1468-4004.2010.51428.x

Glatt, I., Dörnbrack, A., Jones, S., Keller, J., Martius, O., Müller, A., Peters, D.H.W., Wirth, V.: Utility of Hovmöller diagrams to diagnose Rossby wave trains. Tellus **63A**, 991–1006 (2011)

Harlander, U., von Larcher, T., Wang, Y., Egbers, C.: PIV- and LDV-measurements of baroclinic wave interactions in a thermally driven rotating annulus. Exp. Fluids **51**, 37–49 (2011). doi: 10.1007/s00348-009-0792-5

Hoinka, K.P.: The tropopause: Discovery, definition and demarcation. Meteor. Z. **6**, 281–303 (1997)

Hoinka, K.P.: Statistics of the global tropopause pressure. Mon. Wea. Rev. **126**(12), 3303–3325 (1998)

Madden, R.A.: Large-scale, free Rossby waves in the atmosphere—an update. Tellus **59A**, 571–590 (2007)

McIntyre, M.E., Palmer, T.N.: The surf zone in the stratosphere. J. Atmos. Terr. Phys. **46**, 825–849 (1984)

Pitts, M.C., Poole, L.R., Dörnbrack, A., Thomason, L.W.: The 2009-2010 Arctic Polar Stratospheric Cloud Season: A CALIPSO Perspective. Atmos. Chem. Phys. **11**, 2161–2177 (2011)

Thorpe, A.J., Volkert, H.: Potential vorticity: A short history of its definitions and uses. Meteor. Z. **6**, 275–280 (1997)

Tuck, A.F.: Atmospheric turbulence: A molecular dynamics perspective. Oxford University Press, New York (2008)

Tuck, A.F., et al.: The Brewer-Dobson circulation in the light of high altitude in situ aircraft observations. Quart. J. Roy. Meteor. Soc. **123**, 1–69 (1997)

Vallis, G.K.: Atmospheric and Oceanic Fluid Dynamics, p. 745. Cambridge University Press, Cambridge (2006)

Chemical Composition of the Atmosphere

2

Hans Schlager, Volker Grewe and Anke Roiger

Abstract

Atmospheric trace gases have an important impact on Earth's radiative budget, the oxidative or cleansing ability of the atmosphere, the formation, growth and properties of aerosols, air quality, and human health. During recent years, the coupling between atmospheric chemistry and climate has received particular attention. Therefore, research is now focused on the composition and processes in the upper troposphere and lower stratosphere, a key region in this respect. In this chapter the chemical composition of the atmosphere is addressed and selected examples of significant advances in this field are presented.

2.1 Introduction

Earth's *atmosphere* is mainly composed of N_2, O_2, Ar, and H_2O (chemical formulas are explained in Table 2.1). In addition, it includes a multitude of trace gases with concentrations ranging from a few μmol/mol (see Box 2.1) to fmol/mol (Fig. 2.1). The *atmospheric lifetime* of the chemical compounds varies from a few seconds to

H. Schlager (✉) · V. Grewe · A. Roiger
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Hans.Schlager@dlr.de

V. Grewe
e-mail: Volker.Grewe@dlr.de

A. Roiger
e-mail: Anke-Elisabeth.Roiger@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_2,

Table 2.1 Molecular formulas and chemical compounds

Formula	Compound	Formula	Compound
Ar	argon	NH_3	ammonia
CH_4	*methane*	N_2O	*dinitrogen oxide*
$(CH_3)_2CO$	*acetone*	NO	*nitric oxide*
ClO	*chlorine monoxide*	NO_2	*nitrogen dioxide*
$ClONO_2$	chlorine nitrate	NO_x	*nitrogen oxides*
CO	*carbon monoxide*	NO_y	*total reactive nitrogen*
CO_2	*carbon dioxide*	O_2	molecular oxygen
DMS	dimethyl sulfide	O_3	*ozone*
HCl	*hydrogen chloride*	OH	*hydroxyl radical*
HNO_2	*nitrous acid*	OCS	carbonyl sulfide
HNO_3	*nitric acid*	PAN	*peroxyacetyl nitrate*
H_2O	*water*	PFC	*perfluorocarbons*
H_2SO_4	*sulfuric acid*	SO_2	*sulfur dioxide*
N_2	molecular nitrogen	VOC	volatile organic compounds

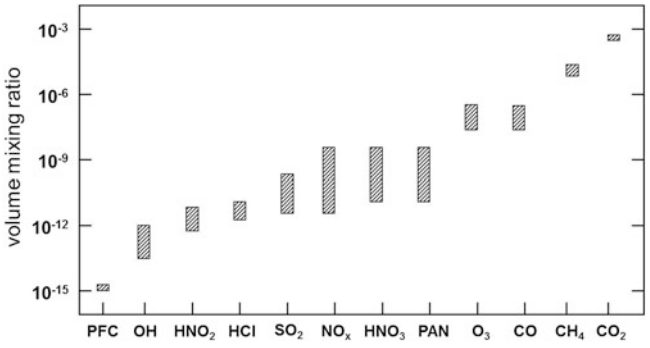

Fig. 2.1 Range of concentrations for atmospheric trace gases. All denoted compounds are measured by the Institute of Atmospheric Physics (IPA) with airborne *in situ* instruments using a variety of techniques including thermo desorption *gas chromatography* (PFC is used for active tracer experiments because of the very low atmospheric concentrations), *chemiluminescence* (NO_x), *chemical ionization mass spectrometry* (OH, SO_2, HCl, HNO_2, HNO_3, PAN), *ultraviolet* (UV) *absorption* (O_3), UV *fluorescence* (CO), and *cavity ring-down spectrometry* (CH_4, CO_2)

more than 1 000 years. Most of the long-lived trace gases, including CO_2, CH_4, N_2O and halocarbons, are important greenhouse gases and have large effects on the *radiation budget* of the atmosphere. Many trace constituents are directly emitted into the atmosphere. Others, most importantly O_3, are chemically or *photochemically*

produced from *precursor* gases. The short-lived chemical species are very reactive and control the so called *oxidation capacity* of the atmosphere, the ability of the atmosphere to oxidize chemical compounds to water-soluble gases which can be scavenged by *precipitation*. The most prominent oxidizing compound is the highly reactive *hydroxyl radical* (OH) with a typical atmospheric lifetime of less than one second, e.g. Brasseur et al. (2003).

Some trace gases also have an indirect impact on *climate* by destroying or generating species which are important for the radiative balance. Examples are *chlorine*- and *bromine*-containing gases causing *ozone depletion*, and aerosol precursor gases, e.g. sulfuric acid (H_2SO_4), which can form new particles in the atmosphere. Furthermore, some primary emitted gases or photochemically produced trace gases are toxic and thus have an impact on *human health*.

The regional and global distributions of trace gases are determined by a variety of factors including the distribution and strength of their *sources* and *sinks*, transport and mixing in the atmosphere, *chemical reactions*, and the interaction with *radiation*, *aerosols* and *clouds*. Major sources *include fossil fuel combustion* and *biomass burning* (important gas emissions: CO_2, SO_2, NO_x, hydrocarbons), gas and oil production (CH_4, VOC), *traffic* (CO_2, NO_x, CO), agriculture (CH_4, N_2O, NH_3), *vegetation* (*volatile* organic carbon compounds), *volcanoes* (SO_2, *halogens*), and *oceans* (N_2O, OCS, DMS, halocarbons). Important sinks of trace gases include oxidation by OH, *photolysis*, and *wet removal*.

Understanding the present and future distribution of trace gases and their impact on climate is one of the major objectives in atmospheric science. Research on the *chemistry* of the atmosphere started at the Institute of Atmospheric Physics (IPA) in the early 1980s with studies on the formation of O_3 from precursor pollutants near the surface during high pressure conditions in summer (Paffrath 1990). At the beginning of the 1990s, the impact of *air traffic* on the *composition* of the atmosphere became a major research topic at IPA. Therefore, experimental investigations and numerical model simulations were focused on the region of the *upper troposphere* and *lower stratosphere* (UTLS), extending from about 8–25 km altitude (depending on latitude). As a consequence, climate-chemistry model developments were enforced (Steil et al. 1998; Hein et al. 2001; Dameris et al. 2005) to better analyze processes in the UTLS. A schematic of processes affecting this region is depicted in Fig. 2.2, including the vertical distribution of the O_3 mixing ratio for September conditions, as simulated with the Chemistry-Climate Model EMAC (Jöckel et al. 2006). Further examples of trace gas distributions are shown in Fig. 2.3.

The upper troposphere and lower stratosphere are dynamically and chemically distinct regions. Characteristics of the lower stratosphere are slow vertical transport and a chemistry initiated by the *photodissociation* of O_2 into atomic oxygen by ultraviolet radiation at *wavelengths* less than 242 nm, e.g. Brasseur and Solomon (1984). The chemistry in the lower stratosphere is also dominated by *catalytic* ozone destruction cycles (Box 2.2). The largest production rates for ozone and reactive nitrogen are found in tropical regions, caused by a maximum in *solar radiation*. However, also ozone loss rates are largest in that region, leading to a very

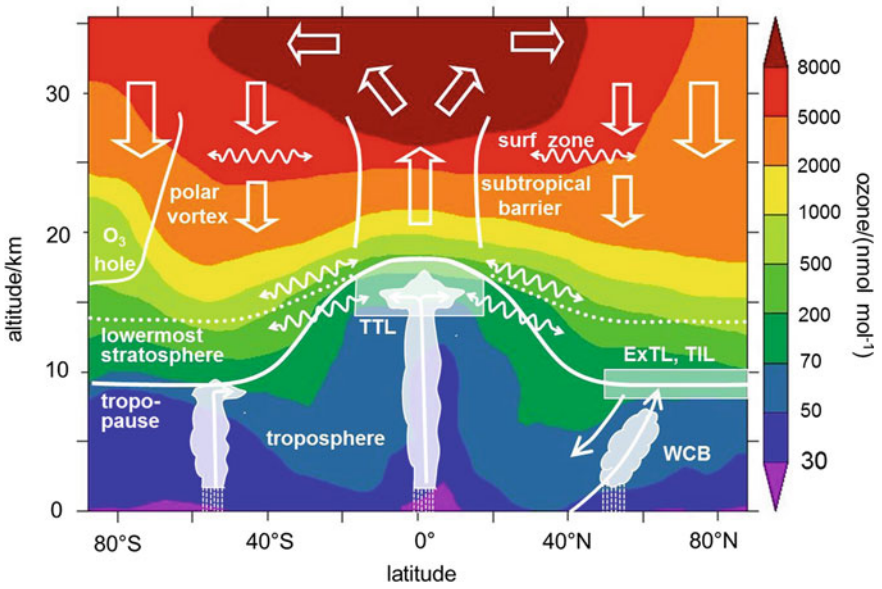

Fig. 2.2 *Zonal* (longitudinal) averaged ozone distribution for September (EMAC simulation) and schematic of distinct regions and important transport processes in the upper troposphere and lower stratosphere (UTLS). The mean *tropopause* (TP) is plotted (solid white curve). The region near the TP represents a transition layer (TTL—*Tropical Transition Layer*, ExTL—*Extratropical Tropopause Layer*) with characteristics intermediate to the UT and LS. The arrows indicate long-term mean transport pathways. In the lower stratosphere, slow downward transport occurs poleward of the subtropical *barrier*, while ascent of air masses occurs in the *tropics*. In the so called *surf-zone*, efficient horizontal exchange of air masses takes place. Fast upward transport of air from the Earth surface is caused by deep *convective clouds* in the tropics and mid-latitudes and by *warm conveyor belts* (WCB) associated with the warm *fronts* of *mid-latitude cyclones*. Quasi-horizontal air mass exchanges occur between the TTL and the LS. The stratosphere in the tropics and subtropics is a main ozone production region. The ozone concentration in the LS is higher in the spring hemisphere (here the southern hemisphere) due to air mass transport by the large-scale *meridional circulation*. However, in the polar vortex ozone is depleted due to *anthropogenic ozone destruction* (*ozone hole* in the *Antarctic* region). The lowest ozone concentrations are found in the *atmospheric boundary layer* in the tropics and mid/high latitudes in the southern hemisphere. Convective vertical transport in these regions also causes low ozone abundances in the UT and TTL

short lifetime of ozone. Detailed modeling studies show that less than 50 % of the ozone found at high latitudes is actually produced at tropical latitudes and transported poleward by the large scale meridional circulation. The larger contribution arises from mid- and high-latitude ozone production (Grewe 2006). *Water vapor* is removed in the upper troposphere by condensation processes in the ascending air masses and at the cold tropopause. The descending air in the polar lower stratosphere stems from the tropical upper stratosphere after a transport time of a few years. Thus, this aged air is rich in ozone and NO_y and poor in water vapor (see Fig. 2.3). In the winter *polar vortex*, ozone is significantly destroyed by chlorine and bromine chemistry in the presence of *polar stratospheric clouds*. Depletion of

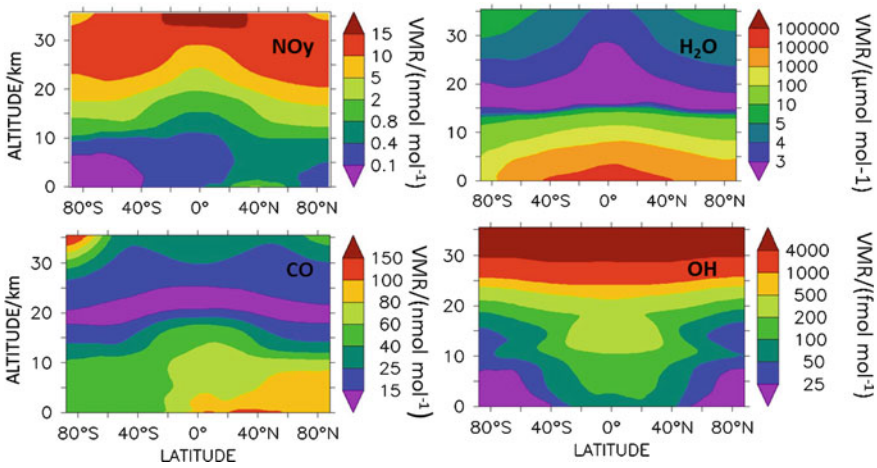

Fig. 2.3 Zonal and annual mean distributions of total reactive nitrogen (NO_y), CO, H_2O, and OH as calculated with the EMAC global chemistry-climate model of IPA. Note the nonlinearly distributed contour intervals

ozone in the lower stratosphere has a significant impact on the surface UV flux. The highest OH concentrations are present in the tropical middle stratosphere. There are also relatively high OH values in the TTL due to convective uplift of OH precursors. H_2O and CO increases in the stratosphere above about 25 km due to their production by the oxidation of CH_4. NO_y increases in the stratosphere because of the reaction of N_2O with excited oxygen atoms, a major source of nitrogen oxides.

In contrast, the upper troposphere is characterized by fast trace gas transport driven by weather systems and convection. The upper troposphere contains significantly lower O_3, higher H_2O and lower reactive nitrogen mixing ratios than the lower stratosphere. A key reaction in the upper troposphere is the photodissociation of NO_2 into NO and O by visible radiation at wavelengths less than 400 nm. The latter immediately recombines with O_2 and forms O_3. In the upper troposphere, atomic oxygen cannot be formed by the photodissociation of O_2 since the required UV radiation is shielded by the stratospheric *ozone layer*. The role of NO_x as catalytic agent in tropospheric ozone formation is described in Box 2.3. During daytime NO and NO_2 are in a chemical steady state involving the formation and destruction of NO_2 by the reaction of NO with O_3 and photolysis, respectively. Formation of NO_2 by reactions of NO with peroxy radicals leads to O_3 production. The O_3 production rate depends on the NO_x concentration in the atmosphere. Photodissociation of ozone in the upper troposphere may lead to hydroxyl radicals (OH), the main oxidizing species in the upper troposphere. They react with most trace gases in the troposphere and are therefore often called cleaning agents.

Instrument platforms used to measure the trace gas composition in the UTLS include *research aircraft*, in-service aircraft, *balloons*, and *satellites* and ground-based stations equipped with *remote sensing* instruments. Airborne measurements

are particularly well suited for studying the UTLS because of the high spatial resolution allowing the observation of small-scale horizontal and vertical structures in the trace gas distributions.

Satellite measurements are an important tool for monitoring the large scale distribution of chemical compounds in the atmosphere and their temporal evolution with global coverage. Important satellite instruments for chemical measurements during recent years were HALOE, GOME, MLS, MIPAS, SCIAMACHY, MOPITT, OMI, and IASI (Burrows et al. 2011). Observations of chemical species from space are discussed elsewhere in this book.

A range of models is used to describe the chemical and transport processes in the UTLS and the future evolution of the chemical composition including trajectory chemistry box models, *Lagrangian* particle dispersion models, *chemistry-transport models*, and *climate-chemistry models*. A recent activity at IPA combines all these model types in a consistent way, enabling an analysis from a box model to a chemistry-climate model using an efficient interface (Jöckel et al. 2006).

Box 2.1: Measure of atmospheric composition

The principle measure of atmospheric composition is the *volume mixing ratio* (also called the mole fraction). It is defined as the number of moles of a trace gas per mole of air (unit mol/mol). It is equivalent to the number density of a trace gas per total number density of air at a given pressure. Volume mixing ratios are commonly given in the following units:

$$\begin{array}{ll} \text{mmol/mol } (10^{-3}) & \text{pmol/mol } (10^{-12}) \\ \mu\text{mol/mol } (10^{-6}) & \text{fmol/mol } (10^{-15}) \\ \text{nmol/mol } (10^{-9}) & \end{array}$$

Box 2.2: Ozone production and destruction in the stratosphere

(a) Pure oxygen chemistry (*Chapman mechanism*):

$$\begin{array}{ll} & \mathrm{O_2 + h\nu \rightarrow O + O \ (\lambda < 242\ nm)} \\ & \mathrm{2\,(O + O_2 + M \rightarrow O_3 + M)} \\ \hline \text{Net:} & \mathrm{3\ O_2 + h\nu \rightarrow 2\ O_3} \text{ (ozone production)} \end{array}$$

$$\begin{array}{ll} & \mathrm{O_3 + h\nu \rightarrow O + O_2 \ (\lambda < 1200\ nm)} \\ & \mathrm{O_3 + O \rightarrow O_2 + O_2} \\ \hline \text{Net:} & \mathrm{2\ O_3 + h\nu \rightarrow 3\ O_2} \text{ (ozone destruction)} \end{array}$$

(hν are photons with a wavelength λ, M denotes a collision partner, e.g. N_2 or O_2)

(b) Gas-phase catalytic ozone destruction

$$\begin{aligned} X + O_3 &\rightarrow OX + O_2 \\ OX + O &\rightarrow X + O_2 \\ \hline \text{Net:} \quad O_3 + O &\rightarrow O_2 + O_2 \text{ (ozone destruction)} \end{aligned}$$

(X denotes a radical, e.g. Cl, Br, NO, OH)

(c) Heterogeneous ozone destruction (example)

$$\begin{aligned} ClONO_2 + HCl\,(s) &\rightarrow HNO_3(s) + Cl_2 \\ Cl_2 + h\nu &\rightarrow Cl + Cl \\ 2(Cl + O_3 &\rightarrow ClO + O_2) \\ ClO + ClO + M &\rightarrow Cl_2O_2 + M \\ Cl_2O_2 + h\nu &\rightarrow Cl + ClO_2 \\ ClO_2 + M &\rightarrow Cl + O2 + M \\ \hline \text{Net:} \quad 2\,O_3 + h\nu &\rightarrow 3\,O_2 \text{ (ozone destruction)} \end{aligned}$$

(s denotes a compound in the solid state in polar stratospheric clouds)

Box 2.3: *Ozone production* and destruction in the troposphere

(a) Photostationary state between O_3, NO and NO_2

$$\begin{aligned} NO + O_3 &\rightarrow NO_2 + O_2 \\ NO_2 + h\nu &\rightarrow NO + O\ (\lambda < 400\text{ nm}) \\ O + O_2 + M &\rightarrow O_3 + M \end{aligned}$$

Net: no change in ozone

(b) *Peroxy-radical*-assisted ozone production

$$\begin{aligned} NO + RO_2 &\rightarrow NO_2 + RO \\ NO_2 + h\nu &\rightarrow NO + O\ (\lambda < 400\text{ nm}) \\ O + O_2 + M &\rightarrow O_3 + M \end{aligned}$$

Net: ozone production

(RO_2 denotes a peroxy radical, e.g. HO_2, CH_3O_2; peroxy radicals are formed by reactions of OH with CO, CH_4 and other hydrocarbons)

(c) Major *ozone destruction* reactions

$$O_3 + h\nu \rightarrow O(^1D) + O_2(\lambda < 310\ \text{nm})$$
$$O(^1D) + H_2O \rightarrow 2\ OH$$

These two reactions are an important sink of ozone and source of highly reactive OH radicals (O^1D denotes an oxygen atom in an excited state)

$$OH + CO \rightarrow H + CO_2$$
$$H + O_2 + M \rightarrow HO_2 + M$$
$$HO_2 + O_3 \rightarrow OH + 2O_2$$

This sequence of reactions constitutes another important sink of ozone in the troposphere.

2.2 Recent Achievements in Understanding the Chemical Composition of the Atmosphere

Important scientific issues relevant for a further understanding of the human impact on the *chemical composition* of the atmosphere include ozone loss in the lower stratosphere, quantification of important sources of reactive nitrogen including *air traffic* and *lightning*, vertical transport of pollution from surface sources by *deep convection* and frontal uplift, *long-range transport*, and exchange of trace gases between the upper troposphere and the lower stratosphere. Selected results on these topics with a focus on IPA investigations are presented in this section for polar, mid-latitude and tropical regions. These results were obtained in a variety of national and international projects which are addressed here and in other chapters of this book (Fig. 2.4). The main tools used by IPA were in situ and *lidar* measurements aboard the DLR research aircraft *Falcon* and the high-altitude research aircraft *Geophysica*, long-term measurements aboard an in-service Airbus 340 of Lufthansa in the frame of the CARIBIC project, and model simulations with microphysical chemical box models and global climate-chemistry models of the ECHAM family.

2.2.1 Advances in Understanding the Composition of the Polar Stratosphere and Troposphere

A recovery of the ozone hole in the winter polar stratosphere is expected for about the year 2050 when the anthropogenic chlorine loading in the stratosphere will have decreased to the same level as in the early 1980s. However, the *recovery* may be delayed because of the impact *of climate change* (Dameris et al. 1998a). These

Fig. 2.4 Aircraft campaigns including chemical measurements performed during the years 2000–2011 with coordination by or participation of IPA. Also included are campaigns foreseen in 2012/2013. The campaign acronyms are explained in Table 2.2

predictions are still highly uncertain because some key processes important for the destruction of ozone are not yet fully understood, in particular for the *Arctic* region. The extent of ozone loss in the Arctic winter polar stratosphere depends, among other factors, on the degree of *denitrification* (removal of reactive nitrogen oxides). In denitrified air masses, ozone-destroying ClO molecules cannot be bound into nonreactive $ClNO_3$ molecules by reaction with NO. It is now generally conceived that the denitrification in the winter Arctic polar stratosphere occurs mainly via a vertical redistribution of HNO_3 by sedimentation of large *nitric acid trihydrate* (NAT: HNO_3. 3 H_2O) particles with diameters in the range 10-20 μm and number densities of 10^{-4}–10^{-3} cm^{-3} (so called *NAT rocks*). These NAT rocks may form at stratospheric temperatures below about 195 K (−78 °C) and distinctly differ from other typical polar stratospheric clouds (PSC) containing much smaller particles (< 6 μm) with a higher density (10^{-2} cm^{-3}). However, until recently there was only one single observation where these large HNO_3 containing particles were detected in the Arctic (Fahey et al. 2001). Also, the formation mechanism for these particles is not yet understood.

The formation and distribution of NAT rocks and the denitrification in the Arctic vortex were key objectives of the projects EUPLEX and RECONCILE in the Arctic winters 2002/2003 and 2009/2010, respectively. Here, IPA performed measurements of NO_y contained in the gas-phase and in *ice particles* aboard the Geophysica high altitude aircraft. During EUPLEX, for the first time NAT rocks could be detected in a very early state of formation at temperatures above the *frost point*. Thus, these particles did not form on ice but potentially on meteoritic *condensation nuclei* (Voigt et al. 2005).

Table 2.2 Airborne campaigns in the years 2000–2013 with chemical measurements involving IPA as coordinator or participant

2000	EXPORT	European Export of Precursors and Ozone by Long-range Transport
	INCA	Interhemispheric Differences in Cirrus Properties from Anthropogenic Emissions
2001	CONTRACE	Convective Transport of Trace Gases into the Middle and Upper Troposphere over Europe: Budget and Impact on Chemistry
	MINOS	Mediterranean Intensive Oxidant Study
2003	EUPLEX	European Polar stratospheric cloud and Lee-wave Experiment
2004	TROCCINOX	Tropical Convection, Cirrus, and Nitrogen Oxides Experiment
	ICARTT	International Consortium for Atmospheric Research on Transport and Transformation
2005	SCOUT-O3	Stratospheric-Climate Links with Emphasis on the UT and LS
2006	AMMA	African Monsoon Multidisciplinary Analysis
	INTEX	The Intercontinental Chemical Transport Experiment
2007	ASTAR	Arctic Study of Aerosol, Clouds and Radiation
	QUANTIFY	Quantifying the Climate Impact of Global and European Transport Systems
2008	POLARCAT	Polar Study using Aircraft, Remote Sensing, Surface Measurements and Models, of Climate, Chemistry, Aerosols, and Transport
	CONCERT	Contrail and Cirrus Experiment
2010	RECONCILE	Reconciliation of essential process parameters for an enhanced predictability of arctic stratospheric ozone loss and its climate interactions
2011	SHIVA	Stratospheric Ozone: Halogen Impacts in a Varying Atmosphere
2012	ACCESS	Arctic Climate Change, Economy and Society
	DC3	Deep Convective Clouds and Chemistry
	TACTS	Transport and Composition in the UT/LMS
	ESMVal	Earth System Model Validation
2013	ML-CIRRUS	Formation, Lifetime, Properties and Radiative/Chemical Impact of Mid-Latitude Cirrus Clouds
	OMO	Oxidation Mechanism Observations

In winter 2010, the Geophysica successfully measured in polar stratospheric clouds. Figure 2.5 shows a photograph taken from the cockpit of the Geophysica during the approach of a PSC on 24 January 2010. The measured NO_y concentrations within the PSC are depicted in Fig. 2.6. NAT rocks were observed embedded in PSC. During ascent and descent of the Geophysica, evaporating NAT rocks could also be detected below the PSC in the renitrification layer.

Fig. 2.5 Polar stratospheric clouds in the lower stratosphere south of Spitzbergen on 24 January 2010 as seen from the Geophysica during a RECONCILE mission flying southbound from Spitzbergen. The PSC layers (*dark lines*) above the bright horizon are marked with an *arrow*

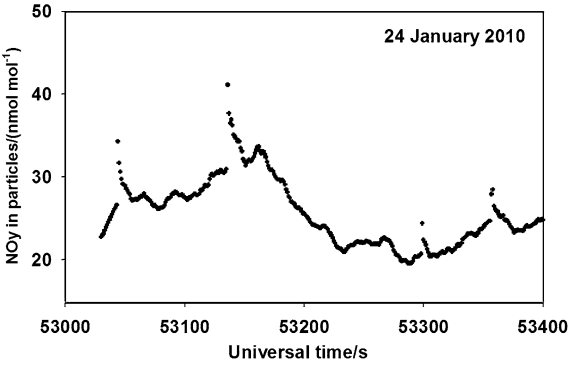

Fig. 2.6 Measurements of total (gas and condensed phase) reactive nitrogen (NO_y) from the Geophysica on 24 January 2010 in a polar stratospheric cloud (PSC). Four NAT rocks were detected during this flight sequence. The background signal is due to NO_y contained in smaller PSC particles with higher number density

Chemical fingerprints of human activities in the Arctic can also be found at lower altitudes. The Arctic troposphere is a receptor of anthropogenic pollution from North America and Eurasia. Import of emissions from boreal forest fires from North America and Siberia significantly increased during the last decade because of an increase in the frequency and area of the *fires*. Deposition of imported *black carbon* decreases the snow *albedo* and accelerates warming in the Arctic. This may cause unique regional responses including melting of ice sheets and *permafrost*. Import of pollution into the Arctic was studied during the ASTAR and POLARCAT campaigns in Spitzbergen and Greenland, respectively. The chemical composition in the Arctic during spring and summer was found to be significantly disturbed at altitudes from the lower troposphere to the tropopause region by emissions from *forest fires* in Canada and Siberia, and to a lesser extent by anthropogenic pollution originating in North America and Asia (Roiger 2011). A multitude of individual *pollution plumes* could be detected in the European sector of the Arctic during POLARCAT, as well as significantly enhanced background profiles of the pollution tracers PAN and CO (Fig. 2.7). Also, the import of East Asian pollution into the lower stratosphere could be detected for the first time (Roiger et al. 2011).

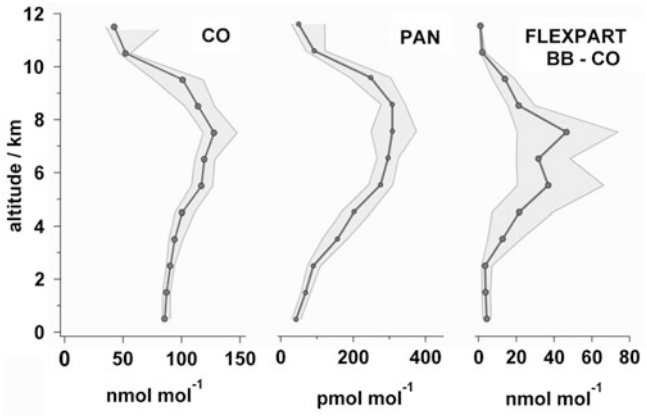

Fig. 2.7 Mean profiles of CO and PAN volume mixing ratios averaged over all Falcon flights in the European sector of the Arctic in summer 2008 during the POLARCAT campaign, including standard deviations (*shaded area*). Also shown are simulations with the FLEXPART model of a biomass burning (BB) CO tracer interpolated onto the Falcon flight tracks. The FLEXPART results indicate that the observed enhanced CO and PAN mixing ratios at altitudes between 4 and 10 km can be explained by the import of biomass burning emissions from Canada and Siberia (Roiger 2011)

2.2.2 Advances in Understanding the Composition at Mid-Latitudes

In order to understand the anthropogenic influence on the composition and chemistry of the UTLS over Europe, the impact of surface emissions from Europe and upstream continents, air traffic and lightning need to be quantified. The uplift of emissions from the boundary layer in Europe and Northern America to the upper troposphere by warm conveyor belts and convection was studied in a number of projects including CONTRACE, EXPORT, and ICARTT. Figure 2.8 shows an example of elevated *pollution layers* of different origin observed during EXPORT in the middle and upper troposphere over Germany (Schlager et al. 2003).

Depicted is a vertical profile of CO, a tracer of polluted boundary layer air, measured over Hanover. Three distinct pollution layers were observed at altitudes between 4 and 8 km originating from different European sources and uplifted either by local *thunderstorms* in the Hanover area or by a warm front between Geneva and Hanover. The CO mixing ratios measured in the outflow layers are very similar to the mixing ratios observed in the boundary layer near Hanover, indicating little dilution during the fast vertical transport. The pollution layer observed in the upper troposphere between 9.5 and 10.5 km was imported from the USA after being uplifted by a warm conveyor belt over the western North Atlantic.

The impact of *aircraft emissions* on the composition and chemistry of the UTLS was studied during the projects "*Schadstoffe in der Luftfahrt*" ("Pollutants from Air Traffic") and POLINAT (Schumann et al. 2000). Emissions from air traffic are concentrated in major *flight corridors* over Europe, North America, and

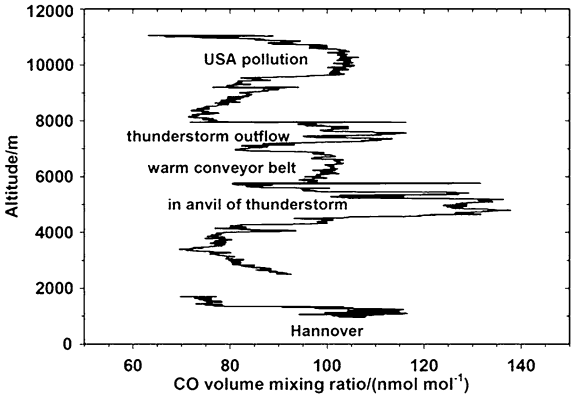

Fig. 2.8 Elevated pollution layers in the troposphere of various origins, measured with the DLR Falcon during EXPORT. Depicted is the volume mixing ratio of CO, a tracer of pollution from combustion sources at the surface (Schlager 2003)

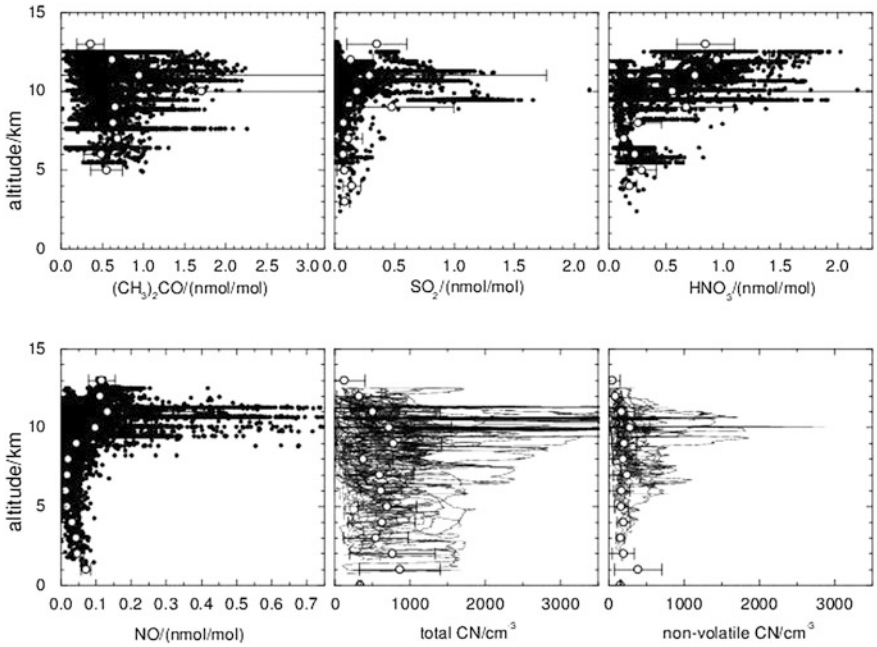

Fig. 2.9 Profiles of chemical species and condensation nuclei particles measured in the UTLS over the North Atlantic (composites of all measurements during the POLINAT Falcon campaigns in 1994, 1995 and 1997). Shown are individual measurements and mean values averaged over 1 km intervals (*white dots*) with standards deviations (Schumann et al. 2000)

the North Atlantic at altitudes between 9 and 12 km. A series of Falcon campaigns was performed in the eastern North Atlantic flight corridor. The objective of these

Fig. 2.10 An intense thunderstorm system (known as Hector) over the Tiwi Islands north of Darwin, Australia (picture taken from the Falcon during SCOUT-O3)

measurements was to determine the distribution of nitrogen oxides, ozone and tracer species for analysis of the air mass origin (Fig. 2.9). Distinct enhancements of primary and secondary gas emissions from air traffic (NO, NO_y, HNO_3, SO_2) and condensation nuclei (CN) particles could be detected in the flight corridor (Schlager et al. 1997; Schlager et al. 1999; Ziereis et al. 1999). Accompanying model simulations revealed that up to 60 % of the NO_x abundance in the upper troposphere is caused by aircraft emissions, depending on the time of the year and the meteorological situation (Sausen and Köhler 1994; Köhler et al. 1997; Dameris et al. 1998b; Grewe et al. 1999; Meijer et al. 2000; Schumann et al. 2000; Grewe et al. 2002; Grewe et al. 2001).

2.2.3 Advances in the Understanding of the Composition in the Tropics

The source of nitrogen oxides produced by lightning in the tropics represents an important uncertainty in the global *budget* of *reactive nitrogen*. According to model simulations, *lightning-induced nitrogen oxides* contribute about 70 % to the NO_x abundance in the tropical upper troposphere and is responsible for about 30 % of the ozone in this region (Grewe 2004). In a series of aircraft campaigns in the tropics (TROCCINOX, SCOUT-O3, AMMA) IPA investigated the NO_x production by *electrified* thunderstorms (Huntrieser et al. 2007; Schumann and Huntrieser 2007; Huntrieser et al. 2008; Huntrieser et al. 2009; Huntrieser et al. 2011). Measurements of nitrogen oxides and related chemical compounds were performed in anvils and in the vicinity of strong deep *convective systems*, as for

Fig. 2.11 Research aircraft deployed during the second TROCCINOX campaign in Aracatuba, Brazil in 2005 (*left* DLR Falcon, *center* Geophysica of Myasishchev Design Bureau, Russia, *right* Bandeirante of Brazilian partners) and the campaign team

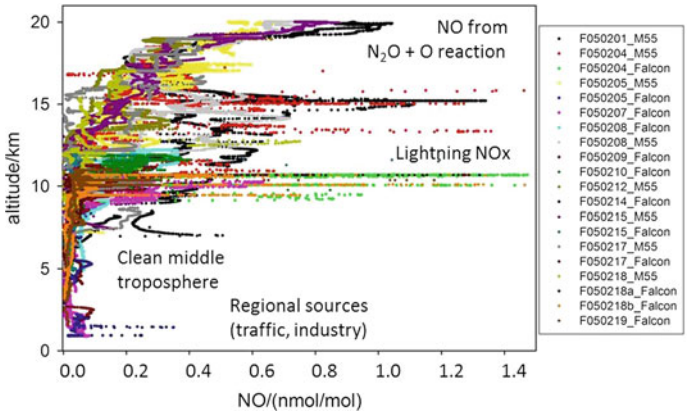

Fig. 2.12 Composite of all NO measurements aboard the Falcon and Geophysica M55 during the TROCCINOX campaign in 2005

example the *Hector* thundercloud (Fig. 2.10). A combination of two or three aircraft with different maximum cruising altitudes were deployed during the tropical campaigns to cover the inflow and outflow regions as well as the top of the thunderclouds (Fig. 2.11).

A composite of all nitric oxide profiles measured with the Falcon and Geophysica aircraft during the TROCCINOX campaign in Brazil in 2005 is depicted in Fig. 2.12. Distinct signatures of fresh lightning-produced NO_x were observed in the outflow region of deep convection at altitudes between 9 and 16 km. From these kinds of measurements and accompanying observations of lightning by ground- and space-based systems, a best estimate of the global

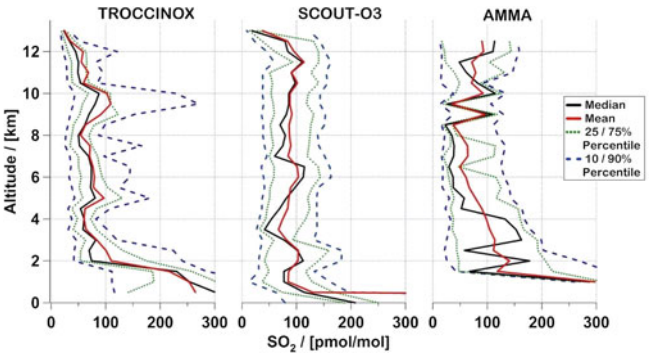

Fig. 2.13 Median (*black*) and mean (*red*) SO_2 volume mixing ratios from Falcon campaigns in the tropics. Also included are 10, 25, 75 and 90 % percentiles (Fiedler 2007)

nitrogen mass source rate of 5 ± 3 Tg per year could be determined (Schumann and Huntrieser 2007). These measurements were also used to evaluate the simulation of lightning NO_x in climate-chemistry simulations (Grewe 2009).

Another important *trace gas*, which has adverse effects on climate, is sulfur dioxide, e.g. Seinfeld and Pandis (1998). SO_2 may undergo conversion to sulfuric acid, a major aerosol precursor. However, observations of SO_2 distribution in the UTLS are very sparse. During recent years IPA performed extensive high-quality SO_2 measurements covering polar, mid-latitude and tropical regions (Fiedler 2007; Fiedler et al. 2009a, b) using chemical ionization mass spectrometry with on-line calibration employing isotopically labeled SO_2 (Speidel et al. 2007). Figure 2.13 shows mean and median SO_2 profiles measured during TROCCINOX, SCOUT-O3, and AMMA. The median mixing ratios are mostly constant in the *free troposphere* at the different measurement sites; however, the values observed in the tropics (70–90 nmol/mol) are interestingly higher than in Europe (30–40 nmol/mol). This indicates more effective vertical transport of SO_2 in the tropics by more frequent convection. These SO_2 measurements represent the most comprehensive SO_2 data set for the UTLS presently available.

2.3 Outlook

With the extended capabilities of the new research aircraft *HALO* in terms of payload and range, new opportunities will arise for airborne field studies on the composition of the UTLS. Chemical and *particle measurement systems* can be combined in one payload in a more comprehensive way than possible so far. Thereby, heterogeneous and multiphase processes involving ice and liquid particles that have the potential to influence the composition of the atmosphere can be investigated in future missions. The impact of large scale dynamic atmospheric features like the Asian monsoon can be studied during long-range flights. The air

composition in the Asian *monsoon* is affected by emissions from India and eastern China, regions with largely enhanced and increasing levels of anthropogenic pollutants. The transport of these emissions into the lower stratosphere may impact its composition on a global scale. Another important topic is understanding the ozone trends in the upper troposphere and, in this respect, the impact of NO_x caused by *aviation*.

The present highly coupled 3D chemistry-climate models enable global simulations over decades at a resolution of roughly 300 km. Many processes are interactively coupled, like chemistry with radiation, clouds and transport. This complexity of the climate-chemistry model systems often hinders an interpretation of results. Future model developments will have to include specific diagnostics, enabling an analysis of causes and effects (Grewe et al. 2010; Jöckel et al. 2010; Garny et al. 2011). In addition, improvements are required, including a better representation of physical processes and the inclusion of other domains, such as the *biosphere*.

References

Brasseur, G., Solomon, S.: Aeronomy of the middle atmosphere. D. Reidel Publishing Company, Dordrecht (1984)

Brasseur, G.P., Prinn, R.G., Pszenney, A.P. (eds.): Atmospheric chemistry in a changing world. Springer, Berlin (2003)

Burrows, J.P., Platt, U., Borrell, P. (eds.): The remote sensing of tropospheric composition from space. Springer, Berlin (2011)

Dameris, M., Grewe, V., Hein, R., Schnadt, C., Brühl, C., Steil, B.: Assessment of the future development of the ozone layer. Geophys. Res. Lett. **25**, 3579–3582 (1998a)

Dameris, M., Grewe, V., Köhler, I., Sausen, R., Brühl, C., Grooß, J.-U., Steil, B.: Impact of aircraft NO_x-emissions on tropospheric and stratospheric ozone. Part II: 3-D model results. Atmos. Env. **32**, 3185–3200 (1998b)

Dameris, M., Grewe, V., Ponater, M., Deckert, R., Eyring, V., Mager, F., Matthes, S., Schnadt, C., Stenke, A., Steil, B., et al.: Long-term changes and variability in a transient simulation with a chemistry-climate model employing realistic forcing. ACP **5**, 2121–2145 (2005)

Fahey, D.W., Gao, R.S., Carslaw, K.S., et al.: The detection of large HNO_3-containing particles in the winter Arctic stratosphere. Science **291**, 1026–1031 (2001)

Fiedler, V.: Atmospheric SO_2: global measurements using aircraft-based CIMS. Dissertation, DLR-Forschungsbericht, Ruprecht-Karls Universität Heidelberg, pp. 1434–8454 (2007)

Fiedler, V., Arnold, F., Schlager, H., Dörnbrack, A., Pirjola, L., Stohl, A.: East Asian SO_2 pollution plume over Europe—Part 1: evolution and potential impact. Atmos. Chem. Phys. **9**, 4729–4745 (2009a). doi:10.5194/acp-9-4729-200

Fiedler, V., Nau, R., Ludmann, S., Arnold, F., Schlager, H., Stohl, A.: East Asian SO_2 pollution plume over Europe—Part 2: Airborne trace gas measurements and source identification by particle dispersion model simulations. Atmos. Chem. Phys. **9**, 4717–4728 (2009b). doi:10.5194/acp-9-4717-2009

Garny, H., Grewe, V., Dameris, M., Bodeker, G.E., Stenke, A.: Attribution of ozone changes to dynamical and chemical processes in CCMs and CTMs. Geosci. Model Dev. **4**, 271–286 (2011). doi:10.5194/gmd-4-271-2011

Grewe, V.: The origin of ozone. Atmos. Chem. Phys. **6**, 1495–1511 (2006). doi:10.5194/acp-6-1495-2006

Grewe, V., Dameris, M., Hein, R., Köhler, I., Sausen, R.: Impact of future subsonic aircraft NO_x emissions on the atmospheric composition. Geophys. Res. Lett. **26**, 47–50 (1999). doi: 10.1029/1998GL900249

Grewe, V., Dameris, M., Fichter, C., Sausen, R.: Impact of aircraft NO_x emissions. Part 1: Interactively coupled climate-chemistry simulations and sensitivities to climate-chemistry feedback, lightning and model resolution. Meteor. Z. **11**, 177–186 (2002)

Grewe, V.: Technical note: A diagnostic for ozone contributions of various NO_x emissions in multi-decadal chemistry-climate model simulations. Atmos. Chem. Phys. **4**, 327–342 (2004)

Grewe, V.: Impact of lightning on air chemistry and climate. In: Betz, H.D., Schumann, U., Laroche, P. (eds.) Lightning: principles, instruments and applications review of modern lightning research, pp. 524–551. Springer, Dordrecht (2009)

Grewe, V., Tsati, E., Hoor, P.: On the attribution of contributions of atmospheric trace gases to emissions in atmospheric model applications. Geosci. Model Dev. **3**, 487–499 (2010)

Grewe, V., Brunner, D., Dameris, M., Grenfell, J.L., Hein, R., Shindell, D., Staehelin, J.: Origin and variability of upper tropospheric nitrogen oxides and ozone at northern mid-latitudes. Atmos. Env. **35**, 3421–3433 (2001). doi:10.1016/S1352-2310(01)00134-0

Hein, R., Dameris, M., Schnadt, C., Land, C., Grewe, V., Köhler, I., Ponater, M., Sausen, R., Steil, B., Landgraf, J., et al.: Results of an interactively coupled atmospheric chemistry-general circulation model: Comparison with observations. Ann. Geophys. **19**, 435–457 (2001)

Huntrieser, H., Schlager, H., Roiger, A., Lichtenstern, M., Schumann, U., Kurz, C., Brunner, D., Schwierz, C., Richter, A., Stohl, A.: Lightning-produced NO_x over Brazil during TROCCINOX: airborne measurements in tropical and subtropical thunderstorms and the importance of mesoscale convective systems. Atmos. Chem. Phys. **7**, 2987–3013 (2007). doi:10.5194/acp-7-2987-2007

Huntrieser, H., Schumann, U., Schlager, H., Höller, H., Giez, A., Betz, H.-D., Brunner, D., Forster, C., Pinto Jr, O., Calheiros, R.: Lightning activity in Brazilian thunderstorms during TROCCINOX: implications for NO_x production. Atmos. Chem. Phys. **8**, 921–953 (2008). doi: 10.5194/acp-8-921-2008

Huntrieser, H., Schlager, H., Lichtenstern, M., Roiger, A., Stock, P., Minikin, A., Höller, H., Schmidt, K., Betz, H.-D., Allen, G., et al.: NO_x production by lightning in Hector: first airborne measurements during SCOUT-O_3/ACTIVE. Atmos. Chem. Phys. **9**, 8377–8412 (2009). doi:10.5194/acp-9-8377-2009

Huntrieser, H., Schlager, H., Lichtenstern, M., Stock, P., Hamburger, T., Höller, H., Schmidt, K., Betz, H.-D., Ulanovsky, A., Ravegnani, F.: Mesoscale convective systems observed during AMMA and their impact on the NO_x and O_3 budget over West Africa. Atmos. Chem. Phys. **11**, 2503–2536 (2011). doi:10.5194/acp-11-2503-2011

Jöckel, P., Tost, H., Pozzer, A., Brühl, C., Buchholz, J., Ganzeveld, L., Hoor, P., Kerkweg, A., Lawrence, M., Sander, R., et al.: The atmospheric chemistry general circulation model ECHAM5/MESSy1: consistent simulation of ozone from the surface to the mesosphere. Atmos. Chem. Phys. **6**, 5067–5104 (2006). doi:10.5194/acp-6-5067-2006

Jöckel, P., Kerkweg, A., Pozzer, A., Sander, R., Tost, H., Riede, H., Baumgaertner, A., Gromov, S., Kern, B.: Development cycle 2 of the modular earth submodel system (MESSy2). Geosci. Model Dev. **3**, 717–752 (2010). doi:10.5194/gmd-3-717-2010

Köhler, I., Sausen, R., Reinberger, R.: Contributions of aircraft emissions to the atmospheric NO_x content. Atmos. Environ. **31**, 1801–1818 (1997)

Meijer, E.W., van Velthoven, P.F.J., Thompson, A.M., Pfister, L., Schlager, H., Schulte, P., Kelder, H.: Model calculations of the impact of NO_x from air traffic, lightning, and surface emissions, compared with measurements. J. Geophys. Res. **105**, 3833–3850 (2000). doi: 10.1029/1999JD901052

Paffrath, D.: Airborne chemistry measurements. In: Proceedings of Remote Sensing and The Earth's Environment, ESA, Alpbach Summer School, Alpbach, Austria, 26 July–4 August 1989, pp.43–49 (1990)

Roiger, A.: Biomass burning pollution in the summer time Arctic atmosphere: development and deployment of a novel CI-ITMS instrument for PAN detection. Dissertation, Universität Heidelberg (2011)

Roiger, A., Schlager, H., Schaefler, A., Huntrieser, H., Scheibe, M., Aufmhoff, H., Cooper, O.R., Sodemann, H., Stohl, A., Burkhart, J.: In situ observation of Asian pollution transported into the Arctic lowermost stratosphere. Atmos. Chem. Phys. **11**, 10975–10994 (2011)

Sausen, R., Köhler, I.: Simulating the global transport of nitrogen oxides emissions from aircraft. Ann. Geophys. **12**, 394–402 (1994)

Schlager, H., Konopka, P., Schulte, P., Schumann, U., Ziereis, H., Arnold, F., Klemm, M., Hagen, D., Whitefield, P., Ovarlez, J.: In situ observations of air traffic emission signatures in the North Atlantic flight corridor. J. Geophys. Res. **102**, 10739–10750 (1997)

Schlager, H., Schulte, P., Flatoy, F., Slemr, F., van Velthoven, P., Ziereis, H., Schumann, U.: Regional nitric oxide enhancements in the North Atlantic flight corridor observed and modeled during POLINAT 2—a case study. Geophys. Res. Lett. **26**, 3061–3064 (1999). doi:10.1029/1999GL900349

Schlager, H., Huntrieser, H., Heland, J., Ziereis,H.: Analysis of convective and long-range transport of reactive nitrogen species over Europe using aircraft observations. EXPORT-E2 Final Report, EUROTRAC-2 ISS, Penkett, S., Law, K.S., Platt, U., Volz-Thomas, A. (eds.) GSF—National Research Center for Environment and Health, Munich, Germany, pp.151–155 (2003)

Schumann, U., Schlager, H., Arnold, F., Ovarlez, J., Kelder, H., Hov, Ø., Hayman, G., Isaksen, I.S.A., Staehelin, J., Whitefield, P.D.: Pollution from aircraft emissions in the North Atlantic flight corridor: Overview on the POLINAT projects. J. Geophys. Res. **105**, 3605–3631 (2000). doi:10.1029/1999JD900941

Schumann, U., Huntrieser, H.: The global lightning-induced nitrogen oxides source. Atmos. Chem. Phys. **7**, 3823–3907 (2007). doi:10.5194/acp-7-3823-2007

Seinfeld, J.H., Pandis, S.N.: Atmospheric chemistry and physics. Wiley, New York (1998)

Speidel, M., Nau, R., Arnold, F., Schlager, H., Stohl, A.: Sulfur dioxide measurements in the lower, middle and upper troposphere: Deployment of an aircraft-based chemical ionization mass spectrometer with permanent in-flight calibration. Atmos. Env. **41**, 2427–2437 (2007). doi:10.1016/j.atmosenv.2006.07.047

Steil, B., Dameris, M., Brühl, C., Crutzen, P.J., Grewe, V., Ponater, M., Sausen, R.: Development of a chemistry module for GCMs: first results of a multiannual integration. Ann. Geophys. **16**, 205–228 (1998)

Voigt, C., Schlager, H., Luo, B.P., Dörnbrack, A., Roiger, A., Stock, P., Curtius, J., Vössing, H., Borrmann, S., Davies, S., et al.: Nitric Acid Trihydrate (NAT) Formation at Low NAT Supersaturation in Polar Stratospheric Clouds (PSCs). Atmos. Chem. Phys. **5**, 1371–1380 (2005). doi:10.5194/acp-10-9039-2010

Ziereis, H., Schlager, H., Schulte, P., Köhler, I., Marquardt, R., Feigl, C.: In situ measurements of the NO_x distribution and variability over the eastern North Atlantic. J. Geophys. Res. **104**, 16021–16032 (1999). doi:10.1029/1999JD900175

Aerosols in the Atmosphere

3

Andreas Petzold and Bernd Kärcher

Abstract

This overview introduces atmospheric aerosol, including basic definitions and properties. Types of aerosols, sources, transformation processes and sinks are described. The spatial distribution of atmospheric aerosol is discussed on the basis of field observations conducted at the DLR Institute of Atmospheric Physics. The chapter concludes with a brief overview of aerosol impacts on the global climate.

3.1 Introduction

Airborne particles, or *aerosols*, form an intrinsic constituent of the global *atmosphere*. They affect *global climate* via their interaction with sunlight, influence the formation of clouds, and participate in many atmospheric *chemical reactions*. Figure 3.1 provides an overview on the various processes and interactions. These particles can travel over distances of several thousands of kilometers and transport particulate material such as dust, sea salt, spores, pollen and bacteria. Most of us think of atmospheric aerosols in terms of fine dust, *soot*, *air pollution*, reduced

A. Petzold (✉) · B. Kärcher
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: a.petzold@fz-juelich.de

B. Kärcher
e-mail: Bernd.Kaercher@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_3,

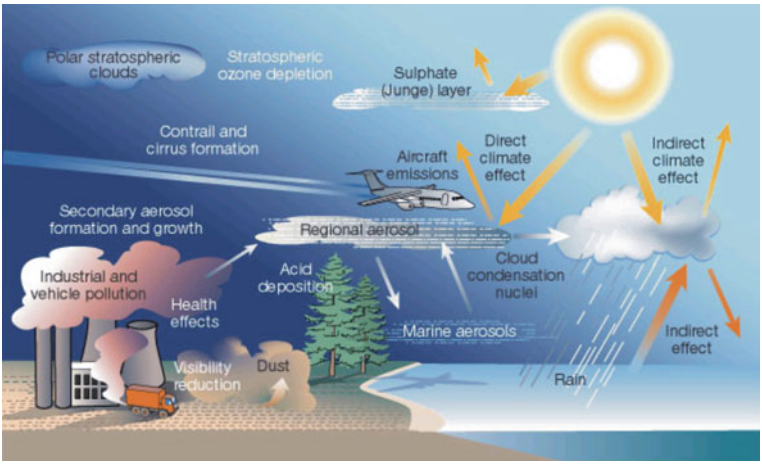

Fig. 3.1 Major sources of atmospheric aerosols and their impacts on the global climate and air quality; reprinted by permission from Macmillan Publishers Ltd:[NATURE] (Kolb 2002), copyright (2002)

visibility, and adverse *health* effect. However, spectacular atmospheric effects such as colored sunsets are also associated with aerosol particles; oceanic algae are fertilized by dust-bound iron transported from the great deserts of Earth, and the polar *stratospheric ozone holes* are also triggered by aerosol particles.

The ocean is one major *source* of natural aerosols with an estimated annual emission of 1 000 to 10 000 Tg (= 10^{12} g = 1 megaton) of material (Raes et al. 2000; Seinfeld and Pandis 2006). The exchange of particulate matter between air and sea contributes to the global cycles of carbon, nitrogen, and sulfur particles such as dimethyl sulfide (DMS), produced by phytoplankton. Ocean water and sea salt are transferred to the atmosphere through sea spray and air bubbles at the sea surface. As a droplet of water evaporates, the salt is left suspended in the atmosphere, forming maritime aerosols, or sea salt. Another major natural source of particulate mass is windblown *mineral dust* from dry continental regions like deserts or semi-arid areas. Estimates of the global mass of dust particles released per year into the atmosphere vary between 1 000 and 5 000 Tg (Raes et al. 2000). Other significant sources of airborne particles are *biomass burning* from forest or savannah *fires*, *volcanic eruptions*, and *anthropogenic pollution* from industrial *emissions* and *traffic*.

Corresponding to the major sources, the global distribution of airborne particles in terms of mass has its hot spots over or near the world's large deserts, along the storm tracks of the southern hemispheric oceans south of the Americas and Africa, over tropical rainforests and over the highly industrialized regions of North America, Europe and South East Asia. Figure 3.2 shows the global distribution of *aerosol optical depth* AOD (see below for a definition) as a measure of the atmospheric aerosol load, as recorded by the NASA *satellite* instrument MODIS (Moderate Resolution Imaging Spectroradiometer, see http://modis.gsfc.nasa.gov/).

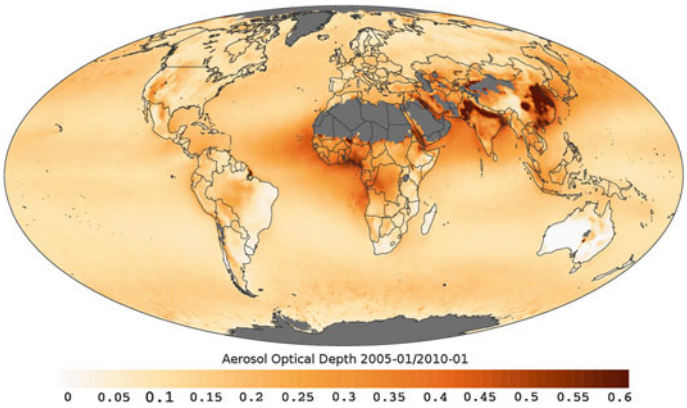

Fig. 3.2 Global distribution of aerosol optical depth (AOD) for the years 2005 to 2010 taken from the MODIS satellite instrument; since the MODIS instrument cannot resolve AOD over highly reflecting surfaces like deserts or snow, no MODIS data are available for these regions. Adapted from http://en.wikipedia.org/wiki/File:Modis_aerosol_optical_depth.png

3.2 Definitions and Properties

Atmospheric aerosols are defined as a suspension of fine particles consisting of solid or liquid compounds of various origins in air. Water-based airborne particles like cloud drops or *ice crystals*, however, are not considered as aerosols but as *hydrometeors*. While gaseous atmospheric constituents are usually characterized by their *mixing ratio* compared to air, a description of aerosols requires the consideration of a set of physical and chemical properties.

One key property is the size of the particles, given as particle diameter D or as typical dimension in the case of irregularly shaped particles. Particle size spans over multiple orders of magnitude from a few nanometers (1 nm $= 10^{-9}$ m) for particles freshly produced from gaseous *precursors* by *gas-to-particle conversion* up to almost 1 mm for large dust particles. For comparison, a typical "air" molecule has a diameter of $\sim$0.4 nm. Atmospheric aerosol always contains particles of different sizes and is therefore classified as polydisperse. In contrast, an aerosol composed of particles of a single size would be classified as monodisperse. Generally, the distribution of particle sizes present in an aerosol is given as the probability of occurrence of particles with diameters at certain size intervals. This property is called the particle size distribution function.

In general terms, key properties required for a full characterization of an aerosol are:

- *particle size distribution*
- total number of particles per reference volume (*number concentration N* given in cm^{-3})
- total mass of suspended *particulate matter* per reference volume (mass concentration M given in $\mu\mathrm{g\ m}^{-3}$ or $\mathrm{ng\ m}^{-3}$

- particle density, ranging from 1.0 g cm^{-3} for organic aerosol to 2.6 g cm^{-3} for mineral dust
- particle *shape* (spherical, irregularly shaped, agglomerates)
- particle *phase* (liquid, solid)
- chemical *composition*

Clean continental air often contains less than a few thousand particles per cubic centimeter (cm^{-3}), while polluted continental air may reach concentration values as high as 50 000 cm^{-3}, and in urban air, particle number concentrations may exceed 10^5 cm^{-3}. Natural sources of *combustion* particles like forest fires may also exceed particle number concentrations of 10^5 cm^{-3} (Andreae and Merlet 2001). Desert air has about 1 000 cm^{-3} on average. Clean marine air generally has a few 100 s cm^{-3} but may reach values well above 10 000 cm^{-3} in cases of new particle formation events. These are all near-surface concentration values. In the free *troposphere* a typical background number concentration is 300 cm^{-3}. Polar aerosols usually do not exceed number concentrations of 200 cm^{-3} on average above 1 km altitude, strongly decreasing to values of several 10 cm^{-3} near the surface (Seinfeld and Pandis 2006). The background *stratospheric aerosol* layer is characterized by a typical number concentration and mean diameter of accumulation mode particles of 1 000 cm^{-3} and 0.14 μm, respectively. In honor of its first investigator (C. Junge, 1912–1996), the layer of stratospheric aerosol is also named the *Junge layer*, see Fig. 3.1. For comparison, 1 cm^3 of air at sea level contains about 3×10^{19} gas molecules.

A common representation of the particle size distribution function is the so-called log-normal distribution (Hinds 1999)

$$n(\log D) = \frac{N}{\sqrt{2\pi}\,\log\,\sigma_g}\,\exp\left(-\frac{1}{2}\frac{(\log D - \log D_{med})^2}{\log^2 \sigma_g}\right)$$

which assumes a *Gaussian* (bell-curve shaped) or normal distribution of the logarithm of the particle diameter, log D, around the logarithm of its *count median* (or number median) diameter, log D_{med}, with a *geometric standard deviation* of the distribution of σ_g. The distribution function can also be expressed on the basis of the natural logarithm. It is characterized by three parameters: total number concentration N, D_{med} and σ_g. For its simple parameterization and its straightforward character this distribution is widely used in *Earth system models*.

Atmospheric aerosol can be regarded as a superposition of four particle modes, each described by a single *log-normal size distribution*. Figure 3.3 is a schematic representation of the modes of a generic atmospheric aerosol of the *atmospheric boundary layer*, the size ranges covered by the different aerosol modes, the major chemical compounds and the respective size distribution.

The nucleation mode ($D \leq 20$ nm) contains particles freshly formed from gaseous precursors by gas-to-particle conversion, or homogeneous particle nucleation; most aerosol particles in the nucleation mode are composed of sulfuric compounds and hydrocarbon compounds.

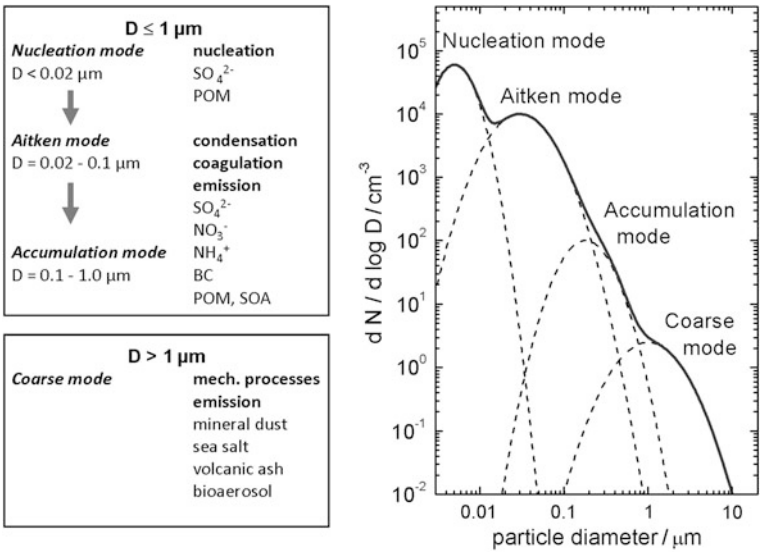

Fig. 3.3 Schematic representation of the modes of a typical atmospheric aerosol of the planetary boundary layer for different particle diameters *D Left* key chemical compounds and particle generation processes; *right* size ranges

The *Aitken mode* (kursiv setzen = 20–100 nm) includes particles formed by coagulation of nucleated particles and condensation of vapors on already existing particles, or emitted directly into the atmosphere; it was named after the Scottish meteorologist and physicist John Aitken (1839–1919), who was one of the founders of aerosol science. Nucleation mode and Aitken mode as defined here are combined by some authors into a single Aitken mode which then covers all particles smaller than 100 nm in diameter.

The *accumulation mode* ($D = 0.1$–1.0 µm) consists of particles formed from the Aitken mode by particle coagulation and particles emitted directly from primary sources like combustion of vegetation or fossil fuels; for atmospheric aerosols the accumulation mode forms the sink of particles growing from nucleation via the Aitken mode into the accumulation mode.

The *coarse mode* contains all particles larger than 1 µm in diameter; these particles are generated mainly by mechanical processes like wind-blown dust, sea spray, or plant debris, or are emitted from volcanoes or large fires. Because of the different formation pathways, coarse mode particles are separated from smaller particles with respect to their chemical composition. Bioaerosols like spores, bacteria or pollen are also mainly part of the coarse mode, but may be part of the accumulation mode as well.

The major chemical constituents of sub-µm aerosol in the lower troposphere are *sulfates* (SO_4^{2-}), nitrates (NO_3^-), *ammonium* (NH_4^+), *black carbon* (BC), and particulate *organic matter* (POM) partially emitted directly from the source or

formed as *secondary organic aerosol* (SOA). Super-μm particles in the coarse mode are mainly associated with *mineral dust*, *sea salt*, *volcanic ash* and *biogenic particles* like spores, bacteria or plant debris. In the free troposphere and above, the concentration of coarse mode particles in general and of source-related compounds like POM and BC is much smaller than in the boundary layer. Stratospheric aerosol is predominantly composed of *sulfuric acid* droplets and volcanic ash particles.

3.3 Sources, Atmospheric Processing, and Sinks

Atmospheric aerosol particles arise from a variety of sources and formation processes, as discussed in the previous section. They can be classified as being of primary or secondary origin according to their history of formation. Primary particles are emitted directly from the source while secondary particles form from gaseous precursors which may have been emitted from ground-based sources or formed in the atmosphere by chemical reactions. Primary particles include soil material and minerals, sea salt, biogenic particles such as pollen or bacteria, and soot (elemental carbon including organic material). Examples for secondary particles are sulfates originating from DMS (dimethylsulfide) or SO_2, nitrates formed from NO_x, and secondary organic aerosol (SOA) particles formed from biogenic or anthropogenic gas-phase hydrocarbons.

Source processes for primary particles are manifold (e.g., anthropogenic fossil fuel combustion, biomass burning, volcanic eruptions, wind-blown mobilization of soil material, *desert dust* storms, wind-driven *sea spray* production) but are mostly related to particle emissions from Earth's surface, with the exception of emissions from cruising *aircraft*, which is the only source of primary aerosol particles in the *upper troposphere* (Petzold et al. 1999). The formation of secondary particles by gas-to-particle conversion is solely associated to nucleation. Secondary particle formation can occur in the *marine boundary layer*, over rural continental sites, in the polluted *urban atmosphere*, and in the free troposphere (Raes et al. 2000; Kulmala et al. 2001).

Secondary particle formation by *homogeneous nucleation* always includes H_2SO_4, organic vapors and H_2O (Kulmala et al. 2001). New particle formation takes place in the form of bursts in clean air characterized by low particle number concentrations. Correspondingly, the aerosol surface area provided to condensable gases by these few airborne particles is low and *supersaturation* with respect to those gases can develop. In contrast, at high levels of available aerosol, surface area condensation of molecules on pre-existing particles is favored over new particle production, and this quenching of nucleation is proportional to the aerosol surface area in most cases. Recent speculations about a link between galactic cosmic rays associated to the solar cycle and the formation of new particles in the upper troposphere could not be validated by observations performed during the European field study EUCAARI, to which the DLR's *Falcon* aircraft significantly contributed (Kulmala et al. 2010).

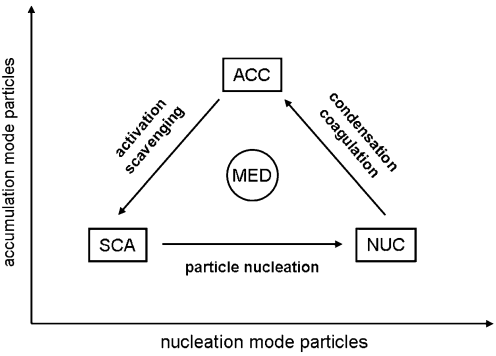

Fig. 3.4 Schematic illustrating the map of principal states of tropospheric aerosols: *NUC* fresh, nucleated state, *ACC* aged, accumulated state, *SCA* cloud-processed state. The MED state represents typical background aerosol; adapted from Schröder et al. (2002)

Once released into the atmosphere or formed by gas-to-particle conversion, small particles further grow by condensation of molecules from the gas phase or by the collision of small and large particles (so-called *coagulation*), which leads to the evolution of larger but fewer aerosol particles of mixed chemical composition. Figure 3.4 shows a map of states of atmospheric aerosol as proposed by Schröder et al. (2002). Particle nucleation is favored by low number concentrations of aged particles in the accumulation mode, i.e., low surface area of already existing particles. Aging by coagulation/accumulation drives the aerosol to a well-aged state characterized by a high number concentration of accumulation mode particles and a medium number concentration of fresh nucleation mode particles. If these aged aerosol particles interact with clouds, preferably larger particles from the accumulation mode with $D > 0.1$ μm act as nuclei for cloud drop or ice formation. These particles are then removed from the atmosphere via *scavenging* by cloud droplets and *wet deposition*, leaving behind an aerosol characterized by low number concentrations of both fresh and aged particles, which then favors the formation of new particles by nucleation. The scheme shown in Fig. 3.5 summarizes how aerosol particles are processed in the atmosphere (Hinds 1999).

One specific mechanism of reversible particle growth in the atmosphere is associated with the uptake of *water vapor* by water-soluble particle constituents in a humid atmosphere. This so-called hygroscopic growth of aerosol particles is attracting increasing attention because it governs the activation of aerosol particles into *cloud condensation nuclei* (CCN), initiating cloud formation. It also enhances the climate impact (see Sect. 3.6) of aerosol particles by significantly increasing the particle size and thus the *scattering* of sunlight. To give an example, sulfuric acid and ammonium sulfate particles, which are quite common constituents of atmospheric aerosol, more than double their diameter at 95 % *relative humidity* compared to the dry state simply by taking up water from the surrounding air (Seinfeld and Pandis 2006). Contemporary concepts for understanding *hygroscopic* growth and cloud droplet formation are sufficiently explained by the Köhler theory (Pruppacher and Klett 1997).

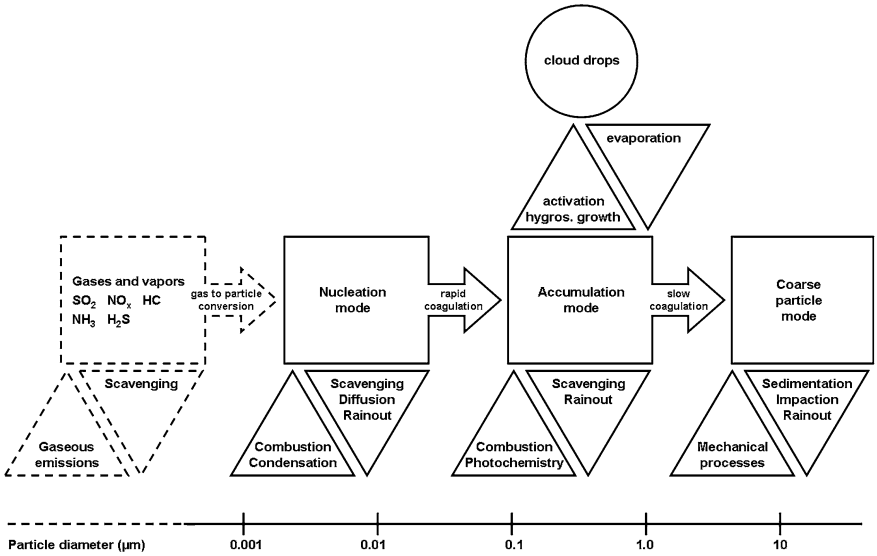

Fig. 3.5 Scheme of the microphysical processes that influence the size distribution and chemical composition of atmospheric aerosol; adapted from Hinds (1999)

3.4 Spatial Distribution of Aerosols

Driven by the spatial distribution of sources, sinks and transport, atmospheric aerosol exhibits a certain vertical number concentration profile. Over more than a decade of intense field study work at the DLR Institute of Atmospheric Physics, a large set of data on the vertical distribution of aerosol particles was collected. Figure 3.6 shows examples of measurements taken at *mid-latitudes* (a) in the clean background atmosphere over the southern tip of South America near Punta Arenas, Chile, (b) in the polluted background aerosol over the northern hemisphere near Prestwick, Scotland, with both data sets emerged from the European field study INCA (Minikin et al. 2003), and (c) over continental Europe related to the German field study LACE (Petzold et al. 2002) near Berlin, Germany. Number concentration data are given for standard conditions (STP: 273.15 K, 1013.25 hPa) since in this case number concentrations are conserved for the vertical motion of the respective air parcel.

In all cases, total particle number concentrations are highest near Earth's surface, in accordance with the dominating aerosol sources there. Through the lowest part of the tropospheric column, the planetary boundary layer, the vertical profile is usually characterized by a strong negative gradient. Across the transition from the boundary layer into the free troposphere, the particle number concentration develops into a well-mixed free tropospheric aerosol with only weak vertical structure. In the polluted northern hemisphere these profiles show minimum values

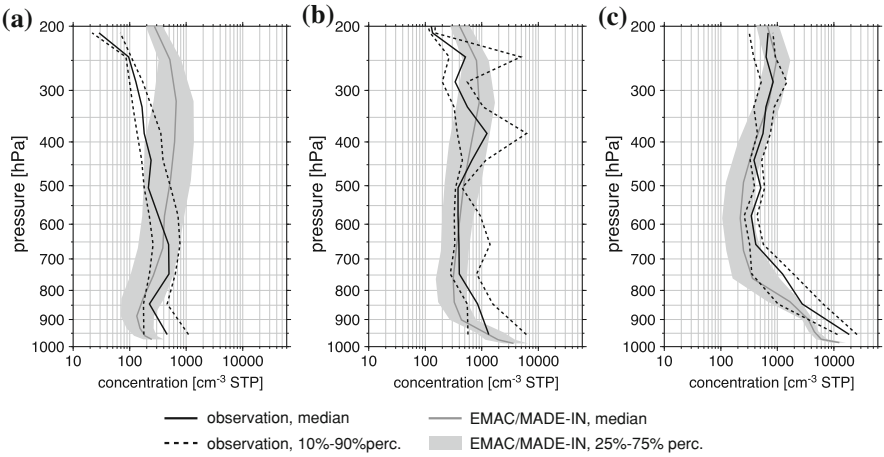

Fig. 3.6 Vertical profiles of the number concentration (converted to standard temperature and pressure (STP) conditions: 273 K, 1 013 hPa) of aerosol particles larger than 5 nm as measured by Minikin et al. (2003) during INCA over Punta Arenas (Chile) in March–April 2000 and over Prestwick (Scotland) in October 2000 (**a** and **b**, respectively) by Petzold et al. (2002) during LACE over central Europe in August 1998 (**c**), and from corresponding climate model simulations (Aquila et al. 2011)

in the middle free troposphere at pressure levels around 800 to 500 hPa, which is the altitude range for intense aerosol-cloud interaction. Near the *tropopause* the particle number concentration increases again, which can be attributed to particle nucleation in the upper troposphere and the lack of efficient aerosol removal processes. As can be taken from Fig. 3.6, global aerosol *models*, see, e.g., Aquila et al. (2011), successfully reproduce aerosol vertical profiles in the polluted northern hemisphere, while they produce less good agreement for the less polluted southern hemisphere.

Specific primary aerosol particles emitted from surface-bound sources and transported upward by lifted air parcels show even more pronounced vertical profiles than total aerosol, as shown in Fig. 3.6. For example, black carbon mass concentrations observed on average in the free troposphere above the planetary boundary layer are lower by two orders of magnitude and beyond compared to the respective surface data (Petzold et al. 2007; Aquila et al. 2011; Hamburger et al. 2011).

Another aerosol type of particular interest is mineral dust, which was extensively studied during the German Saharan Mineral Dust Experiment SAMUM where DLR was responsible for the airborne observations using aircraft. Figure 3.7 shows an illustrative example of the vertical distribution of mineral dust transported from the African mainland to the Cape Verde islands where the measurements were performed. Data shown here originate from a *high spectral resolution lidar* which is capable of actively *sounding* aerosol *extinction* from an *airborne platform*. The plot shows the effect of an island on vertical dust distribution.

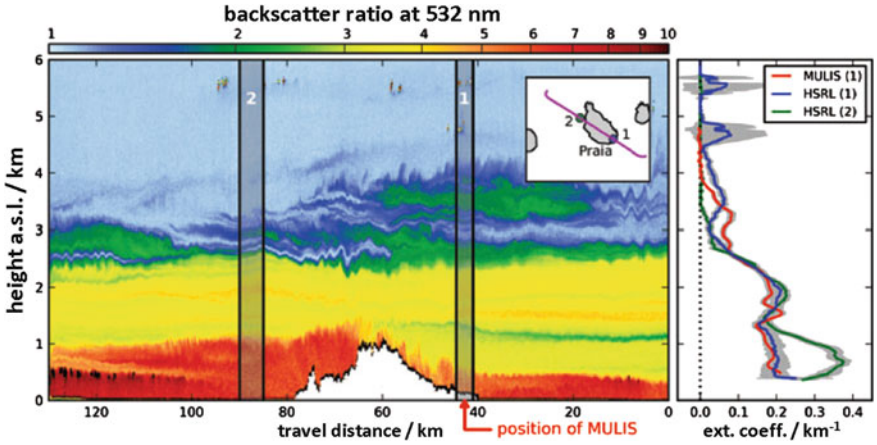

Fig. 3.7 Illustration of island effects on mineral dust layers over the ocean. The *left* plot shows the backscatter ratio at 532 nm measured by the high spectral resolution lidar (HSRL) during an overpass over Santiago Island (Cape Verde Islands) between 16:05 and 16:22 UTC 25 January 2008. MULIS refers to a ground-based lidar operated at the airport of Praia. The *inset* shows the flight track during the sounding; the flight direction was from SE to NW (Weinzierl et al. 2011)

The mineral dust layer extends from sea level to about 2.5 km and is topped by a thin layer of biomass burning aerosol.

At the upwind side (the main wind direction is from left to right) the island lifts the dense part of the mineral dust layer over the top of the island's mountains orographically, while at the downwind side a clear three-layer structure develops. SAMUM also demonstrated that over the desert mineral dust may extend up to an altitude of 4 km above ground (Weinzierl et al. 2009) with a sharp upper bound where total particle number drops off by more than one order of magnitude over $\sim$100 m altitude. If dust particles are lifted to higher altitudes they serve as efficient *ice nuclei* initiating the formation of *mixed-phase* clouds containing both liquid drops and ice crystals (Ansmann et al. 2008). SAMUM was among the first intense field studies which provided this kind of information on the vertical distribution of mineral dust (Ansmann et al. 2011).

3.5 Atmospheric Lifetime and Long-Range Transport

The *lifetime* of atmospheric aerosol particles depends on their chemical nature and size and on the altitude range. Airborne particles in the troposphere have typical lifetimes of 3–10 days on average (Andreae 2007). After 30 days, 95 % of the particles are removed by coagulation with other particles or cloud drops (so-called scavenging), and wet deposition via *precipitation*. Aerosol particles in the *lower stratosphere* have a much longer lifetime of up to 1 year before they penetrate the tropopause and undergo tropospheric loss processes. Figure 3.8 shows the

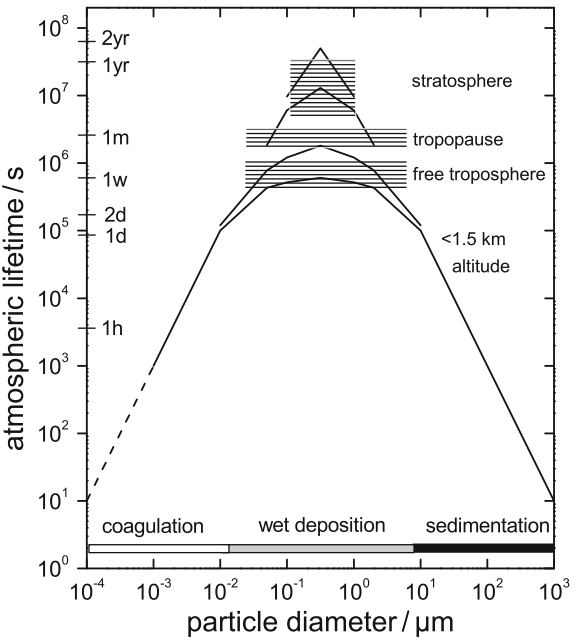

Fig. 3.8 Atmospheric lifetime for individual particles with indicated dominant loss processes in the troposphere; adapted from Journal of Aerosol Science, 11, Jaenicke, R., Atmospheric aerosols and global climate, 577–588, copyright (1980) with permission from Elsevier

atmospheric lifetime of individual particles for the different altitude ranges of the atmosphere (Jaenicke 1980). As can be deduced from this figure, smaller particles are efficiently removed by coagulation and large particles by sedimentation, while particles of the accumulation mode size range are efficiently removed only by wet deposition via aerosol-cloud interaction processes.

During its atmospheric lifetime, aerosol can easily be transported on intra-hemispheric scales during its residence time of more than one week in the free troposphere. Intercontinental transport of particles is observed frequently associated with lifting in *warm conveyor belts* associated with low pressure systems and *long-range transport* in the free troposphere. In 10 years of regular measurements of aerosols over Leipzig, Germany, by means of *Raman lidar* (Mattis et al. 2008), the authors observed pollution layers above the continental boundary layer in 43 % of the cases. The lofted layers largely resulted from intercontinental long-range transport. In 19 % of all regular measurements, free-tropospheric pollution was advected from North America. Forest-fire smoke from Canada and anthropogenic pollution from urban areas of the USA and Canada were the sources of the particle layers.

From a European perspective, extreme cases of lifted aerosol layers in the free troposphere are strong fire plumes transported from North American sources (Petzold et al. 2007), Saharan dust layers advected from North Africa (Mattis et al. 2008), and volcanic ash plumes (Schumann et al. 2011). Even air pollution transportation over the Mediterranean Sea (Lelieveld et al. 2002) and pollution plumes from southern and eastern Asia, including regions in India and China

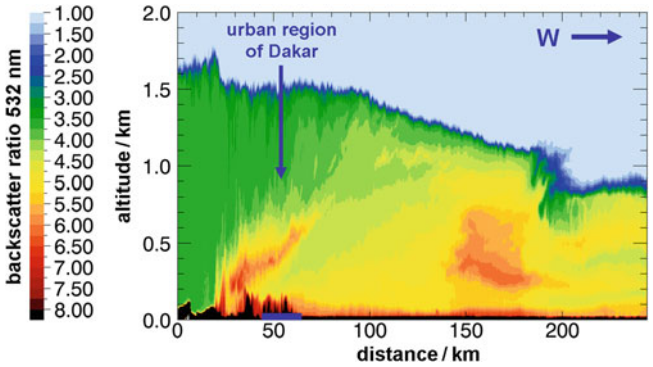

Fig. 3.9 High Spectral Resolution Lidar (HSRL) cross-section of the Dakar plume mixing into the mineral dust layer; the city of Dakar is situated at 50 km distance from the arbitrarily chosen point of origin of the distance scale, whereas the plume appears to originate from an area 30 km inland of Dakar; this local shift is caused by sea-breeze circulation (see text for details); at distance 150 km a residual plume is observed. Viewing direction is south, direction to west is indicated (Petzold et al. 2011)

(Stohl et al. 2007), were observed over Europe in measurements of the DLR Falcon *research aircraft* after more than one week of atmospheric transport.

We observed a particularly spectacular example of aerosol transport and mixing during the SAMUM project when investigating the mixing of the urban pollution plume over the city of Dakar (Senegal) with mineral dust. Figure 3.9 shows a cross-section of the Dakar plume with clean dust moving in from the east (left part of the plot) and mixing with urban pollution during westerly transport from the continent out to the ocean. The origin of the distance scale is arbitrarily set to a point west of Dakar. The shoreline is situated at distance 60 km. Signatures of urban pollution were still found 100 km offshore after several hours of atmospheric transport. However, the data analysis clearly showed that the impact of urban emissions on mineral dust via mixing with black carbon (soot) is limited to less than one day of the plume's lifetime (Petzold et al. 2011).

3.6 Aerosols in the Climate System

Anthropogenic aerosol particles influence the *radiative budget* of the *Earth-atmosphere system* in two different ways. The first is the *direct effect*, whereby particles scatter and absorb solar and thermal infrared radiation and thereby alter the radiative balance of the Earth-atmosphere system or, equivalently, the planetary *albedo* (= reflectivity of the Earth-atmosphere system). The second are called the *indirect effects*, since particles modify the microphysical and hence the radiative properties and lifetime of clouds (Haywood and Boucher 2000). Figure 3.10

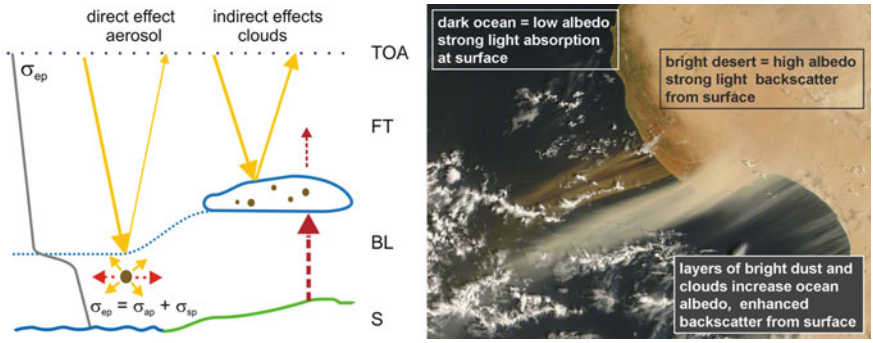

Fig. 3.10 *Left* schematic of direct and indirect radiative effects of aerosols and clouds on the global radiative balance in terms of coefficients σ for particle scattering (sp), absorption (ap), and extinction (ep). S surface, BL planetary boundary layer, FT free troposphere, TOA top of atmosphere. *Right* satellite image of a mineral dust outbreak from the northern Sahara towards the Mediterranean Sea; seen on October 26, 2007 by the MODIS instrument on NASA's Terra satellite (http://earthobservatory.nasa.gov/NaturalHazards/view.php?id=19264)

schematically summarizes the effects of aerosols and clouds on Earth's radiative balance; see also Fig. 3.1 for illustration.

Radiative effects are evaluated in terms of radiative fluxes in the *solar* (solid arrows in Fig. 3.10) and *thermal infrared* (broken arrows in Fig. 3.10) spectral regions at the *top of the atmosphere* (TOA). If the anthropogenic aerosol or the clouds modified by the anthropogenic aerosol increase the net (i.e., the solar plus thermal infrared) radiative flux at TOA, more energy is reflected back into space than in the unperturbed case without anthropogenic aerosol in the atmosphere, and the overall effect is a cooling. If the net radiative flux at TOA is reduced by light-absorbing aerosol particles or by clouds, the resulting effect is a heating of the Earth-Atmosphere system because more radiative energy remains in the system compared to the unperturbed case.

According to Fig. 3.10 the effects of atmospheric aerosol particles on planetary albedo can be sketched as follows: Incoming sunlight is partially scattered (solid arrows) and partially absorbed (broken arrows) and subsequently transferred as heat to the surrounding air. These processes are described by the respective coefficients σ_{sp} for particle scattering and σ_{ap} for particle *absorption*, which are reported as the removed fraction of light by the respective processes per unit length of the light path. The sum of both is the *extinction coefficient* σ_{ep}, which exhibits a distinct vertical profile with its major contribution arising from atmospheric boundary layer aerosol; see also the discussion on the vertical distribution of aerosols in Sect. 3.4. The free troposphere contributes significantly only if *lofted* aerosol layers from long-range transport events are embedded.

The fundamental aerosol parameters governing the aerosol impact on climate forcing are the aerosol optical depth and the ratio of particle scattering to extinction at a *wavelength* of 550 nm, the so-called *single-scattering albedo* ω_0 (Haywood and Shine 1995). Aerosol optical depth (AOD) is defined as the

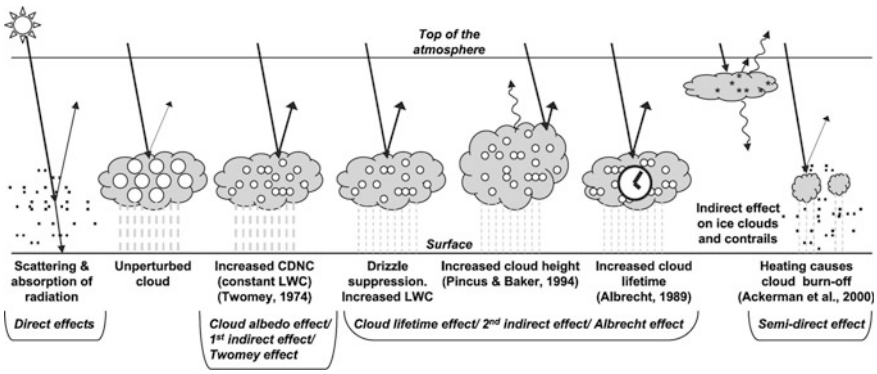

Fig. 3.11 Schematic diagram showing the various radiative mechanisms associated with cloud effects that have been identified as significant in relation to aerosols (modified from Haywood and Boucher, 2000). The small black dots represent aerosol particles; the larger open circles cloud droplets. Straight lines represent the incident and reflected solar radiation, and wavy lines represent terrestrial radiation. The filled white circles indicate cloud droplet number concentration (CDNC). The unperturbed cloud contains larger cloud drops as only natural aerosols are available as cloud condensation nuclei, while the perturbed cloud contains a greater number of smaller cloud drops as both natural and anthropogenic aerosols are available as cloud condensation nuclei (CCN). The vertical grey dashes represent rainfall, and LWC refers to the liquid water content; Fig. 2.10 of IPCC report 2007 (IPCC 2007)

extinction coefficient integrated from the surface to TOA for a unit cross-section footprint. AOD is a common product delivered by space-borne sensors for the global mapping of atmospheric aerosol; see also Fig. 3.2 for a global distribution of AOD.

Exclusively light-scattering particles are characterized by $\omega_0 = 1.0$, while partially light-absorbing aerosol particles are associated with values of $\omega_0 < 1$. As a rule-of-thumb, the critical value ω_0^* which separates net heating from net cooling aerosol is $\omega_0^* \cong 0.87$ (Hansen et al. 1997). Strongly light-absorbing aerosol particles with $\omega_0 < \omega_0^*$ have a net-heating effect, while weakly absorbing aerosols with $\omega_0 > \omega_0^*$ still have a net cooling effect dominated by their contribution to light-scattering. Tropospheric aerosols that have a substantial anthropogenic component include sulfate particles, black carbon, organic carbon, mineral dust, and nitrate aerosol. From these aerosol components black carbon and mineral dust are the only constituents which absorb *solar radiation* in the visible spectrum (Petzold et al. 2011), while sulfate and nitrate particles are purely light-scattering. Since, in particular, the efficiency of particles for scattering light strongly increases with increasing particle size (Hinds 1999), the above-mentioned humidity-related growth of particles considerably enhances the net cooling effect of water-soluble aerosols in a humid atmosphere (Seinfeld and Pandis 2006). Furthermore, the role of organic aerosols, also named "brown carbon," in the *climate system* is still widely unknown (Andreae and Gelencsér 2006).

Figure 3.11 shows a schematic diagram of the various radiative mechanisms associated with aerosol and cloud effects that have been identified as significant (Haywood and Boucher 2000). Clouds interact with radiation both in the *solar spectrum* and in the thermal infrared spectrum. In the solar spectrum clouds increase planetary reflectivity and reduce the fraction of radiation reaching the surface. The unperturbed cloud contains larger cloud drops as only natural aerosols are available as cloud condensation nuclei. The perturbed cloud contains a greater number of smaller cloud drops as both natural and anthropogenic aerosols are available as cloud condensation nuclei. In the infrared spectrum clouds efficiently absorb thermal radiation emitted from the surface and re-emit thermal radiation corresponding to their actual temperature. Therefore, thermal effects are largest for cold high-altitude *cirrus* clouds, while warm low-level clouds mainly contribute to an increase in planetary albedo. Investigation of *aerosol-cloud interactions* and resulting impacts on the global climate system is one of the hot spots of current research.

References

Andreae, M.O., Merlet, P.: Emission of trace gases and aerosols from biomass burning. Global Biogeochem. Cy. **15**, 955–966 (2001)

Andreae, M.O., Gelencsér, A.: Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols. Atmos. Chem. Phys. **6**, 3131–3148 (2006)

Andreae, M.O.: Aerosols before pollution. Science **315**, 50–51 (2007)

Ansmann, A., Tesche, M., Althausen, D., Müller, D., Seifert, P., Freudenthaler, V., Heese, B., Wiegner, M., Pisani, G., Knippertz, P., et al.: Influence of Saharan dust on cloud glaciation in southern Morocco during the Saharan Mineral Dust Experiment. J. Geophys. Res. **113**, D04210 (2008). doi:10.1029/2007JD008785

Ansmann, A., Petzold, A., Kandler, K., Tegen, I.N.A., Wendisch, M., Müller, D., Weinzierl, B., Müller, T., Heintzenberg, J.: Saharan mineral dust experiments SAMUM–1 and SAMUM–2: What have we learned? Tellus B **63**, 403–429 (2011). doi:10.1111/j.1600-0889.2011.00555.x

Aquila, V., Hendricks, J., Lauer, A., Riemer, N., Vogel, H., Baumgardner, D., Minikin, A., Petzold, A., Schwarz, J.P., Spackman, J.R., et al.: MADE-in: A new aerosol microphysics submodel for global simulation of insoluble particles and their mixing state. Geosci. Model Dev. **4**, 325–355 (2011). doi:10.5194/gmd-4-325-2011

Hamburger, T., McMeeking, G., Minikin, A., Birmili, W., Dall'Osto, M., O'Dowd, C., Flentje, H., Henzing, B., Junninen, H., Kristensson, A., et al.: Overview of the synoptic and pollution situation over Europe during the EUCAARI-LONGREX field campaign. Atmos. Chem. Phys. **11**, 1065–1082 (2011). doi:10.5194/acp-11-1065-2011

Hansen, J., Sato, M., Ruedy, R.: Radiative forcing and climate response. J. Geophys. Res. **102**, 6831–6864 (1997). doi:10.1029/96JD03436

Haywood, J.M., Shine, K.P.: The effect of anthropogenic sulfate and soot aerosol on the clear sky planetary radiation budget. Geophys. Res. Lett. **22**, 603–606 (1995)

Haywood, J.M., Boucher, O.: Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: a review. Rev. Geophys. **38**, 513–543 (2000). doi:10.1029/1999RG000078

Hinds, W.C.: Aerosol technology: Properties, behaviour and measurement of airborne particles. Wiley, New York (1999). p. 483

IPCC.: Climate change 2007: the scientific basis, p. 940. Cambridge University Press, Cambridge (2007)

Jaenicke, R.: Atmospheric aerosol and global climate. J. Aerosol Sci. **11**, 577–588 (1980). doi:10.1016/0021-8502(80)90131-7

Kolb, C.E.: Iodine's air of importance. Nature **417**, 597–598 (2002). doi:10.1038/417597a

Kulmala, M., Dal Maso, M., Mäkela, J.M., Pirjola, L., Väkeva, M., Aalto, P., Miikkulainen, P., Hämeri, K., O'Dowd, C.D.: On the formation, growth and composition of nucleation mode particles. Tellus B **53**, 479–490 (2001)

Kulmala, M., Riipinen, I., Nieminen, T., Hulkkonen, M., Sogacheva, L., Manninen, H.E., Paasonen, P., Petäjä, T., Dal Maso, M., Aalto, P.P., et al.: Atmospheric data over a solar cycle: No connection between galactic cosmic rays and new particle formation. Atmos. Chem. Phys. **10**, 1885–1898 (2010). doi:10.5194/acp-10-1885-2010

Lelieveld, J., Berresheim, H., Borrmann, S., Crutzen, P.J., Dentener, F.J., Fischer, H., Feichter, J., Flatau, P.J., Heland, J., Holzinger, R., et al.: Global air pollution crossroads over the Mediterranean. Science **298**, 794–799 (2002). doi:10.1126/science.1075457

Mattis, I., Müller, D., Ansmann, A., Wandinger, U., Preissler, J., Seifert, P., Tesche, M.: Ten years of multiwavelength Raman lidar observations of free-tropospheric aerosol layers over central Europe: Geometrical properties and annual cycle. J. Geophys. Res., 113 (2008). doi:10.1029/2007jd009636

Minikin, A., Petzold, A., Strom, J., Krejci, R., Seifert, M., van Velthoven, P., Schlager, H., Schumann, U.: Aircraft observations of the upper tropospheric fine particle aerosol in the northern and southern hemispheres at midlatitudes. Geophys. Res. Lett. **30**, L10809 (2003). doi:10.1029/2002gl016458

Petzold, A., Döpelheuer, A., Brock, C.A., Schröder, F.: In situ observations and model calculations of black carbon emission by aircraft at cruise altitude. J. Geophys. Res. **104**, 22171–22181 (1999)

Petzold, A., Fiebig, M., Flentje, H., Keil, A., Leiterer, U., Schröder, F., Stifter, A., Wendisch, M., Wendling, P.: Vertical variability of aerosol properties observed at a continental site during the Lindenberg aerosol characterization experiment (LACE 98). J. Geophys. Res. **107** (2002). doi:10.1029/2001JD001043

Petzold, A., Weinzierl, B., Huntrieser, H., Stohl, A., Real, E., Cozic, J., Fiebig, M., Hendricks, J., Lauer, A., Law, K., et al.: Perturbation of the European free troposphere aerosol by North American forest fire plumes during the ICARTT-ITOP Experiment in summer 2004. Atmos. Chem. Phys. **7**, 5105–5127 (2007). doi:SRef-ID:1680-7375/acpd/2007-7-4925

Petzold, A., Veira, A., Mund, S., Esselborn, M., Kiemle, C., Weinzierl, B., Hamburger, T., Ehret, G., Lieke, K., Kandler, K.: Mixing of mineral dust with urban pollution aerosol over Dakar (Senegal): impact on dust physico-chemical and radiative properties. Tellus B **63**, 619–634 (2011). doi:10.1111/j.1600-0889.2011.00547.x

Pruppacher, H.R., Klett, J.D.: Microphysics of clouds and precipitation, 2nd edn. Kluwer Academic Publishers, Dordrecht (1997)

Raes, F., Dingenen, R.V., Vignati, E., Wilson, J., Putaud, J.-P., Seinfeld, J.H., Adams, P.: Formation and cycling of aerosols in the global troposphere. Atmos. Env. **34**, 4215–4240 (2000)

Schröder, F., Kärcher, B., Fiebig, M., Petzold, A.: Aerosol states in the free troposphere at northern midlatitudes. J. Geophys. Res. **107**, (2002). doi:10.1029/2000jd000194

Schumann, U., Weinzierl, B., Reitebuch, O., Schlager, H., Minikin, A., Forster, C., Baumann, R., Sailer, T., Graf, K., Mannstein, H., et al.: Airborne observations of the Eyjafjalla volcano ash cloud over Europe during air space closure in April and May 2010. Atmos. Chem. Phys. **11**, 2245–2279 (2011). doi:10.5194/acp-11-2245-2011

Seinfeld, J.H., Pandis, S.N.: Atmospheric chemistry and physics: from air pollution to climate change, 2nd edn. Wiley, New York Inc (2006)

Stohl, A., Forster, C., Huntrieser, H., Mannstein, H., McMillan, W.W., Petzold, A., Schlager, H., Weinzierl, B.: Aircraft measurements over Europe of an air pollution plume from Southeast Asia—aerosol and chemical characterization. Atmos. Chem. Phys. **7**, 913–937 (2007)

Weinzierl, B., Petzold, A., Esselborn, M., Wirth, M., Rasp, K., Kandler, K., Schütz, L., Koepke, P., Fiebig, M.: Airborne measurements of dust layer properties, particle size distribution and mixing state of Saharan dust during SAMUM. Tellus **61B**, 96–117 (2009). doi:10.1111/j.1600-0889.2008.00392.x

Weinzierl, B., Sauer, D., Esselborn, M., Petzold, A., Veira, A., Rose, M., Mund, S., Wirth, M., Ansmann, A., Tesche, M., et al.: Microphysical and optical properties of dust and tropical biomass burning aerosol layers in the Cape Verde region—an overview of the airborne in situ and lidar measurements during SAMUM-2. Tellus B **63**, 589–618 (2011). doi:10.1111/j.1600-0889.2011.00566.x

Earth's Radiation Budget: The Driver for Weather and Climate

4

Ralf Meerkötter and Margarita Vázquez-Navarro

Abstract

Earth's structure today is manifested in the known distribution of continents and oceans, in a variety of surface types, as well as in a surrounding atmosphere consisting of gaseous constituents, aerosols, and clouds. This complex composition defines a framework that determines the specific amounts of solar energy our planet currently absorbs, reflects, and emits back to space via radiation. The components of Earth's radiation budget and how they are measured are described, the primary energy flows within the Earth system in an equilibrium state are explained and the imbalances of the radiation budget when zooming into smaller temporal and spatial scales are discussed.

4.1 A Few Physical Basics

In order to facilitate reading of the following sections and for a better understanding of the principles of *Earth's radiation budget*, this chapter begins with an introduction of two important physical quantities and three basic laws.

The quantity representing the intensity of *radiation* is called the spectral *radiance*. The spectral radiance is the radiative energy that passes per time unit through an area

R. Meerkötter (✉) · M. Vázquez-Navarro
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Ralf.Meerkoetter@dlr.de

M. Vázquez-Navarro
e-mail: Margarita.Vazquez@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_4,

unit into a solid angle at a distinct *wavelength*. The solid angle is a two-dimensional angle spanned by an area on the surface of a sphere and its center point. It is measured in *steradian*; one steradian corresponds to a solid angle valid for a sphere of radius r and area of r^2 on its surface. The unit of radiance is therefore watts per square meter, steradian, and wavelength unit, i.e. W/(m^2 nm). The wavelength unit of 1 nm corresponds to 10^{-9} m.

Even more frequently used in the context of radiation budget aspects is the spectral flux density or the spectral *irradiance* F_λ. The spectral irradiance is the hemispherical integral of the spectral radiance, in other words it summarizes the projected amount of radiative energy coming from or going into all directions of a hemisphere per unit area and time. The unit of spectral irradiance is therefore watts per square meter and per wavelength unit W/(m^2 nm). Mathematically, spectral irradiance is defined as:

$$F_\lambda = \int_0^{2\pi} \int_0^{\pi/2} I_\lambda \cos\theta \sin\theta \, d\theta \, d\varphi \tag{4.1}$$

where I_λ denotes the spectral radiance, θ and φ the zenith and azimuth angles, respectively. Finally, the irradiance or flux density, which is important for the integral energy budget of the Earth, is the integral of F_λ over defined wavelength intervals.

A law describing the relation between *absorption* and *emission* of real bodies in thermal equilibrium was formulated in 1859 by the German physicist Gustav Robert Kirchhoff (1860). The law states that at a wavelength λ, the amount of spectral radiation K_λ emitted by a body (or surface) equals the amount of radiative energy the body absorbs from the emission L_λ of another *blackbody* at the same temperature. A blackbody is an idealized physical body that absorbs all incident *electromagnetic radiation* and is therefore also the best possible emitter of *thermal radiation*. If a substance absorbs only part of the incident radiation, its spectral *absorptivity* α_λ at wavelength λ is less than 1. Thus, *Kirchhoff's law* is formulated as:

$$K_\lambda = \alpha_\lambda L_\lambda \tag{4.2}$$

Later, in the year 1900, it was Max Planck (1900) who proposed a mathematical formulation describing how a blackbody emits radiative energy as a function of temperature T and wavelength λ:

$$B(\lambda, T) = 2hc^2 / \left(\lambda^5 \left(e^{(hc/(\lambda kT))} - 1 \right) \right) \tag{4.3}$$

This is *Planck's law*, with $B(\lambda,T)$ given in units of spectral radiance. The factor k is the Boltzmann constant, c the speed of light, and h the Planck constant. Figure 4.1 shows calculated curves of spectral irradiances derived from $B(\lambda,T)$. A characteristic feature of irradiances is their strong increase with temperature

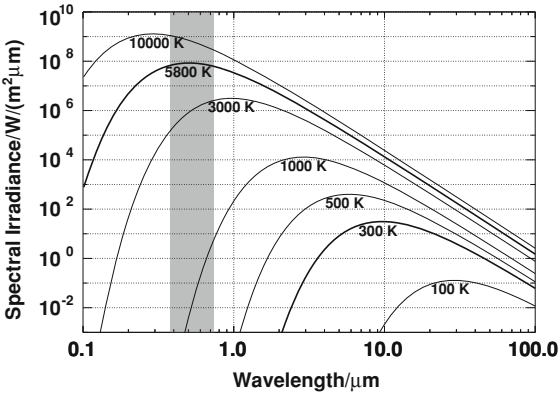

Fig. 4.1 Spectral irradiances calculated for different blackbody temperatures. The spectral range of visible light is *shaded*

T and the concurrent displacement of curve maxima towards shorter wavelengths. In Fig. 4.1, the curve for T = 5 800 K corresponds approximately to the surface temperature of the sun. Note that a part of *solar radiation* lies in the spectral range of visible sunlight (shaded in Fig. 4.1). The location of the curve for T = 300 K represents typical temperatures on Earth.

Integrating the spectral blackbody emission after Planck over the hemisphere and over all wavelengths results in *Stefan-Boltzmann's law* (Boltzmann 1884) showing that the total irradiance of an emitting body is proportional to the fourth power of temperature:

$$E = \varepsilon \sigma T^4 \tag{4.4}$$

with $\sigma = 2\,\pi^5\,k^4/(15\,h^3\,c^2)$ as the Stefan-Boltzmann constant and ε the *emissivity*. The emissivity ε is 1 in case of a blackbody and between 0 and 1 for substances emitting as 'gray' bodies.

4.2 The Sun and Solar Radiation Reaching Earth

The sun, as the star at the center of the solar system, is a powerful nuclear fusion reactor transforming *hydrogen* (H) into *helium* (He), thereby producing tremendous amounts of energy. The sun emits energy by radiation from its photosphere, a layer several hundred kilometers thick below its surface.

The distance between the sun and the earth is about 149 596 000 km. Thus, only a very small part of the radiative energy emitted to space by the sun reaches Earth. To give an illustrative example: *solar energy* captured by the earth over a period of 1 000 years is equal to the energy produced by the sun in just 14 s. This small part of solar energy is the key driver for all physical and chemical processes occurring in the Earth system, finally also responsible for the origin of life and, further on, fueling life on Earth.

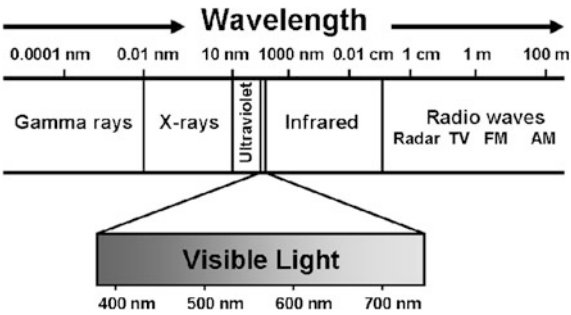

Fig. 4.2 The spectrum of electromagnetic waves with a zoom into the visible part

The irradiance of solar radiation reaching Earth at the top of the atmosphere is known as the *solar constant*. According to recent measurements (Trenberth et al. 2009), the average solar constant amounts to 1 365 W/m^2. In fact, the solar output is not constant. During one year, the solar irradiance at the top of the atmosphere varies by around 6.9 %, i.e., from 1 412 W/m^2 in early January to 1 318 W/m^2 in early July, due to Earth's varying distance to the sun, caused by the eccentricity of Earth's orbit. Over longer time periods, magnetic storms inducing local temperature fluctuations at the sun's surface are responsible for periodical changes of the solar constant. These temperature fluctuations can be observed as dark spots at the sun's surface (sunspots). Their occurrence depends on internal solar activity and presents an 11-year solar sunspot cycle. Over the last three 11-year cycles, the solar output is measured to vary by approximately 0.1 %. Weaker but long-enduring solar fluctuations on the time scale of million of years are still important for paleoclimatological variability. The solar constant is an irradiance integrated over the wavelength. The total radiative energy summarized in the solar constant is actually distributed over a broad spectrum of electromagnetic waves, ranging from gamma rays to *X-rays* to *ultraviolet* radiation to visible light and down to infrared and *radio waves* (Fig. 4.2).

Considering solar radiation as the emission of a blackbody with a temperature of 5777 K, Planck's law shows that more than 95 % of the radiative energy emitted by the sun is contained within the wavelength interval between 0.3 μm and 3.0 μm. This is called the *solar spectral range* or the *shortwave spectral range*. Likewise, considering Earth as an emitting blackbody, an effective average temperature of 255 K results. The maximum of the Planck's curve valid for 255 K is shifted towards longer wavelengths, and more than 95 % of the radiative energy emitted by Earth is located within the interval between 3.0 μm and 100 μm. This interval, denoted as *thermal infrared*, is hereafter also denoted as the *terrestrial spectral range*, or the *longwave spectral range*.

4.3 Earth's Radiation Budget in Equilibrium

The exchange of energy between Planet Earth and space is governed by the gain of solar radiation and losses by thermal emission. With respect to long-term and global averages of incoming solar and outgoing thermal radiation, Earth's energy budget can be regarded as being in equilibrium.

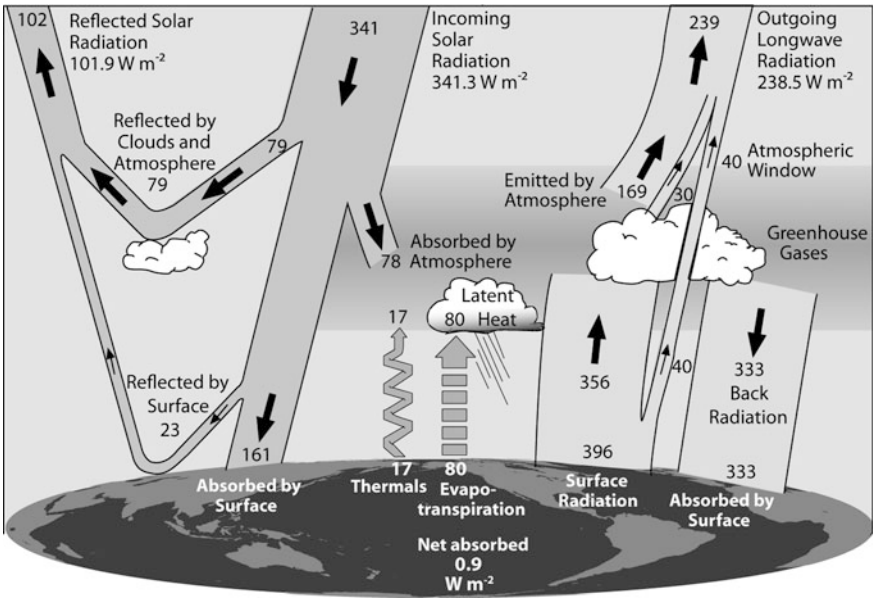

Fig. 4.3 Earth's global annual mean energy budget for the period March 2000 to March 2004. *Arrow thickness* is proportional to the importance of fluxes, given in W/m^2 (image courtesy of K. E. Trenberth, National Center for Atmospheric Research, Boulder, Colorado, USA)

Before presenting the quantification of *energy fluxes* within the Earth system, a geometrical fact has still to be considered. Earth receives a total amount of radiation determined by its cross section, πR_E^2, with R_E as the Earth radius. Due to rotation, this energy is distributed across the entire surface area $4\ \pi\ R_E^2$. Hence, the average incoming solar radiation, taking into account the angle at which the rays strike and that at any time half of the planet does not receive any solar radiation, is one-fourth of the solar constant. As a consequence, 341 W/m^2 is the annual average of solar radiative energy finally available and responsible for all interaction and energy conversion processes taking place within the Earth system. Figure 4.3 depicts how this amount quantitatively consists of different energy flow components.

The net gain from solar radiation is 239 W/m^2 and corresponds to the difference between the average amount of incoming solar radiation at the top of the atmosphere of 341 W/m^2 and the total outgoing amount of solar radiation reflected by Earth of 102 W/m^2. The ratio of both these values, i.e., the ratio between reflected irradiance and incident irradiance, is called the *albedo* of Earth. Measurements reveal that the average Earth albedo is about 0.3. Physically, the albedo results from all *scattering* processes taking place at Earth's land and ocean surfaces and occurring on *water cloud droplets*, *ice cloud crystals*, *aerosol* particles, as well as on atmospheric molecules. The energy flux of 239 W/m^2 remaining in the Earth system is absorbed by the surface and atmospheric constituents and determines their temperature and state of *aggregation* as solid, liquid, or gaseous. Heating of the surface triggers

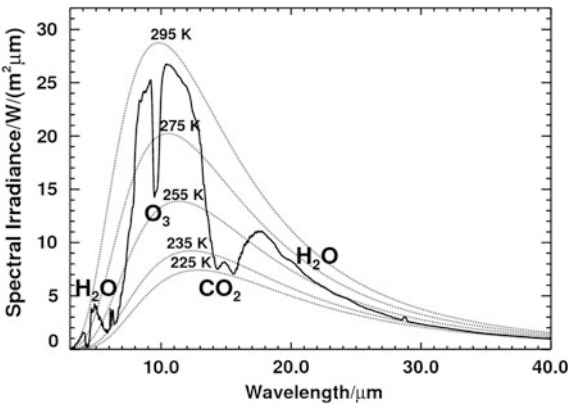

Fig. 4.4 Spectral irradiances at the top of the atmosphere in the terrestrial spectral range for a midlatitude summer atmosphere (*solid curve*) and curves calculated for a blackbody emitting at temperatures typically occurring in the Earth system (*dotted curves*). Irradiances are computed using the *libRadtran* software package (Mayer and Kylling 2005). Indicated are the regions of absorption bands of the gases H_2O, CO_2, and O_3

upward directed energy fluxes known as the transports of *latent heat* and *sensible heat*, amounting to 80 W/m^2 and 17 W/m^2, respectively. The flux of latent heat is related to phase changing processes, *evaporation* and *condensation* of water, which we observe as surface drying or cloud formation in the atmosphere. The sensible heat flux or thermal *convection* is a consequence of buoyancy forces affecting warmed layers near the surface and subsequent lifting of air parcels into upper atmospheric levels. The average net energy release of the surface via emission of longwave radiation is 63 W/m^2. It is the difference between the temperature dependent surface emission of 396 W/m^2, according to the Stefan-Boltzmann's law (Eq. 4.4), and the average downward directed irradiance of 333 W/m^2 (Fig. 4.3). The latter value summarizes the entire re-emission of longwave radiation originating from absorbing atmospheric *trace gases*, aerosols, and clouds.

As mentioned, Fig. 4.3 represents the equilibrium state of the Earth system in which gains and losses of energy are balanced at each altitude, which means that they are balanced at the surface as well as at all levels of the atmosphere including its upper boundary. On a long-term average, such a balanced energy budget conditions a specific stratification of constant temperatures. Measurements reveal that the global mean *surface temperature* of the Earth is currently 288 K at 2 m altitude. However, when measuring Earth's blackbody temperature from space, a value of only 255 K is obtained. This is because the surface emits longwave radiation, which to a certain amount is absorbed by atmospheric trace gases such as *water vapor* (H_2O), *carbon dioxide* (CO_2), *methane* (CH_4), *carbon monoxide* (CO), *ozone* (O_3), and *nitrous oxide* (N_2O), by aerosol particles, and by water and ice clouds. All these constituents are able to re-emit longwave radiation as a function of their local temperatures. Furthermore, the ability of a specific constituent to absorb or emit thermal radiation depends on the wavelength within the terrestrial spectral range. This is illustrated in Fig. 4.4, showing

the characteristic shape of spectral irradiances simulated for the top of a mid-latitude summer atmosphere. As can be seen, irradiances are clearly reduced in the absorption bands of H_2O, CO_2, and O_3.

A *cloud-free* atmosphere is almost transparent in the spectral region between 8 μm and 13 μm, the so-called window in the thermal infrared. In this region, most of the longwave radiation originating at the surface can be emitted to space, whereas the atmosphere is completely opaque in the strong absorption bands of H_2O around 6.3 μm, of O_3 around 9.7 μm, and of CO_2 around 15 μm. These are the spectral regions where trace gases are emitting from very high and cold atmospheric layers into space. In semitransparent spectral regions, i.e., between the window and the regions of maximum absorption, the layers of maximum emission are located somewhere between the surface and the top of the atmosphere. Such effective emission altitudes also vary with trace gas concentration. *Water clouds* in the atmosphere predominantly act as blackbodies over the whole range of the thermal infrared. As a consequence, they close the window and re-emit longwave radiation from their base to the surface and from their colder tops to space. When assuming a cloud layer with a top temperature of 255 K, the maximum values of simulated irradiances in Fig. 4.4 would be shifted down to the curve for 255 K. On the contrary, thin ice clouds, also known as cirrus clouds, appear as semitransparent clouds usually transmitting a certain amount of thermal emission from the surface and from atmospheric layers below.

4.4 Imbalances in Earth's Radiation Budget

As explained in the preceding section, Earth's energy budget in its equilibrium state constitutes long-term and global averages. However, when zooming into smaller temporal and spatial scales, Earth shows imbalances between irradiances of *absorbed shortwave* and *outgoing longwave radiation* at the surface and in all atmospheric layers up to the top of the atmosphere. On the one hand, there are simple geometrical factors causing inhomogeneities, such as the latitude-dependent elevation of the sun. Due to the fact that the axis of the earth is tilted, the amount of solar radiation is also a function of season. On the other hand, imbalances result from inhomogenous distributions and temporal fluctuations of all components contributing to absorption, scattering, and emission processes. An illustration of imbalances detectable within a time-scale of one year is given in Fig. 4.5, presenting global maps of annual averages of the absorbed shortwave radiation, the outgoing longwave radiation, and the *net radiation* at the top of the atmosphere.

As can be seen in Fig. 4.5a, the amount of absorbed solar radiation reaches maximum values higher than 300 W/m^2 in regions of the tropical oceans, where the surface albedo is low ($\approx$5 %) and the sun elevation is high. In contrast, minimum values lower than 100 W/m^2 are found in *polar regions*. At these latitudes, the albedo is high due to extended areas of snow and ice cover, but also due to enhanced *cloud cover*. In addition, the sun elevation in polar regions remains low throughout the year.

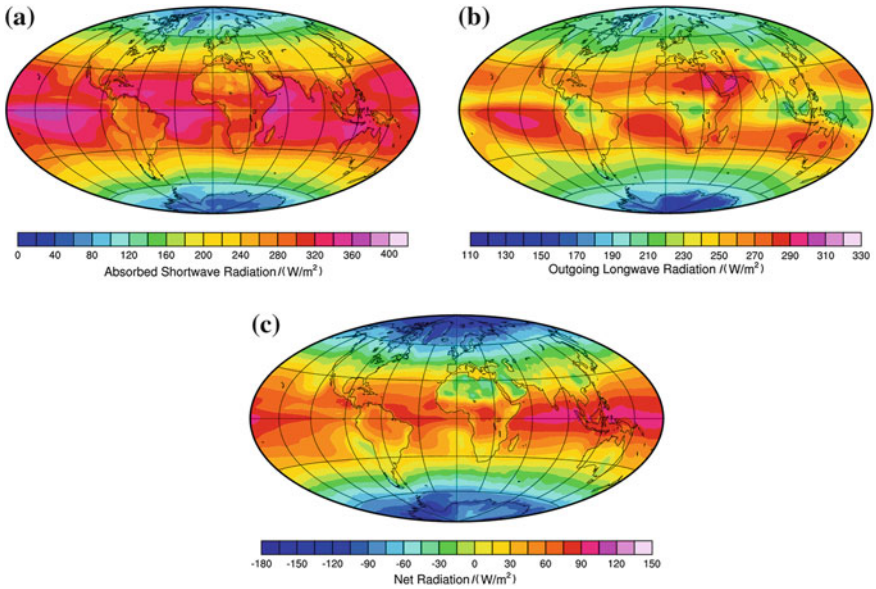

Fig. 4.5 Global distributions of annual averages of **a** absorbed shortwave radiation, **b** outgoing longwave radiation, and **c** net radiation at the top of the atmosphere derived from data of the NASA satellite mission CERES in the period from 2003 to 2006 (images courtesy of D. L. Hartmann, University of Washington, Washington, USA)

According to Eq. 4.4, the outgoing longwave radiation depends strongly on surface temperature and on the temperature of atmospheric constituents. As shown in Fig. 4.5b, maximum values of about 300 W/m^2 occur in the cloud-free regions of warm *subtropical oceans* and over hot *deserts*. In such areas, surface emission is radiated into space, whereas the amounts of absorbed and re-emitted radiation by cold clouds or layers of enhanced water vapor concentration remain small. Lower values of outgoing longwave radiation, in the order of 200 W/m^2, are distributed over the equatorial land masses. Here, thermal convection processes generate vertically extended clouds which reach high altitudes in the atmosphere, where their tops emit at very low temperatures.

The pattern of net irradiances, resulting from the difference of absorbed shortwave radiation and outgoing longwave radiation, is presented in Fig. 4.5c. Subtropical deserts, i.e., regions of low cloud cover, high surface temperature, and high albedo produce negative values of net irradiance, which means that net fluxes are directed upward. In contrast, maximum positive values are distributed along the equatorial oceans, where the absorption of solar radiation is the dominant factor. However, the strong equator-to-pole gradient of net radiation, which results in a corresponding imbalance of net heating, is striking. Net heating imbalances between the equator and pole act as the central driving mechanism for compensation processes occurring in the Earth system, which is recognized as the large-

scale oceanic and atmospheric *circulations*. These circulations are rather stable and transport energy from low to high latitudes. On shorter time scales, the circulations are accompanied by various dynamic processes, locally experienced as *weather* phenomena like wind and turbulence, temperature changes, cloud forming processes, precipitation events, and other well-known features. In climate science, this entire complex is described as the '*heat engine of the Earth*'.

It should be mentioned that there are further natural and *anthropogenic* factors which lead to imbalances in the Earth radiation budget, even on a global scale. Natural influences are, for example, due to changes of the solar constant, so-called *Milankovitch* cycles, representing changes in Earth's orbit which are responsible for the major glacial periods. Another natural phenomenon are *volcanic eruptions* on Earth, ejecting large amounts of aerosol particles which are able to reduce incoming solar radiation over several years. Anthropogenic factors that influence radiative balances are the emission of *greenhouse gases*, which alter the amount of heat radiated back to space, and the emission of pollution aerosols, which influence reflection and absorption of incoming sunlight. Changes in the surface albedo can also be induced by significant deforestation activities; in the atmosphere the introduction of additional cloudiness such as *contrails* may affect the amount of radiation reflected and emitted to space.

4.5 Relevance of Satellites for Monitoring the Earth Radiation Budget

Although the estimation of the Earth radiation budget relies on many sources of measured data and additional numerical computations, *satellites* are essential for *monitoring* its components because they enable measurements with global coverage. In the 1970s, NASA monitored Earth's radiation budget with an instrument called Earth Radiation Budget Experiment (ERBE) aboard the Earth Radiation Budget Satellite (ERBS), launched in 1984 (Barkstrom and Hall 1982). ERBE provided continuous measurements of outgoing longwave and reflected shortwave irradiances, as well as the global Earth albedo and the incoming solar radiation. The ERBE mission officially ended in 2005.

The spatial resolution of the ERBE instruments was 40 km at the subsatellite point. Since the instruments measured angular dependent radiances, a conversion from radiances into irradiances via certain angular dependency models had to be applied. In principle, angular dependency models are based on the ratio between the radiance measured in a given direction and the irradiance integrated over the whole hemisphere (see also Eq. 4.1). Therefore, to accurately provide irradiance measurements, not only the radiance measurements must be accurate, but also the corresponding model assumptions must be correct. Angular dependency models are usually developed by using a combination of measurements and numerical *radiative transfer models*. Their accuracy relies on a proper classification of scenes like different land surface types, water surfaces, cloud fields, *snow* and ice covered regions, and more.

A more recent NASA mission studying the Earth radiation budget is Clouds and the Earth Radiant Energy System (CERES) (Wielicki et al. 1996). It not only focuses on the irradiances of reflected shortwave and emitted longwave radiation, but also on accurately describing cloud and aerosol properties to understand the role clouds and aerosols play in the surface-atmosphere system. CERES has a scan mode which allows radiances from the same scenes to be measured with varying incident solar radiation and observing geometry, hereby improving the angular dependency models and finally the accuracy of derived irradiances.

The first CERES instrument was launched aboard the Tropical Rainfall Measuring Mission (TRMM) satellite in 1997. Additional CERES instruments orbit Earth aboard the Terra and Aqua satellites launched in 1999 and 2002, respectively. The spatial resolution is 10 km for the CERES instrument on TRMM and 20 km for the CERES instruments on the other two platforms. Due to a higher spatial resolution and more precise angular dependency models, CERES provides the best measurements to date of the Earth radiation budget from polar orbiting platforms.

The main studies of the Earth radiation budget focus on its long-term variations. This is why the scarce temporal resolution of *polar orbiting satellites*, providing only one or a few measurements per day for a given location, is not a problem. Usually, seasonal and annual means are derived. While a better temporal sampling is always desirable, these data suffice for most studies, since the response to variations in the Earth radiation budget components usually take place relatively slowly.

However, strong diurnal variations in the radiation budget, particularly over land in response to the diurnal variation of solar heating, cannot be carefully monitored by polar-orbiting satellites, since they lack the temporal resolution necessary to resolve processes on shorter time-scales.

An instrument with a higher temporal resolution is the Geostationary Earth Radiation Budget Experiment (GERB) aboard the European *geostationary satellite* platform Meteosat launched in 2004 (Harries et al. 2005). GERB provides measurements of reflected shortwave and outgoing longwave irradiances with a temporal resolution of only 15 min. Unfortunately, its spatial resolution is low, approximately half a degree in the subsatellite point, which corresponds to about 45 km $\times$ 39 km. This resolution is again sufficient for radiation budget studies over longer time periods, but it is not quite adequate when considering the radiative effect of small structures such as clouds during the day.

An instrument providing radiation measurements of Earth combining both a good spatial and an excellent temporal resolution is the Spinning Enhanced Visible and Infrared Imager (SEVIRI) aboard Meteosat (Schmetz et al. 2002), operating since 2004. Compared to GERB, the spatial resolution of SEVIRI is 3 km $\times$ 3 km at the subsatellite point with data acquisition every 15 min. However, instead of delivering the desired broadband irradiance measurements, SEVIRI retrieves information in twelve narrowband channels ranging from visible to thermal infrared wavelengths. At the DLR Institute of Atmospheric Physics a narrowband-to-broadband-conversion has been developed to obtain shortwave and longwave

irradiances from the narrowband SEVIRI channels, as well as algorithms providing relevant radiative parameters of clouds and cloud fields (Meerkötter and Bugliaro 2009; Bugliaro et al. 2011). These methods are targeted at an assessment and interpretation of Earth's radiation budget as well, and will provide information about atmospheric processes, independent of size.

Figure 4.6 shows irradiances of absorbed shortwave, outgoing longwave, and net radiation. They are derived from instantaneous high resolution Meteosat data applying the aforementioned DLR algorithm. For comparison, the corresponding irradiances forecasted by the numerical model of the European Centre for Medium-Range Weather Forecasts (ECMWF), which represent averages over a 3 h time interval including the Meteosat observation time, are also shown. On the sunlit part of the Earth disk, the lowest values of absorbed solar radiation occur where clouds reflect much of the incoming solar radiation, whereas the highest absorption of solar radiation corresponds to the dark ocean surfaces. In the longwave range, the lowest values are found in regions of enhanced cloud cover at higher latitudes and especially along the equator where cloud tops usually reach the highest and coldest atmospheric levels. The maximum outgoing longwave emission (negative values) is found over the cloud-free hot desert regions of the southern Sahara. Cirrus and water clouds visible as spots over the northern and western Sahara reduce the surface emission and may even replace it if the clouds are opaque. Net irradiances are defined as the sum of absorbed shortwave and outgoing longwave radiation (negative values). The distribution of net irradiances is dominated by the elevation of the sun, resulting in a pattern with maximum positive values in cloud-free regions over the western part of the disk, whereas negative values are found in the nighttime area, where only longwave emission is acting. Some structures are again marked by clouds increasing reflection in the shortwave and decreasing thermal emission in the longwave range. When comparing Meteosat-derived data from SEVIRI and forecasted data from ECMWF, good agreement is obvious in general. Striking are fine scale structures in the Meteosat results, whereas the model forecasts show somewhat smoother distributions, which result from the averaging over three hours, incorporating the effects of changing solar elevation, as well as from moving and vertically growing clouds.

4.6 Ongoing Research

In view of the many kinds of natural and anthropogenic influences affecting the Earth radiation budget and due to the fact that variations in the Earth radiation budget feedback on the composition and physical state of the surface-atmosphere system, our planet is permanently exposed to changes that in turn affect past, present, and future climate.

One of the great challenges is to identify, understand, and reduce the sources of errors in estimating the components of the radiation and energy budgets. The different procedures for estimating these budgets still reveal a spread of results and uncertainties. To give some numbers: at present, top of atmosphere fluxes are afflicted with uncertainties in the order of ± 3 %, whereas errors for the different surface fluxes range from about ± 5 % to ± 10 %.

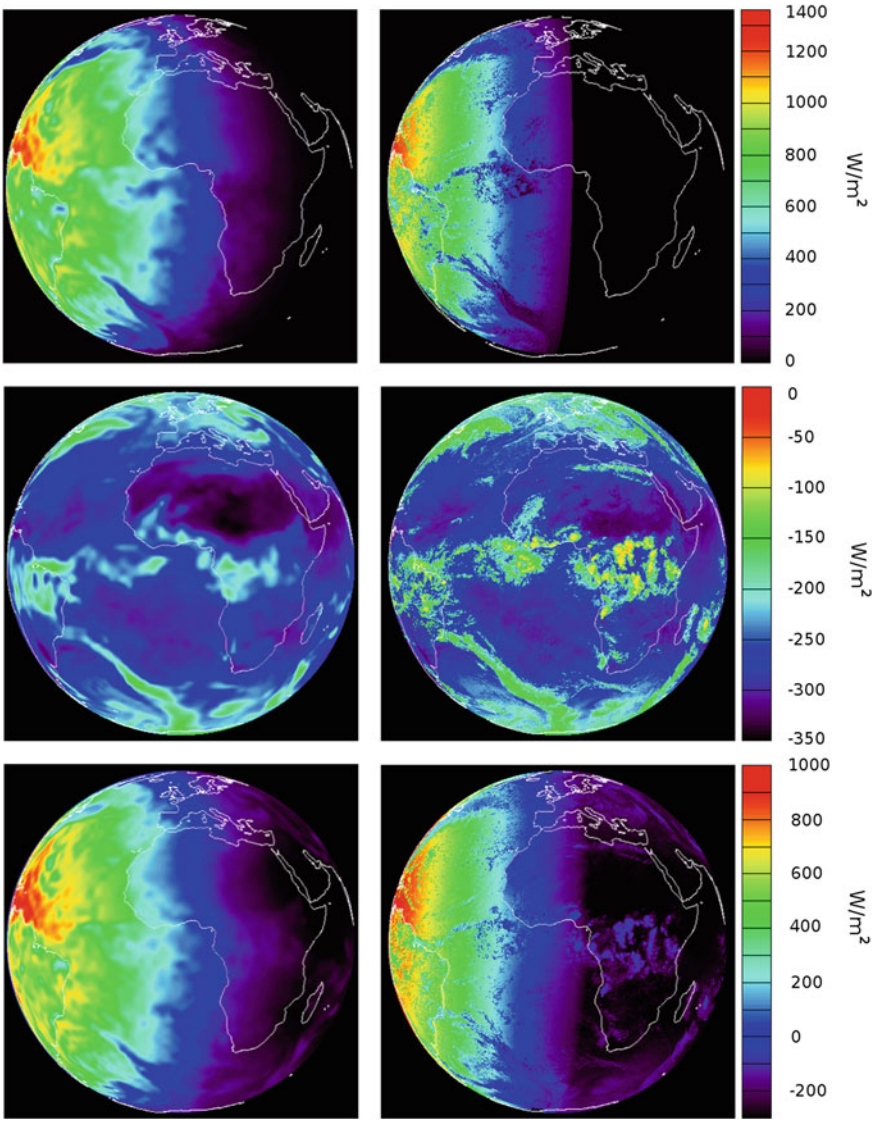

Fig. 4.6 *Top row* absorbed shortwave, *middle row* emitted longwave, *lower row* net irradiances at the top of the atmosphere on 25 March 2008. *Left* forecasted by a numerical model (ECMWF) valid for the time interval from 15:00 to 18:00 UTC, *right* derived at DLR from Meteosat measurements gained at 17:30 UTC

Therefore, direct observations of the Earth radiation budget by satellites have to be continued in future. This is why NASA, for example, launched an additional CERES instrument in October 2011 which is orbiting the Earth aboard the Suomi-NPP satellite. The European Space Agency ESA is preparing to launch EarthCARE

(Earth Clouds, Aerosols and Radiation Explorer) carrying four different instruments which should allow to determine horizontal and vertical vertical distributions of various cloud parameters as well as shortwave and longwave irradiances at the top of the atmosphere. At the DLR Institute of Atmospheric Physics, algorithms are being applied to retrieve Earth radiation budget and cloud parameters from the imager data of the Meteosat Second Generation satellite (MSG) currently in orbit and operating since 2004. It should be noted that the institute's remote sensing products also serve as input to a DLR project aimed at evaluating an Earth system model. This work will continue with data from Meteosat Third Generation (MTG), which is planned for 2020 as a subsequent mission to MSG. Among other instruments, MTG will carry an imager with further improved spatial and temporal resolution and finer spectral resolution.

References

Barkstrom, B.R., Hall Jr, J.B.: Earth radiation budget experiment (ERBE): an overview. J. Energy **6**, 141–146 (1982)

Boltzmann, L.: Ableitung des Stefan'schen Gesetzes, betreffend Würfel Abhängigkeit der Wärmestrahlung von der Temperatur aus der electromagnetischen Lichttheorie. Ann. Phys. Chem. **22**, 291–294 (1884)

Bugliaro, L., Zinner, T., Keil, C., Mayer, B., Hoffmann, R., Hollmann, R., Reuter, M., Thomas, W.: Validation of cloud property retrievals with simulated satellite radiances: a case study for SEVIRI. Atoms. Chem. Phys. **11**, 5603–5624 (2011). doi:10.5194/acp-11-5603-2011

Harries, J.E., Russell, J.E., Hanafin, J.A., Brindley, H., Futyan, J., Rufus, J., Kellock, S., Matthews, G., Wrigley, R., Last, A., et al.: The geostationary earth radiation budget project. Bull. Am. Meteorol. Soc. **86**(7), 945–960 (2005)

Kirchhoff, G.: Über das Verhältnis zwischen dem Emissionsvermögen und dem Absorptionsvermögen der Körper für Wärme und Licht. Ann. Phys. Chem. (Leipzig) **109**, 275–301 (1860)

Meerkötter, R., Bugliaro, L.: Diurnal evolution of cloud base heights in convective cloud fields from MSG/SEVIRI data. Atoms. Chem. Phys. **9**, 1767–1778 (2009). doi:10.5194/acp-9-1767-2009

Planck, M.: Über eine Verbesserung der Wienschen Spektralgleichung. Verhandl. der Deutschen Physikalischen Gesellschaft **2**, 202–204 (1900)

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., Ratier, A.: An introduction to Meteosat Second Generation (MSG). Bull. Am. Meteorol. Soc. **83**(7), 977–992 (2002)

Trenberth, K.E., Fasullo, J.T., Kiehl, J.: Earth's global radiation budget. Bull. Am. Meteorol. Soc. **90**(3), 311–323 (2009). doi:10.1175/2008BAMS2634.1

Wielicki, B.A., Barkstrom, B.R., Harrison, E.F., Lee, R.B., Smith, G.L., Cooper, J.E.: Clouds and the earth's radiant energy system (CERES): an earth observing system experiment. Bull. Am. Meteorol. Soc. **77**(5), 853–868 (1996)

Light Scattering on Molecules in the Atmosphere

5

Benjamin Witschas

Abstract

The scattering of laser light on air molecules is a promising tool in atmospheric science. Atmospheric quantities like air composition, temperature, pressure and wind speed can be derived from the spectrum of molecular scattered light by analyzing the measured spectra with respective physical models of the underlying scattering processes. This article gives an overview of the various kinds of molecular scattering processes and their utilization in atmospheric science.

5.1 Introduction

Amazed by the deep blue and polarized *radiation* coming from Earth's daytime sky, scientists of the nineteenth century started to investigate the origin of this phenomenon. In 1871 John William Strutt, better known as Lord Rayleigh, explained the blue color of the sky by assuming that particles much smaller than the wavelength of light were responsible for *light scattering* (Rayleigh 1871). It took almost a further 30 years before he suggested that not particles but molecules in air are actually responsible for the scattering process (Rayleigh 1899). This awareness was the beginning of molecular light scattering investigations.

B. Witschas (✉)
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Benjamin.Witschas@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_5,

From that time, theoretical and experimental research on molecular light scattering has been an inherent part of physical science. However, the great scientific breakthrough was driven by the invention of the *laser* by Theodore Maiman in 1960. Although many of the basic scattering mechanisms were theoretically suggested or rather known before, only then, with the power, purity and high resolution available from laser light, was one able to properly examine the spectrum of molecular scattered light.

These days, molecular laser light scattering is a widespread diagnostic tool in atmospheric applications such as *lidar* measurements (Weitkamp 2005) as well as in aerospace applications such as combustion measurements (Elliott et al. 2001) and flow measurements (Seasholtz et al. 1997). The concept of all these measurement techniques is based on accurately measuring the characteristics of the molecular scattered light and relating them to theoretical models of the underlying scattering process. In doing so, gas properties like temperature, density, pressure and gas flow velocity (atmospheric *wind speed*) can be derived, given an accurate understanding of the respective scattering process.

Within the last few decades, the lidar department of the Institute of Atmospheric Physics at DLR developed and utilized various lidar systems for measuring atmospheric *wind* (Reitebuch et al. 2009), aerosol extinction and backscatter coefficients (Esselborn et al. 2008), water vapor (Wirth et al. 2009), ozone, and carbon dioxide (Amediek et al. 2008), to name just a few of the applications. In addition, the Institute was involved in a laboratory research study recently carried out at the laser lab of VU University Amsterdam, aimed at measuring the so-called Cabannes line using laser light scattered in air (see Sect. 5.2.3) with high accuracy and for the first time ever (Witschas et al. 2010). The results of this study have implications for fundamental research on molecular gas dynamics as well as for the future space lidars on the *Atmospheric Dynamics Mission* (ADM-Aeolus) and EarthCARE, two missions within ESA's Earth Explorer program.

In this article, the basic light scattering mechanisms arising in molecular gases, namely *Rayleigh scattering* (Sect. 5.2.1), *Raman scattering* (Sect. 5.2.2) and *Brillouin scattering* (Sect. 5.2.3), are introduced as to their origin and influence on the spectrum of scattered light. Additionally, some examples of utilizing molecular light scattering for probing the atmosphere are presented.

5.2 Classification of Light Scattering Mechanisms in Molecular Gases

When light travels through matter, various scattering processes occur. Independent of the aggregate state of the matter, the origin of these scattering processes is a result of inhomogeneities in the optical properties of the propagation medium. For atmospheric applications the propagation medium is air, which is mainly composed of the diatomic gases nitrogen (78.09 vol.%) and oxygen (20.95 vol.%). In such molecular gases the inhomogeneities mentioned are caused by statistical

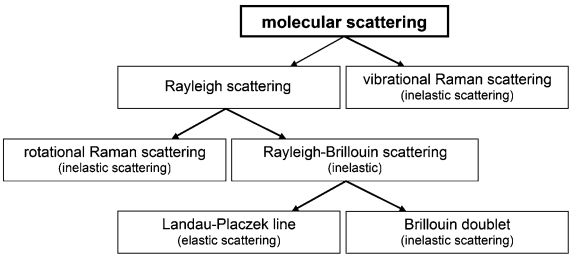

Fig. 5.1 Overview of the various scattering mechanisms in molecular gases according to Young (1981)

fluctuations in the arrangement of the molecules, which lead to different kinds of scattering processes. Thereby, each physical contribution to the generation of scattered light makes its characteristic imprint on the intensity, *polarization* and spectral composition.

For the characterization of different scattering mechanisms, a chief distinction is drawn between elastic and inelastic scattering. Elastic scattering denotes light scattering without, and inelastic scattering with, a change in the scattered light's frequency (or rather wavelength). Furthermore, a distinction can be made between light scattering caused by single molecules (e.g., Raman scattering, Sect. 5.2.2) and light scattering induced by the collective effects of many molecules (e.g., Rayleigh–Brillouin scattering, Sect. 5.2.3).

The most prominent scattering mechanisms in molecular gases are Rayleigh scattering, Raman scattering and Brillouin scattering, all of them named after their discoverer. The detailed terminology of the different molecular scattering processes is sketched in Fig. 5.1.

In Fig. 5.1, it can be seen that molecular light scattering can be basically divided into Rayleigh scattering (which consists of elastic and inelastic light scattering mechanisms with little wavelength shifts smaller than 5 nm; see Sect. 5.2.1) and vibrational Raman scattering (which is an inelastic light scattering mechanism with large wavelength shifts in the order of 30 nm; see Sect. 5.2.2). In more detail, the mechanisms that were (unknowingly) described by Lord Rayleigh are rotational Raman scattering and Rayleigh–Brillouin scattering. Concerning its origin, the spectrum of Rayleigh–Brillouin scattered light can additionally be divided into the Landau-Placzek line and the Brillouin doublet which results from Brillouin scattering (see Sect. 5.2.3). In the following, each of the previously mentioned scattering mechanisms is explained in more detail.

5.2.1 Rayleigh Scattering

The invention of the concept of molecular light scattering traces back to Lord Rayleigh, who indeed used the word scattering (instead of *reflection*) for the first time (Rayleigh 1871). By considering light scattering on small (with respect to the light's wavelength) spherical particles, he derived the λ^{-4} dependence of the intensity of the scattered light (see also Eq. (5.5)). This finding enabled him to

explain the blue color of the sky, as it shows that the short, blue wavelengths of sunlight are scattered much more strongly (to the fourth power) towards Earth's surface than the longer, red ones. Ten years later, Rayleigh re-derived this result using the laws of Maxwell, still considering spherical particles as the origin of the scattering mechanism. Finally, in 1899, Rayleigh suggested that not particles but molecules in air are responsible for light scattering within the terrestrial atmosphere (Rayleigh 1899).

In the case of light scattering on molecules, the scattering bodies are much smaller than the wavelength of light. For example, the diameter of a nitrogen molecule ($\sim$0.3 nm) is about 3 orders of magnitude smaller than the wavelength of visible light ($\sim$500 nm). If one further assumes spherical molecules without any vibrational or rotational internal states, the Rayleigh scattering process can be described by the model of induced dipole radiation. That is, the electrons in the molecules radiate like dipole antennas when they are forced to oscillate by the applied electromagnetic field $\mathbf{E_i}$ of the incident light wave. If the scattering sources (the molecules) are furthermore assumed to be stationary (which means without any proper motion), the dipoles re-emit radiation in all directions and with the same frequency as that of the incident light. The electric field $\mathbf{E_s}$ and the corresponding intensity I_s which is emitted by a dipole is given according to (Miles et al. 2001) as

$$|\mathbf{E_s}(r,\Phi)| = \frac{\omega^2 \cdot p}{4 \cdot \pi \cdot \varepsilon_0 \cdot c^2 \cdot r} \sin(\Phi) \quad ; \quad I_s(r,t) = \frac{1}{2} \varepsilon_0 \cdot c \cdot |\mathbf{E_s}(r,\Phi)|^2 \tag{5.1}$$

where $\omega = 2{\cdot}\pi{\cdot}f = 2{\cdot}\pi{\cdot}c/\lambda$ is the oscillation angular frequency, f the frequency and λ the wavelength of the emitted/incident light wave, c the velocity of light in vacuum, p the magnitude of the oscillation dipole moment induced by the incident electrical field $\mathbf{E_i}$, r the distance from the dipole, ε_0 the vacuum permittivity, and Φ the angle of observation with respect to the dipole vector.

With the simplified assumption of radially symmetric molecules, the dipole moment $\mathbf{p}$ is induced in the same direction as the incident field polarization and is linearly proportional to the incident electrical field $\mathbf{E_i}$. The scalar proportionality is termed the polarizability α

$$\mathbf{p} = \alpha \cdot \mathbf{E_i} \tag{5.2}$$

By inserting Eq. (5.2) in Eq. (5.1), and considering the transformation between electrical field $\mathbf{E_i}$ and intensity I_i as given with Eq. (5.1), the scattering intensity induced by one molecule can be calculated as

$$I_s = \frac{\pi^2 \cdot \alpha^2}{\varepsilon_0^2} \cdot \frac{1}{\lambda^4 \cdot r^2} \cdot I_i \cdot \sin^2(\Phi) \tag{5.3}$$

For further discussion of the spectral contribution of the various scattering processes, it is useful to define the differential scattering cross section $\partial\sigma/\partial\Omega$,

where σ is the scattering cross section (which describes the likelihood of light being scattered by a molecule) and Ω is the solid angle. The differential scattering cross section is defined according to

$$I_s = \frac{\partial\sigma}{\partial\Omega} \cdot \frac{1}{r^2} \cdot I_i \tag{5.4}$$

Using Eq. (5.4) in combination with Eq. (5.3) it is possible to calculate the differential scattering cross section for spherically symmetric molecules. However, real molecules are not spherically symmetric but slightly elongated. This fact leads to a more complicated description of the scattering process. More precisely, the elongated shape of real molecules leads to rotational and vibrational Raman scattering (a further scattering process, explained in Sect. 5.2.2) and a slight *depolarization* of the scattered light, as the induced dipole moment is not necessarily in the direction of the applied electrical field. To describe the scattering process on real molecules, it is thus necessary to express the polarizability α mentioned in Eq. (5.2) as a tensor $\boldsymbol{\alpha}$, which is composed of the mean polarizability a and the anisotropy of the polarizability γ. Using these two quantities, and considering a random orientation of molecules, the differential Rayleigh scattering cross section is given by Miles et al. (2001).

$$\left(\frac{\partial\sigma}{\partial\Omega}\right)_{\text{Rayleigh}} = \frac{\pi^2}{\varepsilon_0^2 \cdot \lambda^4}\left(a^2 + \frac{7}{45}\gamma^2\right) \cdot \sin^2(\Phi) \tag{5.5}$$

With Eq. (5.5) it is obvious that the differential Rayleigh scattering cross section is proportional to λ^{-4}, a fact which was used by Lord Rayleigh to explain the blue color of the sky. Furthermore, it can be seen (see also Fig. 5.1) that Rayleigh scattering is composed of inelastic rotational Raman scattering, which is covered by the $(7/45) \times \gamma^2$ term in Eq. (5.5), and of quasi elastic Rayleigh–Brillouin scattering (Cabannes line), which is covered by the a^2 term in Eq. (5.5). For an estimate of the ratios of the last-mentioned scattering mechanisms within the atmosphere, one can use the mean polarizability $a^2 = 3.92 \times 10^{-80}$ m^6 and its anisotropy $\gamma^2 = 6.44 \times 10^{-81}$ m^6 for molecular nitrogen N_2 (Wandinger 2005). In doing so, one can immediately see that the contribution of the entire rotational Raman scattering to Rayleigh scattering is only about 2.5 % (see also Fig. 5.5).

5.2.2 Raman Scattering

The phenomenon of Raman scattering was discovered and first described by the Indian physicists Chandrasekhara Raman and Kariamanickam Krishnan in 1928 (Raman and Krishnan 1928). Similar to Rayleigh scattering, the discovery of Raman scattering was driven by a natural phenomenon, namely the attempt to explain the deep blue color of oceanic water. Nowadays we know that it is partly due to sunlight scattering on water molecules, the stronger absorption of red

wavelengths, which primarily arises from vibration modes of the water molecule's O–H bonds, and the reflection of the color of the (blue) sky (Dickey et al. 2011).

Raman scattering in molecular gases originates from light interaction with resonant modes of the molecule, and thus a quantum mechanical description of the scattering process is helpful and even needed. If a photon of incidenting light is scattered on a molecule which thereby simultaneously undergoes a transition between internal energy states (vibrational or rotational), its frequency changes. If the molecule thereby absorbs energy during the scattering process, i.e., if a higher energy level is excited, the frequency of the scattered photon is decreased (equivalent to a wavelength shift to longer wavelengths, the so-called red shift). This process is called Stokes Raman scattering. If the molecule transfers energy to the scattered photon by decreasing its internal energy, the frequency of the scattered light is increased (equivalent to wavelength shift to shorter wavelengths, the so-called blue shift). This process is called anti-Stokes Raman scattering. Thus, Raman scattering is an inelastic scattering process, since the frequency of scattered light changes. To be able to compare the different spectral components of the various scattering mechanisms, some details of the Raman scattering process on diatomic molecules are discussed in the following. An extensive overview of Raman scattering is given by Long (2002).

If f_i and f_s are the frequencies of the incident and the scattered light, respectively, the frequency shift Δf which occurs during the Raman scattering process is given by

$$\Delta f = |f_i - f_s| = \frac{\Delta E}{h} \tag{5.6}$$

where h is the Planck's constant and ΔE the energy difference between the involved molecular energy levels (transitions between different rotational and vibrational states). Thus, to be able to calculate the occurring frequency changes within the scattering process, one has to calculate the possible energy levels. In the case of diatomic molecules like N_2 and O_2, this can be done by modeling the molecules as freely rotating harmonic oscillators. In doing so, the Raman scattering cross section can be determined with a few characteristic molecular quantities, namely the mean polarizability a, the anisotropy of the polarizability γ, and their derivatives with respect to the normal coordinate of vibration at the equilibrium positions a' and γ', which characterize the changing distance between the vibrating atoms. For example, the differential rotational Raman scattering cross section is described by the anisotropy of the polarizability γ, as already described with the $(7/45) \times \gamma^2$ term in Eq. (5.5). Details about the calculation of the differential vibrational Raman scattering cross section using a' and γ' can be found in (Wandinger 2005).

For an even more detailed derivation of the spectral distribution of Raman lines, it is furthermore necessary to consider the temperature-dependent population distribution of internal energy levels, the degeneracy of the initial rotational energy level, the nuclear spin degeneracy, and the spin statistical weight factor. If this is

done, the Raman differential scattering cross section, and therefore the contribution to the spectral distribution of molecular scattered light, can be calculated. An example of a highly resolved Raman spectrum for nitrogen molecules at ambient temperature 300 K and scattering angle $\theta = 180°$ is illustrated in Fig. 5.5 and further discussed in Sect. 5.3.

5.2.3 Rayleigh–Brillouin Scattering

To be able to analyze the spectral distribution of Rayleigh–Brillouin scattered light, one has to use an approach which goes back to Albert Einstein, who suggested that light scattering can be viewed as originating from density fluctuations in inhomogeneous media. The spectral component of the scattered light which is caused by these density fluctuations is called the Cabannes line, named after Jean Cabannes, who first derived the intensities of the central and rotational Raman shifted lines and showed that the sum of both is exactly Rayleigh's result (given with Eq. (5.5)). The scattering process which yields the Cabannes line is commonly termed Rayleigh–Brillouin scattering.

Density fluctuations in gases arise from the thermal motion of molecules and lead to fluctuations in the optical dielectric constant ε, which in turn leads to fluctuations in the index of refraction n_r ($n_r = (\varepsilon)^{1/2}$ in a nonmagnetic medium with unit relative permeability), and finally, to the scattering of light. The density fluctuations in a gas can thereby be divided into thermally generated *entropy* fluctuations (at constant pressure) and pressure fluctuations (at constant entropy). In addition, the pressure fluctuations can be viewed as random local compressions and rarefactions which, as a consequence of the elastic properties of the medium, do not remain fixed in position but travel throughout the volume of the medium. Thus, the numerous random compressions and rarefactions can be considered as elastic *sound waves* (acoustic phonons) of various frequencies, which propagate with the corresponding velocity of sound in all possible directions. When laser light enters the medium, every sound wave scatters the incident light wave at angles for which constructive *interference* occurs (Bragg condition). Furthermore, the frequency of the scattered light is shifted by the Doppler effect due to the motion of the sound wave responsible for light scattering. This type of inelastic scattering mechanism was independently described by Leon Brillouin in 1922 and Leonid Mandelstam in 1926 and is therefore commonly called Brillouin scattering in the literature. It is worth mentioning that the term Brillouin scattering is also used in connection with light scattering in solid states, where lattice fluctuations (e.g., acoustic phonons) are the origin of the scattering process. For further analysis of the Brillouin scattering process, the Bragg scattering analogy is considered, as pictured in Fig. 5.2.

Incident and scattered light are defined by their wave vectors $\mathbf{k_i}$, $\mathbf{k_{sc}}$ and frequencies f_i, f_{sc}, respectively, and are pictured with solid black and dashed black arrows. The scattering angle is denoted by θ. The traveling sound wave which is responsible for the scattering of $\mathbf{k_i}$ is defined by the wave vector $\mathbf{k_s}$, frequency f_s and

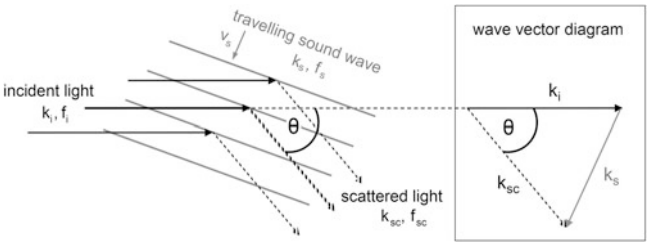

Fig. 5.2 Schematic diagram illustrating the Brillouin scattering process. The *black arrows* indicate incident light, the *black dashed arrows* scattered light, and the *gray lines* the traveling sound wave caused by pressure fluctuations and responsible for the Brillouin scattering process. The *box* on the *right* shows the corresponding wave vector diagram

phase velocity v_s, and is indicated with gray lines. Because of momentum conservation within the scattering process ($\mathbf{k_s = k_i - k_{sc}}$), the wave vector of the acoustic phonons can be calculated using the cosine law (Fig. 5.2, wave vector diagram)

$$|\mathbf{k_s}|^2 = |\mathbf{k_i}|^2 + |\mathbf{k_{sc}}|^2 - 2 \cdot |\mathbf{k_i}| \cdot |\mathbf{k_{sc}}| \cos(\theta) \tag{5.7}$$

As the change in frequency of the scattered light is quite small (e.g., $2{\cdot}10^{-6}$ for an incidenting wavelength of 355 nm and a scattering angle of 180°), one can approximate $|\mathbf{k_{sc}}| \sim |\mathbf{k_i}|$. Using $1 - \cos(\theta) = 2{\cdot}\sin^2(\theta/2)$, Eq. (5.7) reduces to

$$|\mathbf{k_s}| = \pm 2 \cdot |\mathbf{k_i}| \cdot \sin\left(\frac{\theta}{2}\right) \tag{5.8}$$

This gives the possibility to calculate the frequency shift f_s between incident and scattered light, and therefore the spectral location of the Brillouin peaks (Brillouin doublet, see also Fig. 5.5, bottom, gray line). By definition, $|\mathbf{k_i}| = 2{\cdot}\pi{\cdot}n_r/\lambda_i$, with λ_i the wavelength of the incident light, n_r the index of refraction of the medium, and $\lambda_s = c_s/f_s$, with c_s and f_s the sound velocity and sound frequency, respectively. Thus, the frequency shift f_s due to Brillouin scattering is

$$|f_s| = 2 \cdot \frac{c_s \cdot n_r}{\lambda_i} \cdot \sin\left(\frac{\theta}{2}\right) \tag{5.9}$$

From Eq. (5.9), it can be seen that f_s reduces to zero in the forward direction ($\theta = 0°$) and that it is at a maximum in the backward direction ($\theta = 180°$). Therefore, the frequency shift due to Brillouin scattering is largest for lidar applications in backscattering geometry. For example, by considering a velocity of sound in air of about 340 m/s, the frequency shift of Brillouin scattered light is about 1.9 GHz for lidars operating at a wavelength of 355 nm ($n_r \sim 1$, $\theta = 180°$).

In addition to Brillouin scattering, which arises from pressure fluctuations in the medium, there is a scattering mechanism associated with thermally generated entropy fluctuations. These entropy fluctuations can likewise be viewed as waves

which lead to a scattering of light if the Bragg condition is fulfilled. However, entropy fluctuations are static in space and therefore do not produce a frequency *Doppler shift* for the scattered light. This part of the Cabannes line was first resolved by Eugene Gross in 1930 and theoretically described by Lev Landau and Georg Placzek in 1934. It is therefore commonly called the Gross or Landau-Placzek line (see also Fig. 5.1) and leads to the actual elastic part of the spectrum of molecular scattered light.

As is obvious from the discussion above, there are two different kinds of density fluctuations (entropy and pressure) leading to different parts of the spectrum of molecular scattered light, namely the Brillouin doublet and the Landau-Placzek line. Whereas the entropy fluctuations are caused by the thermal motion of the molecules, pressure fluctuations arise from collisions between different molecules. Thus, the effect of Brillouin scattering becomes more prominent when the number of collisions between molecules increases, or, rather, when the mean free path between molecules decreases. A key parameter for describing which of these two scattering mechanisms is dominant is the ratio y of the scattering wavelength $2 \cdot \pi / |\mathbf{k_s}|$ to the mean path of collisions

$$y = \frac{p}{\mathbf{k_s} \cdot v_0 \cdot \eta} = \frac{n \cdot k_B \cdot T}{\mathbf{k_s} \cdot v_0 \cdot \eta} \tag{5.10}$$

with n the number density, T the gas temperature, p the gas pressure, v_0 the most probable thermal velocity, and η the shear viscosity. For the description of the spectral distribution of the Cabannes line, one uses the y parameter to define different scattering regimes in order to apply approximations in the calculation of the respective spectral line shape.

For $y >> 1$, the mean free path between collisions is much smaller than the scattering wavelength, as is true for dense gases. In that case, the molecular gas can be treated as a continuum and the density fluctuations can be calculated using the *Navier–Stokes equations*. The resulting spectrum is well approximated by the sum of three Lorentzian functions displaced by f_s (Eq. (5.9)). Here, the central, unshifted peak (Landau-Placzek) arises from scattering on entropy fluctuations, whereas the shifted peaks (the Brillouin doublet) are due to scattering on pressure fluctuations. It is worth mentioning that the width of each of the three Lorentz peaks is proportional to the lifetime of the respective density fluctuation (connected via Fourier's Law). The spectrum for dense gases is indicated by the black region in Fig. 5.3.

For $y << 1$, the mean free path is much larger than the scattering wavelength, as is true for very dilute gases. In that case, scattering is solely due to individual molecules and the scattering from pressure fluctuations is neglectable. The line profile of the scattered light is described by a *Gaussian* function derived from the Maxwell velocity distribution, as shown by the light gray area in Fig. 5.3.

For visible light and for conditions relevant for atmospheric temperatures and pressures, however, the y parameter is in the order of 1 (Eqs. 5.7–5.10). In this case, neither the Gaussian function nor the sum of Lorentzians leads to a satisfying description of the spectrum of the scattered light. To calculate the spectrum for

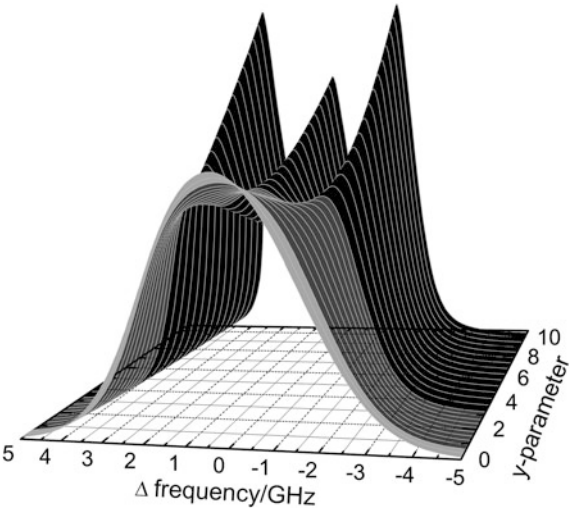

Fig. 5.3 Evolution of the Cabannes line shape of Rayleigh–Brillouin scattered light in nitrogen N_2 (wavelength $\lambda = 355$ nm, scattering angle $\theta = 180°$), calculated using the Tenti model (Tenti et al. (1974))

atmospheric conditions it is necessary to find solutions of the Boltzmann transport equation, which statistically describes the distribution of molecules in a gas. However, because of the complexity of the Boltzmann equation, no analytical solution is possible for molecular gases, and thus, approximative models have to be derived. The most popular and currently best model for that purpose was derived by Tenti et al. (1974) and is therefore called the *Tenti model*. The Cabannes line for atmospheric relevant conditions ($y \sim 1$), calculated by using the Tenti model, is shown by the dark gray area in Fig. 5.3.

It is worth mentioning that the Tenti model was derived for molecular gases of single species. However, atmospheric air is a mixture of several gases, mainly nitrogen and oxygen. That is, collisions between molecules of different species can occur, and thus, it cannot be assured that the Tenti model is still applicable. To investigate the applicability of the Tenti model in air, the first validation measurements have recently been performed at VU University Amsterdam in collaboration with DLR's Institute of Atmospheric Physics. By accurately measuring the Cabannes line shape of UV light scattered in air, it was shown that the accordance between Tenti model and measurement is better than 2 % (Witschas et al. 2010). This result confirms that the Tenti model can be used to accurately describe the Cabannes line of light scattered in air. An example of the measured Cabannes line for air at an ambient pressure of 500 hPa and a temperature of 298 K (gray dots) is shown in Fig. 5.4.

The black line shows the corresponding line shape calculated with the Tenti model, whereas the gray line shows the spectrum calculated using the Gaussian function (only considering the thermal motion of molecules and neglecting the effect of Brillouin scattering). It can be clearly seen that the measured Cabannes line is described by the Tenti model with deviations smaller than 2 % with respect to the measured peak intensity, whereas the deviations to the Gaussian function are

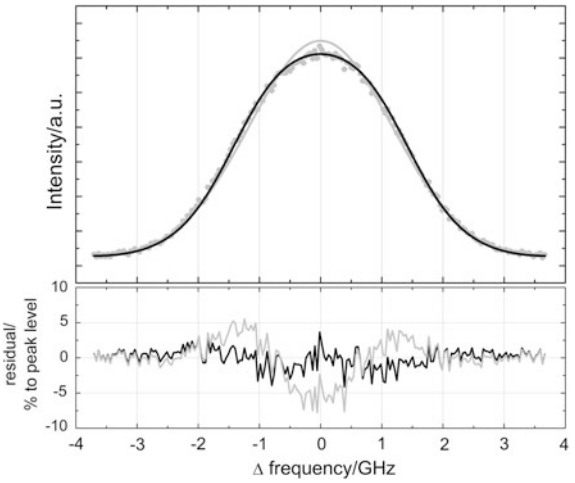

Fig. 5.4 Measured Cabannes line in air (*gray dots*) compared to the Tenti model (*black line*) and the Gaussian function (*gray line*) for a wavelength of $\lambda = 366.5$ nm, a pressure of $p = 500$ hPa and a temperature of $T = 298$ K (Witschas et al. 2010)

larger than 5 %. This fact clearly demonstrates that the effect of Brillouin scattering has to be taken into account for atmospheric applications, eventhough the mathematical description of the spectrum using the Tenti model is more complicated than one using an ordinary Gaussian function.

To eliminate the mathematical complexity of the Tenti model but still benefit from its correctitude, an analytical model of (almost) the same quality was recently developed by Witschas (2011) at the Institute of Atmospheric Physics. This analytical model enormously simplifies the calculation of the Rayleigh–Brillouin line shape in air without losing validity. For y-parameters between 0 and 1, the deviation between the analytical model and the Tenti model is smaller than 0.85 % (proof of that statement can be found in Witschas 2011).

5.3 The Different Spectral Components of Molecular Scattered Light and Its Application in Atmospheric Science

In the previous section it was shown that light propagating in molecular gases is simultaneously scattered due to several processes. Thereby, each of these scattering processes makes its characteristic imprint on the intensity, the polarization and the spectral composition of the scattered light. In other words, the spectrum of molecular scattered light contains information about molecular state variables like temperature and density, and the flow velocity, as they are influencing each scattering mechanism. Thus, molecular light scattering offers a unique possibility to nonintrusively measure these quantities in gases, e.g., in atmospheric air. To illustrate this circumstance, the spectral components of the different scattering mechanisms are discussed for the example of ultraviolet laser light scattered in nitrogen N_2 ($\lambda = 355$ nm, $T = 300$ K, $\theta = 180°$). The corresponding spectrum,

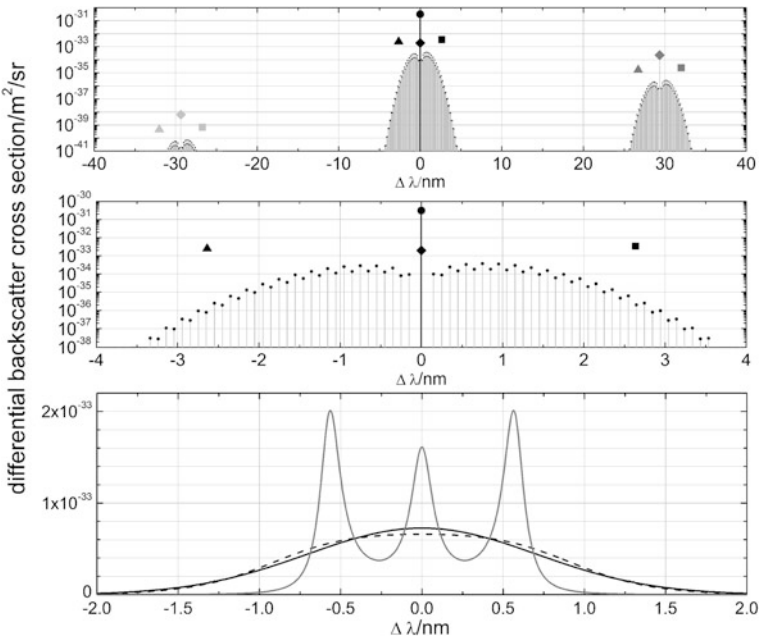

Fig. 5.5 Simulated differential backscattering ($\theta = 180°$) cross section of laser light ($\lambda = 355$ nm) scattered in nitrogen N_2 (T = 300 K). Each graph shows the same data but with a different resolution. *Top*, *middle* resolved vibration–rotation and rotation Raman lines (log-scale). *Bottom* resolved Cabannes line for different pressure values (linear scale). The integral of the Cabannes line represents the scattering cross section already given in the *top* and *middle* graph

or more precisely the differential scattering cross sections of the different scattering processes, is calculated as explained in Sect. 5.2 and plotted in Fig. 5.5.

The upper graph in Fig. 5.5 shows resolved Stokes and anti-Stokes vibration–rotation Raman lines (black dots, gray lines) as well as their sum (light gray and dark gray squares, triangles and diamonds, respectively). It is obvious that they are shifted by about 30 nm from the incident wavelength of 355 nm. This quite large wavelength shift is due to the large energy difference between the vibrational energy levels involved. Furthermore, it can be seen that the scattering cross section of the anti-Stokes vibration–rotation Raman lines is about six orders of magnitude smaller than the one of the Stokes vibration–rotation Raman lines. This is characteristic for light scattering in molecular gases with atmospherically relevant temperatures where almost all molecules are in the vibrational ground state. Hence, vibrational Raman scattering mainly occurs when molecules absorb energy to excite a higher vibrational level, but not vice versa, as there are almost no vibrational levels excited.

Due to the low scattering cross section, anti-Stokes vibration–rotation Raman lines do not play any role in atmospheric applications. However, the Stokes vibration Raman lines can be used to determine the atmospheric water vapor mixing ratio and the aerosol optical extinction coefficient (Wandinger 2005).

In the middle graph (zoom of the top graph by a factor of 10), the resolved Stokes and anti-Stokes rotation Raman lines (black dots, gray lines) as well as their sum (black square, triangle, and diamond) are shown. In contrast to the vibration–rotation Raman lines, the scattering cross section of the Stokes and anti-Stokes rotation Raman lines is almost equal. This is because of the lower energy difference between rotational energy levels. Hence, there are enough excited rotational energy levels and the Raman scattering process can occur in both directions. Furthermore, the wavelength shift is smaller compared to vibrational Raman scattering. The wavelength shift of the first rotational Raman line is about 0.15 nm with respect to the fundamental wavelength of 355 nm, and therefore, narrowband filters can be used to separate the Cabannes line from rotational Raman scattering. Moreover, it is obvious that there is a frequency-unshifted rotational Raman signal (black diamond), which is called the Q-branch of rotational Raman scattering. This part of the spectrum arises from transitions between degenerated states with different magnetic quantum numbers. The scattering cross section of the unshifted rotation Raman Q-branch is about two orders of magnitude smaller than the one of the Cabannes line (black dot). In particular, for the example shown, the contribution of the rotation Raman Q-branch to the Cabannes line is 0.64 %.

A common atmospheric application of the rotational Raman signal is, e.g., atmospheric temperature measurement with *Raman lidar* (Weitkamp 2005). As the spectral distribution of the single rotational Raman lines depends on the population distribution of internal energy levels, and therefore on temperature, it is possible to derive the atmospheric temperature from the resolved rotational Raman lines (Fig. 5.5, middle graph, gray lines between ± 0.2 and ± 3.5 nm). Nowadays, with the improvement of laser sources and filter techniques, atmospheric temperature can be determined with uncertainties in the order of $2-3$ K for altitudes between 0 and 25 km. However, the quite low rotational Raman scattering cross section requires high power laser systems and averaging times of several 10's of minutes. To avoid disturbances from sunlight, most Raman lidars operate during nighttime.

The lower graph (zoom of the middle graph by a factor of 2 000) in Fig. 5.5 shows the Cabannes line for different gas pressure values (compare also with Fig. 5.3). The black solid curve is representative for very dilute gases ($p = 10^{-6}$ hPa). The spectrum of the scattered light is described by a Gaussian line shape according to the Maxwell velocity distribution. The gray curve shows the line shape representative of dense gases ($p = 10\,000$ hPa). At such high pressures, the effect of Brillouin scattering is strongly pronounced, and the central Landau-Placzek line can be easily distinguished from the wavelength-shifted Brillouin doublet. The wavelength shift of the Brillouin peaks is approximately 0.55 nm, which corresponds to a frequency shift of 1.8 GHz and thus to a sound velocity of 320 m/s. The black dashed curve shows the line shape representative of gases at atmospheric conditions ($p = 1\,000$ hPa). In this intermediate region, molecular gas dynamic equations have

to be used to calculate the spectrum of the scattered light. The best model for that purpose is the Tenti model (Tenti et al. 1974), which is also used here.

Due to the relatively large scattering cross section, Rayleigh–Brillouin scattering is a widely used tool in atmospheric applications such as, for example, atmospheric temperature measurements (Shimizu et al. 1983) and measurements of aerosol optical properties (Esselborn et al. 2008; Liu et al. 2009) using *high spectral resolution lidars*. Thereby, high resolution filters are used to resolve the Cabannes line shape, which is subsequently compared to the Tenti model.

Another atmospheric application that utilizes Rayleigh–Brillouin scattering is the measurement of atmospheric wind using *Doppler wind lidars* (Flesia and Korb 1999; Reitebuch et al. 2009). This technology uses the fact that the proper motion of air molecules, caused by the wind velocity, causes a frequency shift of the entire Cabannes line shape. Thus, if this frequency shift is carefully determined, wind speeds can be derived with an accuracy of 1 m/s. This technology is straightforward and developed to the extent that the first Doppler wind lidar will operate from space in 2014, in the frame of ESA's ADM Aeolus mission.

References

Amediek, A., Fix, A., Wirth, M., Ehret, G.: Development of an OPO system at 1.57 μm for integrated path DIAL measurement of atmospheric carbon dioxide. Appl. Phys. B **92**, 295–302 (2008). doi:10.1007/s00340-008-3075-6

Dickey, T., Kattawar, G., Voss, K.: Shedding new light on light in the ocean. Phys. Today **44–49** (2011). doi:10.1063/1.3580492

Elliott, G., Glumac, N., Carter, C.: Molecular filtered Rayleigh scattering applied to combustion. Meas. Sci. Technol. **12**, 452–466 (2001). doi:10.1088/0957-0233/12/4/309

Esselborn, M., Wirth, M., Fix, A., Tesche, M., Ehret, G.: Airborne high spectral resolution lidar for measuring aerosol extinction and backscatter coefficients. Appl. Opt. **47**, 346–358 (2008). doi:10.1364/AO.47.000346

Flesia, C., Korb, C.: Theory of the double-edge molecular technique for Doppler lidar wind measurement. Appl. Opt. **38**, 432–440 (1999). doi:10.1364/AO.38.000432

Liu, B.-Y., Esselborn, M., Wirth, M., Fix, A., Bi, D.-C., Ehret, G.: Influence of molecular scattering models on aerosol optical properties measured by high spectral resolution lidar. Appl. Opt. **48**, 5143–5154 (2009). doi:10.1364/AO.48.005143

Long, D.: The Raman Effect: A Unified Treatment of the Theory of Raman Scattering by Molecules, p. 584. Wiley, New York (2002). doi:10.1002/0470845767

Miles, R.B., Lampert, W.R., Forkey, J.N.: Laser Rayleigh scattering. Meas. Sci. Technol. **12**, R35–R51 (2001). doi:10.1088/0957-0233/12/5/201

Raman, C., Krishnan, K.: A new type of secondary radiation. Nature **121**, 501–502 (1928). doi:10.1038/121501c0

Rayleigh, L.: On the scattering of light by small particles. Philos. Mag. **41**, 447–454 (1871). doi:10.1080/14786447108640507

Rayleigh, L.: On the transmission of light through an atmosphere containing small particles in suspension, and on the origin of the blue of the sky. Philos. Mag. **47**, 375–384 (1899). doi:10.1080/14786449908621276

Reitebuch, O., Lemmerz, C., Nagel, E., Paffrath, U., Durand, Y., Endemann, M., Fabre, F., Chaloupy, M.: The Airborne Demonstrator for the Direct-Detection Doppler Wind Lidar ALADIN on ADM-Aeolus. Part I: Instrument Design and Comparison to Satellite Instrument. J. At. Ocean. Technol. **26**, 2501–2515 (2009). doi:10.1175/2009JTECHA1309.1

Seasholtz, R., Buggle, A., Reeder, M.: Flow measurements based on Rayleigh scattering and Fabry-Perot interferometer. Opt. Laser. Eng. **27**, 543–570 (1997). doi:10.1016/S0143-8166(96)00063-2

Shimizu, H., Lee, S., She, C.: High spectral resolution lidar system with atomic blocking filters for measuring atmospheric parameters. Appl. Opt. **22**, 1373–1381 (1983). doi:10.1364/AO.22.001373

Tenti, G., Boley, C.D., Desai, R.C.: On the kinetic model description of Rayleigh–Brillouin scattering from molecular gases. Can. J. Phys. **52**, 285–290 (1974). doi:10.1139/p74-041

Wandinger, U.: Raman Lidar. In: Weitkamp, C. (ed.) Range-Resolved Optical Remote Sensing of the Atmosphere, pp. 241−271. Springer, New York (2005)

Weitkamp, C.: Lidar: Range-resolved optical remote sensing of the atmosphere, p. 455. Springer, New York (2005). doi:10.1007/b106786

Wirth, M., Fix, A., Mahnke, P., Schwarzer, H., Schrandt, F., Ehret, G.: The airborne multi-wavelength water vapor differential absorption lidar WALES: system design and performance. Appl. Phys. B **96**, 201–213 (2009). doi:10.1007/s00340-009-3365-7

Witschas, B., Vieitez, M.O., van Duijn, E.-J., Reitebuch, O., van de Water, W., Ubachs, W.: Spontaneous Rayleigh–Brillouin scattering of ultraviolet light in nitrogen, dry air, and moist air. Appl. Opt. **49**, 4217–4227 (2010). doi:10.1364/AO.49.004217

Witschas, B.: Analytical model for Rayleigh–Brillouin line shapes in air. Appl. Opt. **50**, 267–270 (2011). doi:10.1364/AO.50.000267

Young, A.T.: Rayleigh scattering. Appl. Opt. **20**, 533–535 (1981). doi:10.1364/AO.20.000533

Greenhouse Effect, Radiative Forcing and Climate Sensitivity

6

Michael Ponater, Simone Dietmüller and Robert Sausen

Abstract

Temperature conditions and climate on Earth are controlled by the balance between absorbed solar radiation and outgoing terrestrial radiation. The greenhouse effect is a synonym for the trapping of infrared radiation by radiatively active atmospheric constituents. It generally causes a warming of the planet's surface, compared to the case without atmosphere. Perturbing the radiation balance of the planet, e.g., by anthropogenic greenhouse gas emissions, induces climate change. Individual contributions to a total climate impact are usually quantified and ranked in terms of their respective radiative forcing. This method involves some limitations, because the effect of the external forcing is modified by radiative feedbacks. Here the current concept of radiative forcing and potential improvements are explained.

M. Ponater (✉) · S. Dietmüller · R. Sausen
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Michael.Ponater@dlr.de

S. Dietmüller
e-mail: Simone.Dietmueller@dlr.de

R. Sausen
e-mail: Robert.Sausen@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_6,

6.1 The Greenhouse Effect as Part of Earth's Radiation Balance

Earth is continuously receiving energy from the *Sun*, and it is the balance between incoming *solar* (*shortwave*) *radiation* and emitted terrestrial (*longwave*) *radiation* that sets the framework for the *climate* on our planet. The solar radiative input to Earth is almost constant on long-term average, as its fluctuations are controlled mainly by periodic processes like the annual variation in the distance between Sun and Earth or the 11-year sun-spot cycle. This causes a quasi-stationary global climate, as long as some terrestrial parameters like Earth's rotation, land-sea distribution, and the *composition* of the *atmosphere* remain unchanged, too. The notion of a radiative flux balance between *absorbed solar radiation* and *outgoing thermal radiation* (at the top of the atmosphere) and a corresponding quasi-stationary climate only makes sense in the global and long-term mean. For certain regions on Earth, there are always imbalances and also extensive variability, not only on the seasonal (intra-annual) *time scale* but also from year to year (inter-annual time scale).

The terrestrial component of Earth's radiation balance is governed by the *Stefan-Boltzmann law* of radiation

$$E = \varepsilon \sigma T^4 \tag{6.1}$$

that indicates a linear dependence of the radiation energy emitted per unit surface area and time (E) on the *emissivity* (ε) of the absorber and a nonlinear dependence on the *temperature* (T) of the absorber, with σ being the Stefan-Boltzmann constant. A so-called *black body* ($\varepsilon = 1$) absorbs all radiation it receives and re-emits it with its own temperature, while a grey body ($0 < \varepsilon < 1$) can be regarded as being semitransparent for incoming radiation. ε in Eq. 6.1 indicates an integrated emissivity, often comprising wavelength ranges within the electromagnetic spectrum where the absorber is opaque and others where it is transparent.

The existence of an atmosphere is of great importance for the radiation balance of Earth and has essential implications for the temperature at the planet's surface. Without an atmosphere solar *insolation* is the only radiative source for the Earth's surface. If an atmosphere containing absorptive components surrounds the Earth, part of the *thermal* radiation emitted by the surface (due to its temperature and emissivity) is absorbed and re-emitted by the atmosphere, increasing the energy gain of the surface and inducing a temperature rise there. This mechanism, called the *greenhouse effect*, is of substantial relevance for Earth, whose atmosphere contains several gases absorbing in the thermal spectral range, such as *water vapor*, *carbon dioxide*, *ozone*, *methane*, and *nitrous oxide*. Water vapor makes the largest contribution to the greenhouse effect. Some gases, notably ozone, also absorb sunlight, an effect which warms the atmosphere but (instantaneously) cools the Earth's surface. Furthermore, the atmosphere contains some liquid and solid components (*cloud* water, cloud ice, *aerosols*) which contribute to the greenhouse warming, but which also absorb or *backscatter* incoming solar radiation. Despite such partial

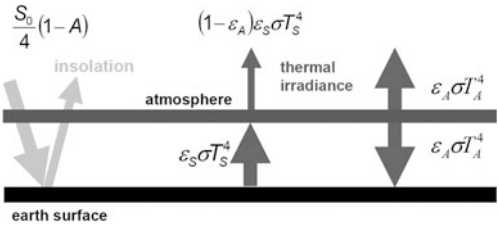

Fig. 6.1 Simplified model of Earth's global *climate system* containing two layers (atmosphere and surface)

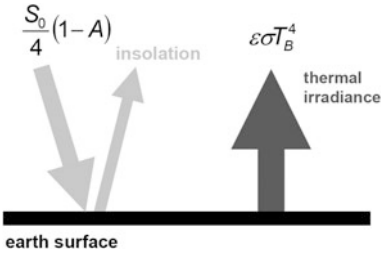

Fig. 6.2 Simplified global one-layer model of a fictive Earth without atmosphere

compensation the net effect of the Earth's atmosphere acts to warm the surface. As illustrated in Fig. 6.1, this is very instructively described by a simple analytical *climate model* consisting of just two layers: an atmosphere with global mean temperature T_A and emissivity ε_A and a surface with global mean temperature T_S and emissivity ε_S.

Insolation is represented by the *solar constant* $S_0 = 1365$ W/m^2 (to be reduced by a factor 4 as only one fourth of the Earth's actual surface is effectively irradiated at a given time), of which only a fraction of $(1 - A)$ is absorbed within the surface-atmosphere system. A is the planetary *albedo*, which indicates the fraction of total solar *irradiance* that is reflected. It amounts to approximately 30 % according to current knowledge from *satellite* and other measurements. The thermal irradiance emitted by the atmosphere and the surface is described by Eq. 6.1. This simplified model thus reduces the climate system to two equations for the two unknowns, T_S and T_A:

$$(1 - A)\, S_0/4 + \varepsilon_A \sigma\, T_A^4 = \varepsilon_S \sigma\, T_S^4 \tag{6.2}$$

$$(1 - \varepsilon_A)\varepsilon_S \sigma T_S^4 + 2\varepsilon_A \sigma T_A^4 = \varepsilon_S \sigma T_S^4 \tag{6.3}$$

.

These can be solved analytically and deliver reasonable results if some realistic assumptions for the emissivity of atmosphere and surface are made. Using values of $\varepsilon_S = 0.95$ and $\varepsilon_A = 0.75$, the results are $T_S = 14$ °C and $T_A = -33$ °C for *surface temperature* and atmospheric temperature, respectively. This comes surprisingly close (given the simplicity of the model) to the global mean values yielded through present-day observations from space and in situ.

Removing the atmosphere from the model gives an impression of how important the greenhouse effect is for the surface temperature. What remains without the atmosphere is a simple global mean radiation balance model for the Earth's surface (Fig. 6.2). Retaining $A = 0.3$ and $\varepsilon = \varepsilon_S = 0.95$, as above, the calculated surface temperature becomes $T_S = -15$ °C, indicating a drop of about 30 K as a result of removing the atmosphere's greenhouse effect.

It is important to note that this theoretical estimate can be regarded as reasonable in a qualitative and illustrative sense, but not as a value describing the impact of the greenhouse effect in the real world exactly and completely. Due to the role of clouds and changes in water, ice and land surface *coverage* for an Earth without atmosphere, both A and ε_S would certainly be different from the case with atmosphere. Moreover, even for the observed planet the estimate of representative global long-term averages for A, ε_S, and ε_A (as used in Eq. 6.1) (Clough and Iacono 1995) forms a difficult problem because of spatial, temporal, and spectral variability. In nature, the global temperature and *humidity* distribution is influenced by dynamics and is also strongly correlated. Air is warmer and moister in summer compared to winter, in tropical compared to extra-tropical latitudes, and near the surface compared to altitudes in the free and *upper troposphere*.

For the purpose of realistic calculations of irradiances, climate, *climate variability*, and *climate change* on Earth, complex three-dimensional (3D) models (evaluated with respective 3D observations) must be used that cover the variability of the relevant parameters and processes as well as the dynamic, physical, and chemical interactions between them. In principle, higher concentration of an atmospheric greenhouse gas simply increases ε_A in Eqs. 6.2 and 6.3, initiating further surface warming. However, just calculating the respective ε_A change would not suffice to determine a reliable change of T_S from the simple model. As will be discussed in Sect. 6.3, additional knowledge of all radiative *feedbacks* (also modifying ε_A, and A as well) would be necessary, too.

6.2 Climate Variability and Anthropogenic Climate Change

Climate on Earth is not in equilibrium and has never been so. First, Earth's radiation balance always changes because of natural variations in the external boundary conditions (insolation, change of atmospheric composition from strong *volcanic eruptions*, meteorite impacts, etc.). Second, even without such events internal natural variability occurs on a variety of time scales. It is caused by the complexity of dynamic, physical and chemical processes in the atmosphere as well as by interactions and feedbacks between components of the Earth system (*ocean*-atmosphere interaction, *vegetation*-atmosphere interaction, etc.) and determines climate as a basically chaotic system. Hence, the annual mean conditions observed in certain locations and regions, but also on the global scale, may undergo significant changes from year to year (inter-annual climatic variability). A prominent feature of the latter is the ENSO cycle (variation between so-called *El-Niño* and *La-Niña*

events) in the tropical Pacific, a result of ocean internal dynamics and ocean–atmosphere interaction.

Distinguishing between climate variability and climate change is not straightforward. Even strong changes in climate conditions such as those occurring on paleoclimatologic time scales (e.g., in connection with ice ages) might be viewed as part of natural climate variability around a mean value defined over millions of years. In recent decades, there has been growing concern that *anthropogenic* activities will induce climate change (with respect to a quasi-stationary climate state over the past several centuries) strong enough to jeopardize the evolution of Earth's biosphere and human civilization in particular. The extent of socioeconomic impacts is not only a matter of the strength of changes in climate parameters such as surface temperature or *precipitation*, but also of the speed by which the changes develop. Taking this perspective it is sensible to quantify current climate change (observed as well as modeled) relative to a pre-industrial climate assumed as quasi-stationary, and to establish significant *climate signals* by statistical analysis based on some 30 or 50 years. In this way, differences between the mean values of relevant climate parameters (the climate signal) are assessed in comparison to background variability (the *climate noise*) by contemporary research on anthropogenic climate change.

6.3 Component Effects, Radiative Forcing of Climate Change, and Climate Feedbacks

The *Intergovernmental Panel on Climate Change* (*IPCC*) provides regular reviews of research results on global climate change. The reviews cover observations and model *simulations* of climate in the past as well as model-simulated *projections* of climate for the future. The most recent *assessment* (IPCC 2007) states that the current change observed for a considerable number of relevant climate parameters since pre-industrial times is very unlikely to be of natural origin. It expects that these changes will continue and increase for the twenty-first century and beyond if anthropogenic *emissions* of some *greenhouse gases* and greenhouse gas *precursors* are not severely reduced. Total anthropogenic climate change (i.e., the effects induced by men) is, hence, strong enough to form a significant signal emerging from the background noise, already now. While many respective climate variables have been analyzed, surface temperature has remained the parameter most frequently considered, both with respect to its global mean and its geographical distribution (pattern change).

Total anthropogenic climate change consists of various component effects, i.e., contributions induced by individual sectors (industry, *transport* systems, agriculture, etc.) or agents (ozone, aerosols, etc.). The component signals are often small in comparison to the background noise level, e.g., for the contributions to surface temperature change. Hence, the individual components are usually quantified in terms of the *radiative forcing*. This is defined as the radiation flux change induced by the concentration change of a greenhouse gas or an aerosol, or by cloud changes related

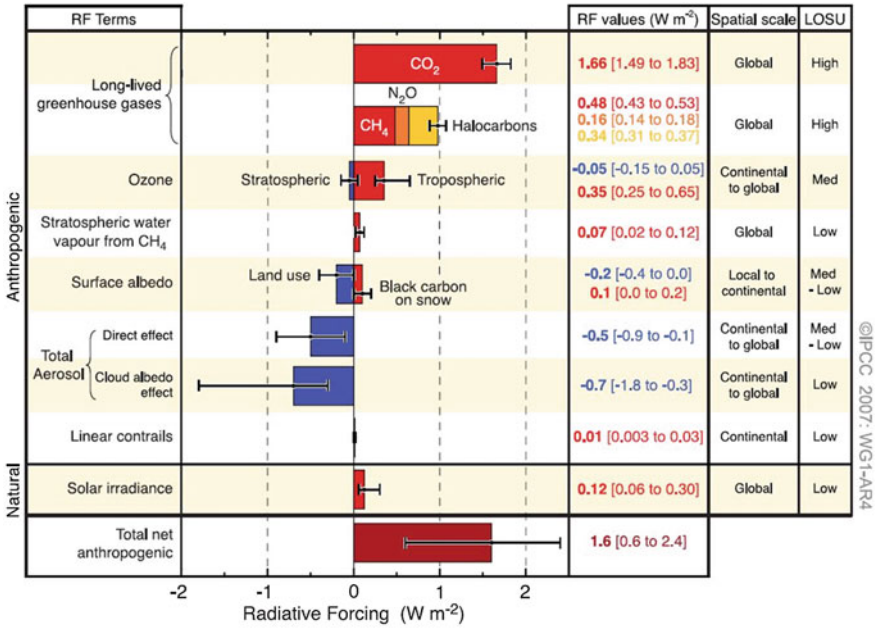

Fig. 6.3 Radiative forcing in 2005. The figure is a reproduction of Figure TS.5a from IPCC (2007) given there with the following description: "Global mean radiative forcings (RF) and their 90 % confidence intervals in 2005 for various agents and mechanisms. Columns on the right hand side specify best estimates and confidence intervals (RF values); typical geographical extent of the forcing (Spatial scale); and level of scientific understanding (LOSU) indicating the scientific confidence level …. Errors for CH_4, N_2O and halocarbons have been combined. The net anthropogenic radiative forcing and its range are also shown. Best estimates and uncertainty ranges can not be obtained by direct addition of individual terms due to asymmetric uncertainty ranges for some factors; the values given here were obtained from a Monte Carlo technique…. Additional forcing factors not included here are considered to have a very low LOSU. Volcanic aerosols contribute an additional form of natural forcing but are not included due to their episodic nature. The range for linear contrails does not include other possible effects of aviation on cloudiness"

(directly or indirectly) to anthropogenic activity. In most cases, radiative forcing can be determined with nearly the same statistical accuracy as the perturbation itself.

Figure 6.3 represents an example from the last IPCC report (IPCC 2007), quantifying the radiative forcing contribution of individual agents. The largest single positive component comes from carbon dioxide increase (CO_2) induced by fossil fuel burning, followed by methane and tropospheric ozone increase. Aerosols provide substantially negative contributions.

The radiative forcing of a component can be used as a proxy of the expected climate change caused by that component. This notion originates from a simple theory of forcing, feedback, and *response* (Hansen et al. 1984, see also Sect. 6.5 below). It relates the externally induced radiative balance change (i.e., the radiative forcing, RF) to the expected equilibrium surface temperature change (ΔT_S) forced by

this change, culminating in a fundamental linear equation linking the global means of these two variables:

$$\Delta T_S = \lambda RF. \tag{6.4}$$

The *climate sensitivity* parameter λ comprises the impact of the radiative feedbacks within the climate system. These denote the modifications to the radiation balance that are induced by the temperature change initiated by the external forcing. Feedbacks are called positive (negative), if they amplify (dampen) the initial effect. The positive water vapor feedback comes from enhanced uptake of water vapor by warmer air, causing an additional greenhouse effect. The ice albedo feedback is also positive. It is created by melting of *snow* and *ice* at the surface as temperature rises, increasing the energy gain of the climate system by reduced *reflection* of solar irradiance. The most important negative radiative feedback is caused by the temperature change itself: temperature changes counteract the disturbance of the radiative balance induced by the forcing by increasing or decreasing *thermal emission* to space.

The actual point of Eq. 6.4 is that the climate sensitivity parameter appears, in first-order approximation, to be constant. This means that λ is largely independent of the amount, distribution (homogeneous or inhomogeneous) or nature (shortwave or longwave, atmospheric or surface directed) of the radiative forcing. With this concept, the global climate response of a component perturbation need not be simulated individually, but can be derived from its radiative forcing. The validity of Eq. 6.4 with constant λ should not be viewed as a physical law but rather as an empirical relationship. It forms an approximation and must be subject to re-evaluation for any new forcing mechanism considered. The concept was established in the early days of climate modeling by a series of simulations (Wetherald and Manabe 1975) that suggested the universal character of λ by comparing the radiative forcing from a CO_2 increase with that of a solar constant change. The climate sensitivity parameters for both types of forcing were nearly equal, if determined from equilibrium climate change simulations with the same model. This apparent equivalence of solar and CO_2 forcing has been repeatedly confirmed by climate model generations up to recent times, as results of model runs performed at DLR (using the ECHAM4 climate model, Fig. 6.4) indicate. The equivalence is so remarkable, because the forcings are very different indeed: the CO_2 effect impacts mainly in the longwave (LW) spectral range and on the atmosphere, the solar effect mainly in the shortwave (SW) spectral range and at the planet's surface.

The approximation of constant λ holds if Eq. 6.4 is evaluated for various forcing mechanisms in one and the same climate model framework. It is important to note, however, that the climate sensitivity parameter may substantially differ between different climate models, mainly because the cloud feedback varies strongly from model to model (Bony et al. 2006). In fact, there are climate models for which the cloud feedback to a CO_2 doubling is positive, but also some where it is negative.

The actual climate sensitivity parameter of the observed climate system is extremely difficult to determine for both principal and practical reasons. At present, it is not known with an accuracy sufficient to constrain the quality of climate models.

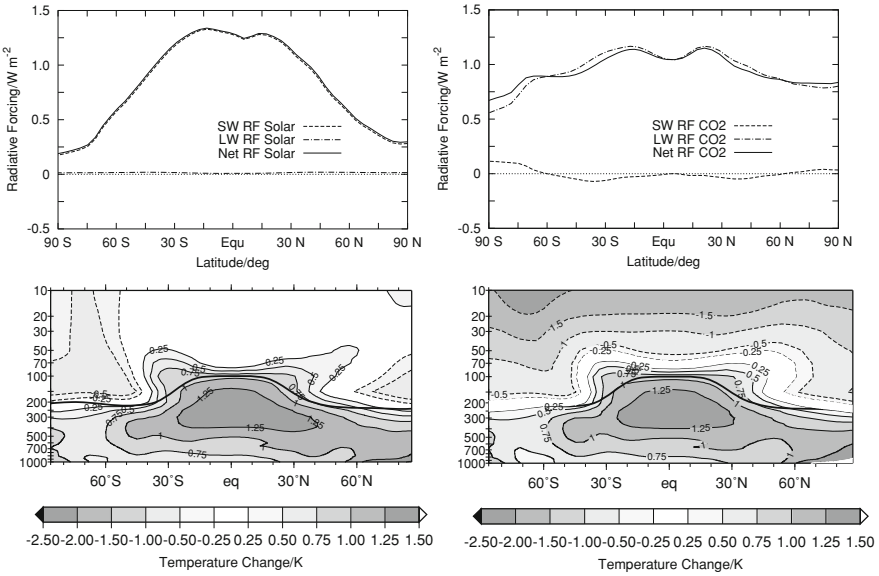

Fig. 6.4 *Zonal* (average over all longitudes) and annual mean radiative forcing from a solar insolation increase by 0.445 % (*top left*) and from a CO_2 concentration increase by 68 μmol/mol (top right, both in W/m^2). The bottom panel shows the resulting zonal mean equilibrium temperature change. The climate sensitivity parameters are 0.74 and 0.73 K/ (W m^{-2}) for solar and CO_2 forcing, respectively. The heavy line indicates the *tropopause*. All results are from simulations with the ECHAM4 climate model (Roeckner et al. 1996; Stuber et al. 2005)

The value from observational studies is uncertain by a factor of 2.5, which nearly coincides with the range of climate sensitivity parameters produced by the various climate models worldwide (Knutti and Hegerl 2008).

The simulations displayed in Fig. 6.4 have been run with forcings normalized to a global annual mean value of 1 W/m^2 for a most straightforward comparison. Some typical features are obvious from the results: besides the shortwave/longwave distinction mentioned above, the zonal profile of both forcing types is also remarkably different, as the equator-to-pole gradient of the solar forcing is stronger than that caused by the homogeneous CO_2 concentration change (due to the characteristic latitudinal dependence of solar insolation). Nevertheless, not only the global mean surface temperature change (and, thus, the climate sensitivity parameter, Eq. 6.4) comes out almost equal. The temperature change pattern within the *troposphere* is also closely similar, both zonally and vertically. Only in the *stratosphere* there are obvious differences, as the stratospheric cooling caused by the spectral characteristics of CO_2 *absorption* (Clough and Iacono 1995) has no counterpart to solar forcing.

Initially, the RF concept was based on instantaneous radiative flux changes at the top of the atmosphere. Here, the radiative imbalance is calculated as soon as the forcing perturbation starts, i.e., before the climate has time to adapt to the disturbance

and create any feedback. However, such a definition of RF does not account for the different response times of forcings above and below the tropopause. In the troposphere/surface system (below the tropopause) the climate response develops slowly on time scales of decades, being delayed by the large heat capacity of the ocean. Strong convective mixing within the troposphere distributes any temperature change swiftly all over the domain, independent of where the primary heating source is actually located. Layers above the tropopause, however, are less influenced by the delay from oceanic thermal inertia. If the radiative heating source partly affects the stratosphere, this domain returns quickly, within few months, to radiative equilibrium. As a consequence, λ values derived according to this concept vary, depending on the vertical profile of radiative heating.

This finding gave rise to the concept of the *stratosphere-adjusted radiative forcing*. Here, the RF as used in Eq. (6.4) is determined after the stratospheric temperature has adjusted to the external radiative perturbation. The radiative feedback from stratospheric temperature change thus becomes part of the forcing. This revised RF concept (also known as the fixed dynamic heating approximation) has been used for about 20 years with considerable success. It proved best suited to guarantee that Eq. 6.4 is valid with constant λ for a wide range of forcing agents. Most RF values in Fig. 6.3 have been calculated according to this definition. We note that, within this concept, the net RF (sum of the shortwave and longwave components) is the same at the top of the troposphere (the tropopause) and at the top of the atmosphere, while this need not to be true for the spectral components: For example, an ozone increase in the stratosphere causes an increase in the shortwave RF at the top of the atmosphere because more radiation is absorbed within the atmosphere-surface system. In contrast, the shortwave RF component decreases at the tropopause because energy absorbed above the tropopause will no longer reach the troposphere-surface system. Differences between instantaneous and stratosphere-adjusted radiative forcing are most distinct if a perturbation includes absorption (solar or terrestrial) above the tropopause.

6.4 Assessing Tradeoffs; Climate Sensitivity and Efficacy

When about 20 years ago the DLR Institute of Atmospheric Physics adopted *aviation climate impact* as one of its main research missions, it faced precisely the problems discussed in the previous section: components of total anthropogenic climate change were to be quantified due to their high relevance in the socio-economic context. However, they proved much too small to be derived from complex 3D climate model simulations because of an insufficient signal-to-noise ratio. As a consequence, research focused on determining aviation climate impact and its individual components in terms of radiative forcing. Eq. 6.4 was then used to derive the global temperature response.

Total aviation impact is composed of various forcing components which are partly of different sign (some warming, some cooling). A frequent issue is, therefore, to assess a net effect involving *trade-offs* between two or more

components. Especially when evaluating the effects of technical or operational climate *mitigation* measures, the combined effect can be much smaller than that of the individual components (i.e., CO_2, *contrails*, etc.). Such situations require ensuring the representativity of Eq. 6.4 for using the component radiative forcings as a proxy for the individual climate effect contributions. It must be tested whether the climate sensitivity is actually equal for the effects involved. This test is not always a success, as results from the institute have frequently indicated (Ponater et al. 2005; Stuber et al. 2005).

Possible deviations of the climate sensitivity parameter from that of CO_2 (λ_{CO2}) for certain forcing agents (λ_i) were first suspected some 15 years ago, mainly for forcings caused by a spatially nonhomogeneous distribution of the perturbation (e.g., Stuber et al. 2001; Joshi et al. 2003; Stuber et al. 2005). Since then, the notion of one universally valid climate sensitivity parameter has come under increasing scrutiny. Hansen et al. (2005) have coined the term of an individual *efficacy*, r_i, for each climate perturbation (i), with the definition:

$$r_i = \lambda_i / \lambda_{CO2}. \tag{6.5}$$

With this concept, Eq. 6.4 takes the more flexible form

$$\Delta T_S = \lambda_{CO2} r_i \, RF. \tag{6.6}$$

This implies that r_i and RF become independent descriptors of one and the same perturbation, both of which have to be known to determine its global climate impact in terms of global mean surface temperature change. Efficacies, like feedbacks, can only be calculated from simulations with complex climate models. Hence, the efficacy concept only makes sense if typical efficacy values can be attributed to certain classes of forcings, retaining the original purpose of Eqs. 6.4 and 6.6: to estimate ΔT_S without the necessity to run a dedicated climate model simulation for each single perturbation. Distinctive efficacies have been proposed for quite a number of forcing types occurring in the field of transport climate impact, in particular for aviation. For contrails, most notably in this context, current estimates indicate an r_{cont} significantly less than unity (i.e., less than for CO_2) by 40 % or more (e.g., Ponater et al. 2005; Ponater 2010).

An anomalous efficacy value, different from unity, indicates that the perturbation induces feedbacks within the climate system that are characteristically different from those induced by more homogeneous forcings. For such cases, a crucial methodological question arises: should this anomaly be envisaged as being part of the forcing such that r_i RF in Eq. 6.6 describes some kind of effective radiative forcing to be used instead of RF itself? Or would it be more physical to consider it as part of the climate sensitivity in the sense of Eq. 6.5, retaining the stratosphere-adjusted forcing as the appropriate measure for RF, but abandoning the notion of constant climate sensitivity? This boils down to defining the most appropriate separation between forcing and feedback.

A similar problem is well known from the field of aerosol climate impact, where conceptual differences are made between the direct aerosol radiative forcing (which is consistently described by the RF definition of Eq. 6.4), several kinds of indirect aerosol radiative impacts (see, for example, Fig. 6.3) that refer to changes of *cloud radiative forcing* arising from the presence of aerosols (cloud brightness, cloud lifetime), and the so-called semi-direct effect. The latter is also related to clouds but describes the *cloud cover* response due to the quasi-instantaneous changes forced by the aerosol perturbation on the static *stability* of the troposphere (i.e., on the vertical temperature profile). Semi-direct effects are mainly induced by absorbing aerosols like *soot*. Both semi-direct and indirect aerosol forcing form another kind of swift feedback, similar to the stratospheric temperature adjustment discussed above. Thus, they have to be distinguished, conceptually, from climate feedbacks that evolve slowly as a reaction to a gradually developing surface temperature response.

As noted before, the determination of efficacy values entering Eqs. 6.5 and 6.6 requires long equilibrium climate change simulations with complex models. In many cases, the forcing has to be increased in order to ensure the statistical significance of the resulting ΔT_S and r_i values. Establishing significant and reliable efficacies for certain "exotic" forcings is still a matter of ongoing research. Current assessments of non-CO_2 forcings, e.g., in the transport sectors (Fuglestvedt et al. 2008) so far used Eq. 6.4, implicitly assuming $r_i = 1$ (Eq. 6.5) for all forcings. Doing so may be justified as long as efficacy values are not sufficiently well established because of a still limited number of climate model studies. However, it also conceals an important gap in scientific knowledge which renders the results of previous trade-off studies uncertain.

6.5 Methodology Prospects of an Improved Quantification of Component Effects

There are some innovative ideas with considerable potential to improve the current ability to quantify small component effects in climate research. Gregory et al. (2004) have proposed a new definition of radiative forcing and climate sensitivity that returns to the roots of Eq. 6.4. They assume that the development of radiative feedbacks in time, H(t), occurs linearly in ΔT_S so that the return of the net radiative flux at the top of the atmosphere, N, to equilibrium (after N has been brought out of balance by a forcing perturbation) can be written as a basically linear relationship:

$$N(t) = RF - H(t) \approx RF - \alpha \Delta T_S(t) \tag{6.7}$$

Clearly, ΔT_S has to be zero at $t = 0$, so that $N(0) = RF$. After a sufficient time for the climate (atmosphere–ocean) system to respond, theoretically at infinity, ΔT_S reaches its equilibrium value, ΔT_S^{equ}, while N will have returned to its long-term balanced state, $N(\infty) = 0$:

$$0 = N(\infty) = RF - \alpha \Delta T_S^{equ} = RF - 1/\lambda \Delta T_S^{equ} \tag{6.8}$$

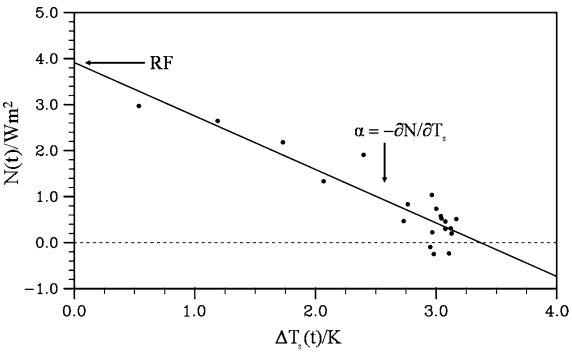

Fig. 6.5 Linear regression of the N-ΔT_S-relation in a simulation of equilibrium climate change for a doubling of CO_2 concentration (relative to 348 μmol/mol). The radiative forcing (at the top of the atmosphere) is given by the intersection point (RF) of the regression line with the ordinate. The slope of the regression line (α) is the inverse of the climate sensitivity parameter (λ). Circles indicate individual annual mean values

This implies that RF and the feedback factor, α, can be determined by linear regression, according to Eq. 6.7, if N and ΔT_S have been simulated close to their equilibrium state. Because $\lambda = 1/\alpha$, as is obvious from Eqs. 6.4 and 6.8, this determines the climate sensitivity parameter, too.

Figure 6.5 illustrates, for example, the N-ΔT_S-dependency for a CO_2 doubling. This case has been simulated at DLR using the ECHAM5 global climate model (Roeckner et al. 2006) with a *mixed layer* ocean. The "spin-up" period, i.e., the simulation time needed to approximate the new equilibrium, is about 20 years in this model version.

Unlike the values determined for stratosphere-adjusted RF, the values determined by regression possess a marked statistical uncertainty (both for RF and for λ), indicated by the scattering of the (N, ΔT_S) values around the regression line. Here, uncertainties are given by the standard regression error (Gregory et al. 2004). For the example in Fig. 6.5, the values for RF_{regr} and λ_{regr} are 3.92 (±0.33) W m^{-2} and 0.86 (±0.09) K/ (W m^{-2}), respectively. In this case, RF_{regr} is almost equal to RF_{adj}, which amounts to 3.91 W m^{-2}. However, *significant* deviations could occur (and indeed occur for other forcing types), because the regression method according to Gregory et al. (2004) constitutes a self-consistent separation of forcing and feedback, much more straightforward than the classical method of stratosphere-adjusted forcing. The latter approach singles out one radiative feedback mechanism (viz., the stratospheric temperature adjustment) for pragmatic reasons and assigns it to the forcing. In the regression method, all radiative feedbacks developing more swiftly in time than the surface temperature response become parts of the forcing, because H(t) in Eq. 6.7 only describes slow feedbacks. From the point of view of practical usefulness, RF_{regr} and λ_{regr} have proven similarly or even better suited to fulfill Eq. 6.4 with a constant λ than the values according to the classical definition. The physical origin of efficacy differences for different forcing agents is thus easier to explore.

Table 6.1 Radiative forcing RF_{regr} and climate sensitivity parameter λ_{regr} derived by means of regression for five equilibrium climate change simulations ("Sim-n"), each forced by a CO_2 concentration increase of 72.3 μmol/mol over a background of 348 μmol/mol. Sim-mean indicates regression values based on the average of Sim-1, Sim-2, Sim-3 Sim-4, and Sim-5

	$RF_{regr}/(Wm^{-2})$	$\lambda_{regr}/(K/Wm^{-2})$
Sim-1	0.96 ± 0.21	0.79 ± 0.21
Sim-2	0.96 ± 0.16	0.81 ± 0.17
Sim-3	0.64 ± 0.18	1.39 ± 0.60
Sim-4	0.99 ± 0.18	0.85 ± 0.18
Sim-5	0.84 ± 0.12	0.94 ± 0.18
Sim-mean	0.92 ± 0.08	0.85 ± 0.10

Unfortunately, the advantages of the regression method are gained at the expense of the statistical uncertainty of the regression values, as mentioned above. For forcings lower than 1 W m^{-2} the climate sensitivity parameter can be determined with a statistical accuracy of not more than 20 %. Such an extent of uncertainty makes the interpretation of efficacy differences once again problematic. The statistical noise can be reduced by simulating more than one spin-up period and by computing the regression for mean values of the simulation *ensemble*. In Table 6.1 another example is shown from simulations run at DLR (this time using the ECHAM4 climate model). Both features, the large primary uncertainty and the reduction potential through an ensemble evaluation, can be noticed. Of course, the necessity to run ensembles substantially increases the computational demands to yield robust values of RF_{regr} and λ_{regr}.

The determination of small component forcings is most difficult with respect to the total aerosol impacts (Fig. 6.3) due to a complex superposition of different physical effects, as mentioned in the previous section. The level of scientific understanding concerning aerosol-cloud interaction is unsatisfactory and is particularly poor for the aerosol impact on ice clouds. Gradual progress in process understanding can be expected to diminish these physical uncertainties.

However, there are also conceptual problems with consistently determining indirect aerosol forcings. As explained in Sect. 6.3, much of the usefulness of the radiative forcing concept stems from the absence of a significant statistical uncertainty. This holds if a radiative transfer calculation can be done with and without the forcing perturbation, with all other parameters (apart from stratospheric temperature adjustment) left unchanged. However, the latter procedure is extremely difficult to realize if forcing, feedback, and response are as intimately linked as they are in the case of aerosol-cloud-interactions: it is hard to envisage how an aerosol effect on cloud lifetime could be disconnected from the cloud effect on static stability, *convection*, dynamics, and the whole *hydrological cycle* if all these processes feedback on the clouds and the aerosols themselves. Unlike the other effects displayed in Fig. 6.3, the indirect radiative forcing for anthropogenic aerosol emissions has thus

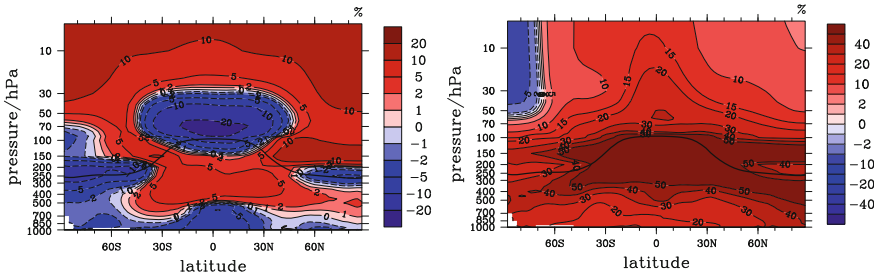

Fig. 6.6 Zonal and annual mean concentration change of ozone (*left*) and water vapor (*right*), given in percentage of the respective background concentration. Equilibrium changes induced by the radiative forcing of a CO_2 doubling are shown. Most changes are significant on a 99 % level. The heavy line indicates the tropopause. All results are from simulations with the EMAC chemistry-climate model (Dietmüller 2011)

been derived by calculating differences between dedicated simulations with different background atmospheres. This involves the usual signal-to-noise detection problems and can only be successful if the signal is large enough (as it is for the impact of all anthropogenic aerosols). Attempts to use the regression method could help to overcome part of the conceptual problems, as it also allows separating a clear-sky from a cloudy-sky component (Gregory and Webb 2008). However, the difficulties in providing accurate estimates in the statistical sense would still remain. A satisfactory method to quantify the component effects contributing to the total aerosol forcing has, thus, still to be demonstrated.

6.6 Concluding Remarks and Outlook

Climate models have consisted of a fixed-composition atmosphere and an ocean component over almost two decades of research. These models do not cover the full complexity of the climate system, as some feedbacks of potentially large importance are not accounted for. Including interactive *atmospheric chemistry* into climate models is a relatively new endeavor, as is the coupling to models of vegetation and ocean biogeochemistry. Recently, a coupled *chemistry-climate model* (EMAC, Jöckel et al. 2006) has been used at DLR to identify the radiative feedback induced by interactive ozone changes in equilibrium climate change simulations (Dietmüller 2011). Figure 6.6 (left) illustrates the ozone feedback in a simulation where the greenhouse effect of CO_2 doubling provides the radiative forcing of climate change.

The ozone concentration feedback is mainly induced by changes in tropical upwelling (causing the minimum in the lower tropical stratosphere) and by impacts of CO_2-induced stratospheric cooling (Fig. 6.4) on *catalytic* chemical ozone destruction (causing the ozone increase above 30 hPa). Various regional contributions of different signs result in a small global mean radiative feedback of ozone by $-0.022\ \mathrm{Wm}^{-2}/\mathrm{K}$, which is only about 1 % of the water vapor radiative feedback. The latter is mainly caused by an overall humidity increase in the troposphere (displayed in the right panel of Fig. 6.6). Nevertheless, due to impacts of the

interactive ozone changes on other feedbacks, the effect on climate sensitivity is significant. The new simulations indicate that the efficacies of various forcings differ less when chemical feedbacks are included than when fixed prescribed ozone distributions are used.

There are other interactive chemical feedbacks, e.g., due to methane, which are still only rudimentarily included in these recent model extensions. An even higher level of complexity will be achieved in the upcoming generation of *Earth system models*. It is currently an open question whether quantification and ranking of anthropogenic greenhouse effects by their radiative forcing will remain a viable option in these extended model frameworks.

Throughout this article the focus has been on global mean surface temperature response as the most representative parameter for global climate change. It has been pointed out that, from this perspective, radiative forcing is a suitable measure to quantify impacts arising from greenhouse gas increases or from other changes of Earth's long-term radiative balance. If changes in other climate features, such as the hydrological cycle, are chosen to represent climate change, this concept may be less appropriate. Bala et al. (2008) have run a global climate simulation in which CO_2 induced positive radiative forcing was compensated by a negative forcing from reduced solar insolation. The global mean surface temperature change was indeed close to zero in this simulation, consistent with what was pointed out in Sect. 6.3. Nevertheless, a significant change in global mean precipitation occurs. This may be taken as another hint that understanding of the climate system in its full complexity is still not complete.

References

Bala, G., Duffy, P.B., Taylor, K.E.: Impact of geoengineering schemes on the global hydrological cycle. Proc. Nat. Acad. Sci. USA **105**, 7664–7669 (2008). doi:10.1073/pnas.0711648105

Bony, S., Colman, R., Kattsov, V.M., Allan, R.P., Bretherton, C.S., Dufresne, J.-L., Hall, A., Hallegatte, S., Holland, M.M., Ingram, W., et al.: How well do we understand and evaluate climate change feedback processes? J. Clim. **19**, 3445–3482 (2006)

Clough, S.A., Iacono, M.J.: Line-by-line calculation of atmospheric fluxes and cooling rates, 2. Application to carbon dioxide, ozone, methane, nitrous oxide and the halocarbons. J. Geophys. Res. **100**, 16519–16535 (1995)

Dietmüller, S.: Relative Bedeutung chemischer und physikalischer Rückkopplungen in Klimasensitivitätsstudien mit dem Klima-Chemie-Modellsystem EMAC/MLO. Forschungsbericht 2011–19, Deutsches Zentrum für Luft- und Raumfahrt, pp. 124 (2011)

Fuglestvedt, J.S., Berntsen, T.K., Myhre, G., Rypdal, K., Skeie, R.B.: Climate forcing from the transport sectors. Proc. Nat. Acad. Sci. USA **105**, 454–458 (2008). doi:10.1073/pnas.0702958104

Gregory, J., Ingram, W., Palmer, M., Jones, G., Stott, P., Thorpe, R., Lowe, J., Johns, T., Williams, K.: A new method for diagnosing radiative forcing and climate sensitivity. Geophys. Res. Lett. **31**, L03205 (2004). doi:10.1029/2003GL018747

Gregory, J., Webb, M.: Tropospheric adjustment induces a cloud component in CO2 forcing. J. Clim. **21**, 58–71 (2008). doi:10.1175/2007JCLI1834.1

Hansen, J., Lacis, A., Rind, D., Russell, G., Stone, P., Fung, I., Ruedy, R., Lerner, J.: Climate sensitivity: analysis of feedback mechanism. Geophys. Monogr. **29**, 130–163 (1984)

Hansen, J., and 54 co-authors.: Efficacy of climate forcings. J. Geophys. Res. **110**, D18804 (2005)

IPCC: Climate Change 2007: the scientific basis, p. 940, Cambridge University Press, Cambridge (2007)

Jöckel, P., Tost, H., Pozzer, A., Brühl, C., Buchholz, J., Ganzeveld, L., Hoor, P., Kerkweg, A., Lawrence, M., Sander, R., et al.: The atmospheric chemistry general circulation model ECHAM5/MESSy1: consistent simulation of ozone from the surface to the mesosphere. At. Chem. Phys. **6**, 5067–5104 (2006). doi:10.5194/acp-6-5067-2006

Joshi, M., Shine, K., Ponater, M., Stuber, N., Sausen, R., Li, L.: A comparison of climate response to different radiative forcings in three general circulation models: Towards an improved metric of climate change. Clim. Dyn. **20**, 843–854 (2003)

Knutti, R., Hegerl, G.: The equilibrium climate sensitivity of the Earth's temperature to radiation changes. Nat. Geosci. **1**, 735–743 (2008). doi:10.1038/ngeo337

Ponater, M., Marquart, S., Sausen, R., Schumann, U.: On contrail climate sensitivity. Geophys. Res. Lett. **32**, L10706 (2005). doi:10.1029/2005GL022580

Ponater, M.: Distinctive efficacies of the components contributing to total aviation climate impact, Sausen, R., et al., (ed.) Proceedings of the 2nd International Conference on Transport, Atmosphere and Climate (TAC), DLR-Forschungsbericht No. 2010–10, Aachen, Maastricht, June 2009, pp. 227–232 (2010)

Roeckner, E., Arpe, K., Bengtsson, L., Christoph, M., Claussen, M., Dümenil, L., Esch, M., Giorgetta, M., Schlese, U., Schulzweida, U.: The atmospheric general circulation model ECHAM-4: model description and simulation of present-day climate. MPI Meteorol. **218**, 90 (1996)

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., S.U, : Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model. J. Clim. **19**, 3771–3791 (2006)

Stuber, N., Ponater, M., Sausen, R.: Is the climate sensitivity to ozone perturbations enhanced by stratospheric water vapor feedback? Geophys. Res. Lett. **28**, 2887–2890 (2001)

Stuber, N., Ponater, M., Sausen, R.: Why radiative forcing might fail as a predictor of climate change. Clim. Dyn. **24**, 497–510 (2005)

Wetherald, R.T., Manabe, S.: The effects of changing the solar constant on the climate in a general circulation model. J. Atmos. Sci. **32**, 2044–2059 (1975)

Thunderstorms: Thermodynamics and Organization

7

Tobias Zinner and Pieter Groenemeijer

Abstract

Thunderstorm research is strongly motivated by the wish to reduce the harm they do to people and their property. Thunderstorms are a global phenomenon, although some areas in the mid-latitudes and tropics are particularly at risk. They form where and whenever the ingredients for their formation come together: instability, moisture and lift. Especially upon interaction with vertical wind shear, they may develop into well-organized systems that produce hazards such as large hail, severe winds, heavy precipitation, and tornadoes.

7.1 Introduction

Thunderstorms are among the most fascinating objects of study in atmospheric physics, not only because of their interesting scientific aspects, but also because everyone can relate to them. Almost everyone has a personal and typically intense thunderstorm experience, such as an early childhood night of fear, a memory of an historic hail or flash flood event, or terrifying turbulence in an area of thunderstorm

T. Zinner (✉)
Ludwig-Maximilians-Universität München (LMU),
Meteorological Institute Munich (MIM), Theresienstraße 37,
80333 München, Germany
e-mail: tobias.zinner@lmu.de

P. Groenemeijer
European Severe Storms Laboratory e.V. (ESSL), c/o DLR-IPA,
Münchner Strasse 20, 82234 Oberpfaffenhofen, Germany
e-mail: pieter.groenemeijer@essl.org

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_7,

activity during the last long-haul flight, when we were reminded to "keep the seatbelts fastened at all times."

In some regions of the world the strong impact of thunderstorms on society during certain seasons causes great alertness to the phenomenon, for example across the North American Great Plains where thunderstorms and tornadoes in particular threaten the population each spring, or the northern Australian and northern Mediterranean regions which experience an annually returning season of heightened flash flood risk, or the tropical thunderstorm season when particularly dangerous *storms* develop from clusters of individual thunderstorms, the so-called tropical disturbances. The latter eventually even provide the birthplace for a much larger, weather scale (*synoptic* scale) tropical cyclone formation.

Although the general deeply-rooted fascination for thunder and lightning probably motivates students and more established scientists alike to get involved in thunderstorm research, there are more far-reaching motivations driving the research. First, thunderstorms are a main contributor to both the local and global water cycle. Second, they play an important role in the global electric circuit and in atmospheric chemistry. However, the most important driver of thunderstorm research is undoubtedly the hazard they pose to human life and infrastructure through lightning, extreme precipitation, hail, wind gusts, and tornadoes. The economic impact of thunderstorms is considerable: the inflicted losses are estimated to amount to almost 1 billion euro per year in Germany alone (Munich Re 2008), and by extension, to many billions across Europe.

The DLR Oberpfaffenhofen Institute of Atmospheric Physics is located in a less dramatically affected region of the world. However, like large areas in Europe it still experiences active convective seasons with occasional damaging events. Sitting on the northern foothills of the Alps, the proximity to the mountains and the availability of moisture suffice for occasional active stretches of convective weather. That way, an almost daily laboratory of thunderstorm activity with occasional extreme events is created, such as large hail and flash floods. Thunderstorm systems usually move along the mountain range west to east or detach from it and head east-north-east over the foothills of the Alps.

The Institute of Atmospheric Physics has been active in the field of *convection* research since its foundation in the year 1962. Already at its former location at München-Riem airport a precipitation radar was installed which, e.g., was used to investigate showers and thunderstorms in the region (Müller and Singler 1972). Historic local events like the Munich hailstorm of July 12, 1984 have been analyzed in detail in the institute using radar data, satellite data and numerical cloud models (Höller and Reinhardt 1986). In more recent years, new severe weather research efforts in Europe were initiated from within the institute and led to the foundation of the *European Severe Storms Laboratory* in 2006 (Dotzek et al. 2009). Other thunderstorm related topics have been studied in the institute and are covered in other chapters of this book, including the scientific challenge and economical value of an accurate thunderstorm forecast especially with respect to aviation, the use of polarimetric Doppler radar in the observation of details of thunderstorm cloud dynamics and physics, the physics and observation of

Fig. 7.1 Photographs of single cell thunderstorm taken near Waltenheim sur Zorne, France looking southeast towards the Black Forest during the Convective and Orographically-induced Precipitation Study (COPS) between 14:43 and 15:26 UTC 15 July 2007. Initiation of convection (*top*, very *left*). Development of a single thunderstorm cell from growth over mature stage (*center* of *top* row) to decay leaving behind a thin cirrus ice cloud layer (lower *right*) (courtesy H. Höller)

lightning flashes, and the production of NO_x and other trace gases by lightning, which affects tropospheric ozone chemistry.

A recent study by the institute suggests that thunderstorm frequencies and intensities may change because of global *climate change*. Climate models suggest a slight reduction of convective precipitation and a higher likelihood of extreme precipitation over Europe in summer. Hence one may expect fewer but possibly more intense thunderstorm events (Sander 2011).

Here the focus is on the thunderstorm as a weather system and on the conditions under which it develops or produces particular hazardous phenomena. First, the basic concepts of convection physics are addressed, and then the different levels of organization of storm systems are described. Subsequently, the dynamics of particularly hazardous phenomena like downbursts and tornadoes are discussed, and an introduction is provided to methods to detect thunderstorms.

7.2 Thermodynamics

Thunderstorms are weather systems that feature strong vertical motions on the scale of a few kilometers in horizontal diameter and are primarily driven by buoyancy, i.e., free convection. Upward and downward motions in convective storms are among the most intense observed in Earth's atmosphere, reaching wind speed values of many tens of m/s, and they may develop within much less than an hour (Fig. 7.1). This is in stark contrast with the weak vertical motions that occur on large scales, the synoptic scale. There, weather systems of hundreds to thousands of kilometers evolve over many hours to days, exhibiting vertical velocities of less than 0.1 m/s.

There are a few ingredients required to initiate convective storms in the atmosphere. To start, two ingredients are needed that create the so-called *convective available potential energy* (CAPE) that storms transform into motion: a rapidly

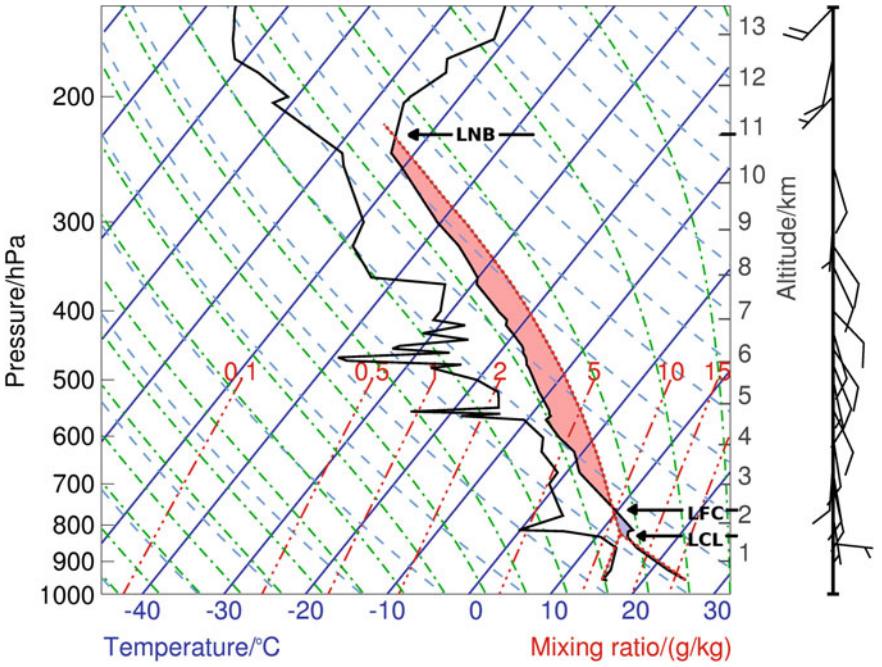

Fig. 7.2 So-called Skew-T-log-P thermodynamic diagram for radiosonde data at München-Oberschleissheim 12 UTC 2 July 2008. Axes are logarithmic pressure ("log-P") on the y-axis (equivalent to linear height on the *left*) and skewed temperature on the x-axis ("skew-T," straight *blue* lines). Additional axes include the water vapor mass mixing ratio labeled with the *red dash-dotted* lines and so-called dry and wet adiabats labeled with *green dash-dotted* and *blue dashed* lines. Displayed in this diagram are data from the radiosonde ascent. Using *black* lines the *dew point temperature* (*left*), the environmental air temperature (*right*) as well as the horizontal wind speed and direction using wind barbs to the *right* of the diagram are indicated. The wind barbs show wind speed in knots using "flags" (full flag = 10 kts, half flag = 5 kts). Significant levels are indicated: LCL—lifting condensation level; LFC—level of free convection; LNB—level of neutral buoyancy

decreasing temperature with height and an abundance of moisture in the lower troposphere. In the *atmospheric boundary layer*—the layer affected by the surface—moisture content and temperature typically increase as a result of solar heating. If the temperature higher up in the troposphere is low enough, CAPE is created. Such a situation is called one of latent *instability*. The only additional ingredient needed to create a storm is a source of lift, also called a trigger mechanism.

These concepts can be illustrated by a *thermodynamic diagram* (Fig. 7.2). In such a diagram (also called *tephigram*), the vertical profiles of moisture, temperature, and the horizontal wind components measured at a given geographic position and height are displayed. These measurements are typically obtained by releasing a weather balloon that carries a thermometer, hygrometer and barometer—a *radiosonde*. Here measurements taken at 2 July 2008 are discussed. During the afternoon, strong thunderstorms developed in this atmospheric environment.

The temperature and moisture measurements are displayed as two separate thick black lines. The temperature at any pressure (or height) level can be read by following the respective isotherm to the x-axis. If the (left) humidity line is read in the same way, one obtains the dew point temperature. This is the temperature at which the given air parcel would be saturated. If air temperature and dew point temperature of the environment are equal, i.e., the two lines touch, the air is saturated. In this case the *relative humidity*, which is the ratio of the *saturation* vapor pressure at dew point temperature and the saturation vapor pressure at the measured air temperature, is at 100 %.

Vertical upward motion of a parcel of air in the atmosphere means that the parcel will cool if no heat is exchanged with the environment, i.e., if the motion is *adiabatic.* This is a prediction of the *second law of thermodynamics*. Sinking pressure is compensated by decreasing temperature and/or density, and vice versa. The rate of temperature change with height is the so-called lapse rate. If the air in the parcel is not saturated with water vapor, that is, for relative humidity below 100 %, the rate is called the *dry adiabatic lapse rate* and equals a constant 9.8 K/km. In case the parcel is saturated, i.e., for 100 % relative humidity, a *moist adiabatic lapse rate* applies, which corresponds to a smaller temperature decrease with height. The reason for this is the *condensation* of water during its ascent. The *latent heat* released by the condensation process compensates part of the cooling during ascent. The moist adiabatic lapse rate typically lies between 5 and 9 K/km. The two lapse rates are reflected in Fig. 7.2 by the light blue dashed isolines (*dry adiabats*) and the green dash-dotted lines (*moist adiabats*) running both from lower right to upper left and describing the change of air temperature during vertical motion.

Important characteristics of potential convection in this situation can be derived and an estimate of the thunderstorm risk can be obtained with such a diagram. Instability of an air parcel at a given height level is simply readable. If the dry adiabat starting at a given environmental temperature decreases slower than the environmental air temperature above this point, the air is absolutely unstable there. Instant vertical motion is expected as the rising air would always be warmer than the environment. If this is not the case for dry adiabatic ascent, but the moist adiabat still leads to warmer temperatures compared to the environment directly above (e.g., the area above 850 hPa in Fig. 7.2), the air parcel is latently unstable. Only a saturated cloudy air parcel could rise. If the environmental temperature gradient is even smaller than the moist adiabatic lapse rate, the situation is absolutely stable. Rising air would always be colder than the environment and would return to the starting level.

The *lifting condensation level* (LCL), i.e., the *cloud bottom height*, is found by tracing the dew point and air temperature change of a rising nonsaturated surface air parcel. Such an air parcel's temperature will decrease along a dry adiabat on the one hand. On the other hand, its dew point temperature will change. The rate of change can be calculated and is plotted with red lines for a constant moisture mixing ratio in the diagram. Where such a line meets the dry adiabat for the surface parcel, i.e., dew point and temperature of the rising air parcel become equal, saturation and the lowest level of cloud formation is reached, if air is lifted to this level by some process.

After condensation at the LCL, further vertical temperature change of the now saturated air parcel has to follow a moist adiabat. For the example in Fig. 7.2, the air parcel is still absolutely stable, no further vertical motion from this hypothetic cloud bottom is possible without a further lifting process. If such a process is present, the temperature decreases along the moist adiabat until the environmental air temperature curve is crossed. The *level offree convection* (LFC) is reached. If a saturated air parcel rises above this point, it will always be warmer than the environment, the atmosphere is unstable, the air parcel becomes buoyant. Following the further rise of the air parcel along the moist adiabat, the ambient air temperature is eventually reached again at the *level of neutral buoyancy* (LNB). This point marks the possible *cloud top* for the convective ascent of the surface air parcel. This level is usually close to or at the *tropopause*, i.e., at the level at which the temperature gradient changes from tropospheric cooling with height to stratospheric warming with height (visible in Fig. 7.2 at a height of about 10.5 km).

The light red area in Fig. 7.2 between LFC and LNB is the convective available potential energy (CAPE). The light blue area between LCL and LFC is the *convective inhibition energy* (CIN) that has to be provided by a *trigger* process to release the instability above. Strong thunderstorms need a sufficiently large value of CAPE, and the likelihood of convective initiation rises if CIN is small. In the sounding of Fig. 7.2, the CAPE is 1 300 J/kg—a rather large value—while the inhibition is only weak (CIN $= -20$ J/kg). Strong thunderstorms seem already likely at the time of the sounding at noon.

Reasons for a growing CAPE, a shrinking CIN, and the final initiation of convection can be increasing surface moisture, e.g., over a water body, or surface warming during the course of the day, as was the case for 2 July 2008 shown in Fig. 7.2. In addition to these surface driven processes, mid-tropospheric cooling by advection or infrared radiation has similar effects. Very common in mid-latitudes is the destabilization and the growing thunderstorm potential ahead of an approaching *front* when colder air arrives at elevated levels ahead of the surface front. The forced lifting of atmosphere layers is the second important trigger process. Most commonly this happens for latently unstable situations lifted by a cold front or mountain orography. Additional beneficial impact on the formation of *deep convection* is provided by synoptic scale forcing, which results in upper tropospheric divergence and related large-scale lifting, such as in the area of the jet stream maxima in mid-latitudes and similarly in the formation of tropical thunderstorms. For further details on thunderstorm thermodynamics see, for example, Emanuel (1994).

7.3 Storm Organization

7.3.1 Single Cell Storms

Once the latent instability is released by some trigger mechanism, a convective cell is formed. In the absence of strong horizontal winds such a cell features a narrow updraft with strong upward motion in its center, visible as the typical cauliflower-like cloud, surrounded by a much larger area of weak subsidence in its surroundings. Although there even might have been a trigger acting on a wider area in the first place (e.g., surface heating, mid-troposphere cooling), there are a few spots where convective initiation takes place first. When air starts to rise there, convergence below these spots and divergence above (at the level of neutral buoyancy, usually close to the tropopause) suppresses further ascent nearby. Such a single cell can last for 15 min up to an hour.

The stage of initiation and growth described above is characterized by increasing updraft speeds and rising cloud top height (top left in Fig. 7.1). The convective cloud is called *cumulus* (Latin for "heap") or *cumulus congestus* ("massive heap") during this phase. Eventually the single cell reaches its maximum vertical extent when the air from the lower troposphere reaches its level of neutral buoyancy. The mature stage is reached. For thunderstorms this level is typically located at an altitude of about 10 km or more, where the air temperature is normally below −40 °C. At such a temperature all cloud water that has condensed during ascent has frozen to ice particles. The top of the thunderstorm cloud develops its characteristic *anvil* shape and the cloud at this stage is called *cumulonimbus* (Latin: "raining heap"). During this stage the storm's updraft is strongest and exceeds the fall speed of all the water particles which it contains. Within the updraft a mix of ice and liquid water particles of the most diverse sizes is present. In such a regime formation of precipitation (e.g., hail) is most efficient. Some particles grow quickly and eventually fall out. The onset of precipitation initiates the decay or dissipation stage of a single cell. Falling precipitation forms a compact *downdraft* region, mainly due to the mass of precipitation within the sinking air and also due to evaporative cooling when part of this mass evaporates as it falls through drier air. When such a downdraft reaches the surface, cool air spreads out below the cloud and eventually stops the supply of converging warm and moist air towards the cloud bottom.

Figure 7.3 shows a satellite image of thunderstorms in the Black Forest area in southern Germany taken by the Meteosat-9 weather satellite again for 2 July 2008. On this day, latent instability was released in the early afternoon to form many different forms of convection and thunderstorms (cf. the Fig. 7.2 analysis). Little white speckles over the clear land surfaces of the area mark fair weather cumulus convection which, in some places, develops into strong thunderstorms. On this day the convective development was mainly triggered by local surface heating: no synoptic forcing was present (no strong winds or approaching fronts). Concentrating on the *single cell storm* in the center of the image between 12:30 and 13:15

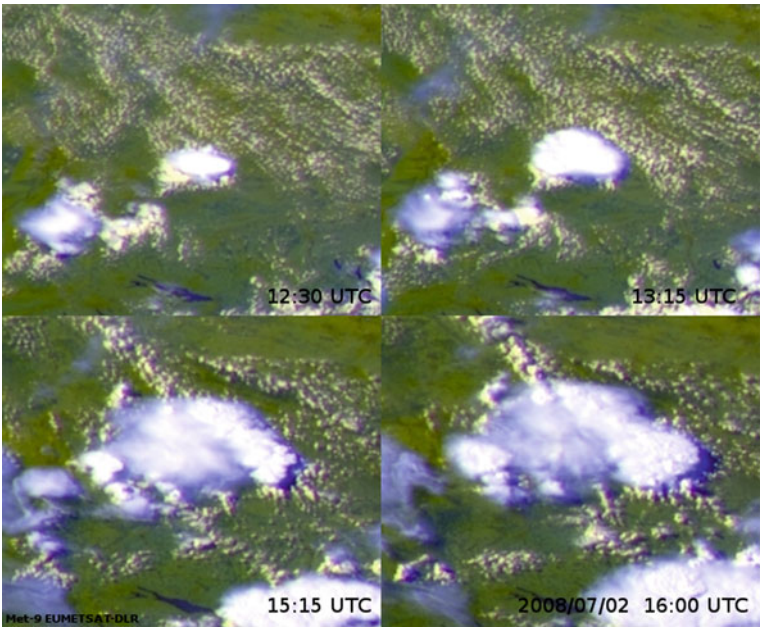

Fig. 7.3 *Satellite* images of thunderstorms over southwestern Germany on 2 July 2008 from the *Meteosat*-9 satellite

UTC, there is a clear suggestion that it suppresses convective clouds in its vicinity. Cumulus convection is only present at some distance from the central cell. The cell's mature stage ends around 13:15 UTC when the large white ice cloud anvil reaches its largest extent while it is still thick (white). After this point in time the anvil still widens, but becomes darker (more transparent), because the updraft of the cell has stopped. Precipitation and related downdraft are probably still active at the time, but the supply of fresh moisture from the lower atmosphere and the related new cloud formation by condensation has ceased. At 15:15 UTC, only a thin *cirrus* (Latin: "tuft of feathers/hair") cloud is left of the decayed original cell. At the same time a fascinating new stage of convective development has started that marks the evolution from single- to multi-cell thunderstorms. While the first are rather rare in reality, the second is the most common type of thunderstorm. At 15:15 UTC, new cells pop up from the background convection in an almost circular formation around the original thunderstorm cell's position. The reason for this is the pool of cold air formed by the downdraft air which spreads out at the surface—in this case in an almost circular form around the original position. The cool air is denser than the ambient air and lifts the surrounding air towards the LFC, forming the trigger for new secondary thunderstorm cells, which themselves reach a mature stage around 16:00 UTC.

7.3.2 Multicell Storms

Multicell storms consist of several thunderstorm cells moving as a group. The different cells are all in different stages of development (compare also Fig. 7.1). The connecting factors between cells are the downdrafts triggering the newer developments. When a horizontal wind component interacts with a multicell motion, new cells usually form upwind (in Fig. 7.1, the left side of the images). On this side, the spreading cold pool created by the downdraft of one cell is decelerated and new cells can develop in the direct neighborhood of old cells, which results in closely interacting updrafts and long-living convective systems (up to several hours). On the downwind side, the spreading cold pool creates less strong updrafts on its leading edge and usually moves too fast to allow for the development of nearby new cells with steady warmth and moisture supply. Multicells can generate large amounts of precipitation and large hail.

A special form of a multicell, not uncommon in Western Europe, is the so-called *squall line*. Due to its horizontal extent it is classified as a *mesoscale convective system* (mesoscale denoting the spatial scale between local scale and the synoptic, the weather scale). In other words, such systems have a typical length scale of >100 km and can last for several hours up to a full day with tracks up to 1 000 km. The squall line is a multicell event which is arranged in a narrow line of individual storms with lengths of >100 km and a width of several 10s of km. The system moves perpendicular to the line orientation. The heaviest precipitation is found at the leading edge of the line, while a large area of more steady, less heavy rain is found behind the leading edge. The formation of a squall line is usually connected to a nearby front. An example of a short-lived squall line is given in Fig. 7.4. Around 17:30 UTC 16 June 2011, a squall line formed just west of Munich ahead of a front approaching from the west and moved eastward for about 100 km before it finally decayed after about two hours. Figures 7.4a–d shows the moving line of precipitation as a horizontal cross section seen by the DLR polarimetric Doppler precipitation radar POLDIRAD over 90 min. Figure 7.4e shows a vertical cross section through the radar data along the moving direction (west–east) perpendicular to the line.

The development of a squall line is decisively dependent on the vertical wind shear perpendicular to the line. Usually winds from the west are stronger in the upper troposphere than in the lower. The thunderstorm line moves approximately at mid-troposphere level wind speed. Because of the motion of the line to the east, a storm-relative wind component exists toward the line from the east at the surface, while the opposite is true for the upper troposphere. New cells form at the leading edge of the line. If the shear is weak, i.e., wind speed increases only slightly with height, the squall line is short-lived as in the case presented. If the shear is strong, squall lines can live for up to 12 h as long as the environment in its path is still favorable to thunderstorms. The reason is the relative speed of the surface *cold pool's* high density air at the leading edge of the line, which is fed by the outflow of the different cells. If the outflow is faster than the thunderstorm line itself, it outruns the line and cuts off the

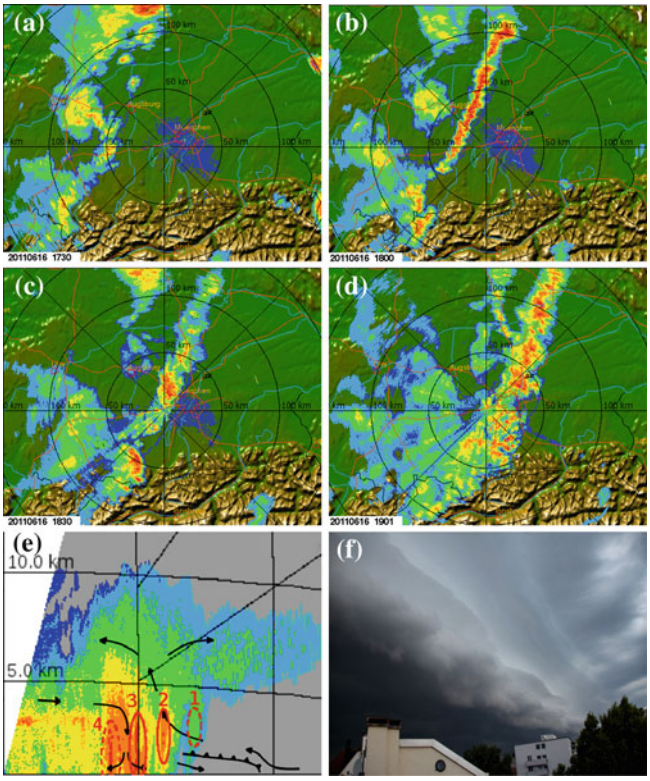

Fig. 7.4 Radar data from the DLR Institute of Atmospheric Physics *precipitation radar POLDIRAD* at Oberpfaffenhofen (at 0 km in both displays) between 17:30 and 19:01 UTC 16 June 2011. Shown is the formation of a squall line in *horizontal* cross sections **(a–d)** and in a vertical cross section for 19:01 UTC **(e)**. Low radar reflectivity related to low precipitation intensity is shown in *blue* and *green*, increasing values in *orange* and *red*. The horizontal and vertical distance to the location of POLDIRAD is labeled in **(a–e)**. **(f)** Photograph of the squall line's shelf cloud moving over Munich just *west* of Oberpfaffenhofen, view to the *north* (courtesy B. Mayer)

line from further warm inflow from in front of the line. If the cold pool front moves at a speed matching the displacement of the whole front (the mid-level winds), new thunderstorm cells are created at the leading edge exactly at the speed of displacement and a quasi-continuous line of precipitation is formed.

The schematic plotted on top of the radar cross section in Fig. 7.4e highlights the general structure of and the motions in the squall line. The cold outflow air that spreads out to the east, marked using the line with added triangles, lifts warmer air ahead of it. This leads to a new cell development (cell labeled "1"). An older cell (2) is already developed further, showing increased intensity of precipitation (orange to red) at a height of a few kilometers. Cell 3's precipitation core has already reached the ground, and with it most likely cold air from aloft reaches the ground. A fourth cell

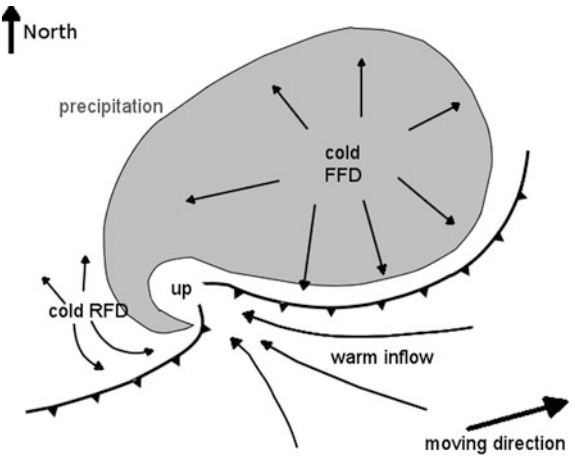

Fig. 7.5 Schematic of a northern hemisphere supercell thunderstorm. Shown are the ground level precipitation field, the storm-relative horizontal wind components caused by storm motion and up/downdrafts, and the separation of *cold* outflow and *warm* inflow air. The rotating updraft region, the mesocyclone, is labeled "*up*." This is the area where tornadoes are most likely. FFD—forward flank downdraft; RFD—rear flank downdraft

further west is still identifiable, while all older cells are merged into an area of much weaker precipitation intensity to the west of the leading edge. The maximum of backward inflow happens at mid-levels and is typically labeled the rear inflow jet. In certain situations this upper rear inflow jet is amplified by a convex curvature of the squall line structure and can even touch ground. Due to the appearance of typical precipitation cross sections in radar data, this situation is called a bow echo. Consequently, squall lines show dangerously high wind speeds on the one hand, at the so-called gust front of the leading edge of the cold outflow, and on the other hand in bow echo regions behind the line. A photograph of the cloud topped gust front (the shelf cloud) on 16 June 2011 is shown in Fig. 7.4f.

Strong convection may also be organized as a mesoscale convective system (MCS), which can be related to squall lines again but are less regular in appearance. An MCS already exhibits characteristics of synoptic low pressure systems in the sense that the surface convergence region and related divergences regions near the tropopause start to rotate. A similar step in convective organization are convective mesoscale disturbances in the tropics which evolve to tropical cyclones (hurricanes, typhoons), large scale rotating low pressure systems with pronounced thunderstorm rain bands, when they move over warm ocean.

7.3.3 Supercell Storms

The most astonishingly organized convective system of the mid-latitudes, which has a disproportional damage potential, is the *supercell* thunderstorm (schematic in Fig. 7.5). Supercells do not only bring heavy precipitation, but also hail, with hailstone diameters in excess of 5 cm. They are often accompanied by strong tornadoes. A supercell storm does not consist of multiple cells, but features a single deep, persistently rotating updraft. Although they vary greatly in size, their

radar representation (i.e., their precipitation pattern) typically has a diameter on the order of 10 km.

Only a small percentage of all thunderstorms are supercells. They form in an environment with strong vertical wind shear in the lowest 5–8 kms of the atmosphere, i.e., a situation in which the wind changes rapidly with height. The strong shear allows for a spatial separation of updraft and downdraft areas which prolongs the lifetime of supercells to up to several hours. The rotation of the supercell updraft—the *mesocyclone*—originates from the upward tilt of horizontally oriented rotation. A horizontal vortex tube, generated by strong vertical wind shear, is bent into the vertical by a convective updraft. This way, a clockwise and a counterclockwise rotating flank of the updraft are created. When precipitation and related downdraft start to develop in this updraft, it is split into two counter rotating separate updrafts. If the wind turns clockwise with height (veers), the counterclockwise rotating updraft is favored, and vice versa. Since the conditions in which storms occur in the northern hemisphere typically feature a veering wind with height, the counterclockwise rotating updraft typically survives, whereas the clockwise one tends to decay.

Rotation and vertical shear lead to the mentioned separation of precipitation region from the updraft core: precipitation mainly falls in the left forward flank of the supercell (northern hemisphere) where it initiates the cold forward flank downdraft (Fig. 7.5). In the right cloud-free forward flank, the thunderstorm's warm inflow area is located. Later in the supercell life cycle, precipitation can extend to a rear flank downdraft region close to the rotating updraft. In this situation, radar images show the so-called bounded weak *echo region* of low radar reflectivities surrounded by high precipitation reflectivities marking the dry updraft, which is too strong (up to 40 m/s) to provide enough time for the formation of precipitation. The interaction of the downdrafts and their spreading cold pools with the rotating supercell mesocyclone most importantly leads to the development of tornadoes.

7.4 Tornadoes and Other Severe Winds

Severe winds are an important hazard of thunderstorms. Thunderstorm-related winds can be distinguished as those associated with a tornado, and other winds, typically referred to as straight-line winds. *Tornadoes* are vortices, typically between a few meters up to as much as 1–2 km in diameter, which extend between a convective cloud and Earth's surface. Pictures of the devastating impact of tornadoes in the "tornado alley" of the North American Great Plains are probably the most prominent examples of wind damages that come to mind, but tornadoes have been observed on all continents except Antarctica. For example, the *European severe weather database* lists about 150 tornadoes in western Europe each year (Dotzek et al. 2009).

Tornadoes are usually marked by a funnel cloud due to droplet condensation in the vortex center and, once they touch ground, by a dust and debris cloud at their lower end. The highest wind speeds in the world have been observed in a tornado, the record measurement at about 30 m above the surface being 135 m/s (480 km/h), observed by a mobile Doppler radar. Most tornadoes are substantially weaker, having wind speeds of 30–80 m/s. In fact, the frequency of tornadoes decreases exponentially with increasing intensity (Dotzek et al. 2003). Weak tornadoes are usually short-lived phenomena and may occur with relatively weak thunderstorms or even cumulus congestus clouds. The much rarer strong tornadoes that produce a disproportional share of all tornado damage occur almost exclusively with supercells. Within the supercell storm, they are located within the mesocyclone at the interface of the updraft and the forward and rear-flank downdrafts. The presence of a supercell, however, does not guarantee that a tornado will occur. Estimates of the percentage of supercells that produce tornadoes lie somewhere between 5 and 20 %.

It is of particular importance for supercell tornado formation that the storm has a low cloud base, and that vertical wind shear is especially strong in the lowest kilometer above ground. The reasons are only partially understood. Recent findings suggest that the air flowing into a tornado predominantly originates from the rear-flank downdraft and recurves into the storm's updraft (Markowski and Richardson 2009). The downdrafts of the supercell generate intense horizontal vortex lines as the cold surface outflow undercuts the environmental air, leading to very strong low-level wind shear. These lines are then bent into the vertical in the mesocyclone updraft where they can intensify to form a tornado. If the cloud base is high, this downdraft is likely to be too cold, because *evaporation* of the precipitation contained in the downdraft leads to strong cooling of the air. Consequently, the air is not buoyant enough to be easily "sucked back" into the updraft, hindering tornado formation.

Most strong tornadoes occur in association with supercells, within the mesocyclone of the storm. Whereas a tornado is clearly related to low-level winds directed into a storm and its updraft, most other severe winds are closely associated with downdraft air that descends out of the storm. The areas affected by these winds differ greatly in size. On one end of the spectrum, intense winds may occur within a local *downburst* in connection with an isolated thunderstorm, which affects an area of perhaps merely a square kilometer. On the other end are large, linearly organized mesoscale convective systems like the mentioned squall lines. Their swaths of very variable, damaging downburst-related winds can be 100 km in width and over 1 000 km in length, the strongest of which are called *derechoes* (Spanish for "straight line").

Downbursts develop when rain or small graupel evaporates in sub saturated air below a cloud's base, causing this air to cool strongly and thus become much denser than the air around it. As a result, it starts to accelerate downward to Earth's surface. Strong downdrafts develop when the cloud base is high. In addition to evaporational cooling, the mere mass of rain, graupel and hail adds to the density of the air–water mixture, enhancing downdraft strength. Within tropical storms,

this effect even dominates over the evaporation effect. Strong wind gusts occur when the downbursts transport strong horizontal momentum down to the surface.

Individual downdrafts in thunderstorms that subsequently spread out horizontally near the surface typically merge and form pools of cold air under the storm cloud. Where they are associated with propagating storm systems such as multicell clusters, squall-lines or supercells, the leading edge of the cold pool, the gust front, is typically well-marked by a line of very dark shelf clouds (see Fig. 7.4f). Typical wind speeds in damaging nontornadic winds are 25–40 m/s, but occasionally speeds up to 50 m/s are observed. For further reading on storm organization, tornadoes and severe winds, the reader is especially referred to Doswell (2001).

References

Doswell, C.A. (ed.): *Severe Convective Storms-an Overview*, Vol. 28, No. 50, pp. 561 Meteor. Mon. (2001)

Dotzek, N., Grieser, J., Brooks, H.E.: Statistical modeling of tornado intensity distributions. Atmos. Res. 67–68, 163–187 (2003). doi:10.1016/S0169-8095(03)00050-4

Dotzek, N., Groenemeijer, P., Feuerstein, B., Holzer, A.M.: Overview of ESSL's severe convective storms research using the European severe weather database ESWD. Atmos. Res. **93**, 575–586 (2009). doi:10.1016/j.atmosres.2008.10.020

Emanuel, K.A.: Atmospheric Convection. Oxford University Press, Oxford (1994)

Höller, H., Reinhardt, M.E.: The Munich Hailstorm of July 12, 1984—convective development and preliminary hailstone analysis. Beitr. Phys. Atmos. (Contr. Atm. Phys.) **59**, 1–12 (1986)

Markowski, P.M., Richardson, Y.P.: Tornadogenesis: our current understanding, forecasting considerations, and questions to guide future research. Atmos. Res. **93**, 3–10 (2009). doi: 10.1016/j.atmosres.2008.09.015

Müller, H.G., Singler, A.: Radarbeobachtungen an konvektiven Niederschlägen. Meteorol. Rundschau **25**, 45–51 (1972)

Munich Re.: Highs and Lows: Weather risks in Central Europe, Munich Re Group Publication, order number 302–05482, pp. 60 (2008)

Sander, J.: Extremwetterereignisse im Klimawandel: Bewertung der derzeitigen und zukünftigen Gefährdung, p. 125. Faculty of Physics, Dissertation, LMU München (2011)

Thunderstorms: Trace Species Generators

8

Heidi Huntrieser, Hartmut Höller and Volker Grewe

Abstract

In the upper troposphere, both natural and anthropogenic processes control the budget of nitric oxide (NO), a highly reactive and pollutant trace gas. The main local NO sources in the upper troposphere are emissions from aircraft and production by lightning. In the past 20 years, DLR studied the latter source in airborne field experiments accompanied with model simulations. The global lightning NO source is found to be distinctly larger than that from aircraft (factor ~5, uncertainty ~50–100 %). Lightning flashes in tropical regions seem to produce less NO per flash compared to other regions.

8.1 Introduction

Deep convective clouds have an important impact on the *trace gas* composition in the *upper troposphere* and *lower stratosphere*, but also affect the *planetary boundary layer* (Dickerson et al. 1987). Strong updrafts and downdrafts within

H. Huntrieser (✉) · H. Höller · V. Grewe
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Heidi.Huntrieser@dlr.de

H. Höller
e-mail: Hartmut.Hoeller@dlr.de

V. Grewe
e-mail: Volker.Grewe@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_8,

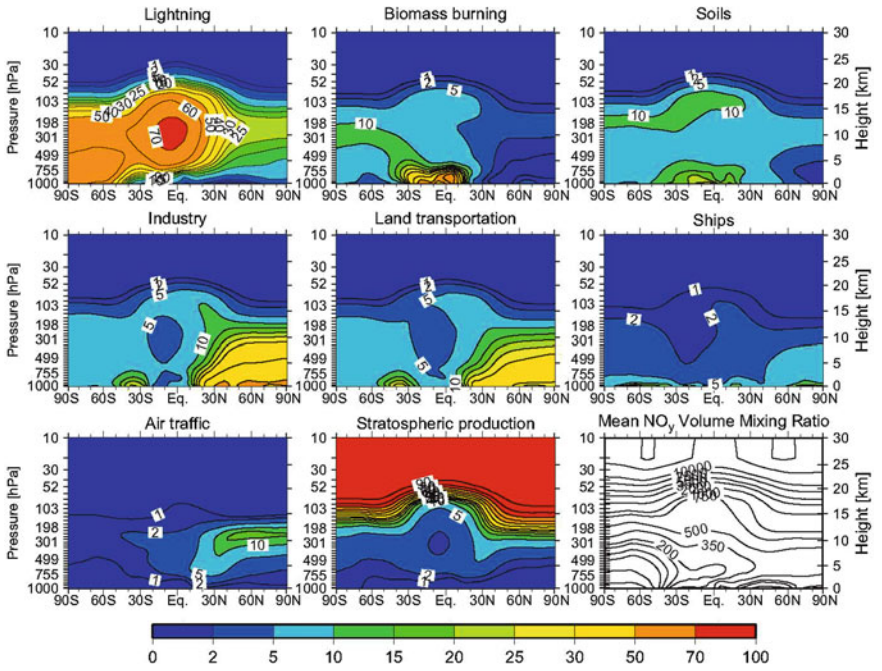

Fig. 8.1 Relative contributions (%) to the NO_y budget from various natural and anthropogenic sources as a function of latitude and pressure-altitude in the annual mean, from Grewe (2007) (with permission from Elsevier)

these clouds efficiently redistribute trace gases on a short time scale. Polluted air masses in the boundary layer may be ingested by the *thunderstorm* (entrainment), rapidly uplifted (convective uplift), and emitted within the *anvil* outflow (detrainment) in the upper troposphere, less than ~ 1 h later. During this rapid and efficient transport within the updraft core, parts of the air mass may reach the upper troposphere almost undisturbed and unaffected by wet scavenging. Due to the short time scale, the initial trace gas composition exhibits almost no change. The sum of nitric oxide (NO) and nitrogen dioxide (NO_2), known as *nitrogen oxide*s (NO_x), is quasi-conserved during this transport. NO_x has a lifetime of several days in the upper troposphere, before, e.g., nitric acid (HNO_3) is formed by reactions with hydroxyl radicals (OH). In comparison, the lifetime of NO_x in the boundary layer is only a few hours.

The sum of NO_x, HNO_3 and a number of other reactive, oxygen-containing nitrogen species is known as *total reactive nitrogen* (NO_y). The NO_x and NO_y budgets are usually measured in terms of the mass of nitrogen contained in the compounds. Besides the *anthropogenic* NO source mentioned above, a variety of other processes are known to impact the global NO_x and NO_y budgets in the upper troposphere (Fig. 8.1). Their source strengths are less compared to NO emitted in the boundary layer (mainly from fossil fuel combustion, biomass burning and soil

microbial activity). An important anthropogenic source in the upper troposphere is the release of NO in the fresh exhaust from *aircraft* (globally <1 Tg a^{-1}). The emissions from air traffic are especially pronounced in the northern hemisphere *mid-latitude*s. Furthermore, in the upper troposphere natural sources have an important impact on the NO_x and NO_y budgets. The injection of NO-rich air masses from the stratosphere is within the same range as the aircraft emissions. However, the strongest natural source in the upper troposphere originates from *lightning*-induced nitrogen oxides (LNO_x). This source is believed to be especially pronounced in tropical regions, coinciding with the global maximum in lightning activity. Almost 200 years ago, von Liebig (1827) first suggested that lightning plays a major role in the global nitrogen cycle. Until recently, a global LNOx nitrogen-mass source strength varying from 1–20 Tg a^{-1} was given by the World Meteorological Organization (WMO), whereas for model simulations a value of 5 Tg a^{-1} was frequently used. In a review paper by Schumann and Huntrieser (2007) summarizing three decades of research activities on this topic (considering theoretical and laboratory studies as well as surface-, airborne-, and satellite-based measurements), the wide range of the LNO_x source strength was slightly reduced. The best estimate of the annual global LNOx source and its uncertainty range was estimated to 5 ± 3 Tg a^{-1}, which is about 10 % of the total NO mass emitted into the atmosphere.

A *flash* of lightning induces a hot channel of air. When it expands, we hear thunder. In this hot lightning channel ($\sim$30 000 K), the nitrogen (N_2) and oxygen (O_2) molecules of the air are dissociated into single N and O atoms. During the rapid cooling of the channel, these atoms and molecules may recombine to form NO (Zeldovich mechanism). At 4 000 K, a higher NO concentration than expected from its thermochemical equilibrium level becomes frozen into the gas ("freeze-out" temperature). During further rapid cooling, by mixing with the ambient air, the lightning-produced NO is prevented from undergoing thermal decay reactions (known as NO content "freezes out").

NO plays an important role in the production of *ozone* (O_3). It acts as a *catalyst* during the photochemical oxidation of hydrocarbons and *carbon monoxide* (CO) producing O_3. On the other hand, ozone plays an essential role in atmospheric *chemistry* since it determines the oxidizing capacity of the troposphere and acts as an important *greenhouse gas*. Therefore, to increase understanding about the *climate impact* of NO_x, it is a crucial task to reduce the large uncertainties in estimates of the global LNO_x source.

8.2 Experimental Set-Up of Airborne DLR Thunderstorm Field Experiments Focusing on LNO_x

Between 1992 and 2006, DLR, under the lead of the Institut für Physik der Atmosphäre (IPA), performed a number of airborne field *experiment*s in thunderstorms focusing on LNO_x. The measurements were carried out over almost all continents (Europe, South America, Australia and Africa). The DLR thunderstorm

field experiments mainly consisted of in situ trace gas measurements aboard an aircraft, ground-based measurements with *radar* monitoring the dynamic evolution and microphysical composition of the thunderstorms, and ground-based measurements with a *lightning detection network*.

8.2.1 Investigated Thunderstorm Types

Different types of thunderstorms were investigated depending on different factors, such as latitude, stability, orography and wind conditions in the regions. A number of smaller single-cell thunderstorms (short-lived storms forming in low vertical wind shear environments), sometimes organizing to more long-lived multicell storm systems, were frequently observed in tropical regions. In environments with significant wind shear, as in the mid-latitudes, thunderstorm systems may organize into squall lines with a longer life cycle maintained by the pronounced wind shear ahead of frontal zones. The most vigorous thunderstorm type is the *supercell*, an isolated, strong rotating thunderstorm associated with extensive vertical updrafts, large *hail*, gusty winds, and tornadoes. The largest thunderstorm type is the *mesoscale convective system*, which can extend over $\sim$500 km (Houze 2004). It is a complex of thunderstorms that becomes organized into a single cloud system with a leading line and is characterized by a huge cirriform cloud structure accompanied by a large region of trailing stratiform precipitation (>100 km in at least one direction). These systems develop overnight, are especially long-lived (up to $\sim$20 h) and have a maximum activity in the late afternoon and evening. It is known that mesoscale convective systems with extensive ice scattering develop over preferred regions in, e.g., the Sahel zone, Central Africa, Argentina, and the southeastern United States.

8.2.2 Airborne In Situ Trace Gas Measurements

For measurements in the thunderstorm anvil cloud, the DLR *research aircraft Falcon* is a suitable platform due to its robust construction and its elevated operating altitude up to $\sim$13 km, covering the main anvil outflow region. The standard in situ trace gas instrumentation during these *campaign*s is mainly used to measure NO, NO_y, CO, CO_2, O_3 and the photolysis rate of NO_2 [$J(NO_2)$], with the high temporal resolution of 1s for most instruments. In addition, standard meteorological parameters such as pressure, height, temperature, humidity and wind were available. The trace gases NO and O_3 in addition to $J(NO_2)$, pressure and temperature measurements were used to calculate the NO_2 and NO_x molar mixing ratios on the basis of the photostationary steady state equation. During the convective transport from the boundary layer to the upper troposphere, NO_x is a conservative tracer, but not NO or NO_2 individually. During anvil penetration, a pronounced enhancement in NO was mostly observed, exceeding background

mixing ratios (in general <0.2 nmol mol^{-1}) by factors of 2 up to 100. Furthermore, a high ratio between NO and NO_y is indicative of fresh NO emissions from, e.g., lightning production, a low ratio indicative of aged emissions.

The most suitable tracers indicative of boundary layer air mass transport to the anvil outflow region are CO (with a lifetime of ~2–3 months in the troposphere) and CO_2 (with a long lifetime, at least in the free atmosphere). For fresh transport from the boundary layer, O_3, with a lifetime in the order of days to months, can partly also be used as tracer. Measurements in thunderstorms have shown that the mixing ratios of tracers in the main outflow region can occasionally be in the same range as measured at the top of the boundary layer, which indicates a very fast and undiluted vertical transport, a "mirroring" of trace gas compositions (Huntrieser 2002). The main inflow region was found to be located at the top of the boundary layer. Model simulations have shown that most LNO_x produced in a thunderstorm is transported into the anvil region, e.g., Fehr et al. (2004). Therefore, calculations based on measurements in the anvil outflow are to some extent expected to give a reliable estimate of the total amount of LNO_x produced in the thundercloud.

8.2.3 Ground-Based Lightning Measurements

During the thunderstorm *field campaign*s carried out between 2004 and 2006 in Brazil, Australia and Burkina Faso, a lightning location system called LINET was set up by DLR in the operation area. This system can detect the flash location, rate, and type (intracloud or cloud-to-ground), the peak current, and in ideal cases also the flash length, e.g. Huntrieser et al. (2007, 2008, 2009, 2011); and Höller et al. 2009). LINET continuously measures the transient magnetic components of lightning discharge very low frequency (VLF)/low frequency (LF) emissions. These signals (strokes) are emitted by certain components of the flashes. Strokes within a certain time period (<1s) and within a close area (<35 km) are combined to so-called "flash components". LINET locates few VLF strokes per flash components (on average 3 and up to 9) (Huntrieser et al. 2008).

8.2.4 Flight Planning Tools and Flight Patterns

For planning the flights, numerical weather forecasts including a variety of different thunderstorm indices and forward trajectory calculations were used in addition to satellite images showing cloud development. Specific forecasts considering LNO_x were developed and applied during some campaigns (examples given below).

A major part of the LNO_x produced in thunderclouds is transported into the anvil outflow region located close to the tropopause, as shown in model studies. The Falcon aircraft probed this region in high detail both close to the convective core with lightning and several 100 km further downwind to investigate chemical processes in the anvil outflow. Ideally, the outflow is also probed at different

altitudes. In the aircraft, the integrated radar system gives pilots information on the radar *reflectivity* within the penetrated thunderstorm. Pilots try to avoid penetrations too close to the convective core with strong turbulence, updrafts and downdrafts, *graupel*, hail and lightning that can be dangerous to the aircraft. However, during some flights the Falcon aircraft was hit by smaller lightning strikes (Huntrieser 2002), luckily without disruptive effects on the aircraft or its instrumentation.

8.2.5 Quantification of LNO_x

A number of different methods to derive estimates for the global LNO_x production rate exist, based on flash and storm extrapolations from laboratory, experimental or theoretical studies or global model fits (Schumann and Huntrieser 2007). Frequently, the global flash frequency value obtained from space by the Lightning Imaging Sensor (LIS) (44 s^{-1}) is used to extrapolate the estimates of LNO_x per flash to a rough annual global estimate for the LNO_x production rate. For the extrapolation to global LNO_x production rates of the DLR in situ *measurements* with the Falcon aircraft in single thunderstorms, a modified version of the method introduced by Chameides et al. (1987) was applied (Huntrieser et al. 2008). The horizontal LNO_x mass flux was calculated for each thunderstorm penetration as a function of the mean LNO_x mixing ratio, the air density depending on temperature and pressure, the difference between the wind vectors in the anvil outflow and at the steering level ($\sim$3 km), and the area of the vertical cross-section perpendicular to the wind direction in the anvil outflow. To estimate the LNO_x production rate per stroke, the horizontal LNO_x mass flux estimated for a thunderstorm was divided by the LINET stroke rate observed for the same storm. For comparison with other published results, the LNO_x production rates estimates per LINET stroke were scaled to estimates per LIS flash and finally multiplied with the global frequency of LIS flashes.

8.3 Results From Airborne DLR Thunderstorm Field Experiments Focusing on LNO_x

In this section, some main results from the DLR airborne field experiments carried out in the past years are summarized. Between 1992 and 2006, DLR participated in six airborne field experiments in mid-latitude, *subtropical* and tropical regions. Field experiment details are given in Table 8.1. Reasons for different findings concerning LNOx in the *tropics* (latitude 0–20°) compared to the subtropics (latitude: 20–35°) and mid-latitudes (35–55°) are discussed. Parameters influencing the LNOx production rate per flash are highlighted.

Table 8.1 Thunderstorm field experiments carried out with the DLR research aircraft Falcon between 1992 and 2006

DLR thunderstorm field experiments	Aircraft base	Mission flight coordinates (longitude)	Mission flight coordinates (latitude)	Month/year
CLEOPATRA Cloud experiment oberpfaffenhofen and transports	Southern Germany	~11°E	~48 °N (mid-latitudes)	July 1992
LINOX Lightning-produced NO_x	Southern Germany	6–13°E	46–50 °N (mid-latitudes)	July–Aug. 1996
EULINOX European lightning nitrogen oxides	Southern Germany	6–19°E	44–50 °N (mid-latitudes)	July 1998
TROCCINOX Tropical convection, cirrus and nitrogen oxides experiment	SE Brazil	38–55°W	10–28 °S (tropics-subtropics)	Feb–March 2004/2005
SCOUT-O3 Stratospheric-climate links with emphasis on the upper troposphere and lower stratosphere	Northern Australia	125–135°E	7–23 °S (tropics–subtropics)	Nov–Dec. 2005
AMMA African monsoon multidisciplinary analyses	Central Burkina Faso	10°W–3°E	4–17 °N (tropics–subtropics)	August 2006

8.3.1 LNO_x Mixing Ratios in Fresh Anvils Over Europe and NO_x Contribution From the Boundary Layer

At DLR, first airborne trace gas measurements in thunderclouds (yet without NO) were performed within the framework of CLEOPATRA (Hauf et al. 1995). During four penetrations in the anvil outflow (8–10 km), a distinct decrease in the range of 13–25 nmol mol^{-1} in O_3 mixing ratios was observed compared to mean background (~65–75 nmol mol^{-1}), while H_2O mixing ratios increased in the range of 0.18–0.30 g kg^{-1} compared to mean background (~0.2–0.6 g kg^{-1}). Accompanying 3D numerical thunderstorm simulations confirmed that boundary layer air with a similar chemical composition was transported upward, almost undiluted, into the anvil region.

During the LINOX experiment, for the first time in Europe, airborne NO_x measurements were performed within anvils of active thunderstorms (Huntrieser et al. 1998; Höller et al. 1999). More than 20 anvil penetrations were carried out with the Falcon aircraft at altitudes between 7 and 10 km. Mean NO_x mixing ratios, averaged during the anvil penetrations, varied between 0.8 and 2.2 nmol mol^{-1} and peak values reached up to 4 nmol mol^{-1}. To account for the amount of NO_x transported

upward from the boundary layer, the correlation between NO_x and CO_2 (a tracer for boundary layer air) was used. In medium-size and larger LINOX thunderstorms the contribution from lightning production dominated (60–75 %) and was on average $\sim 1.0 \pm 0.5$ nmol mol^{-1} NO_x. Based on the average mass flux rate at which air is advected out of single LINOX anvils, a rough mean estimate for the annual global NO_x production rate for lightning was obtained (assuming that LINOX thunderstorms are typical global thunderstorms), resulting in 4 Tg a^{-1}.

The fresh anvil outflow was probed almost 30 times in different deep convective clouds with and without lightning over central Europe during EULINOX (Huntrieser 2002). On average, about 70 % of the anvil-NO_x could be attributed to production by lightning (~ 0.9 nmol mol^{-1} LNO_x). These results agree very well with observations from the previous LINOX field experiment. However, during EULINOX an isolated supercell storm (diameter 40–50 km) with very high lightning activity was successfully penetrated on 21 July 1998 (mainly intracloud flashes). The Falcon aircraft performed measurements close to the convective core. A maximum NO mixing ratio of 25 nmol mol^{-1} was measured (Fig. 8.2a) when the aircraft was hit by a small lightning strike. In addition, the measurements in the supercell storm indicated very efficient vertical transport of polluted boundary layer air from the Munich urban area up to the anvil outflow region. Strongly enhanced CO_2 mixing ratios and condensation nuclei (CN) concentrations at the top of the boundary layer were observed in the same range in the anvil outflow region at 8–10 km altitude (Figs. 8.2a, b). Condensation nuclei concentrations are given for standard pressure and temperature (STP) conditions because this concentration is conserved in vertical motions. The chemical signatures found in the anvil region seem to partly "mirror" the conditions at the top of the boundary layer (located near 2.6 km on this day). To estimate the annual global NO_x production rate by lightning, the same method as for LINOX resulted in ~ 3 Tg a^{-1}. On the global scale, it was estimated that NO_x produced by lightning prevails over aircraft-produced NO_x (0.6 Tg a^{-1}). However, on the regional scale over Europe, NO_x emissions from aircraft of ~ 0.1 Tg a^{-1} dominate in comparison to LNO_x emissions of ~ 0.03 Tg a^{-1}.

The isolated EULINOX supercell storm observed on 21 July 1998 was studied in more detail by Fehr et al. (2004) and Ott et al. (2007). Simulations with three-dimensional (3D) *cloud-resolving model*s were performed. Fehr et al. (2004) introduced a new parameterization for LNO_x, representing the dispersion of LNO_x as Lagrangian particles instead of a Eulerian tracer field as used previously. The introduction of a LNO_x line source is more in line with the actual lightning physics. Airborne NO_x measurements were compared to model results indicating an equivalent or even higher NO_x production rate for intracloud IC flashes compared to cloud-to-ground CG flashes (ratio 1.4). Furthermore, the results indicate that >90 % of the total LNO_x production in the storm was produced by intracloud lightning, which is a much higher contribution than previously assumed. Ott et al. (2007) estimated a production rate of 360 mol NO per cloud-to-ground flash from simulations with a cumulus ensemble model, which is close to the value of 330 mol NO found by Fehr et al. (2004). However, Ott et al. (2007) found a

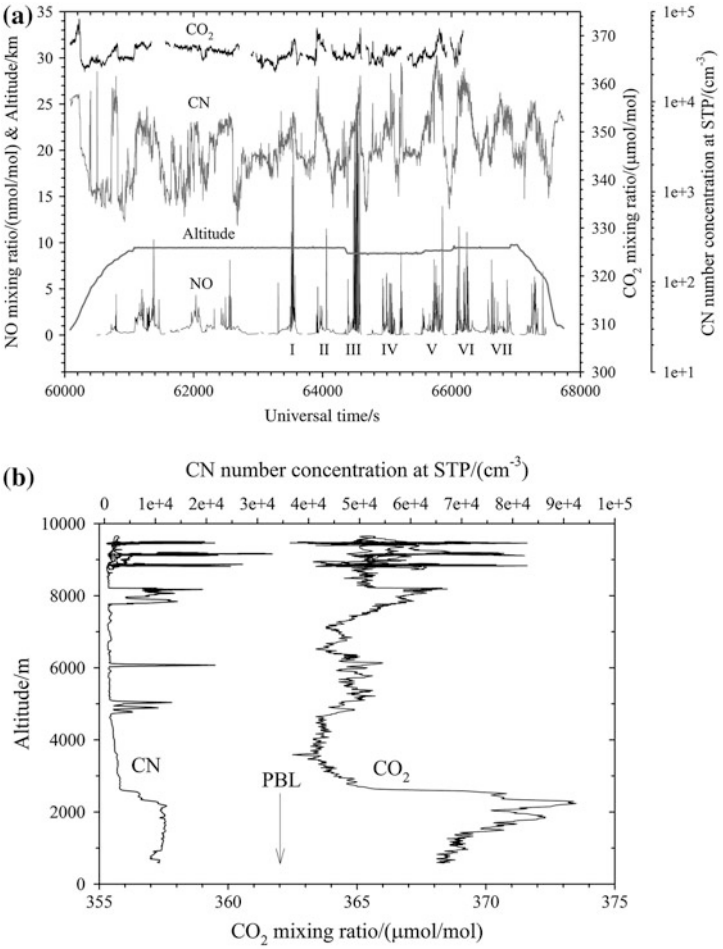

Fig. 8.2 **a** Mixing ratios of NO and CO_2, and particle concentration (condensation nuclei, CN) for standard pressure and temperature (STP) versus time during the EULINOX flight of 21 July 1998. The DLR Falcon repeatedly penetrated an isolated supercell storm, which is clearly visible in NO mixing ratios, labeled I–VII (modified from Huntrieser 2002). **b** Vertical profiles of condensation nuclei (CN) concentration at STP (left) and CO_2 mixing ratio (right) as observed during the EULINOX flight on 21 July 1998 with the DLR Falcon aircraft. The planetary boundary layer (PBL), located below 2.6 km, is discernible by high mixing ratios of CO_2 and high CN concentrations due to urban pollution (modified from Huntrieser 2002)

slightly lower ratio for the NO_x production rate per intracloud flash compared to cloud-to-ground flashes (ratio 1.0), suggesting an equal LNO_x production rate for both types of flash.

Table 8.2 Summary of observations during thunderstorm penetrations and estimate of the lightning-induced nitrogen oxides in units of nitrogen mass (LNOx) and planetary boundary layer (PBL) NO_x contribution in anvils (mixing ratios are given as χ). The LNOx production rates per LIS flash (P_{LNOx}) and the annual global LNOx production rates (G_{LNOx}) are estimated based on selected tropical and subtropical thunderstorms from AMMA, SCOUT-O3 and TROCCINOX and finally averaged over all of these thunderstorms (based on Table 4 in Huntrieser et al. 2011)

Field experiment	Anvil width, km	Mean χ_{LNOx} in anvils, nmol mol^{-1}	Ratio $\chi_{PBL\text{-}NOx}$ to χ_{LNOx}, %	Ratio NO/NO_y	P_{LNOx}, kg per LIS flash	G_{LNOx}, Tg a^{-1}
AMMA	150–220	~0.2	50/50	~0.3	1.0–2.5	1.4–3.5
SCOUT-O3 (except Hector)	30–60	0.4–2.3	10/90	0.6-0.8	1.3–5.4	1.8–7.6
SCOUT-O3 (only Hector)	~100	2.2–2.5	10/90	~0.8	4.1–4.8	5.7–6.6
TROCCINOX	30–50	0.1–1.1	20/80	0.3–0.8	1.0–2.8	1.3–3.9
	Mean tropical					~2
	Mean subtropical					~4
	Mean all (range)					~3 (1–8)

8.3.2 LNO_x in the Tropics and Subtropics: Parameters Influencing the LNO_x Production Rate

First airborne in situ measurements in tropical thunderstorms were conducted with the Falcon aircraft during TROCCINOX (Huntrieser et al. 2007, 2008). In general, the penetrated tropical thunderstorms were of sizes similar to European mid-latitude thunderstorms (Table 8.2). However, the tropical deep convective clouds reached higher altitudes (due to the higher tropopause) and the main anvil outflow was observed from 8 to 13 km. The mean NO_x mixing ratios during anvil penetrations varied between 0.2–1.6 nmol mol^{-1}, a similar or even lower range than observed in thunderstorms over Europe. Initially, higher NO_x mixing ratios in tropical thunderstorms compared to mid-latitude thunderstorms had been expected due to higher flash rates in the tropics. During TROCCINOX, on average ~80–90 % of the anvil-NO_x was attributed to LNOx as derived from the NO_x-CO correlations, which is a higher contribution compared to observations in European mid-latitude thunderstorms.

During TROCCINOX, both tropical and subtropical thunderstorms were investigated depending on the location of the South Atlantic convergence zone. In a case study, the chemical composition in three tropical and one subtropical thunderstorm was compared (Huntrieser et al. 2008). During anvil penetrations, mean NO_x mixing ratios were slightly lower (0.2–0.8 nmol mol^{-1}) in the subtropical thunderstorm compared to the selected tropical thunderstorms (0.7–1.2 nmol mol^{-1}). However, due to the higher flash rate in the tropical thunderstorms, the LNO_x production rate in nitrogen mass per flash was estimated to be lower in the tropical thunderstorms (1.2 kg per LIS flash) compared to the subtropical thunderstorm (2.2 kg per LIS flash). The equivalent mean annual global LNO_x nitrogen mass production rate based on these types of thunderstorms was estimated to be 1.6 and 3.1 Tg a^{-1}, respectively. The results indicate a similar LNO_x production rate for subtropical thunderstorms as observed in mid-latitude European thunderstorms; however, tropical thunderstorms seem to produce less LNO_x per flash (factor ~2). The stroke characteristics in tropical and subtropical thunderstorms were compared in detail and a major difference was observed in horizontal stroke length (also factor ~2). Higher upper level wind velocities in subtropical air masses (15–25 m s^{-1}) located closer to the subtropical jet stream compared to tropical air masses (5–10 m s^{-1}) induce higher vertical wind shear affecting the horizontal spread of deep convective clouds. It is suggested that enhanced vertical wind shear in the vicinity of a thunderstorm causes a horizontal separation of charged regions in the upper parts of a thunderstorm (in general, positive charge) in contrast to the mainly negatively charged region in the lower parts of the thunderstorm. This horizontal extension of charged layers over a wider area may generate flashes with enhanced horizontal lengths. Since LNO_x production is proportional to flash length, a higher NO_x production per flash can be expected in a higher vertical wind shear regime than in a mid-latitude and subtropical air mass compared to a tropical air mass.

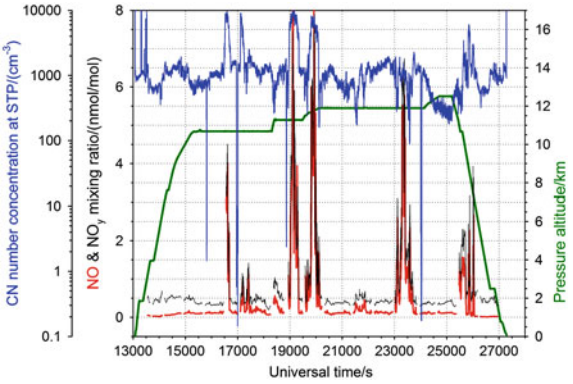

Fig. 8.3 Time series of NO and NO_y mixing ratios, condensation nuclei (CN) concentration at STP, and pressure altitude for the Falcon flight on 19 November 2005 during SCOUT-O3. Anvil penetrations of Hector are labelled "2a" in *black* (adapted from Huntrieser et al. 2009)

During the SCOUT-O3 experiment, the well-known thunderstorm "*Hector*" was investigated. This especially vigorous deep mesoscale convective system develops almost on a daily basis during the transition and monsoon break periods in response to diurnal heating over the Tiwi Islands north of Darwin. During SCOUT-O3, for the first time airborne in situ NO_x measurements were performed in Hector (Huntrieser et al. 2009). The main anvil outflow was located at ~11.5–12.5 km altitude, coinciding with the mean height of intracloud stroke sources at 12.7 km; however, cloud tops reached up to 18 km. The mean NO_x mixing ratios during these anvil penetrations at 11–12 km altitude (~100 km wide) were exceptionally high and varied between 2.2 and 2.5 nmol mol^{-1}, with NO peaks up to 8 nmol mol^{-1} (Fig. 8.3). In this case, more than 90 % of the observed anvil-NO_x was attributed to LNO_x and emissions from other NO_x sources in the boundary layer were almost negligible. The NO/NO_y ratio was also especially high in these thunderstorms (up to 0.8), indicative of fresh NO production by lightning (Table 8.2). In Hector, also the highest average stroke rate was observed of all thunderstorms ever investigated with the Falcon aircraft in the tropics. The stroke distribution in Hector during the Falcon flight is indicated in Fig. 8.4. The LNO_x N production rate was estimated as 4.1–4.8 kg per LIS flash for the well-developed Hector system, corresponding to an equivalent mean annual global LNO_x nitrogen mass production rate of ~5.7–6.6 Tg a^{-1}. This value is distinctly higher compared to the previous results presented here based on thunderstorms in mid-latitude and tropical regions (~2–4 Tg a^{-1}) (Table 8.2). Hector seems to be an exceptional thunderstorm for its vertical and horizontal extension and its lightning activity, producing more LNO_x than expected from a typical global thunderstorm. Therefore, this value can be seen as an upper level of what we may expect based on airborne observations in thunderstorm anvils. The average flash component length was enhanced compared to observations in tropical thunderstorms, even though Hector developed in an environment with low vertical wind shear. It was suggested that the dimension and the cell organization of the thunderstorm system may also play an important role for the flash lengths (e.g., inter-cloud flashes) that may originate. Especially the ice and graupel volume

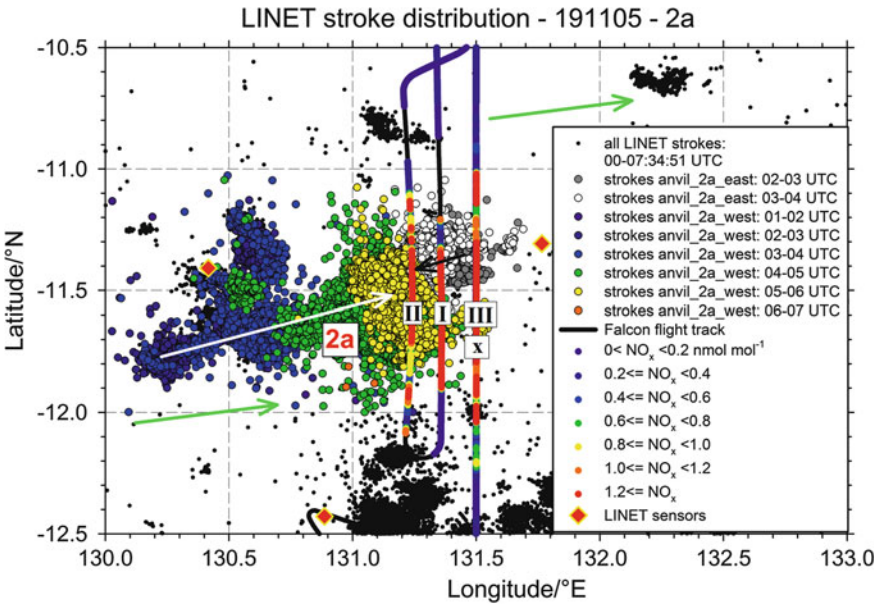

Fig. 8.4 Spatial and temporal LINET stroke distribution during SCOUT-O3, 00:00 and 08:00 UTC on 19 November 2005 (upper color scale). The observed NO_x mixing ratio along the Falcon flight path over Tiwi Island is superimposed (lower color scale). Enhanced NO_x mixing ratios were observed during the passage of the anvil outflow of Hector labelled "2a". All strokes registered before the Falcon penetrated Hector are colored. The *white arrows* indicate the direction of storm motion and the *green arrows* the main wind direction in the anvil outflow (not scaled according to velocity). In addition, the positions of the LINET sensors are indicated (*diamonds*) (adapted from Huntrieser et al. 2009)

mass within a thundercloud seem to be crucial parameters for the evolving flash length. The cold-cloud depth (distance between the 0 °C level and the storm height) seems to be one of the essential parameters impacting the flash characteristics. In the well-developed Hector system, elevated radar reflectivities >40 dBZ (containing graupel and hail) were observed up to 12 km altitude (Fig. 8.5, left storm), which is an exceptionally high altitude compared to ~6 km observed in other tropical thunderstorms, see, e.g., Fig 8.5 (right storm). The concentrations of ice and graupel particles were especially high in the well-developed Hector system, which may initiate lightning over longer distances and produce an especially high amount of LNOx per flash.

The most recent thunderstorm field campaign carried out in tropical regions with the Falcon aircraft was AMMA (Huntrieser et al. 2011). Within the fresh mesoscale convective system outflow during AMMA, mean NO_x mixing ratios were rather low (~0.3–0.4 nmol mol^{-1}) and only slightly enhanced compared to the background. Based on NO_x–CO correlations, it was estimated that LNO_x and boundary layer NO_x contributed about equally (50:50) to this enhancement, which

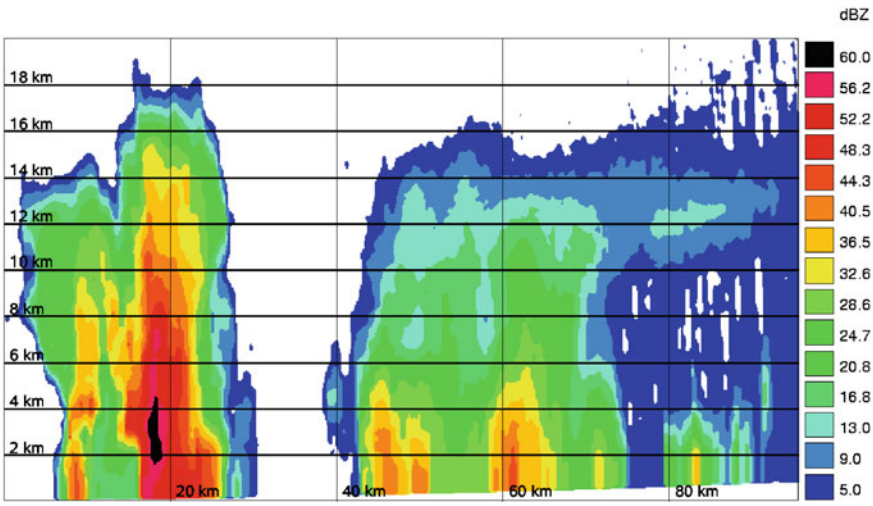

Fig. 8.5 Vertical cross-sections of radar reflectivity for the 19 November 2005 Hector thunderstorm over the Tiwi Islands (*left*) measured with a C-Band radar in Darwin (adapted from Huntrieser et al. 2009)

indicates that NO_x sources in the boundary layer are not negligible in this case. In comparison to the huge extension of these mesoscale convective systems (anvil outflow ~150–220 km wide), the average stroke rate was low and in the same range as observed during previous campaigns in much smaller tropical thunderstorms. The rather low lightning activity in these huge mesoscale convective systems may be affected by the lack of ice and graupel particles in the upper part of the cloud region. The 40-dBZ radar reflectivity level was in general located at altitudes <5–6.5 km. Due to the low cloud base and lifting condensation level (LCL), precipitation particles grow fast and fall out of the cloud at an early stage and only few *ice particle*s reach higher altitudes (known as a "warm-rain process" characterized by low ice water content and no large ice particles), which agrees with the observation of a large fraction of trailing stratiform precipitation in these kinds of mesoscale convective systems.

8.4 Lightning Parameterization in Models and Results from Simulations with DLR Models

For an assessment of the global impact of LNOx on the atmospheric composition, global modeling of lightning and the related LNO_x distribution is an important complement to the random probing of thunderstorms described above. Fundamental for modeling are meteorological fields which represent the *convection* that may develop to thunderclouds with lightning. The LNO_x model parameterizations

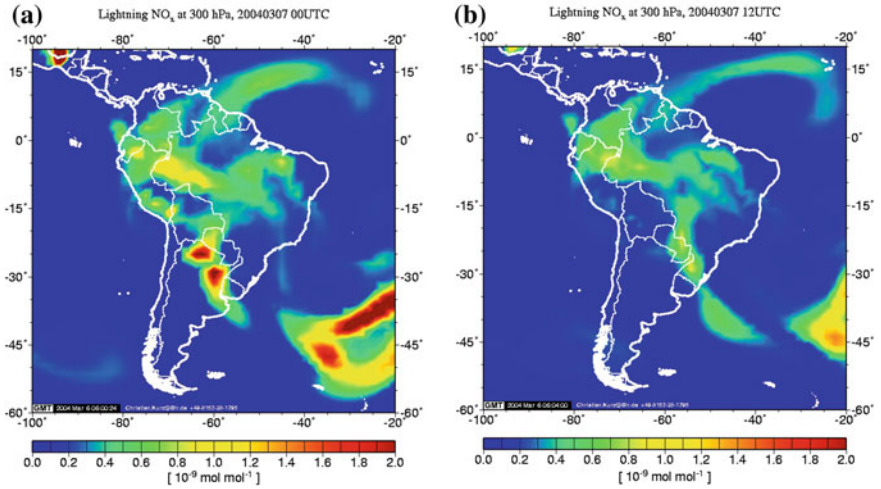

Fig. 8.6 ECHAM5/MESSy model forecasts of LNOx mixing ratio at 300 hPa for 00:00 UTC (*left*) and 12:00 UTC 7 March 2004 (*right*) during TROCCINOX (adapted from Huntrieser et al. 2007)

compute the spatial and temporal distribution of lightning events, the relative contribution from cloud-to-ground and intracloud flashes, and the vertical distribution of the emitted LNO_x.

Overviews of different LNO_x parameterizations are summarized in Schumann and Huntrieser (2007) and Pickering et al. (2009) LNO_x parameterizations have been implemented in 3D global models as General Circulation Models which compute their own meteorology in a climate mode. General circulation models may also use a "nudging" technique to adjust their simulated fields to current meteorological analysis or forecast fields from, e.g., the European Centre of Medium-Range Weather Forecasts (ECMWF). During TROCCINOX, DLR successfully used a nudged version of the coupled atmospheric chemistry general circulation model ECHAM5/MESSy to predict the distribution of regional LNO_x fields over South America, as shown in Fig. 8.6 (Kurz 2006; Huntrieser et al. 2007). A number of different methods have been used to parameterize the spatial distribution of the flash frequency in global general circulation models and regional chemical transport models (the latter use meteorological fields from general circulation models or current meteorological analyses or forecast fields as input). Up to recently, the most common method was to relate the flash frequency to the maximum *cloud-top height* (CTH). Alternative methods are based on such parameters as the convectively induced updraft velocity (UPD), the convective available potential energy (CAPE), or the convective precipitation rate. The latter parameterization was compared to measurements from the EULINOX experiment and showed good agreement (Meijer et al. 2001). A variant of the UPD parameterization, based on the mean upward convective mass flux in a general circulation

model, was introduced by Grewe et al. (2001) and applied by Kurz (2006) to the TROCCINOX data set. A comparison with observational data shows that the UPD parameterisation agreed better with the measured NO_x distributions in the upper troposphere, related to LNO_x, than with the model results from the commonly used CTH parameterization. This is in agreement with other published studies. Recently, the precipitating ice mass has been suggested as an important indicator of lightning activity, which is in agreement with the conclusions from the different DLR thunderstorm field experiments presented above.

Cloud-scale models partly use observed meteorological parameters (such as pressure, temperature, hydrometeor fields, updraft speed and thickness of the cold cloud region between 0 °C and cloud top) to simulate the spatial and temporal occurrence of flashes inside the cloud and the resulting LNOx distribution and transport. Only few two-dimensional (2D) and 3D models exist which explicitly simulate the lightning process as charge separation (when the electric field reaches critical strengths) and the resulting LNO_x formation.

The LNO_x source is distributed vertically within the models, resulting in effective vertical emission profiles. The horizontal distribution of LNO_x in models is sensitive to the shape of these profiles, as the wind speed and direction in the LNOx outflow altitude determines the distribution. Up to recently, it was assumed that the LNO_x emission profile has a C-shaped profile, differing slightly depending on storm type and region. It was assumed that part of the LNO_x emitted within the convective cloud is transported upward to the anvil region; another part is transported downwards with the downdrafts to the boundary layer. More recent results based on simulations with a 3D cloud-scale chemical transport model indicate that a large part of LNO_x remains in the middle and upper troposphere where it originated, suggesting a backward C-shaped profile, in contrast to the initial assumption (Ott et al. 2010).

8.5 Climate Impact of LNO_x

Tropospheric nitrogen oxides largely control the abundance of tropospheric ozone, which is a greenhouse gas contributing to the greenhouse effect. The sources of NO_x in the atmosphere are at very different locations, with most sources in the boundary layer, except for aircraft and lightning NO_x emissions (Fig. 8.1). The height of the emission not only impacts the atmospheric residence time of the emitted NO_x molecule, but also its ozone production potential. Emissions located at Earth's surface produce on average 10–30 molecules O_3 (Dahlmann et al. 2011), whereas upper tropospheric sources like air traffic and lightning produce 50 and 100 molecules O_3 per emitted NO_x molecule, respectively. The radiative efficiency of ozone, i.e., how much ozone contributes to its greenhouse effect, is again dependent on its location. The more equatorwards and the closer to the tropopause the ozone is located (stronger solar radiation), the higher is its impact on the greenhouse effect. Most of the LNO_x-induced ozone is located in the tropical upper troposphere, leading to the

strongest greenhouse effect compared to all other sources. The total *radiative forcing* by LNO_x is estimated to be around 0.5 W/m^2, which is around 35–40 % of the total greenhouse effect resulting from all O_3 produced by NO_x.

Thunderstorms and lightning are affected by climate variability and of course climate change. The interannual variability of annual mean lightning is too large and the measurement time series too short to identify any statistically significant trend. Numerical simulations with climate models have a data set long enough to overcome statistical limitations. However, the picture they provide is rather inconclusive. Models using CTH parameterization basically simulate no increasing lightning frequency, whereas one model based on UPD parameterization simulates a decrease in lightning frequencies. In that study convection becomes stronger with larger mass fluxes but occurs more rarely, which leads to trends of −0.3 % per decade (Grewe 2009).

8.6 Conclusions

Extensive airborne field experiments performed by DLR in the past few years indicate that most thunderstorms in tropical regions produce less LNO_x than previously assumed (Huntrieser et al. 2008, 2009, 2011). The overall mean for the annual global LNOx N-mass production rate, estimated from airborne measurements in selected AMMA, SCOUT-O3 and TROCCINOX thunderstorms, is ~ 3 Tg a^{-1}, with a most-likelihood range of ~ 1–8 Tg a^{-1} (see Table 8.2). These results indicate a slightly lower global LNO_x source strength compared to the value of 5 ± 3 Tg a^{-1} given in Schumann and Huntrieser (2007) as the best estimate of the annual global LNOx production rate based on available literature up to that date. In general, thunderstorms in tropical regions have flash rates similar to those observed in subtropical and mid-altitude regions; however, the NO production rate per flash was estimated to be lower. It is suggested that the vertical wind shear may influence the overall flash length by separating the charged regions in the thunderstorms differently. In tropical regions, the predominantly weak wind shear may induce shorter flash lengths in general, resulting in less LNOx production per flash. Furthermore, the ice and graupel volume mass within a thundercloud seem to be crucial parameters for the evolving flash length. Model simulations performed by DLR indicate that a variant of the UPD parameterization based on the mean upward convective mass flux in a general circulation model can be efficiently used to parameterize LNOx (Kurz 2006; Grewe et al. 2001). Local NO sources in the upper troposphere are distinctly more efficient in producing O_3 compared to sources in the boundary layer. LNOx therefore has an important climate impact, since it contributes up to $\sim$35–40 % of the total greenhouse effect resulting from O_3 production by the different atmospheric NO sources (Dahlmann et al. 2011).

Acknowledgments The work presented here was mainly funded by the European Commission and DLR. During the field campaigns, many colleagues from the DLR Institute of Atmospheric Physics and the DLR Flight Experiments Department contributed to this work.

References

Chameides, W.L., Davis, D.D., Bradshaw, J., Rodgers, M., Sandholm, S., Bai, D.B.: An estimate of the NOx production rate in electrified clouds based on NO observations from the GTE/CITE 1 fall 1983 field operations. J. Geophys. Res. **92**, 2153–2156 (1987). doi:10.1029/JD092iD02p02153

Dahlmann, K., Grewe, V., Ponater, M., Matthes, S.: Quantifying the contributions of individual NOx sources to the trend in ozone radiative forcing. Atmos. Environ. **45**(17), 2860–2868 (2011). doi:10.1016/j.atmosenv.2011.02.071

Dickerson, R.R., Huffman, G.J., Luke, W.T., Nunnermacker, L.J., Pickering, K.E., Leslie, A.C.D., Lindsey, C.G., Slinn, W.G.N., Kelly, T.J., Daum, P.H., et al.: Thunderstorms: an important mechanism in the transport of air pollutants. Science **235**, 460–465 (1987). doi:10.1126/science.235.4787.460

Fehr, T., Höller, H., Huntrieser, H.: Model study on production and transport of lightning-produced NOx in an EULINOX supercell storm. J. Geophys. Res. **109**, (2004). doi:10.1029/2003JD003935

Grewe, V., Brunner, D., Dameris, M., Grenfell, J.L., Hein, R., Shindell, D., Staehelin, J.: Origin and variability of upper tropospheric nitrogen oxides and ozone at northern mid-latitudes. Atmos. Environ. **35**, 3421–3433 (2001). doi:10.1016/S1352-2310(01)00134-0

Grewe, V.: Impact of climate variability on tropospheric ozone. Sci. Total Environ. **374**, 167–181 (2007)

Grewe, V.: Impact of lightning on air chemistry and climate. In: H. D. Betz, U. Schumann, and P. Laroche, (eds.) Lightning: Principles, Instruments and Applications Review of Modern Lightning Research, 524–551. Springer, Heidelberg (2009)

Hauf, T., Schulte,P., Alheit, R., Schlager, H.: Rapid vertical trace gas transport by an isolated midlatitude thunderstorm. J. Geophys. Res. **100** (D11), 22957–22970 (1995). doi:10.1029/95JD02324

Höller, H., Finke, U., Huntrieser, H., Hagen, M.,Feigl, C.: Lightning-produced NOx (LINOX): Experimental design and case study results. J. Geophys. Res. **104** (D11), 13,911–913,922, (1999). doi:10.1029/1999JD900019

Höller, H., Betz, H.-D., Schmidt, K., Calheiros, R.V., May, P., Houngninou, E., Scialom, G.: Lightning characteristics observed by a VLF/LF lightning detection network (LINET) in Brazil, Australia, Africa and Germany. Atmos. Chem. Phys. **9**, 7795–7824 (2009). doi:10.5194/acp-9-7795-2009

Houze, R. A.: Mesoscale convective systems. Rev. Geophys. **42**, RG4003, (2004). doi:10.1029/2004RG000150

Huntrieser, H., Schlager, H., Feigl, C., Höller, H.: Transport and production of NOx in electrified thunderstorms: survey of previous studies and new observations at mid-latitudes. J. Geophys. Res. **103** (D21), 28,247–228,264, (1998). doi:10.1029/98JD02353

Huntrieser, H., Feigl, C., Schlager, H., Schröder, F., Gerbig, C., van Velthoven, P., Flatøy, F., Théry, C., Petzold, A., Höller, H.: Airborne measurements of NOx, tracer species, and small particles during the European lightning nitrogen oxides experiment. J. Geophys. Res., 107 (D11), 4113, (2002) doi:10.1029/2000JD000209

Huntrieser, H., Schlager, H., Roiger, A., Lichtenstern, M., Schumann, U., Kurz, C., Brunner, D., Schwierz, C., Richter, A., Stohl, A.: Lightning-produced NOx over Brazil during TROCCINOX: airborne measurements in tropical and subtropical thunderstorms and the importance of mesoscale convective systems. Atmos. Chem. Phys. **7**, 2987–3013 (2007). doi:10.5194/acp-7-2987-2007

Huntrieser, H., Schumann, U., Schlager, H., Höller, H., Giez, A., Betz, H.-D., Brunner, D., Forster, C., Pinto Jr, O., Calheiros, R.: Lightning activity in Brazilian thunderstorms during TROCCINOX: implications for NOx production. Atmos. Chem. Phys. **8**, 921–953 (2008). doi:10.5194/acp-8-921-2008

Huntrieser, H., Schlager, H., Lichtenstern, M., Roiger, A., Stock, P., Minikin, A., Höller, H., Schmidt, K., Betz, H.-D., Allen, G., et al.: NOx production by lightning in hector: first airborne measurements during SCOUT-O3/ACTIVE. Atmos. Chem. Phys. **9**, 8377–8412 (2009). doi:10.5194/acp-9-8377-2009

Huntrieser, H., Schlager, H., Lichtenstern, M., Stock, P., Hamburger, T., Höller, H., Schmidt, K., Betz, H.-D., Ulanovsky, A., Ravegnani, F.: Mesoscale convective systems observed during AMMA and their impact on the NO_x and O_3 budget over West Africa. Atmos. Chem. Phys. **11**, 2503–2536 (2011). doi:10.5194/acp-11-2503-2011

Kurz, C.: Entwicklung und Anwendung eines gekoppelten Klima-Chemie-Modellsystems, p. 142. Doctoral Dissertation, Ludwig-Maximilians-Universität München (2006)

Meijer, E.W., van Velthoven, P.F.J., Brunner, D., Huntrieser, H., Kelder, H.: Improvement and evaluation of the parameterization of nitrogen oxide production by lightning. Phys. Chem. Earth (C) **26**, 577–583 (2001)

Ott, L.E., Pickering, K.E., Stenchikov, G.L., Huntrieser, H., Schumann, U.: Effects of lightning NOx production during the 21 July European Lightning Nitrogen Oxides Project storm studied with a three-dimensional cloud-scale chemical transport model. J. Geophys. Res. **112**, D05307 (2007). doi:10.1029/2006JD007365

Ott, L.E., Pickering, K.E., Stenchikov, G.L., Allen, D.J., DeCaria, A.J., Ridley, B., Lin, R.-F., Lang, S., Tao, W.-K.: Production of lightning NOx and its vertical distribution calculated from three-dimensional cloud-scale chemical transport model simulations. J. Geophys. Res. **115**, D04301 (2010). doi:10.1029/2009JD011880

Pickering, K. E., Huntrieser, H., Schumann, U.: In: H.-D. Betz, U. Schumann, P. Laroche, (eds.), Lightning and NOx production in global models. lightning: principles, instruments and applications review of modern lightning research, pp. 551–571, Springer Heidelberg (2009)

Schumann, U., Huntrieser, H.: The global lightning-induced nitrogen oxides source. Atmos. Chem. Phys. **7**, 3823–3907 (2007). doi:10.5194/acp-7-3823-2007

von Liebig, J.: Une note sur la nitrification. Ann. Chem. Phys. **35**, 329–333 (1827)

Ice Supersaturation

9

Klaus Gierens, Peter Spichtinger and Ulrich Schumann

Abstract

Ice supersaturation is a frequent phenomenon in cold (below 0 °C) regions of the atmosphere. Ice-supersaturated regions are regions where cirrus clouds can form and where condensation trails are long lasting (up to many hours). This article provides an introduction to the topic of ice supersaturation in the atmosphere. It gives answers to the following questions: What is relative humidity? What is the origin of ice supersaturation? Since when does one know about ice supersaturation in the atmosphere? Where does one find ice supersaturation in the atmosphere and how often? What are the properties of ice supersaturated regions? How important is ice supersaturation for the atmosphere, for the weather and for climate? And finally: Which are the highest values of ice supersaturation in the atmosphere?

K. Gierens (✉) · U. Schumann
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Klaus.Gierens@dlr.de

U. Schumann
e-mail: Ulrich.Schumann@dlr.de

P. Spichtinger
Johannes Gutenberg-Universität Mainz, Institut für Physik der Atmosphäre,
Becherweg 21, 55099 Mainz, Germany
e-mail: spichtin@uni-mainz.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_9,

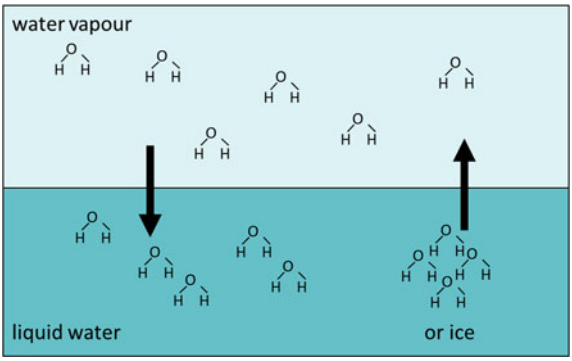

Fig. 9.1 Water molecules in liquid water, or ice, and in the gas space above. The arrows indicate the exchange of water molecules between these reservoirs, in both directions, that is, from the condensed phases into the vapor phase and vice versa. If the concentration of water molecules in the vapor phase has exactly that value needed to make the net exchange rate (in minus out) zero, the water vapor is said to be saturated

9.1 What is Relative Humidity?

Air in Earth's atmosphere contains mainly nitrogen and oxygen; *water vapor* is merely a *trace gas*. But this trace gas is responsible for two-thirds of the natural *greenhouse effect*, for formation of clouds and precipitation in all forms (*rain*, *snow*, *hail*), and for one-sixth of the energy transport from Earth's surface into the atmosphere via evapotranspiration of water at the ground (taking up *latent heat* of fusion) and *condensation* and *freezing* (cloud formation) in the atmosphere (release of the latent heat). The concentration of water vapor can be quantified in a multitude of ways, for instance by the number of water molecules per volume or per mass of air, *partial pressure*, etc.; these measures are various instances of absolute humidity. But the most popular scale, which is shown on all domestic hygrometers, is relative humidity. *Relative humidity* is not only a popular scale; it is used in meteorology and atmospheric science as well because it has certain advantages over absolute measures of water vapor concentration. First: relative humidity is usually in a familiar range between 0 % (totally dry) and 100 % (*saturation*), while an absolute scale must be much wider because the concentration of water molecules decreases from the ground to the *tropopause* by roughly a factor of 10 000. Second: cloud formation (condensation) is controlled by relative humidity, not by absolute humidity.

So what is relative to what in the notion of relative humidity? It is the ratio of the actual concentration of water molecules to the concentration of water molecules at saturation or, alternatively, the ratio of the partial pressure of water molecules to their partial pressure at saturation. Saturation is the expression for a dynamic equilibrium that we consider next, see Fig. 9.1. Consider a large basin of water. As the water molecules are in permanent motion, colliding with each other,

it sometimes happens that a molecule gains enough energy to leave the basin and enter the vapor space above it. In turn, water molecules in the vapor phase are in permanent motion as well and sometimes one of them enters the liquid below. So there is a permanent exchange of water molecules between the liquid in the basin and the vapor space above. While the concentration of water molecules in the liquid varies only slightly with temperature and pressure, the concentration of water molecules in the vapor can vary widely. Therefore, the flux of water molecules out of the liquid into the vapor is almost constant, while the flux from the vapor into the liquid increases with the concentration of water molecules in the vapor. Dynamic equilibrium means that these two opposite fluxes are equal so that the net effect is zero (that is, the water mass in the basin is constant, neither net evaporation nor net condensation). The water vapor in this case is said to be saturated, that is, its relative humidity is 100 %. This concentration is the reference value for the scale of relative humidity. Five things should be noted here.

- First, the explanation was only in terms of water molecules; whether there is pure water vapor or other mixtures of gases above the water basin is irrelevant as long as this mixture behaves as an ideal gas. So relative humidity is a property of water vapor alone whether there are other molecules or not, and it is the water vapor that is saturated; statements like "the air is saturated with water vapor" are incorrect.
- Second, we can make a similar consideration with ice instead of liquid water in the basin. However, as the molecules are much stronger bonded in ice than in liquid water, the molecule flux out of ice is smaller than that out of liquid. Thus a lower concentration of water molecules in the vapor phase can balance that flux, and saturation with respect to an ice surface is achieved with a smaller vapor concentration than saturation with respect to a liquid surface. This means that we have two scales of relative humidity, one with respect to water and one with respect to ice. At subzero temperatures where supercooled liquid water can exist down to about −40 °C, both scales can be applied simultaneously, but at temperatures above the melting point of ice one can only use relative humidity with respect to liquid water, which is usually meant on our domestic hygrometers.
- Third, although these instruments usually have scales from 0 to 100 %, there is nothing that forbids an even higher vapor concentration of water molecules, that is, *supersaturation*. However, supersaturation is not a stable situation. It is out of equilibrium. Then processes will occur (e.g., condensation) that drive the situation back to equilibrium, i.e., to saturation. Similarly, *subsaturation* is unstable as long there is a condensed phase; it induces *evaporation* or *sublimation*.
- Fourth, the concentration at which water vapor is saturated (both versions) depends strongly on temperature (Fig. 9.2). The *Clausius-Clapeyron equation* formulates this fact usually in the partial pressure scale: $\mathrm{d}\ln e/\mathrm{d}T = L/RT^2$, where R is the gas constant and L is the latent heat of evaporation or sublimation and e is the partial pressure of water vapor at either ice or liquid saturation (depending on the choice of L).

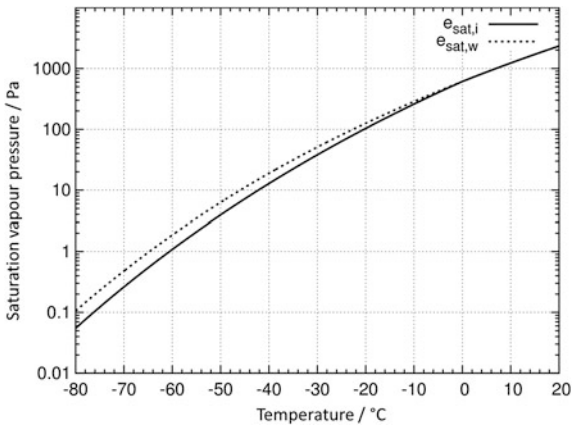

Fig. 9.2 Saturation vapor pressures of water relative to liquid supercooled water (dotted, $e_{sat,w}$) and ice (solid, $e_{sat,i}$) at subzero temperatures as a function of temperature T. These functions solve the Clausius-Clapeyron equation. Note the strong increase of saturation vapor pressure: Over the temperature range from -80 to 0°C the saturation vapor pressure increases by a factor of more than 10 000. Note also that water saturation implies ice supersaturation

- And fifth, although we have explained relative humidity using the concept of dynamic equilibrium, we can use the same scale for vapor concentration in situations where neither a water nor an ice phase is present.

9.2 What is the Origin of Ice Supersaturation?

When air rises in the atmosphere it gets colder because the air pressure decreases and the saturation concentration for water vapor decreases as well, as shown in Fig. 9.2, while the water vapor concentration decreases only in proportion to the decreasing air pressure. Thus the relative humidity increases. In the laws of physics there is nothing that prohibits the concentration of water molecules from increasing above its saturation value. If this happens, it is a *metastable* situation. The state is labelled metastable instead of instable to indicate that it can in principle last for a long time (in contrast, in an instable state processes start immediately to drive the state closer to equilibrium). The process that eventually brings back the vapor concentration to its saturation value (in other words, that reduces relative humidity to 100 %) is condensation. Condensation is the formation of *water droplets*. However, without particles in the air (the so-called condensation nuclei, a certain fraction of the aerosol) the formation of droplets would only commence at relative humidity values of several 100 %! Usually we do not measure such high values in the atmosphere. They do not generally appear because there are plenty of tiny particles in the air that serve as condensation nuclei. Condensation nuclei help water molecules to attach to each other and form a droplet. As liquid water is not a

highly organized material, the conditions a particle has to fulfil as a suitable *condensation nucleus* are weak. Thus there are always enough condensation nuclei around so that water clouds can form almost immediately once the relative humidity with respect to liquid water rises over 100 %. Supersaturation with respect to water can therefore be considered an instable state in the atmosphere, because condensation commences almost immediately in such a state.

Figure 9.2 shows that ice saturation occurs at a lower vapor concentration than does water saturation. This raises the question whether *ice crystals* shouldn't form before water droplets can condense. But this is not observed in the atmosphere. In contrast to the relatively easy formation of water droplets, the formation of ice crystals in the atmosphere is much more complicated, and although a variety of ice forming processes exist, they all proceed at significant supersaturation with respect to ice. The complication of ice formation lies in its crystal structure, a highly organized structure that cannot form on all kinds of (even hydrophilic) particles. In fact, the conditions on particle surface properties which make them suitable as *ice nuclei* are so stringent that only about one of a million (with stark variation) aerosol particles can serve as an ice nucleus. The rarity of ice nuclei explains that ice crystals do not readily form from the vapor as soon as saturation with respect to ice is surpassed. Also, water droplets do not necessarily freeze at subzero temperatures in the atmosphere, so supercooled water clouds (again a metastable state) are frequent in the sky. They are potentially dangerous for air traffic, because supercooled water can freeze on aircraft wings and other surfaces (icing).

In the upper troposphere temperatures are generally so low ($T < -40$ °C) that a pure water droplet would freeze spontaneously. Thus, water clouds in a convective situation where the droplets are lifted with the updraught to these cold regions get icy cloud tops (the ice *anvils* of *thunderstorm* clouds are the most prominent example). However, many *cirrus* clouds (i.e., clouds completely consisting of ice) form in nonconvective situations from liquid aerosol particles that consist of aqueous solutions of sulphuric acids and other substances. Such *solution droplets* do not freeze spontaneously at < -40 °C as pure water does. Again it is the crystalline structure of the ice that needs specific formation conditions. The crystal structure cannot form when too many too large foreign molecules (like a H_2SO_4 molecule) are present; their sheer presence disturbs the formation of the crystal lattice. Freezing can only occur once the solution is highly diluted. In order to reach this state, the solution droplets must gain water molecules from the ambient air and that needs an ambient water vapor concentration corresponding to relative humidity with respect to ice of more than 145 % (but still below water saturation).

Thus we see that ice formation in the atmosphere does not generally proceed at ice saturation; instead, relatively high supersaturation is needed before ice crystals are formed. Thus *ice supersaturation* can be considered a metastable state; processes do not start immediately to drive back ice supersaturation to saturation in the atmosphere. Ice supersaturation can hence last for quite some time (more than a day) and air volumes with ice supersaturated air can be quite extended (see below).

9.3 Since When Does One Know About Ice Supersaturation in the Atmosphere?

The fact that liquid water can be supercooled was discovered by D. Fahrenheit in 1721, almost 300 years ago. He found that one could keep water droplets contained in an evacuated glass bulb liquid at −9 °C (= 15.8 °F) for several hours. The liquid turned into ice only on shaking the glass bulbs. This was the first observation of liquid water in a metastable state. Conjectures that water vapor can exist in a metastable state, i.e., supersaturated with respect to ice, came about in the first decade of the 20th century. The first observations of ice supersaturation were probably made by A. Wegener on the Danmark-Expedition to northeast Greenland in 1906–08. The first of quite a number of recordings of ice supersaturation was obtained on 26 August 1906 in Danmarks Havn using a carefully calibrated hair hygrometer in a ground-based weather hut. Wegener believed his values since he knew that even supersaturation with respect to water had been found in *fog* by A. Wagner with a number of hygrometers on the Hohen Sonnblick (3106 m) in 1908 and he concluded that "it is evident from these values that the air is on average substantially supersaturated with respect to ice in the coldest season, that is, from November to March."[1]

Around 1940 there was a scientific discussion about the formation of ice crystals further up in the atmosphere. While Wegener and W. Findeisen thought that ice would usually form by direct *deposition* onto appropriate solid particles, L. Krastanow and E. Wall gave arguments and presented thermodynamic calculations showing that the ice phase in the atmosphere usually forms via the liquid phase, i.e., by the freezing of supercooled droplets. E. Glückauf and H. K. Weickmann were probably the first who measured ice supersaturation in the tropopause region. In 1945, Glückauf reported from hair hygrometer measurements over southern England that (very high) ice supersaturation frequently occurs in the upper troposphere, and Weickmann, who took microphotographs of ice crystals on research flights, concluded that cirrus clouds form mainly via the water phase and not as soon as ice saturation is reached. He characterized the upper troposphere as a region of high ice supersaturation, albeit a region of low absolute humidity. Only a few years later, 1946–48, B. M. Cwilong and V. J. Schaefer succeeded in supercooling small water drops down to the supercooling limit for pure water at around −40 °C. At even lower temperatures water drops would freeze spontaneously.

It seems that the topic of ice supersaturation was then forgotten until the mid-1980s when studies on weather modification, in particular a study on cloud and precipitation formation in clear air, brought it to a renaissance. The term "ice-supersaturated region" (ISSR) was coined by Detwiler and Pratt in 1984. But it still took another two decades before the existence of ice-supersaturated regions in upper tropospheric clear air was generally accepted. In the late 1980s K. Sassen,

[1] "Aus diesen Zahlen geht zur Evidenz hervor, dass zur kältesten Jahreszeit, nämlich in den Monaten November bis März, im Durchschnitt die Luft in Bezug auf Eis erheblich übersättigt ist"(Wegener 1914).

A. Heymsfield, and their co-workers suggested that nonconvective formation of cirrus clouds at temperatures below −40 °C occurs mainly via freezing of aqueous solution droplets, a process that requires substantial ice supersaturation, as explained above. Research-quality instruments on research aircraft like the DLR Falcon and regular humidity measurements taken from a handful of specially instrumented passenger aircraft (the MOZAIC[2] program) provided evidence in the late 1990s that ice supersaturation occurs quite frequently in the upper troposphere (Gierens et al. 1999, 2000). Theoretical considerations by Khvorostyanov and Sassen, and later Jensen and co-workers showed that ice supersaturation detected even within cirrus clouds could be plausible. Also, the frequent observation of long-lasting condensation trails from aircraft in an otherwise blue sky could only be explained by assuming ice supersaturated conditions. Evidence of the existence of ISSRs was corroborated by many other data sources during the first decade of the 21st century, including various satellite instrument data, calibrated and corrected *radiosonde* data, and data from airborne research campaigns. These data have also been used to determine properties of ISSRs (see below).

The last major step to general acceptance of the reality of ice supersaturation was the implementation of ISSRs into the Integrated Forecast System of the European Centre for Medium-Range Weather Forecasts (ECMWF) in September 2006 (Tompkins et al. 2007). Thus it took almost exactly 100 years from the first observation of ice supersaturation by A. Wegener to the representation of this metastable state in the world's leading weather forecast model.

9.4 Where Does One Find Ice Supersaturation in the Atmosphere and How Often?

As explained above, water saturation implies ice supersaturation at subzero temperatures. Thus all supercooled water clouds are in an ice supersaturated state which leads to an important process, the so-called Wegener-Bergeron-Findeisen process,[3] as soon as ice crystals get into the cloud. Although this is very important, supercooled water clouds are usually not meant by the expression "*ice supersaturated region*." Rather, there is an understanding that ISSRs are cloud-free regions in the cold layers (i.e., below the supercooling limit of pure water) of the upper troposphere where the relative humidity exceeds ice saturation. Ice supersaturation has been detected within cirrus clouds as well, and not all measurement techniques (e.g., radiosondes) are able to discern cloud-free air from cloudy air. Thus a clear distinction between ISSRs as cloud-free regions and ice supersaturation within clouds cannot be guaranteed in all databases. In particular, there is a correlation

[2] Measurement of Ozone and Water Vapour by Airbus In-Service Aircraft, a European project.

[3] Since liquid saturation within a supercooled water cloud implies ice supersaturation, ice crystals that get into such a cloud experience strong growth at the expense of the liquid droplets that start to evaporate as soon as the relative humidity gets below water saturation (i.e., a situation with ice super-, but liquid sub-saturation).

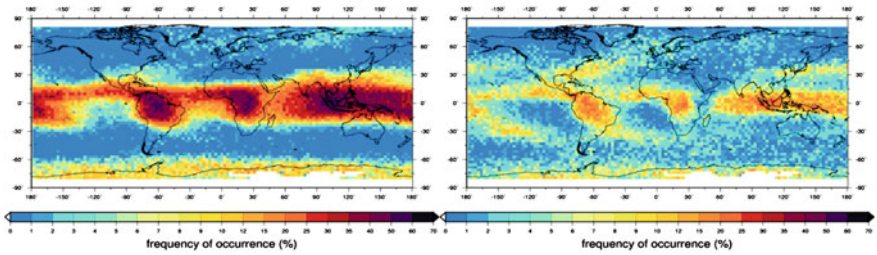

Fig. 9.3 Global distribution of ice-supersaturated regions obtained from MLS data. Shown are ISSRs in two roughly 3 km thick layers, one centered at 147 hPa (around 13.5 km, left) and one centered at 215 hPa (around 11.5 km altitude, right). *(This is a color version of a figure in QJRMS,* Spichtinger et al. (2003b))

between ISSRs and so-called *subvisible* cirrus, which are clouds that are optically too thin to be visible to a human observer. As a strict lower boundary of optical thickness of clouds cannot be given, the transition from a cloud-free ISSR and ice supersaturation within (subvisible) cirrus must be fuzzy.

We have already mentioned the MOZAIC program in which research-quality hygrometers installed on commercial airliners have provided a wealth of humidity data from the *upper troposphere* and *lower stratosphere* since mid-1994. These data were the first to be used to determine where and how often there is ice supersaturation in the upper troposphere. The statistical distribution of relative humidity values and the horizontal extensions of ISSRs were also derived from MOZAIC for the first time (see below). From these data it was determined that the aircraft participating in the MOZAIC program flew about 15 % of the cruising time in ice supersaturated air. Thus it is not a rare phenomenon.

Satellite data are best suited to obtain a global picture of where ISSRs are, although only large ISSRs or ISSRs with high supersaturation can generally be detected by satellite instruments. Two data sources have been used to obtain such global distributions, the Microwave Limb Sounder (MLS) on the UARS satellite (Spichtinger et al. 2003b) and the Atmospheric Infrared Sounder (AIRS) on NASA's AQUA satellite (Gettelman et al. 2006; Lamquin et al. 2012). Data from the Tiros Operational Vertical Sounder (TOVS) have been used as well (Gierens et al. 2004), but only northern mid-latitudes and polar latitudes were covered in that study.

Figure 9.3 shows the ISSR distribution in two roughly 3 km thick layers, centered at 147 and 215 hPa. Obviously, the distributions differ a lot in these two layers. While ISSRs are mainly concentrated in the tropics (in the so-called *tropical transition layer*, TTL) on the 147 hPa layer, they start to appear at mid-latitudes (mainly via storm tracks) in the lower layer. The occurrence frequencies are generally very large in the TTL and can reach 50 %, but as mentioned above, part of this may be thin and subvisible cirrus. A more complete picture in the vertical dimension can be obtained from the AIRS data (Lamquin et al. 2012, their

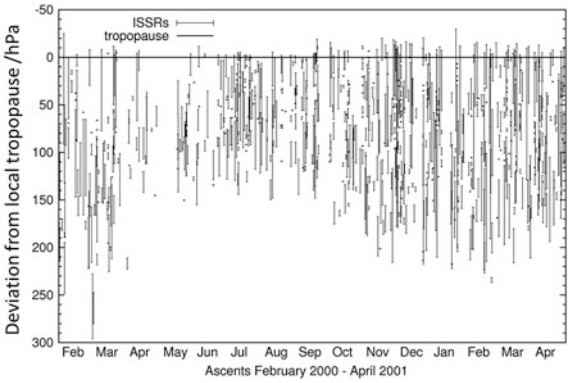

Fig. 9.4 Radiosonde detections of ISSRs (vertical bars) over the Lindenberg meteorological observatory from February 2000 until April 2001. The horizontal bar represents the local tropopause. (*A slightly different color version of this figure appeared in Meteorol. Z., Spichtinger* et al. 2003a.)

Fig. 9). Here we find the following: ISSRs above 150 hPa are confined to the TTL with very high frequencies of occurrence, consistent with the MLS result. Between 150 and 200 hPa ISSRs start to populate extra-tropical latitudes but are still mainly confined to the tropics. From 200 to 250 hPa we see most ISSRs appearing at the mid-latitudes along storm tracks, as indicated already in the MLS data. In turn, ISSRs become rare in the tropics. The so-called subsidence zones are always void of ISSRs because the subsiding air is very dry due to adiabatic compression and heating. Below 250 hPa most ISSRs are located in the mid-latitudes, and at still lower layers (300–500 hPa) they are even more confined to polar zones. This pattern already indicates that ISSRs follow approximately the altitude of the tropopause from the equator to the poles; ISSRs occur only very seldom in the generally too dry stratosphere. However, there is one special feature that does not fit into this pattern; this is the ice supersaturation over Antarctica, which is present far above the local tropopause. This ISSR is a seasonal feature; it occurs during the polar winter when the air, enclosed within a stable *polar vortex*, cools down by irradiating *infrared radiation* to −80 °C and lower. Figure 9.2 shows that the saturation pressure in such cold air is extremely low; this makes it possible to obtain high supersaturation values in excess of 200 % in the MLS data (about water saturation).

The fact that ISSRs are mainly confined to a layer below the tropopause has been demonstrated most clearly using calibrated and corrected radiosonde data from the DWD[4] observatory at Lindenberg (Spichtinger et al. 2003a). There is no information on cloud occurrence in the radiosonde data, so some of the recorded ISSRs could actually be ice supersaturation within cirrus clouds. Figure 9.4 shows the local tropopause as the horizontal line and the detected layers with ice supersaturation as the vertical bars. The data are from 15 months and also show a certain seasonal variation, namely from April to September fewer ISSRs in a narrower layer beneath the tropopause and in the rest of the year more ISSRs within a thicker layer. The vast

[4] DWD: Deutscher Wetterdienst, Germany's national weather service.

majority of ice supersaturation has been detected below the tropopause and only a minor fraction was just above it. Sometimes there are two or more ISSRs stacked on top. A wide variation of vertical extensions of the supersaturated layers is evident from the figure, but this is discussed in the next section.

9.5 What are the Properties of Ice Supersaturated Regions?

For a long time problems with radiosonde humidity sensors in the cold upper troposphere left the existence of ice supersaturation widely unknown even to meteorologists. But sensor improvements and the development of correction methods for known biases from the mid-1990s on made radiosonde data a valuable source of information on ISSRs, an example of which is shown in Fig. 9.4. Radiosonde data like those from Lindenberg have been used to infer the thickness of ice supersaturated layers. It is important to study radiosonde data from many launch sites because the typical thickness of ISSRs may vary with location. The average layer thickness obtained from the Lindenberg data with an altitude resolution of about 30 m was 560 m. Data from Ny Ålesund, Svalbard (Treffeisen et al. 2007) with an altitude resolution of 200 m and southern England (Rädel and Shine 2007) with an altitude resolution of 250–300 m seem to indicate larger average vertical extensions of ISSRs (in Ny Ålesund ca. 600–1100 m, in southern England 1300–2400 m, depending on the season). The larger average thickness of the latter studies may be partly a result of the coarse altitude resolution of the employed radiosonde data. At least, a more recent analysis of British Isles' radiosonde data partly using the data from the same launch stations as in Rädel and Shine (2007), but with high altitude resolution (about 10 m) shows that ISSRs are shallow on average in England as well (Dickson et al. 2010). The thickness distribution obtained from the latter data set is displayed in Fig. 9.5: More than 80 % of the layers are shallower that 1500 m, and roughly 30 % of the layers are less than 100 m deep; 5 % of the ISSRs are over 3 000 m thick (these are the ones that satellite instruments preferentially detect). The distributions are similar for three considered launch stations outside the UK (Gibraltar, St. Helena and Falkland Islands), and the annual and seasonal variations are weak except for the shallowest layers. All studies agree that shallow ISSRs are more frequent than thick ones, but for good estimates of the average vertical extension one should use data of high altitude resolution.

A theoretical explanation of these results can be derived assuming that ice supersaturation mainly originates from lifting in the atmosphere. A humid air parcel with initial humidity below saturation that gets lifted adiabatically first reaches ice saturation at a certain altitude and then reaches liquid saturation somewhat higher up. The vertical depth of the altitude range between ice saturation and liquid saturation typically amounts to 200–500 m. This ISSR depth decreases mainly with temperature, see Fig. 9.6. This explains thermodynamically why ISSRs are often only a few hundred meters thick. This may also explain why observed ISSRs are thicker at higher latitudes, namely because of lower ambient temperature. It further

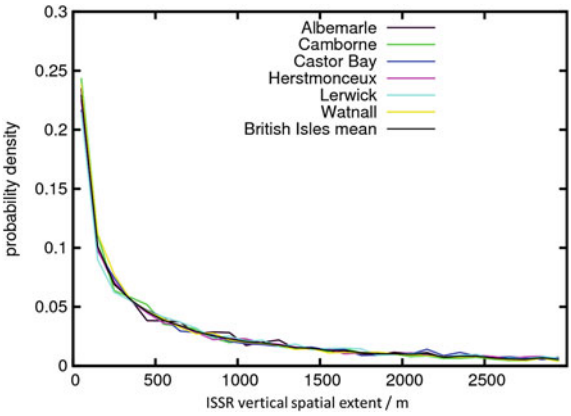

Fig. 9.5 Probability distribution function of vertical layer depths of ice supersaturated layers detected in radiosonde humidity profiles at six launch stations in the UK. Adapted from Dickson et al. (2010)

shows that very fine vertical grid resolution is required to resolve ice supersaturated layers in numerical weather prediction models. Of course air masses usually get lifted not directly in a vertical direction but follow inclined trajectories; they start from different initial conditions, and also experience nonadiabatic processes. Therefore, the adiabatic estimate gives the order of magnitude only and does not exclude thicker and thinner ice supersaturated layers.

The horizontal extension of ISSRs is difficult to measure. Up to now, the only data sources are MOZAIC flights and further research flight data. However, flight data are one-dimensional and what is measured is essentially the length of the flight path through an ISSR that incidentally covers the aircraft route. It is principally impossible to derive the true horizontal size (two dimensional) of an ISSR from one-dimensional data. Gierens and Spichtinger (2000) determined an average path length of 150 ± 250 km (the large standard variation indicates the very strong skewness of the underlying distribution) using data of 15 km horizontal resolution. Higher resolution will reveal smaller structures but they cannot necessarily be interpreted as individual ISSRs. A good way to visualize ISSRs as synoptic systems and to judge their horizontal extension is to use data from weather forecast models that represent ice supersaturation. An example is shown in Fig. 9.7 where the 6 h forecast for 26 August 2011 from the ECMWF model is displayed. But also on smaller scales which are not resolved in the figure there is structure in ISSRs, that is, humidity and temperature vary locally and temporally. In particular atmospheric waves can lead to small scale structures with small ISSRs of relatively short duration.

Very large ISSRs can occur connected to *fronts*. One of the largest specimens (path length >3 000 km) was detected on a MOZAIC flight from Frankfurt, Germany, to Rio de Janeiro, Brazil, when the route was nearly parallel to an approaching warm front. The air in the cruise altitude was already lifted and adiabatically cooled so that the relative humidity could increase beyond 100 %. As it was not yet humid enough for satellite-detectable cirrus formation this was a classic example of an ice-free ISSR.

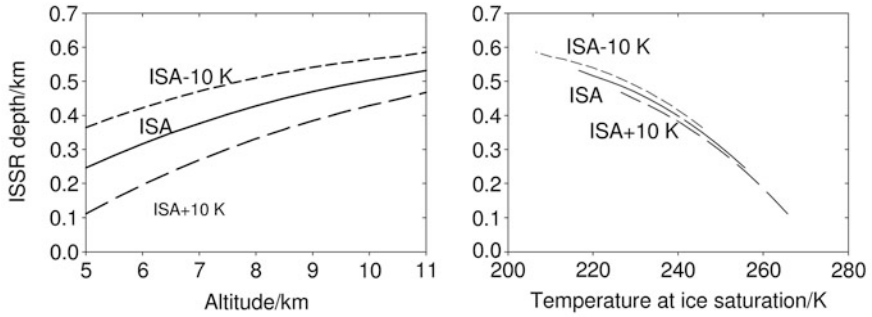

Fig. 9.6 (Left) Depth of the ice supersaturated region versus altitude for the international standard atmosphere (ISA) and 10 K warmer or colder atmospheres. (Right) The same as a function of temperature

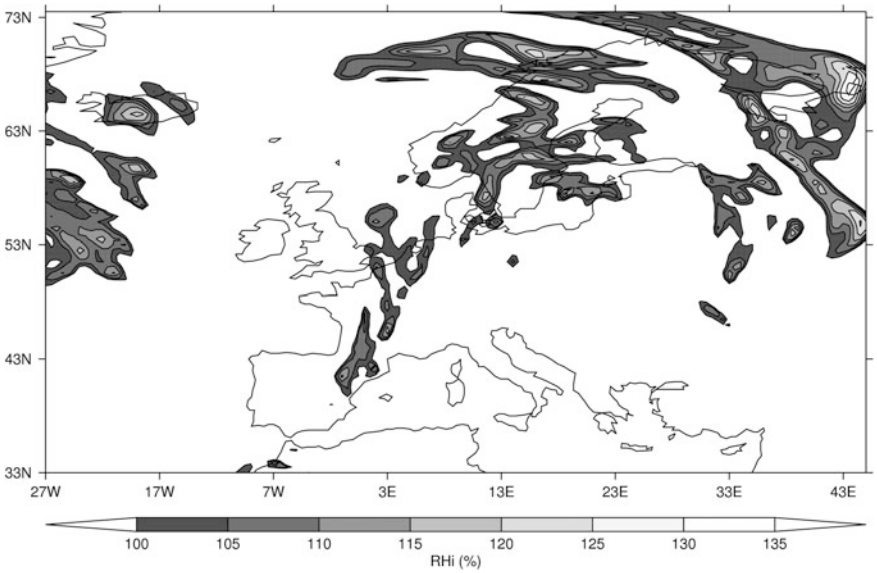

Fig. 9.7 Ice supersaturated regions on the 250 hPa pressure level over Europe as forecasted by the ECMWF model for 06 UTC 26 August 2011. The spatial resolution is 0.5° × 0.5°

Closely related to the spatial extension of ISSRs is their lifetime. Unfortunately, "*lifetime*" is an ill-defined notion for open dynamic systems that incessantly change their shape and size and where air is flowing in and out. There is probably no lower limit to the lifetime; a wave in the atmosphere where the humidity just exceeds saturation in the crest could cause a very small ISSR for a short period of time. Spichtinger et al. (2005a) describe a case where waves superposed to make a small ISSR of perhaps 6 h duration, but shorter periods are certainly possible.

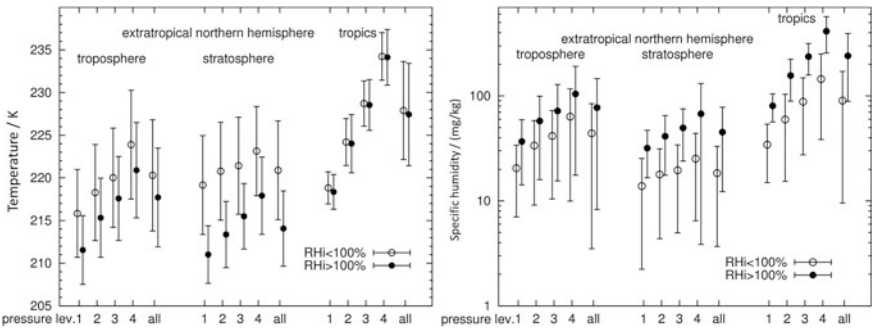

Fig. 9.8 Differences of temperature and specific humidity between ISSRs (filled dots) and subsaturated air masses (open circles) at the same pressure levels in various geographical regions. Average values and standard deviations (bars) are derived from MOZAIC data. The pressure levels are numbered as follows: 190-209 hPa (level 1), 210-230 hPa (2), 231-245 hPa (3), 246-270 hPa (4). The "all" class comprises all pressures from 175 to 275 hPa

On the other hand, true synoptic scale ISSRs can have lifetimes of a day or more. Spichtinger et al. (2005b) describe such a specimen with a lifetime of at least 24 h. "Lifetime" here means essentially how long an air parcel whose trajectory was computed had a relative humidity in excess of ice saturation. However, this can be misleading. Consider a stationary wave pattern with a small ISSR in the crest. An air parcel going through the ISSR may experience supersaturation only briefly, while the ISSR exists as long as the wave pattern exists. These two durations are essentially independent of each other.

In principle, ice supersaturation can result from cooling (decrease of saturation vapor concentration) and convergence of moist air (increase of vapor concentration). Both MOZAIC and MLS data show that ISSRs are on average both colder and moister (higher concentration of vapor) than their subsaturated surroundings (Fig. 9.8). The differences vary with geographical location (latitude and altitude) but are statistically significant. The difference in specific humidity can hardly be a result of moisture convergence alone. Rather, both features can by explained together by vertical transport. Air ascending from the middle or even lower troposphere into the upper troposphere undergoes adiabatic cooling and carries the high vapor concentration of its original altitude unless cloud formation and subsequent precipitation removes the water mass from the ascending air.

9.6 How Important is Ice Supersaturation for the Atmosphere, for the Weather and for Climate?

In order to see the importance of a phenomenon it is a good idea to ask what would happen if it would not exist. So what would happen if ice supersaturation would be, for instance, forbidden by physics and if the maximum relative humidity with respect to ice at subzero temperatures would be 100 %? First, it would be difficult or

impossible to form ice at all, since ice formation, even on good ice nuclei, generally needs at least a small amount of supersaturation. The formation of *supercooled water* would be impossible, because water saturation is a higher concentration of water molecules than ice saturation. If, for any reason, there was ice in a region exactly at ice saturation, the ice crystals could not grow; every small humidity fluctuation would necessarily lead to loss of a part of their mass up, to complete sublimation. Thus ice crystals could not survive a long time in the atmosphere and all cold clouds would disappear from the sky. But if ice supersaturation would be forbidden, water supersaturation would certainly be forbidden for the same reasons, so warm clouds (temperature above 0 °C) would also be impossible. Supersaturation is obviously essential for cloud formation, for crystal and droplet growth and *sedimentation*, and for formation of precipitation via the Wegener-Bergeron-Findeisen process. ISSRs are those regions where condensation trails can grow hours old; outside of ISSRs *contrails* vanish within seconds to minutes, depending on temperature and humidity, after their formation. So, long-lasting contrails in the sky can be used as an indicator of ISSRs. ISSRs have a small effect on the *radiation* flow through the atmosphere as well, but Fusina et al. (2007) calculated that as soon as a thin cirrus cloud forms in the ISSR the radiation effect grows by almost two orders of magnitude (a human observer does not see an ISSR, but as soon as a cirrus appears it can be seen). It is this essential role in cloud formation and evolution that makes supersaturation important for weather and climate.

9.7 Which are the Highest Values of Ice Supersaturation in the Atmosphere?

There is water saturation in supercooled water clouds. The highest value of supersaturation can be found when the supercooled cloud reaches the supercooling limit of water, that is, about −40 °C. At that temperature water saturation is equivalent to about 45 % ice supersaturation. At still lower temperatures, ice formation usually proceeds below water saturation, but at ice supersaturation exceeding 45 %. Supersaturation close to water saturation can be approached during strong uplifts while freezing of aqueous solution droplets proceeds. Very high supersaturation exceeding 100 % can be obtained when freezing is suppressed by certain chemical components or if cooling is very fast. Examples have been found at the tropical tropopause in very cold air (−86 °C, Jensen et al. (2005)), and in Antarctic winter (Spichtinger et al. 2002). The highest values of ice supersaturation exceeding 1 000 % by far are related to the formation of *noctilucent clouds*, which are ice clouds at altitudes of 80–85 km at temperatures around −130 °C (Lübken et al. 2009). Extremely fast cooling of 20 K and more within a millisecond occurs over the wings of aircraft in cruise level altitudes; this can cause supersaturation exceeding 1 000 % as well, albeit for only a very short period of time. *Aerodynamic contrails* (Gierens et al. 2009; Kärcher et al. 2009) can form and become visible under these circumstances, and they remain visible if this should happen within an ISSR.

References

Dickson, N. C., Gierens, K. M., Rogers, H. L., Jones, R. L.: Vertical spatial scales of ice supersaturation and probability of ice supersaturated layers in low resolution profiles of relative humidity. In: Proceedimgs of 2nd International Conference on Transport, Atmosphere and Climate (TAC-2). 22–25 June 2009, DLR-FB 2010-10, Köln-Porz, Germany, ISSN 1434-8454, Aachen, 239–243 (2010)

Fusina, F., Spichtinger, P. Lohmann, U.: Impact of ice supersaturated regions and thin cirrus on radiation in the midlatitudes. J. Geophys. Res. **112**, D24S14, (2007) doi:10.1029/2007JD008449

Gettelman, A., Fetzer, E.J., Elderling, A., Irion, F.W.: The global distribution of supersaturation in the upper troposphere from the Atmospheric Infrared Sounder. J. Clim. **19**, 6089–6103 (2006). doi:10.1175/JCLI3955.1

Gierens, K., Schumann, U., Helten, M., Smit, H., Marenco, A.: A distribution law for relative humidity in the upper troposphere and lower stratosphere derived from three years of MOZAIC measurements. Ann. Geophysicae **17**, 1218–1226 (1999). doi:10.1007/s00585-999-1218-7

Gierens, K., Schumann, U., Helten, M., Smit, H., Wang, P.H.: Ice-supersaturated regions and sub visible cirrus in the northern midlatitude upper troposphere. J. Geophys. Res. **105**, 22743–22753 (2000). doi:10.1029/2000JD900341

Gierens, K., Spichtinger, P.: On the size distribution of ice-supersaturated regions in the upper troposphere and lowermost stratosphere. Ann. Geophysicae **18**, 499–504 (2000). doi:10.1007/s00585-000-0499-7

Gierens, K., Kohlhepp, R., Spichtinger, P., Schroedter-Homscheidt, M.: Ice supersaturation as seen from TOVS. Atmos. Chem. Phys. **4**, 539–547 (2004). doi:10.5194/acp-4-539-2004

Gierens, K., Kärcher, B., Mannstein, H., Mayer, B.: Aerodynamic contrails: Phenomenology and flow physics. J. Atmos. Sci. **66**, 217–226 (2009). doi:10.1175/2008JAS2767.1

Jensen, E.J., Smith, J.B., Pfister, L., Pittman, J.V., Weinstock, E.M., Sayres, D.S., Herman, R.L., Troy, R.F., Rosenlof, K., Thompson, T.L., et al.: Ice supersaturations exceeding 100 % at the cold tropical tropopause: implications for cirrus formation and dehydration. Atmos. Chem. Phys. **5**, 851–862 (2005). doi:10.5194/acp-5-851-2005

Kärcher, B., Mayer, B., Gierens, K., Burkhardt, U., Mannstein, H., Chatterjee, R.: Aerodynamic contrails: microphysics and optical properties. J. Atmos. Sci. **66**, 227–243 (2009). doi:10.1175/2008JAS2768.1

Lamquin, N., Stubenrauch, C.J., Gierens, K., Burkhardt, U., Smit, H.: A global climatology of upper-tropospheric ice supersaturation occurrence inferred from the atmospheric infrared sounder calibrated by MOZAIC. Atmos. Chem. Phys. **12**, 381–405 (2012). doi:10.5194/acp-12-381-2012

Lübken, F.-J., Lautenbach, J., Höffner, J., Rapp, M., Zecha, M.: First continuous temperature measurements within polar mesosphere summer echoes. J. Atmos. Solar Terr. Phys. **71**, 453–463 (2009). doi:10.1016/j.jastp.2008.06.001

Rädel, G., Shine, K.P.: Evaluation of the use of radiosonde humidity data to predict the occurrence of persistent contrails. Q. J. R. Meteorol. Soc. **133**, 1413–1423 (2007). doi:10.1002/qj.128

Spichtinger, P., Gierens, K., Read, W.: The statistical distribution law of relative humidity in the global tropopause region. Meteorol. Z. **11**, 83–88 (2002). doi:10.1127/0941-2948/2002/0011-0083

Spichtinger, P., Gierens, K., Leiterer, U., Dier, H.: Ice supersaturation in the tropopause region over Lindenberg, Germany. Meteorol. Z. **12**, 143–156 (2003a). doi:10.1127/0941-2948/2003/0012-0143

Spichtinger, P., Gierens, K., Read, W.: The global distribution of ice-supersaturated regions as seen by the microwave limb sounder. Q. J. R. Meteorol. Soc. **129**, 3391–3410 (2003b). doi:10.1256/qj.02.141

Spichtinger, P., Gierens, K., Dörnbrack, A.: Formation of ice supersaturation by mesoscale gravity waves. Atmos. Chem. Phys. **5**, 1243–1255 (2005a). doi:10.5194/acp-5-1243-2005

Spichtinger, P., Gierens, K., Wernli, H.: A case study of the formation and evolution of ice supersaturation in the vicinity of a warm conveyor belt's outflow region. Atmos. Chem. Phys. **5**, 973–987 (2005b). doi:10.5194/acp-5-973-2005

Tompkins, A., Gierens, K., Rädel, G.: Ice supersaturation in the ECMWF Integrated Forecast System. Q. J. R. Meteorol. Soc. **133**, 53–63 (2007). doi:10.1002/qj.14

Treffeisen, R., Krejci, R., Ström, J., Engvall, A.C., Herber, A., Thomason, L.: Humidity observations in the Arctic troposphere over Ny-Ålesund, Svalbard, based on 15 years of radiosonde data. Atmos. Chem. Phys. **7**, 2721–2732 (2007). doi:10.5194/acp-7-2721-2007

Wegener, A.: Meteorologische Terminbeobachtungen am Danmarks-Havn. Meddelelser om Grønland, Vol. XLII: Danmark-Ekspeditionen til Grønlands Nordøstkyst 1906-1908 **II**, 125–356 (1914)

Atmospheric Ice Formation Processes

10

Bernd Kärcher

Abstract

This contribution provides a basic introduction to the formation of ice in clouds. Various pathways to ice nucleation and factors controlling ice formation processes are addressed. Challenges in characterizing the atmospheric ice phase and novel approaches to better understanding the fundamental mechanisms involved in ice nucleation are outlined.

10.1 Introduction

Water is the only material that occurs in all three phases in the Earth system. Water vapor is the most effective *greenhouse gas* and has a considerable effect on the *radiation budget*. A fundamental constraint on the abundance of water vapor is the *Clausius-Clapeyron* relationship (Becker 1978) determining the maximum amount of water vapor a given volume of air can hold in *thermodynamic equilibrium*. Saturation denotes the condition in which the vapor pressure of a substance in a condensed phase is equal to its equilibrium *vapor pressure* (see Fig. 10.1 for water). Water vapor in excess of the saturation values is termed *supersaturated* and the excess vapor turns into liquid water or ice. Liquid water is the main component in cloud droplets and most *aerosol* particles, allowing trace species to dissolve, forming aqueous *solution droplets* and initiating chemical reactions. Solid water in the form of *ice crystals* often initiates the formation of *precipitation*

B. Kärcher (✉)
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: bernd.kaercher@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_10,

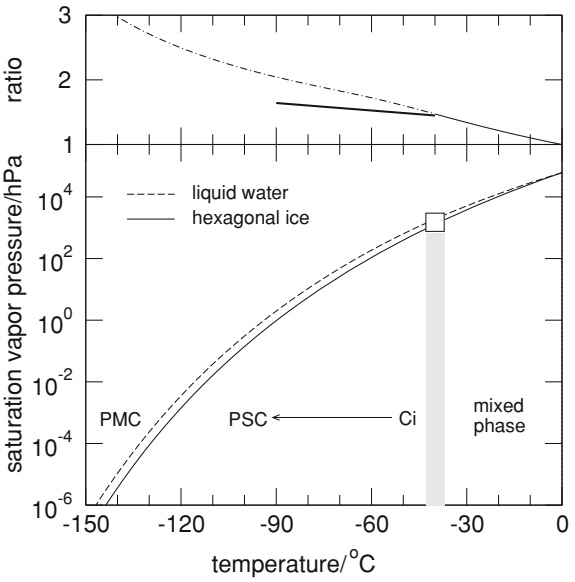

Fig. 10.1 Partial pressures of water molecules in thermodynamic equilibrium with liquid water and hexagonal ice, and their ratio (*top panel*), versus temperature (Murphy and Koop 2005). The spontaneous freezing temperature of supercooled pure water is indicated (*square*), along with the temperature ranges in which the various ice-containing clouds typically form. The water-to-ice pressure ratio, R, is indicated by a solid curve at temperatures where pure liquid water exists. Ice supersaturation (R–1), in mixed-phase clouds reaches a peak value of $\sim$0.47 at –40 °C. The thick solid curve marks the ice supersaturation where supercooled aqueous aerosol droplets freeze homogeneously in cirrus and PSCs (Koop 2004)

in tropospheric clouds and is involved in cloud *electrification*. High altitude clouds composed entirely of ice crystals warm the *climate* by absorbing Earth's *infrared radiation* (*cirrus* and *contrail cirrus* clouds, altitudes $\sim$10 km), are involved in polar *ozone depletion* (*polar stratospheric clouds*, PSCs, $\sim$20 km), and act as sensitive indicators of atmospheric changes (*polar mesospheric clouds*, PMCs, $\sim$85 km). *Ice* is the thermodynamically most stable water phase at subfreezing temperatures.

Ice nucleation plays an important role in diverse areas such as cryobiology and atmospheric cloud formation. In general, nucleation summarizes the processes by which material is transferred from one phase to another. The formation of a new phase, initiated by an external forcing, is driven by a supersaturation of the parent phase (vapor or liquid)—making it more abundant than dictated by thermodynamic equilibrium (i.e., in saturated conditions). The forcing, usually a cooling, is necessary to overcome the energy barrier to nucleation. The parent phase is unstable, exhibiting a tendency to establish saturated (equilibrium) conditions. The new phase has lower bulk energy and is therefore more stable. Once molecular-sized, stable patches of the new phase ('critical clusters') have been created due to

the energy supplied by thermal molecular motion, they grow further, depleting the parent supersaturation. In the case of liquid water/ice particles, this latter growth process is termed condensation/deposition; the reverse processes are *evaporation/sublimation*.

A water phase is *ice-supersaturated* at a certain temperature if the concentration of water molecules exceeds the ice-saturated value. In Earth's *atmosphere*, ice forms in ice-supersaturated conditions mainly within liquid water or aqueous aerosol droplets, or on the solid surfaces of a subset of aerosol particles known as heterogeneous *ice nuclei* (IN). Ice crystals formed this way acquire sizes ranging between a millionth of a centimeter (0.01 μm) and several millimeters (1 000 μm) and can develop into precipitation particles (*snow*, *graupel*, *hail*) in low-altitude clouds. It has long been known that the geometrical forms of snowflakes bear evidence of the hexagonal crystal lattice of water molecules that are tied together by a *hydrogen-bonding* network.[1] Three geometrical arrangements of ice exist in the atmosphere: *hexagonal*, *cubic*, and *amorphous* (noncrystalline) (Hobbs 1974). The two latter form at low temperatures and transform into the more stable, hexagonal, arrangement upon warming. Water ice has also been discovered on other planets in our *solar system*, on some of their moons, in the rings of Saturn, and on asteroidal surfaces.

Hexagonal or cubic ice does not nucleate from the vapor phase alone in the atmosphere. Therefore, aerosol particles must serve as the precursors of ice crystals. This chapter provides a brief overview of aerosol-mediated pathways to ice formation in clouds (nucleation modes) and addresses dynamical controls of cloud ice formation with special emphasis on cirrus.

10.2 Aerosol-Mediated Modes of Ice Nucleation

Cloud particles form by nucleation (i.e., nonequilibrium) processes (Houze Jr. 1993). At *temperatures* below the equilibrium melting point (0 °C or 273.15 K) cloud particles may be composed of either *supercooled water* or ice. Ice crystals may form either homogeneously in cloud water droplets, in water condensed on aerosol particles, or heterogeneously involving water with solid impurities. The exact temperature at which ice forms depends on the nucleation process. The presence of IN in supercooled liquid water droplets has been shown to be necessary for the formation of ice-containing clouds at temperatures above about –30 °C. Incomplete understanding of ice initiation—the first appearance of the ice phase—is a large obstacle for the detailed numerical modeling of ice processes in clouds.

[1] Water is a highly ordered liquid with water molecules conjoined by hydrogen atoms. The exact number of hydrogen bonds formed by a molecule of liquid water (up to four because the oxygen atom of one water molecule has two lone pairs of electrons) depends on the temperature. In crystalline ice, hydrogen bonding creates a lattice structure exhibiting long range order. Amorphous water lacks this long range order yet it is still more 'solid-like' than liquid water.

10.2.1 Homogeneous Ice Nucleation

Small samples of liquid water can be supercooled to a temperature of about –40 °C (≈235 K), the spontaneous (or homogeneous) freezing temperature. *Homogeneous freezing* occurs at a range of temperatures depending on the sample volume (droplet size), the rate of cooling, the degree of purity of the water substance, and on other factors (Pruppacher and Klett 1978). The temperature regime in which supercooled liquid water droplets and ice crystals can coexist in the atmosphere defines the existence region of mixed-phase clouds (Fig. 10.1). Below the spontaneous freezing temperature, ice nucleates in aqueous droplets in a very short time (relative to meteorological time scales), quickly solidifying the entire droplet volume. In such cold conditions, water droplets can no longer exist. Cirrus and contrail cirrus are pure ice clouds occurring across a wide range of temperatures.

Aqueous aerosol particles containing soluble matter (*solutes*) at high humidity are sometimes called '*haze*' droplets. A common solute in atmospheric (particularly stratospheric) aerosol particles is *sulfuric acid*, but organic species are taken up by them as well. *Hygroscopic* material dissolved in aerosol particles attracts water molecules and thereby decreases the freezing temperature of pure water. In this way, ice nucleation can be suppressed to very low temperatures. Enhanced uptake of *nitric acid* in aqueous sulfuric acid particles causes the formation of liquid PSC particles in the winter polar lower stratosphere. These particles freeze homogeneously around –90 °C (≈185 K), forming ice-phase clouds.[2] PSCs are largely responsible for the formation of the '*ozone hole*' (Solomon 1999).

Solutes may precipitate in aqueous solution particles when the solute concentration becomes too high, forming crystalline inclusions (e.g., in *sea salt* particles containing *sodium chloride* as solute). With increasing humidity hygroscopic solution particles absorb more water molecules and those precipitates can melt away. This solid-to-liquid phase transition is termed deliquescence. Often *deliquescence* occurs before freezing. Homogeneous freezing temperatures of fully liquid, aqueous solution droplets have been measured below −40 °C for a wide array of solutes and solute concentrations (Koop 2004). It has been found that homogeneous ice nucleation in aerosols commences very suddenly within a narrow range of high ice supersaturations and is determined by the temperature and *water activity*—a thermodynamic property of the solution droplets equal to the ratio between the water vapor pressures of the solution and of pure water under the same conditions—and not by the chemical nature of the aerosol particles. The rate of freezing is proportional to the droplets' volume owing to the stochastic nature of the homogeneous nucleation process, and also depends on the rate of cooling.

[2] These are known as 'type-2' PSCs, in contrast to type-1 PSCs that form at slightly warmer temperatures and contain nitric acid, either dissolved in the liquid phase or present as solid nitric acid trihydrate (NAT) crystals.

It has been established observationally that soluble *organics* hinder homogeneous freezing (Cziczo et al. 2004) and can reduce the uptake of water molecules, causing organic-rich liquid particles to preferentially remain unfrozen (Kärcher and Koop 2005). A remarkable feature of some *supercooled*, organic-rich aqueous solutions is that they form glasses—highly viscous, amorphous substances behaving like solids—at low temperatures (Murray 2008). The extremely high viscosity of the liquid vitrifying at a transition temperature characteristic to the solution droplets (Koop et al. 2011) impedes molecular motion and therefore homogeneous ice nucleation. *Glassy* aerosol particles may explain the high supersaturations observed at the tropical *tropopause*, although there is an ongoing debate as to the accuracy of humidity measurements at low temperatures and pressures. Below –70 °C cubic ice nucleates in preference to hexagonal ice, thereby enhancing the *dehydration* (*'freeze-drying'*) caused by cold cirrus clouds and PSCs (Murphy 2003). Cubic ice transforms into hexagonal ice within minutes to days, depending on temperature. At warmer temperatures, cubic ice transforms so rapidly into hexagonal ice that the cubic ice is not observed.

Solution droplets containing substances dissolved in the water phase are ubiquitous and abundant in the *troposphere* and *lower stratosphere*. Therefore, the availability of such particles acting as *precursors* for ice crystal formation is not a limiting factor for cirrus and PSC formation. In the temperature range between 0 °C and about –40 °C cloud ice must form by a heterogeneous process since *homogeneous freezing* does not occur.

10.2.2 Heterogeneous Ice Nucleation

Heterogeneous pathways (modes) to ice formation are catalyzed by aerosol particles providing solid surfaces upon which nucleation can take place. It is thought that favorable tropospheric IN should possess one or more of the following properties (Pruppacher and Klett 1978): they should be highly water-insoluble and larger than a ten-thousandth of a millimeter (0.1 μm), have a lattice structure similar to crystalline ice, and contain active sites—distinct surface features that facilitate the attachment and clustering of water molecules into an ice-like arrangement. Other factors must play a role as well, since particles with the same set of these properties may still exhibit different ice-nucleating behavior. Among those factors is the ability of chemical groups at particle surfaces to accept and donate hydrogen bonds.

Perfect IN would exhibit very small energy barriers to nucleation and therefore form ice crystals slightly above ice saturation (contrary to homogeneous freezing of supercooled liquid water or solution droplets), but such particles are hardly found in the atmospheric aerosol. As opposed to the homogeneous case, heterogeneous ice nucleation behavior may vary strongly between individual particles in a sample and can proceed over a wide range of supersaturations or temperatures. This renders the interpretation of experimental data and modeling of this process difficult. Observed

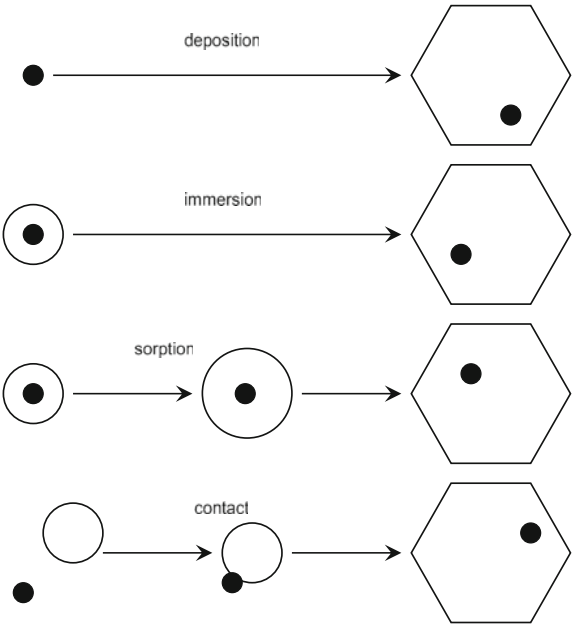

Fig. 10.2 Schematic showing the principal modes of primary heterogeneous ice initiation in clouds. Filled circles indicate IN, open circles are supercooled water droplets or aqueous aerosol particles, and hexagons symbolize ice crystals containing the IN upon which they have formed as inclusions. Terminology follows (Vali 1985). Other modes exist (see text), changing the conditions of the IN and their physical interaction with the droplets

IN constitute only a small subset by number of all tropospheric aerosol particles. Most IN appear to be present or active in the largest aerosol particles and under the most favorable conditions 1 particle out of 1 000 of the total aerosol population can serve as IN (DeMott et al. 2011). While free tropospheric IN concentrations are comparatively low (up to about 10 IN per liter of air), short-term increases in IN concentrations exceeding average values by orders of magnitude ('IN-storms') have been observed. In such cases, IN may control cloud formation and development, particularly mixed-phase clouds (Sassen et al. 2003).

Since supercooled water droplets do not freeze easily at temperatures in the mixed-phase cloud regime, the presence of IN is vital in order to turn liquid water clouds at least partially into ice clouds ('glaciation'). Another mechanism inducing glaciation is the seeding of liquid water-phase clouds by falling ice crystals forming in cirrus clouds at higher altitudes. In the early stages of glaciation, the ice crystals grow quickly by water vapor deposition since the ice supersaturation is large (Fig. 10.1 top, thin solid curve). Thus, glaciation may proceed within tens of minutes in originally purely liquid water-phase clouds, producing a colloidally unstable mixed-phase cloud. The water substance is transferred from the droplets to the ice crystals either by evaporation and diffusion of water vapor through the air or by collection and rapid freezing (riming) of droplets with sufficiently large ice crystals. The glaciation time scale is largely controlled by the temperature and the number and physical properties of the ice crystals. In honor of its discoverers, the cloud glaciation mechanism is termed the Wegener-Bergeron-Findeisen

process. Completely glaciated, purely ice-phase clouds are found at temperatures up to about –10 °C; at higher temperatures cloud particles usually stay liquid.

Figure 10.2 depicts the principal conceptual heterogeneous ice initiation modes. During deposition nucleation, water vapor is transported onto the surface of an IN, taking on a crystalline form and bypassing the formation of liquid water. Immersion nuclei are droplet inclusions initiating the freezing of the particle in which they reside. Sorption nucleation refers to the condensation of water vapor onto a nucleus followed by freezing and is therefore also referred to as condensation freezing. Contact nucleation occurs when IN and droplets come mechanically into contact by collisions. The latter mode is only efficient in the presence of cloud droplets (i.e., in mixed-phase clouds) since they offer a much larger area for collision than the smaller aerosol droplets. Immersion and condensation freezing are not easily distinguishable in measurements. The sets of supersaturations and temperatures at which each of these modes become active are generally different even for the same type of IN.

10.2.3 Other Ice Nucleation Modes

A number of further primary ice nucleation mechanisms are known, often identified in laboratory experiments. Their atmospheric relevance is not well established. Some are related to partially soluble organic and inorganic aerosol phases. Ammonium sulfate crystals may precipitate in aqueous solutions and act as IN before deliquescence occurs. A significant IN activity has been found for oxalic acid when present as a hydrated crystal in aqueous solution droplets. Glassy aerosols impeding homogeneous freezing can actually nucleate ice heterogeneously (Murray et al. 2010). These mechanisms require quite low tropospheric temperatures so that mainly cirrus formation would be affected. Other primary ice formation processes include freezing of evaporating supercooled water droplets and contact freezing ‘inside-out’ (immersed IN collide with the water droplet surface). The presence of electrical charges may also increase the efficiency of nucleation by reducing the nucleation energy barrier and may enhance the contact nucleation process (electro-scavenging).

The ice activity in IN samples usually increases with decreasing temperature and increasing supersaturation. Several mechanisms have been identified modifying this picture. The activity of some IN can be deactivated (damped) by liquid—particularly organic—surface coatings. On the other hand, water droplets covered by certain long-chain alcohols can nucleate ice already at –1 °C. Even without the action of such species, it has been argued that surface crystallization of supercooled water may rival volume-dependent homogeneous freezing. There is a pending discussion on memory effects, meaning that the retention of ice or ice-like liquid layers of water molecules at the surfaces of ice crystal residues—the particles that remain when ice crystals sublimate completely—facilitates ice formation in a subsequent nucleation event (pre-activation).

The temperatures at which the first ice crystals form in liquid water-phase clouds are usually lower than about –10 °C and vary with locale and time. Measurements in mixed-phase clouds reveal that the number of ice crystals sometimes dramatically exceeds the number of available IN. Their ratio can exceed 1 000 around –10 °C and typically decreases with temperature and towards the cloud tops. Such large ratios likely result from secondary ice production processes, such as the fragmentation of large cloud droplets during freezing, the mechanical fracturing of evaporating crystals, and the production of ice splinters during riming. The latter mechanism—referred to as the Hallett-Mossop process—takes place in a narrow temperature window between –3 °C and –8 °C in the presence of cloud droplets larger than a few thousandth of a centimeter (10 μm). It is the only secondary ice formation process active in clouds that has been confirmed experimentally. Ice splinters are collected by cloud droplets that in turn freeze, acting as riming centers to produce new splinters. Hence, this process, also referred to as ice multiplication, is particularly effective as a secondary source of cloud ice.

10.2.4 Ice Nucleation in Clouds

Cloud particle concentrations vary substantially. The formation and development of tropospheric ice clouds are complicated by the many modes of ice nucleation mediated by aerosol particles and supercooled water droplets (Cantrell and Heymsfield 2005), but also depend on the dynamical environment supporting ice nucleation. Only when particles are in close proximity may they physically interact with each other.

The most efficient deposition or immersion nuclei might have already caused ice formation in low-altitude clouds and precipitated before reaching cirrus altitudes (Wiaczek et al. 2010). It is therefore conceivable that IN play only a minor role in cirrus development, except perhaps when they are directly emitted at high altitude (e.g., by aircraft), are rapidly lofted from their source regions in the boundary layer (e.g., by convection), or when cirrus form in slow updrafts. In cold and strongly ice-supersaturated conditions, nitric acid may be taken up in liquid aerosol particles, enhancing the freezing droplet volume and therefore promoting the formation of haze droplets and thin (often invisible) cirrus clouds (Kärcher and Solomon 1999). The water droplets in exhaust contrails, generated by particles emitted from aircraft jet engines, freeze homogeneously within about one wingspan behind airplanes cruising at altitudes where contrail formation is prevalent (Kärcher 1999). It is not necessary to ascribe any particular heterogeneous ice-nucleating property to aircraft-produced particles to explain initial properties of such contrails below the spontaneous freezing temperature.

Very low temperatures are necessary to form PSCs and PMCs owing to the extreme dryness of their atmospheric environment with only few water molecules per million air molecules. These extremely tenuous clouds are susceptible to even small changes in air temperature, in the abundance of water vapor, or in the

radiative characteristics of the atmosphere. The ice clouds in the winter polar stratosphere nucleate homogeneously on supercooled liquid droplets that are a ten-thousandth of a millimeter (0.1 μm) large and constitute the majority of the stratospheric aerosol. Other routes to stratospheric ice cloud formation involving solid hydrates of nitric acid also exist (Peter 1997). The ice clouds in the high-latitude summer mesosphere are thought to form primarily by deposition nucleation on meteoric smoke particles with sizes of only a millionth of a millimeter (0.001 μm). Other candidates for mesospheric ice formation include charged molecular clusters (atmospheric ions) that are abundant at mesospheric altitudes. It has also been proposed that amorphous solid water nucleates homogeneously from the vapor phase if the number of nucleating smoke particles is limited (Murray and Jensen 2010). Observational evidence is insufficient for deciding which nucleation pathway dominates. Water-containing particles in PMCs cause strong radar signals known as polar mesosphere summer echoes. Larger ice crystals at the cloud base give rise to the night-shining (noctilucent) clouds observable during twilight. They are the coldest clouds on Earth, forming at the edge of space below –125 °C ($\approx$150 K) (Rapp and Thomas 2006).

10.2.5 Sources of Ice Nuclei

Heterogeneous ice nuclei originate from natural and man-made sources, predominantly deriving from Earth's surface. *Mineral dust* is perhaps the best documented IN present in the atmosphere. The majority of dust emissions are natural, mostly from the Sahara and Gobi deserts. About one-third of the mineral dust emissions arise from human activities such as land use and desertification. Dust particles are complex mixtures composed of metal oxides or carbonates and come in many chemical variations and sizes, each with different IN behavior. The clay minerals montmorillionite and kaolinite—common weathering products containing sodium, calcium, magnesium, aluminum, and silicon—have been used as surrogates for real mineral dust particles from different source regions in modeling studies examining the alteration of mixed-phase clouds by IN.

A large number of natural and *anthropogenic* incomplete *combustion* processes involving hydrocarbons produce impure *black carbon* (*soot*) particles, a form of amorphous carbon. These particles are associated with various organic and inorganic emissions present as surface *coatings*. Details of the particle surface *morphology*, size, and chemistry may explain the wide variety of ice-nucleating behavior of atmospheric soot samples. The available laboratory and field information regarding the efficiency of these particles for ice formation is not entirely conclusive but suggests that soot particles, while highly variable as a source, are not implicated as primary ice-forming agents in the troposphere. Further anthropogenic IN (e.g., produced by steel and copper smelters) are far less abundant than soot.

High molecular weight organic compounds—such as long-chain alcohols—are produced by open *biomass burning*, i.e., the burning of vegetation over vast areas

including wildfires and deforestation, and can act as IN in such areas. Maritime air masses are thought to be deficient in IN, although marine *biogenic* particles (e.g., phytoplankton) can function as immersion nuclei when present within sea salt droplets. Pollen, leaf litter, fungal spores, strains of bacteria, and even viruses are known to be good ice-nucleating agents and they may affect the precipitation development in clouds near their source regions. *Volcanic ash* particles have rather moderate effects on tropospheric ice formation, although some of them (containing silicate) can become quite ice-active at low temperatures. Silver iodide particles—frequently used in cloud seeding experiments—possess a crystallographic structure similar to ice and are therefore efficient IN, triggering ice formation already at –6 °C. Finally, ablated *meteorite* material represents an extraterrestrial source of deposition nuclei in the upper atmosphere. The experimental characterization of IN types in different locales and seasons is an active area of research.

10.3 Dynamical Controls of Cloud Ice Formation

10.3.1 Generation of Supersaturation

Clouds appear in conjunction with a large number of atmospheric motion systems. The supersaturation required to induce cloud formation is created by dynamical effects, mostly by cooling of pockets of air ('parcels') that are lifted due to vertical air motions. The most common ways to force air parcels to ascend are wind convergence, topographic lifting, and *buoyancy* (recalling Archimedes' principle). The parcels may mix with air from their surroundings by *turbulence* or stay isolated (*adiabatic*). Transport of humid air and subsequent mixing moistens the *air parcels* before or during cooling. The water vapor in the rising parcels is deposited on the nucleated ice crystals in ice-supersaturated conditions, producing cloud ice water. The lifting and associated cooling can be quite vigorous (as in convective, towering clouds with updrafts exceeding several meters per second) or rather gentle (as in stratiform, layered clouds, with updraft speeds of up to tens of centimeters per second). While low- and mid-level clouds usually block the sunlight owing to their large water content and exhibit sharp boundaries because cloud droplets quickly evaporate, high altitude *cirrus* are largely transparent and take the form of fibrous wisps since their water content is low and ice crystals do not readily disappear upon leaving the cloud.

In high-reaching (deep) *cumulus clouds*, ice may initially form in the *mixed-phase* temperature regime and is transported aloft and detrained in the form of *anvils* away from the convective core (convective outflow). In anvil cirrus, further ice nucleation may occur after much of the ice mass has been removed by *sedimentation*. The causes of observed onset of *glaciation* in some low-altitude (shallow) cumulus clouds at temperatures above –10 °C are not well understood. Contrary to most natural clouds, hot and moist *jet aircraft* exhaust plumes cool at constant pressure by mixing with cold environmental air, leading to the formation

of contrails. Similar to anvil cirrus, *contrail cirrus* clouds spread by *wind shear*, increasing their horizontal area and are often assembled in clusters clearly visible in satellite imagery. PSCs and PMCs contain up to 10 000 times less ice water than cirrus. They are visible from the ground only under special conditions. In the *Arctic*, the formation of PSCs preferentially occurs over mountain ridges causing rapid vertical wind perturbations and therefore adiabatic cooling of air. Less is known about the small-scale dynamical environment in which PMCs form.

10.3.2 Cirrus Formation by Homogeneous Freezing

As an illustrative example of how cloud ice particles form, the homogeneous ice formation process within an adiabatic air parcel that is subject to steady cooling and free of supercooled water droplets (containing only liquid aerosol particles) is described below. This simplified scenario replicates common cirrus formation conditions. The *relative humidity* over ice (RHI) is equivalent to the (fractional) ice supersaturation. When cooling starts (e.g., at ice saturation, RHI = 100 %), RHI begins to increase. The amount of supersaturated water vapor condensing onto the haze particles acts to reduce RHI, but this amount is very small. Therefore, in the absence of cloud ice particles, RHI increases with time in proportion to the rate of cooling. Homogeneous ice nucleation in the supercooled aerosol will not commence until high ice supersaturation (RHI $\approx$150 %) is reached (Fig. 10.1 top, thick solid curve). Once ice crystals form, they begin to grow by depositing water vapor, reducing RHI towards ice saturation on a time scale that is inversely proportional to the number concentration of nucleated ice crystals. The nucleated ice crystal number scales approximately with the cooling rate (or the equivalent updraft speed) and is determined by the balance between cooling and depositional growth around the point where RHI reaches its peak value. Peak RHI values are only slightly larger than those characterizing the onset of homogeneous freezing unless the cooling rates become much lower than typically observed. The lower the temperature, the slower the ice growth and water vapor depletion, the longer RHI stays around its peak value, and the more crystals form.

A theoretical analysis of this problem is not straightforward, because nucleation and growth of ice crystals may occur simultaneously. Figure 10.3 shows results of a theory describing this scenario (Kärcher and Lohmann 2002). In many places in the upper troposphere, cloud-scale *cooling rates* lie within 1–10 degrees per hour; nucleated ice crystal *concentrations* should therefore mostly range between 100 and 1 000 particles per liter of air. Atmospheric variability in temperature and other variables as well as further cloud development typically produce a broad spectrum of cirrus ice crystal numbers. These *number concentrations* are often larger than those of IN (DeMott et al. 2003), suggesting that on average IN modify, rather than control, cirrus formation processes. A paucity of IN in the upper troposphere (number concentrations below 10–30 per liter of air) implies a

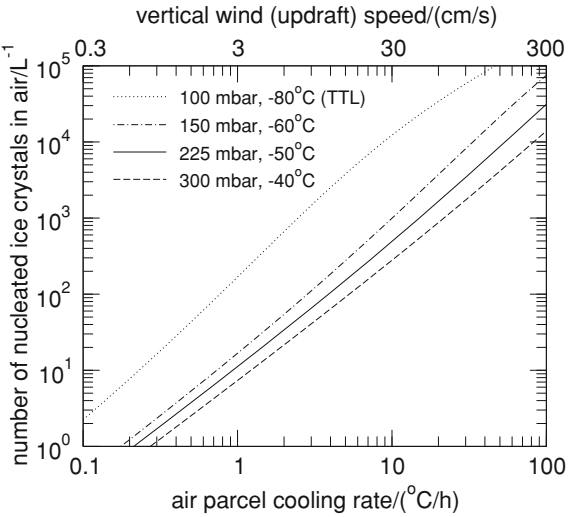

Fig. 10.3 Ice formation from supercooled aerosol particles in an air parcel that is lifted at a constant updraft speed for representative combinations of air pressure and temperature in cirrus conditions. A special case for very cold cirrus clouds forming in the uppermost tropical troposphere (tropical tropopause layer, 'TTL') is included

predominance of homogeneous freezing in cirrus clouds, consistent with the observed high ice supersaturations necessary for homogeneous freezing of haze droplets, at least in extratropical regions (Haag et al. 2003). The theory also predicts that the dependence of nucleated ice crystal number on the size of the freezing aerosol particles is weak. Airborne measurements of ice crystal concentrations can be understood in the way described by this theory, except for the class of cirrus present in the cold upper tropical troposphere, for which significantly smaller ice crystal concentrations have been measured than those shown in Fig. 10.3 at the prevailing cooling rates (Jensen et al. 2010). This may suggest either that ice nucleation processes other than homogeneous freezing are more prevalent in those clouds, or that other mechanisms, such as turbulent mixing with drier air, act towards reducing nucleated ice crystal concentrations.

10.3.3 Role of IN in Cirrus Formation

Having emphasized the importance of homogeneous aerosol freezing for cirrus formation, it is important to note that nucleation on even few IN (1–10 per liter of air) can affect the initial ice crystal number in cirrus clouds and perhaps their further development. The idealized theory has been extended to include effects of IN on the ice formation process. The main effect of efficient IN is to reduce the RHI where ice starts to form. The *deposition* of water vapor on ice crystals nucleating early on the IN slows the subsequent increase in RHI in a cooling event and thereby reduces the number of homogeneously nucleated ice crystals that would have formed in the absence of IN. It may even lead to a complete inhibition of homogeneous freezing. The occurrence and magnitude of this effect depends on the cooling rate as well as on the IN concentration and the temperature and

supersaturation at which the IN nucleate ice. The effect is strong for high (although limited as noted above) IN number concentrations and low cooling rates. On the other hand, sedimentation of ice crystals forming on IN can lead to a removal of those crystals from the nucleation region and thus make homogeneous freezing more likely. The sedimentation effect becomes important for slow cooling causing long nucleation time scales. Although direct ice nucleation from atmospheric aerosols is not the primary determinant for the ice content in anvil and contrail cirrus, ice nucleation may contribute to the ice budgets at some point during their life cycle.

10.3.4 Challenges to Representing Ice Formation in Large-Scale Models

This idealized model framework exemplifies the use of cloud physics theory to describe the interaction of processes in nonconvective lift that eventually lead to the formation of ice in cirrus clouds. The underlying theory has been employed to develop a *parameterization* framework potentially suitable for large-scale models that are not capable of resolving the small spatial and temporal scales pertinent to the atmospheric ice nucleation process in cirrus (Kärcher and Burkhardt 2008). When applying such a parameterization, the challenge is to properly represent the dynamical forcing (cooling rates), the number (ice-active fractions) and nucleation behavior of IN (Hendricks et al. 2011), and the interaction between cirrus and mixed-phase clouds. The forcings driving ice nucleation are not resolved by large-scale models, and the lack of a fundamental understanding of heterogeneous ice nucleation does not allow the prediction of the ice-active fraction of atmospheric IN populations from first principles. Hence, besides the nucleation process, dynamical forcing and ice-active fractions must be parameterized as well (DeMott et al. 2011).

Gravity waves inducing a wide range of cooling rates constitute an important element for the formation of cirrus (Hoyle et al. 2005), Arctic PSCs (Carslaw et al. 1998), and PMCs (Jensen and Thomas 1994). In view of the marked nonlinearity of ice nucleation processes, wave-driven forcing may modify the general picture derived from simple conceptual models of cloud ice formation that are mostly based on the assumption that cooling rates do not vary within the time ice nucleation occurs. Characterizing the cooling rate and resulting supersaturation variability at the spatial scale of clouds probably opens up a novel field of research (Kärcher 2012). Vertical air motion variability that is to a significant degree driven by the release of latent heat during cloud formation is also crucial for the formation of ice in convective clouds. Besides being able to exert a significant effect on the dynamics of mixed-phase clouds, ice particles can also act as seeds for frozen precipitation particles. Therefore, application of cloud physics theory to describe ice formation in low-altitude clouds is arguably more difficult than in the case of cirrus.

Table 10.1 Summary of ice-containing cloud types and selected characteristics related to the ice formation process as discussed in this chapter

Ice cloud type	Altitude (km)	Principal ice-forming agents	Ice supersaturation[a]
Mixed phase	<7	IN, supercooled water droplets	≤0.5[b]
Cirrus[c]	>7	Supercooled liquid aerosol droplets	0.5–0.7[d]
PSC[e]	~20	Ternary solution droplets	0.5–0.7
PMC[f]	~85	Meteoric smoke particles, ions	?

[a] Approximate values at which ice nucleation commences, see also Fig. 10.1
[b] Ice forms at or below liquid water saturation above –40 °C, depending on the ice nucleation mode (Fig. 10.2)
[c] Pure ice clouds may also exist at lower altitudes
[d] Exact value depends on the temperature. Ice nucleation mechanisms in the tropical tropopause region are not well understood
[e] Type-2 polar stratospheric clouds. Supercooled solution particles from the background stratospheric aerosol layer are mainly composed of water and sulfuric acid and take up nitric acid at low temperatures prior to PSC formation. The role of meteoric smoke contributions in ice formation in PSCs is not clear
[f] Sources and ice nucleation behavior of PMC particles, and therefore the supersaturation at which they form, are highly uncertain

10.4 Concluding Remarks

This contribution has been written for a wide readership and therefore provides only a basic introduction to atmospheric ice formation processes. Readers interested in specific details or in-depth discussions might wish to consult the references provided here. Table 10.1 summarizes key features of ice-containing clouds in the atmosphere.

We conclude with a few thoughts on general aspects and future developments. The temperature regime above –20 °C is key to the generation of precipitation often involving the ice phase. Because the corresponding altitudes (a few kilometers) are quite easily accessible, much emphasis was historically put on measuring ice formation processes in situ in this regime. Such measurements have been conducted more frequently at the colder temperatures of the upper troposphere and lowermost stratosphere with the advent of high altitude *research aircraft*, while the mesosphere is studied using remote sensing methods or via rocket-borne sampling. No existing measuring device could probe all the cloud ice formation processes discussed here. Airborne sampling and analysis alter the ice-nucleating properties of the aerosol (e.g., by particle heating) and the mechanical impact of ice at the inlets of particle probes can affect quantitative estimates of crystal number and sizes (e.g., by ice crystal shattering). Furthermore, the fact that small-scale meteorological processes are capable of masking aerosol influences on clouds (Kärcher and Ström 2003) complicates measurements and

their interpretation, the modeling of aerosol-cloud interactions, and the understanding of the behavior of clouds in a changing climate, all of which are topics of present research (Heintzenberg and Charlson 2009).

A theory describing molecular-level nucleation processes in terms of macroscopic quantities developed almost a century ago (Becker and Döring 1935) is often applied to explore general features of the ice nucleation process. Since many of its underlying assumptions cannot be rigorously justified, we should have little confidence in the quantitative predictions of this so-called classical nucleation theory (Hegg and Baker 2009). Nevertheless, it is favorable to have this theory as a framework for first physically-based exploration and discussion, the more so as it seems very unlikely that—in view of the complexity of the problem at hand—a microscopic theory covering the multitude of factors controlling ice nucleation could ever be devised.

The research area of atmospheric ice formation processes has been, and remains, important for the DLR Institute of Atmospheric Physics since a thorough understanding of natural cloud formation is a necessary prerequisite for assessments of atmospheric change caused by human activity. A prime example of such work recently conducted in this institute is the formation of contrail cirrus and the associated alteration of natural cirrus cloudiness (Burkhardt and Kärcher 2011), which has been identified as an important step in better quantifying the overall impact of *aviation* on climate and in developing strategies to mitigate this impact (Boucher 2011). As a technological application, the sudden freezing of supercooled water droplets coming into contact with aircraft surfaces (*aircraft icing*) is relevant to many DLR activities. Adverse effects of aircraft icing in flight—including airfoil and engine stalling—can threaten passenger safety and reduce airplane performance. Measures must be taken to alleviate severe aircraft icing problems by making use of findings regarding atmospheric ice formation, such as estimations of the degree of glaciation in mixed-phase clouds.

Much research remains to better understand the dependencies of ice formation processes on atmospheric conditions. It is of vital importance to untangle the interactions between ice-nucleating aerosols, clouds, and precipitation in the appropriate dynamical environment to provide more reliable simulations of weather and climate. Ice nucleation is only one of many components contributing to atmospheric cloud formation. Nonetheless, a fundamental understanding of molecular processes—in particular those occurring at particle surfaces—would lead to further progress in representing ice formation processes in atmospheric models.

Cloud ice formation processes are closely related to *physical chemistry* and molecular physics as well as *thermodynamics* and statistical physics, reinforcing the notion that *atmospheric physics* is a central subject of the physical sciences. In this way, future research into ice formation will benefit from experimental and theoretical methods established in other areas of physics. Recent examples include the *laser* microscopy of ice crystal surfaces (Sazaki et al. 2010) and first principles calculations of the heterogeneous ice nucleation process (Hu and Michaelides 2007). In conjunction with further advances in the experimental characterization of

the physical and chemical properties as well as the atmospheric variability and abundance of IN, these examples offer promising possibilities to explore open issues regarding the atmospheric ice phase.

References

Becker, R., Döring, W.: Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. Ann. Phys. **24**, 719–752 (1935)

Becker, R.: Theorie der Wärme. Springer-Verlag, Heidelberger Taschenbücher (1978)

Boucher, O.: Seeing through contrails. Nature. Clim. Change **1**, 24–25 (2011)

Burkhardt, U., Kärcher, B.: Global radiative forcing from contrail cirrus. Nature Clim. Change **1**, 54–58 (2011)

Cantrell, W., Heymsfield, A.: Production of ice in tropospheric clouds—A review. Bull. Am. Meteorol. Soc. **86**, 795–807 (2005)

Carslaw, K.S., Wirth, M., Tsias, A., Luo, B.P., Dörnbrack, A., Leutbecher, M., Volkert, H., Renger, W., Bacmeister, J.T., Reimer, E., et al.: Increased stratospheric ozone depletion due to mountain-induced atmospheric waves. Nature **391**, 675–678 (1998)

Cziczo, D.J., DeMott, P.J., Brooks, S.D., Prenni, A.J., Thomson, D.S., Baumgardner, D., Wilson, J.C., Kreidenweis, S.M., Murphy, D.M.: Observations of organic species and atmospheric ice formation. Geophys. Res. Lett. **31**, L12116 (2004). doi:10.1029/2004GL019822

DeMott, P.J., Cziczo, D.J., Prenni, A.J., Murphy, D.M., Kreidenweis, S.M., Thomson, D.S., Borys, R., Rogers, D.C.: Measurements of the concentration and composition of nuclei for cirrus formation. Proc. Nat. Acad. Sci. U.S.A. **100**, 14655–14660 (2003)

DeMott, P.J., Möhler, O., Stetzer, O., Vali, G., Levin, Z., Petters, M.D., Murakami, M., Leisner, T., Bundke, U., Klein, H., et al.: Resurgence in ice nuclei measurement research. Bull. Am. Meteorol. Soc. **92**, 1623–1635 (2011)

Haag, W., Kärcher, B., Ström, J., Minikin, A., Lohmann, U., Ovarlez, J., Stohl, A.: Freezing thresholds and cirrus cloud formation mechanisms inferred from in situ measurements of relative humidity. Atmos. Chem. Phys. **3**, 1791–1806 (2003)

Hegg, D.A., Baker, M.B.: Nucleation in the atmosphere. Rep. Prog. Phys. **72**, 1–21 (2009)

Heintzenberg, J.,Charlson, R.J. (eds.): Clouds in the perturbed climate system: their relationship to energy balance, Atmospheric Dynamics, and Precipitation. vol. 2, Strüngmann Forum Report, The MIT Press, pp. 597 (2009)

Hendricks, J., Kärcher, B., Lohmann, U.: Effects of ice nuclei on cirrus clouds in a global climate model. J. Geophys. Res. **116**, D18206, 1–24, (2011) doi:10.1029/2010JD015302

Hobbs, P. V.: Ice Physics. Clarendon Press (1974)

Houze Jr., R. A.: Cloud Dynamics. Academic Press, London (1993)

Hoyle, C.R., Luo, B.P., Peter, T.: The origin of high ice crystal number densities in cirrus clouds. J. Atmos. Sci. **62**, 2568–2579 (2005)

Hu, X.L., Michaelides, A.: Ice formation on kaolinite: lattice match or amphoterism? Surface Sci. **601**, 5378–5381 (2007)

Jensen, E.J., Thomas, G.E.: Numerical simulations of the effects of gravity waves on noctilucent clouds. J. Geophys. Res. **99**, D2, (1994). doi:10.1029/93JD01736

Jensen, E.J., Pfister, L., Bui, T.-P., Lawson, P., Baumgardner, D.: Ice nucleation and cloud microphysical properties in tropical tropopause cirrus. Atmos. Chem. Phys. **10**, 1369–1384 (2010)

Kärcher, B.: Aviation-produced aerosols and contrails. Surv. Geophys. **20**, 113–167 (1999)

Kärcher, B., Solomon, S.: On the composition and optical extinction of particles in the tropopause region. J. Geophys. Res. **104**, 27441–27459 (1999)

Kärcher, B., Lohmann, U.: A parameterization of cirrus cloud formation: Homogeneous freezing including effects of aerosol size. J. Geophys. Res. **107**, 4698 (2002). doi:10.1029/2001JD001429

Kärcher, B., Ström, J.: The roles of dynamical variability and aerosols in cirrus cloud formation. Atmos. Chem. Phys. **3**, 823–838 (2003)
Kärcher, B., Koop, T.: The role of organic aerosols in homogeneous ice formation. Atmos. Chem. Phys. **5**, 703–714 (2005)
Kärcher, B., Burkhardt, U.: A cirrus cloud scheme for general circulation models. Quart. J. Roy. Meteorol. Soc. **134**, 1439–1461 (2008)
Kärcher, B.: Supersaturation fluctuations in cirrus clouds driven by colored noise. J. Atmos. Sci. **69**, 435–443 (2012)
Koop, T.: Homogeneous ice nucleation in water and aqueous solutions. Z. Phys. Chem. **218**, 1231–1258 (2004)
Koop, T., Bookhold, J., Manabu, S., Pöschl, U.: Glass transition and phase state of organic compounds: dependency on molecular properties and implications for secondary organic aerosols in the atmosphere. Phys. Chem. Chem. Phys. **13**, 19238–19255 (2011)
Murphy, D.M.: Dehydration in cold clouds is enhanced by a transition from cubic to hexagonal ice. Geophys. Res. Lett. **30**, 2230 (2003). doi:10.1029/2003GL018566
Murray, B.J.: Inhibition of ice crystallisation in highly viscous aqueous organic acid droplets. Atmos. Chem. Phys. **8**, 5423–5433 (2008)
Murray, B.J., Jensen, E.J.: Homogeneous nucleation of amorphous solid water particles in the upper mesosphere. J. Atmos. Sol. Terr. Phys. **72**, 51–61 (2010)
Murphy, D.M., Koop, T.: Review of the vapour pressures of ice and supercooled water for atmospheric applications. Q. J. R. Meteorol. Soc. **131**, 1539–1565 (2005)
Murray, B.J., Wilson, T.W., Dobbie, S., Cui, Z., Al-Jumur, S.M.R.K., Möhler, O., Schnaiter, M., Wagner, R., Benz, S., Niemand, M., et al.: Heterogeneous nucleation of ice particles on glassy aerosols under cirrus conditions. Nature Geosci. **3**, 233–237 (2010)
Peter, T.: Microphysics and heterogeneous chemistry of polar stratospheric clouds. Ann. Rev. Phys. Chem. **48**, 785–822 (1997)
Pruppacher, H. R., Klett, J.D.: Microphysics of clouds and precipitation. Kluwer Academic Publishers (1978)
Rapp, M., Thomas, G.E.: Modeling the microphysics of mesospheric ice particles: assessment of current capabilities and basic sensitivities. J. Atmos. Sol. Terr. Phys. **68**, 715–744 (2006)
Sassen, K., DeMott, P.J., Prospero, J.M., Poellot, M.R.: Saharan dust storms and indirect aerosol effects on clouds: CRYSTAL-FACE results. Geophys. Res. Lett. **30**, (2003). doi: 10.1029/2003GL017371
Sazaki, G., Zepeda, S., Nakatsubo, S., Yokoyama, E., Furukawa, Y.: Elementary steps at the surface of ice crystals visualized by advanced optical microscopy. Proc. Nat. Acad. Sci. U.S.A. **107**, 19702–19707 (2010)
Solomon, S.: Stratospheric ozone depletion: a review of concepts and history. Rev. Geophys. **37**(3), 275–316 (1999)
Vali, G.: Nucleation terminology. J. Aerosol Sci. **16**, 575–576 (1985)
Wiaczek, A., Peter, T., Lohmann, U.: The potential influence of Asian and African mineral dust on ice, mixed-phase and liquid water clouds. Atmos. Chem. Phys. **10**, 8649–8667 (2010)

Detection and Analysis of Water Vapor Transport

11

Christoph Kiemle, Andreas Schäfler and Christiane Voigt

Abstract

Water vapor, though a minor constituent of Earth's atmosphere, plays a major role in the atmospheric radiation budget and the global water cycle. Atmospheric water vapor concentrations are highly variable due to the complex interplay between their sources (evaporation) and sinks (condensation and precipitation) in combination with transport and mixing. They strongly decrease with temperature and thus with altitude. Accurate measurement of water vapor is essential for better understanding its transport and cloud formation in the atmosphere and their impact on both weather and climate. To this end the institute develops and deploys lidars and in situ hygrometers onboard aircraft.

C. Kiemle (✉) · A. Schäfler · C. Voigt
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Christoph.Kiemle@dlr.de

A. Schäfler
e-mail: Andreas.Schaefler@dlr.de

C. Voigt
e-mail: Christiane.Voigt@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_11,

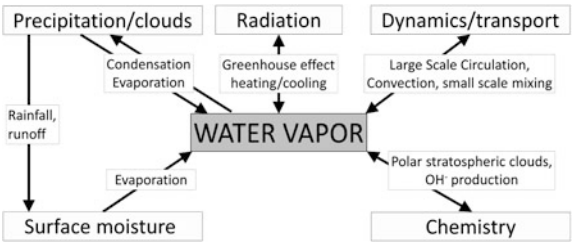

Fig. 11.1 Interactions of water vapor with components of the climate system. The hydrological cycle is sketched on the *left*

11.1 Introduction

The continuous exchange of *water* between the oceans, the continents, and the atmosphere, crucial to life on Earth, is termed the *hydrological cycle* (Fig. 11.1). Although the amount of water in the atmosphere is small [0.25 % of the total air mass; Möller et al. (2011)], the atmospheric branch of the hydrological cycle is of key importance as it distributes *humidity* around the globe and between the different water reservoirs. Driven by *solar radiation*, evaporation from the surface represents the source of tropospheric *water vapor*. The water vapor concentration is highly variable due to variable radiation distribution (which depends on geographic location and cloud conditions) and the strong dependency of the water-holding capacity of air on the temperature: under standard surface conditions, an increase in temperature by 1 °C augments this capacity by 7 % (Trenberth et al. 2007).

Globally, 86 % of the evaporation occurs over the oceans; if all the water in the atmosphere were condensed and deposited onto the surface it would form a 25 mm deep layer (Möller et al. 2011). A rough estimate of the residence time of water in the *troposphere* is 8 days, assuming a global mean *precipitation* rate of 3 mm per day or 1100 mm per year (Trenberth et al. 2003). Net evaporation over the oceans, which provides a continuous water supply to the atmosphere, exceeds the rate of precipitation, which implies a net atmospheric *transport* of water vapor towards the continents. This transport follows global *circulation* patterns: in so-called Hadley circulation (Fig. 11.2), large amounts of subtropical moist air masses are transported toward the equator and converge in the tropics where they ascend into the *upper troposphere* by tropical thunderstorms. At upper levels, there is a poleward reverse flow which descends over the subtropics, causing dry and cloud-free regions that favor evaporation. At higher latitudes, low-pressure systems produce a net poleward transport of water vapor. Conversely, atmospheric circulation is influenced by the transport of humidity, as it carries *latent heat*. The solid (ice) and liquid phases of water have strong intermolecular bonds. Evaporation, the transition to the gas phase, requires a considerable amount of energy to free these bonds. This energy is called latent heat. It is released when atmospheric water vapor condenses into clouds or onto surfaces, and cools the air when condensate (*ice crystals*, cloud or rain droplets) evaporates. A significant portion of the energy

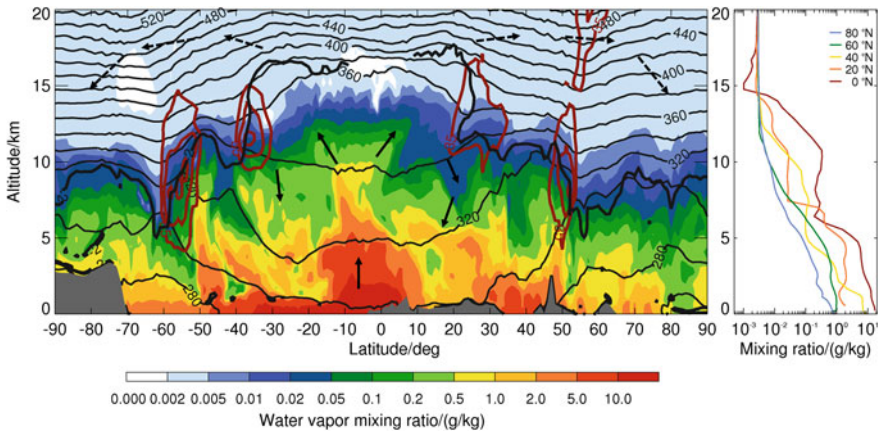

Fig. 11.2 Pole-to-pole meridional humidity cross-section along 11° E (through Oberpfaffenhofen; 48° N) with isentropes (*thin black lines*; K), dynamical tropopause (*thick black line*; 2 PVU, 1 PVU = 10^{-6} m^2/s K/kg) and wind velocity (*red*; m/s; wind direction is perpendicular to the plot's plane). ECMWF analysis for 00 UTC 21 December 2010 at 1° resolution. *Grey shading* is topography: Antarctica, Africa, the Alps, and Norway. Hadley circulation in the tropics, slightly shifted southwards due to northern hemispheric winter time, and Brewer–Dobson circulation in the stratosphere are indicated by *arrows*. The altitude axis is expanded by a factor of 500 and slopes are much steeper than in reality. Selected profiles in the northern hemisphere are shown on the *right*

in extreme weather events such as *fronts*, *thunderstorms* and *cyclones* is provided by the latent heat released when the associated large clouds form and grow.

Water vapor is the most important *greenhouse gas* due to the particular characteristics of its molecular lines that efficiently absorb solar and *terrestrial radiation*, and because it is much more abundant than all other greenhouse gases. Also, water and ice clouds exert an intense influence on both incoming solar and *outgoing longwave radiation*. Climate models reveal that predictions of climate change are very sensitive to water vapor and cloud *feedback*: the water vapor feedback alone, i.e., the humidity increase in a warmer atmosphere, roughly doubles the warming caused by, e.g., increases in other greenhouse gases (Randall et al. 2007). Finally, water vapor is, by its presence or absence, an important component of atmospheric chemistry (Fig. 11.1): The fact that the *stratosphere* is very dry, with only three to six water molecules per million air molecules, is essential to *ozone* chemistry. Driven by Hadley circulation (Fig. 11.2) water vapor mainly enters the stratosphere in the tropics (Fueglistaler et al. 2009). At the particularly cold tropical *tropopause* the air is "freeze-dried." The remaining moisture is transported by the much slower *Brewer–Dobson circulation* within several years towards the poles, where it descends due to radiative cooling.

Figure 11.2 shows the distribution of water vapor on a particular day. In contrast to climatological mean cross sections, such a "global snapshot" preserves the small scale variability while still reproducing the major climatological characteristics on the large

scale. The strong static stability above the tropopause, visible in the closely-spaced *isentropes* in Fig. 11.2, basically suppresses vertical transport into and within the stratosphere. Generally, the tropopause is located at altitudes around 17 km in the tropics and around 10 km elsewhere (Gettelman et al. 2011). At *mid-latitudes* the strong variability of the tropopause is due to intense air mass stirring by frontal systems. Their particular dynamics generate tropopause bends and folds, whereby dry stratospheric air is entrained into the troposphere (Gettelman et al. 2011). The Hadley circulation as well as the subtropical and polar *jet streams* are clearly visible. To display the large concentration differences, the figure has a logarithmic scale. The driest region, the tropical tropopause, has a mass *mixing ratio* of typically 0.002 g/kg, i.e., 2 mg water vapor per kg dry air, or about three water molecules per million air molecules (as the water molecule is lighter). The highest humidity is encountered in the tropics near the surface with up to 25 g/kg, which is more than four orders of magnitude larger (Möller et al. 2011). At the surface the humidity varies by more than one order of magnitude between the equator (>10 g/kg) and the poles (<1 g/kg), as Fig. 11.2 shows. This is mainly due to the strong surface temperature differences and the temperature dependency of the water vapor holding capacity of air, as already explained.

11.2 Observations

Water vapor is routinely measured between Earth's surface and the lower stratosphere with a global *radiosonde* network. While the radiosondes' accuracy is good in the lower and middle troposphere, their precision may require improvement in the *upper troposphere and lower stratosphere* (UTLS) region. Furthermore, few data are collected over the oceans, at high latitudes, and in the southern hemisphere. Water vapor measurements from ground-based remote sensing lidars add up to the radiosonde network. Water vapor is also measured by several satellite instruments with global coverage, yet low *accuracy* and coarse spatial resolution, particularly in the UTLS (Kley et al. 2000). Here, hygrometers installed on commercial and *research aircraft* perform better. For research purposes, upward- or downward-viewing, airborne water vapor *differential absorption lidars* (DIAL) were developed and have been operated by the institute's lidar group for 20 years (Ehret et al. 1993). First deployments in field experiments helped to characterize the variability of water vapor and aerosols in the lower troposphere (Kiemle et al. 1995, 1997) and at the top of the atmospheric boundary layer (Davis et al. 1997, 2000). In 1997 a lidar for the detection of low humidity in the UTLS was developed, and *tropopause folds* (cf. Fig. 11.2) could be observed in unprecedented detail (Ehret et al. 1999; Poberaj et al. 2002; Flentje et al. 2005). Comparisons with other instruments revealed that the lidars' measurement accuracy in the UTLS is ∼8 % (Kiemle et al. 2008). Since 2007 a new DIAL running simultaneously at four wavelengths for high sensitivity throughout the troposphere and the lower stratosphere has been operated (Wirth et al. 2009). By choosing water vapor absorption lines of various strengths, this instrument is able to cope with highly variable concentrations (cf. Fig. 11.2).

Besides water vapor, wind velocity can also be measured by lidar utilizing the Doppler effect. The *Doppler wind lidar* (DWL) is operated either in a scanning operation mode for three-dimensional wind field profiles or in a simple nadir-viewing mode for profiles of vertical wind velocity. Combining a water vapor and a wind lidar on one research aircraft enables the study of a variety of atmospheric transport processes, particularly over remote areas such as oceans, arctic regions, or the tropics. The lidars monitor the water vapor and wind fields in cloud-free areas in a two-dimensional curtain along the flight path. This enables direct measurements of humidity transport. After successful ground measurements (Giez et al. 1999) this unique instrument combination was deployed on board the DLR *Falcon* research aircraft in two field experiments, IHOP_2002 [USA; Kiemle et al. (2007)] and COPS 2007 [Central Europe; Schäfler et al. (2010)], as exemplified below. Thereafter, the lidars participated in the THORPEX-IPY mission in 2008 in Norway, focusing on polar research (Wagner et al. 2011). During T-PARC 2008 in Japan, airborne lidar observations were made in the vicinity of tropical cyclones to investigate the potential for improving cyclone forecasts with supplementary, targeted observations (Harnisch et al. 2011; Weissmann et al. 2012).

In addition to the lidars, the institute operates *dropsondes* and a suite of in situ hygrometers on research aircraft, including a tunable diode laser (TDL) instrument, a frost point mirror and a new water vapor *mass spectrometer*. The TDL instrument WARAN (Water Vapor Atmospheric Analyzer, Spectra Sensors) measures the water vapor absorption by scanning over the 1.3 μm water vapor absorption line. It operates from the surface to the upper troposphere, equivalent to mixing ratios between 12 and 0.03 g/kg, with an accuracy of 10 %. Low humidity in the UTLS region is measured by a frost point hygrometer [CR-2 from Buck research instruments; Busen and Buck (1995)] at an accuracy of 8 % (Voigt et al. 2010; Schumann et al. 2011). The *frost point* is defined as the temperature at which the atmospheric water vapor partial pressure in air equals the water vapor pressure over a flat ice surface. In contrast, the *dew point* is the temperature at which the water vapor partial pressure equals that over a liquid water surface. A *relative humidity* (RH) of 100 % or saturation with respect to water is reached when the ambient temperature equals the dew point. Since ice has a stronger bonding between neighboring water molecules it is more difficult for water molecules to escape a frozen surface than a liquid surface. Hence the frost point is higher than the dew point. In the frost (dew) point sensor, the thickness of a frost (or liquid) layer on a mirror is kept constant by applying cooling/heating cycles to the mirror. The measured frost (dew) point temperature is finally converted to the saturation partial pressure using the *Clausius-Clapeyron* equation.

The in situ water vapor instrumentation is completed by the Atmospheric Ionization Mass Spectrometer AIMS-H2O, a newly developed mass spectrometer system with an in-flight water *calibration* standard. In the AIMS instrument, atmospheric water vapor is ionized in a gas discharge ion source and the resulting protonated water clusters are measured with a quadrupole detection unit. AIMS is optimized for mixing ratios below 0.1 g/kg, and spatial scales down to 40 m horizontally can be resolved with the 4/s time resolution. The in situ hygrometers

were operated on the Falcon aircraft during the CONCERT 2008 and 2011 missions, which focused on measuring absolute and relative humidity in the UTLS region within and outside *cirrus* clouds and *contrails*. The dew point sensor also operated during the *Eyjafjallajökull* mission (Schumann et al. 2011). The TDL hygrometer was deployed during an intercomparison campaign on board the Falcon, in parallel with the water vapor lidar on the high altitude research aircraft *HALO*. Case studies of humidity transport observed by lidar and in-situ instruments on the Falcon aircraft are exemplified in the next sections.

11.3 Vertical Transport in the Atmospheric Boundary Layer

Evaporation over land and especially over complex terrain can be very heterogeneous on small scales that are both under-sampled by observations and unresolved by *numerical weather prediction* (NWP) models. Over a surface heated by the sun, a *convective boundary layer* (CBL) develops, typically reaching heights of 1–2 km above ground in the early afternoon. Buoyant warm air parcels lift the humidity from its surface sources. Cloud *condensation* at the top of the colder boundary layer, and subsequent cloud growth with occasionally heavy precipitation, may occur. Observational deficits due to the heterogeneity of soil and atmospheric moisture, and modeling deficits due to coarse resolution are the main reasons why NWP models fail to accurately forecast continental precipitation in the warm season (Weckwerth et al. 2004). In particular, variations of humidity in the boundary layer have the largest impact on precipitation patterns (Keil et al. 2008). While CBL processes have been the focus of field campaigns, laboratory studies and numerical models for decades, many key questions relevant to the initiation of storms and heavy precipitation persist: What processes govern water vapor distribution in the lower troposphere? How does land cover heterogeneity influence water vapor and cloud growth? How well are CBL processes simulated by models?

Mountains enhance the complexity: As obstacles to the large-scale flow they generate differing local flows. Differences in solar heating and land cover add to the heterogeneity. Consequently, a field experiment called “Convective and Orographically-induced Precipitation Study” (COPS) was conducted in July 2007 over the low mountains of the Black Forest in southwestern Germany, with the participation of the DLR Falcon aircraft. Figure 11.3 exemplarily shows two-dimensional vertical water vapor lidar cross-sections beneath the aircraft flight track. To verify the lidar results, comparisons with ground-based zenith-viewing lidars and with instruments on board other coordinated research aircraft were undertaken. One of the main lidar goals was to measure vertical humidity transport within the CBL over the Rhine valley and the Black Forest on days when thunderstorms, mainly influenced by the local orography and surface humidity, were forecasted. The experiment proved the capability of this novel lidar instrumentation to quantify vertical *moisture transport* over complex terrain. On this exemplary day the lidar-derived latent heat fluxes over the mountains vary between 100 and 500 W/m^2 but are roughly constant with height throughout the CBL (Kiemle et al. 2011). The fluxes moistened the growing CBL by upward transport of humidity

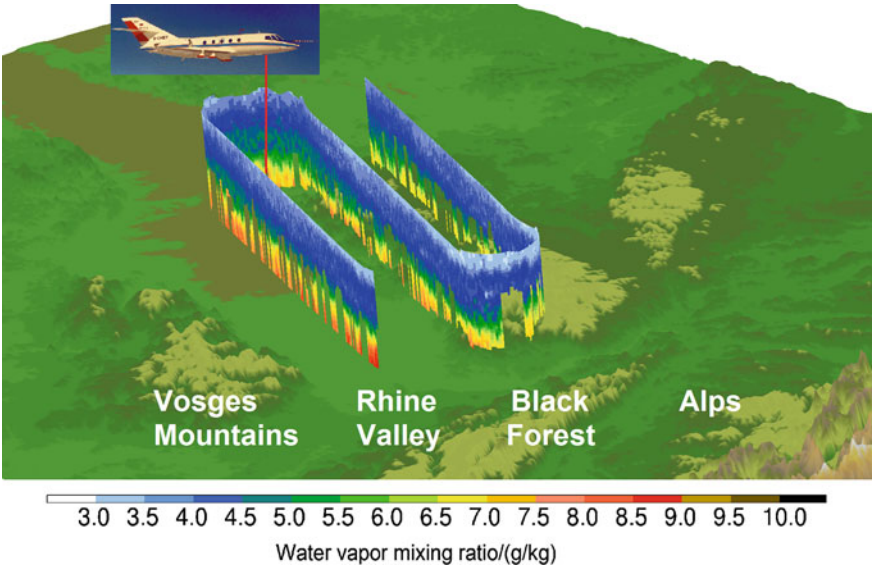

Fig. 11.3 Airborne lidar measurements over complex terrain in southwest Germany on 30 July 2007 at noon as perspective from southwest. *Blue* represents dry air above; *yellow* and *red* show humid air within the convective boundary layer (CBL). Interruptions are due to scattered fair-weather clouds at the CBL top, located between altitudes of about 1.5 and 2.0 km. The length × height of the data curtains is 150 km× 4 km. The topographic heights range from 150 to 1500 m

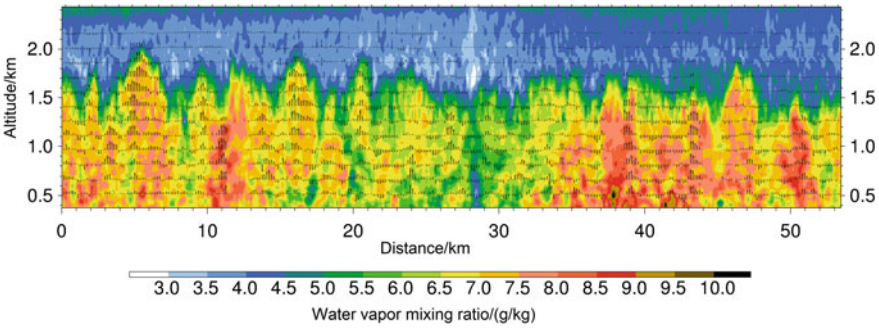

Fig. 11.4 Distance-to-height composite of humidity and vertical velocity (*arrows*) from lidar profiles on 7 June 2002 at noon over southwestern Kansas. Air parcels rising up to 2 km above ground are up to 3 km broad and visibly more humid than the slower downdrafts. The updrafts with a maximum of 7 m/s were very vigorous. The DLR Falcon flew over this cloud-free convective boundary layer (CBL) in 4 km altitude for 6 min, providing an unprecedented snapshot of the flow complexity (from Kiemle et al. 2007)

from surface evaporation due to rain on the previous days. With a CBL air density of about 1 kg/m^3 the net upward water mass flux amounts to (0.04 to 0.20) g/m^2/s or (0.14 to 0.72) kg/m^2/h. For the Black Forest area of roughly 150 km × 50 km,

the hourly water mass uplift equals (1 to 5) Tg. Assuming an average hourly flux of 3 Tg and a 3 h time span of equally vigorous convection, 9 Tg, or 9 million tons of water were lifted through the CBL over the Black Forest on that day.

When humid and warm air parcels rise in the CBL, drier and colder air from aloft sinks in between. The net vertical humidity flux through an idealized horizontal plane can be computed from the vertical wind and water vapor fluctuations. Figure 11.4 illustrates these ups and downs in a vigorous CBL over the U.S. Southern Great Plains during the IHOP_2002 (International H2O Project) field experiment. The observations corroborate previous theoretical and model studies, e.g. Schmidt and Schumann (1989). In this experiment the NOAA High Resolution Doppler Lidar [HRDL; Grund et al. (2001)] was deployed alongside the water vapor DIAL on the DLR Falcon aircraft. The region is subject to devastating summer precipitation. The thunderstorms, occasionally leading to tornadoes, are generated by air mass clashes between the warm and moist air from the Gulf of Mexico and colder air from the north. Water vapor plays a key role since its latent heat provides considerable additional energy, as mentioned before. Consequently, the IHOP_2002 experiment, focusing on this particular trace gas, had four main objectives: (1) determine the degree of improvement in NWP forecast skill from improved knowledge of the initial water vapor fields; (2) understand the processes that determine where and when storms form; (3) examine the relationship between atmospheric water vapor, surface and boundary layer processes and their impact on the initiation of storms; (4) quantify the accuracy of water vapor measurements. The institute's airborne water vapor lidar observations helped advance these issues (Couvreux et al. 2006; Behrendt et al. 2007; Kiemle et al. 2007; Tollerud et al. 2008).

11.4 Vertical Transport in Thunderstorms

Thunderstorms are strong elevators of lower-tropospheric air masses, particles, water vapor, and other trace gases, particularly in the tropics. So-called *deep convection* was a main topic studied during the Tropical Convection, Cirrus and Nitrogen Oxides Experiment (TROCCINOX) that took place in 2004 and 2005 in Brazil. Here the DLR research aircraft Falcon carried the water-vapor DIAL in an upward-pointing direction. Aerosol and water vapor in the tropical upper troposphere were measured in the vicinity of deep convection, as well as in remote areas free of convection, to characterize the atmospheric background state. Figure 11.5 shows lidar results from an underflight of the outflow of tropical thunderstorms into the upper troposphere. An *anvil* cloud with top altitude of 14 km extends horizontally over 150 km and is embedded in a layer of elevated lidar *backscatter ratio*. A zoom into that layer reveals relatively high humidity close to the cloud, between 12 and 13 km altitude and up to a distance of 100 km from the cloud edge. The humidity is about three times higher than the background at the left of the figure and obviously stems from the outflow of one or more anvil clouds.

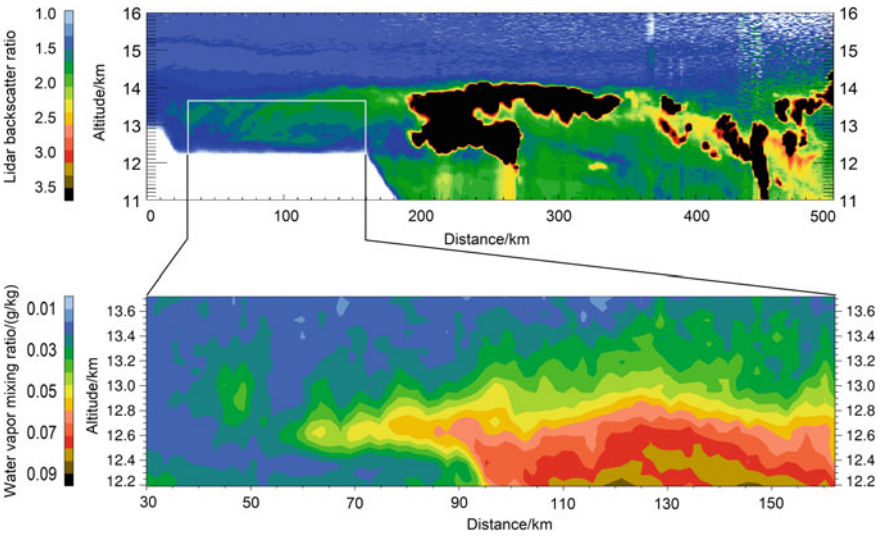

Fig. 11.5 Airborne lidar backscatter (*top*) and humidity (*bottom*) measurements in the outflow of a tropical thunderstorm in Brazil on 3 March 2004. One or possibly more thunderstorms were located beneath the ice clouds (*black*). The border of the water vapor outflow (*bottom*) is 100 km from the ice cloud edge (*top*). The measurements have a horizontal resolution of 2 km and a vertical resolution of 200 m (from Kiemle et al. 2008)

Using the lidar observation in Fig. 11.5 we can estimate the amount of water lifted by the thunderstorm(s). Despite uncertainty about the extent of the outflow and the presence of additional thunderstorms in the perpendicular direction (satellite observations are too coarse), it is reasonable to assume that above 12 km altitude, a volume of at least 200 × 50 × 1 km^3, or 10 000 km^3, was humidified, including the anvil clouds. Figure 11.5 shows that the background excess moisture amounts to at least 0.04 g/kg, which corresponds to a surplus water density of about 10 mg/m^3 and agrees fairly well with an institute's former measurement of a mid-latitude anvil outflow (Hauf et al. 1995). Because of lidar backscatter signal overload, the humidity in the cloud could not be measured, but it is likely that its water content was higher due to the ice particles and to the fact that it was up to 2 km thick. Consequently, the thunderstorms must have lifted at least 100 000 tons or 0.1 Tg of water up to an altitude of 12 km, the lower border of the volume defined above. Assuming a lifetime of 3 h, the vertical uplift amounts to 10^{-6} $kg/m^2/s$ on average over an area of 200 × 50 km^2. This gives a total area flux of 10 tons/s that fits with high-resolution numerical model results of tropical thunderstorms during that campaign (Chaboureau et al. 2007). Rapidly elevated this way into the tropical upper troposphere, water vapor may enter the stratosphere via subsequent slow ascent, with high relevance for climate and ozone chemistry, as mentioned above. Sometimes the updrafts are so strong that the thunderstorms reach the tropopause; yet the amount of water rapidly injected into the stratosphere this way by so-called "overshooting convection" is difficult to quantify (Fueglistaler et al. 2009).

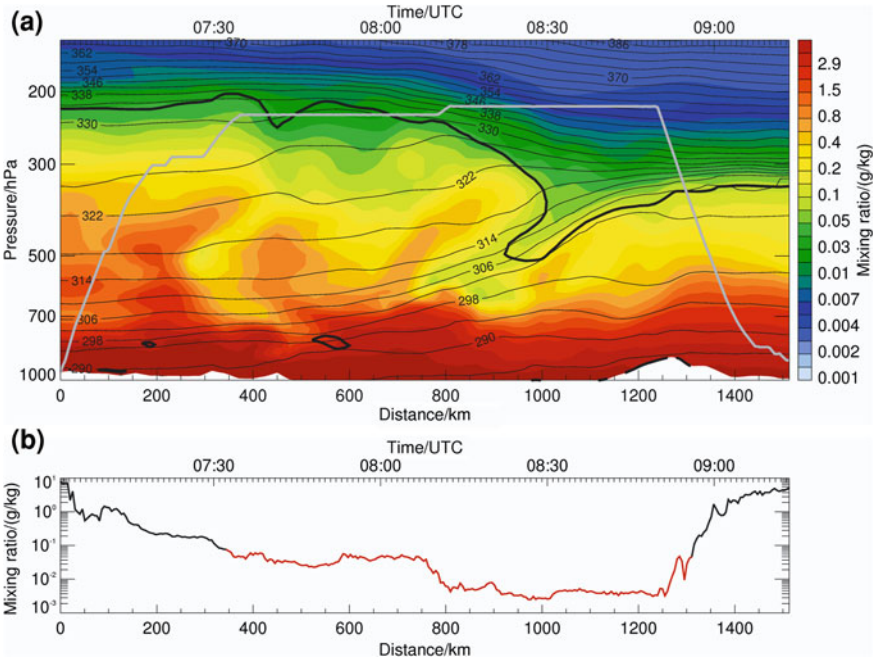

Fig. 11.6 **a** ECMWF analysis for 0600 UTC 22 September 2011 along a south-to-north oriented flight track from Oberpfaffenhofen to Bergen. Dynamical tropopause (2 PVU, *thick black line*), potential temperature (*thin black lines*, K) and the flight altitude (*gray*) are indicated. **b** In situ humidity measurements by the TDL (*black*) and the AIMS-H2O (*red*) during the flight

11.5 Stratosphere-Troposphere Exchange of Water Vapor in a Stratospheric Intrusion

Stratospheric intrusions or tropopause folds (see Fig. 11.2) associated with low pressure systems in the mid-latitudes may lead to enhanced exchange of air between the stratosphere and the troposphere. A stratospheric intrusion has been probed on a flight from Oberpfaffenhofen to Bergen (Norway) on 22 September 2011 during the CONCERT 2011 campaign with in situ water vapor instruments on board the research aircraft Falcon. The intrusion was linked to a low-pressure system above the Norwegian Sea that transported dry polar stratospheric air southward and downward to the lower and mid-troposphere over Denmark. Figure 11.6 shows humidity profiles interpolated from the ECMWF analysis at 06 UTC along the flight track together with the dynamical tropopause that indicates the tropopause fold, bended down to 6 km altitude.

In the fold, dry stratospheric air is transported along the isentropes into the lower troposphere. After ascent through a moist boundary layer below 700 hPa ($\sim$3 km) up to the upper troposphere, the aircraft flew close to the tropopause

between 0730 and 0810 UTC. Here, water vapor and wind fields near the tropopause can be used to investigate small-scale structures and the sharpness of the tropopause. Due to the lowered tropopause the Falcon then entered the stratosphere and was able to reach very low humidity before its descent. Such observations are highly valuable for future comparisons with numerical weather and climate prediction models. Figure 11.6b shows that the humidity varies by more than three orders of magnitude between 8 g/kg in the boundary layer and 0.003 g/kg in the stratosphere. A distinct transition to the stratosphere is observed at 0810 UTC. During the return flight (not shown) the boundary region of the stratospheric intrusion was probed at different altitudes to investigate the dynamics of the tropopause fold and its representation in model simulations.

11.6 Transport of Moisture in a Mid-Latitude Cyclone Observed by Wind and Water Vapor Lidars

In the mid-latitudes between 30 and 60°N, cyclones efficiently smooth differences between polar and subtropical areas that result from differing solar radiation by transporting moist and warm air poleward and cold and dry air toward the equator. Usually the polar and subtropical air masses are separated by the polar front, a narrow band in the mid-latitudes possessing a sharp meridional temperature change. Low pressure systems with typical horizontal scales of (500 to 2000) km evolve along this frontal zone and transport air horizontally and vertically, leading to cloud development and precipitation. Most atmospheric transport processes can here be considered quasi-horizontal due to the minor depth of the layer where weather happens, the troposphere (Fig. 11.2). Intermittent very strong cyclones cause flooding and wind damage particularly challenging for society. Their strength depends on the moisture supply which originates partly from evaporation in the air surrounding the cyclone and partly from the convergence of moist air transported horizontally over large distances. On average, 70 % of the moisture is already in the air before a cyclone develops, while 30 % evaporates from the surface (Trenberth et al. 2003). The release of latent heat critically impacts the development of cyclones and is therefore an important issue of present-day research in dynamical meteorology. Advanced observation techniques like lidar are useful for observing water vapor transport in extra-tropical cyclones. Besides understanding the physical processes, verifying model simulations by independent observations helps to pinpoint insufficiencies in current NWP models. The last example is chosen from the COPS field experiment and highlights the value of high resolution lidar data for observing the complex three-dimensional transport of moisture in the air surrounding a cyclone.

Figure 11.7 shows three-dimensional illustrations of a research flight devoted to mapping the moisture and wind conditions ahead of an approaching low-pressure system (L) located at the western coastline of France. Air circulates counter-clockwise around the low, approximately parallel to the isobars. As mentioned, cyclones form along a frontal zone (i.e., a gradient of temperature and moisture;

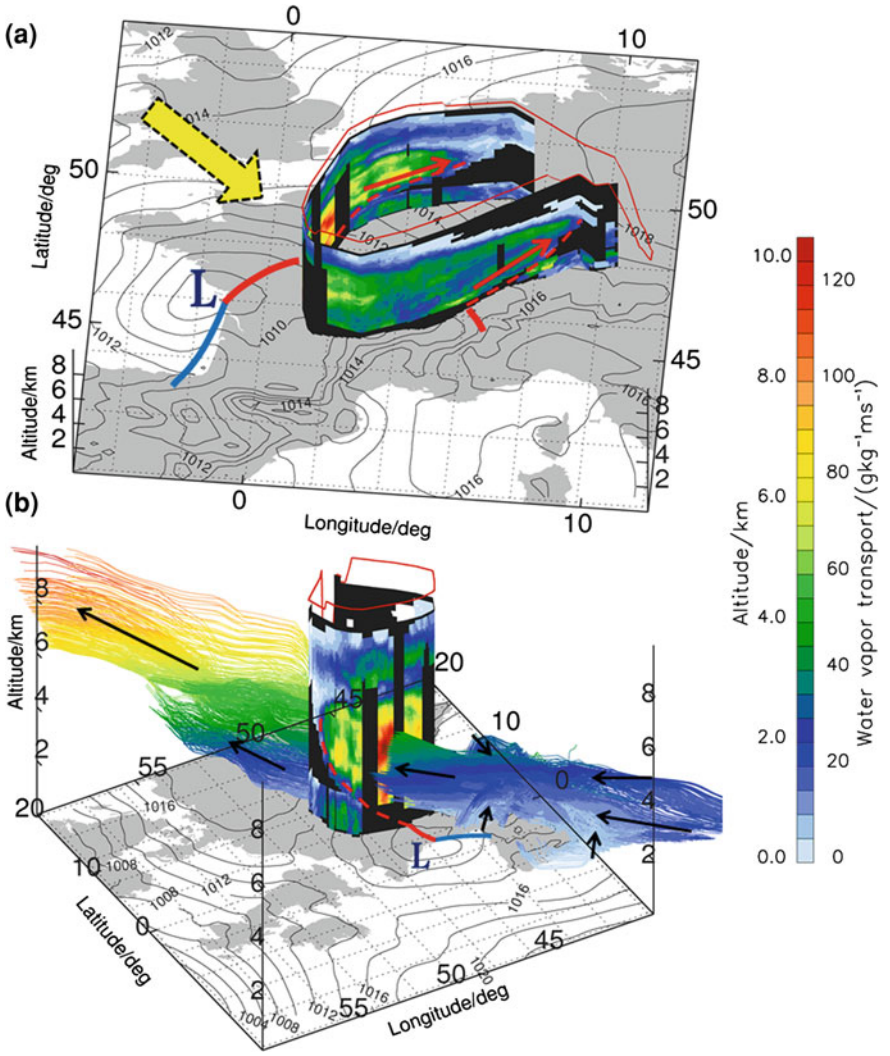

Fig. 11.7 Airborne lidar observations on 1 August 2007 from combined water vapor and wind lidars. **a** Perspective from the south. **b** Perspective from northwest, as indicated by the *yellow arrow* in **a**. The aircraft flew along the *thin red line*. The curtain-like two-dimensional cross-section underneath shows water vapor transport as product of mixing ratio (g/kg) and wind speed (m/s). The *black areas* are data gaps. The *dashed red lines* show the vertical extension of the surface warm front (*solid red line*) along the lidar curtain. **b** Also contains selected trajectories of air parcels related to the maximum observed water vapor transport. Their colors indicate the altitude of the parcels for 2.5 days in advance and 1.5 days after the lidar observations. *Black arrows* show major transport pathways

red and blue lines) that deforms during their evolution. On the rear side of the system, cold dry air from the north moves southward forming a cold front along the French Atlantic coast line. On the leading edge of the system, warm moist air moves northeastward and glides above the cold air, shaping the so-called warm front. The airborne lidars observed this area of enhanced poleward and upward moisture transport in the warm air sector (i.e., the area in between cold and warm fronts) during a twofold overflight of the warm frontal zone. The curtain-like two-dimensional cross section represents a snapshot of the water vapor transport determined by combining wind speed and water vapor measurements. The observations reveal an up to 6 km deep layer of enhanced water vapor transport confined to the warm sector of the cyclone. Maximum values occur on the most western part of the flight facing the cold front, where increased wind speeds and water vapor concentrations coincide. The strong horizontal gradients of moisture transport at the warm front show a tilting with height, accompanying the north-eastward ascending air masses (red dashed lines along the lidar cross section). The tilting is observed on both overpasses of the warm front. The red arrows represent the up-gliding of the warm and moist air above the colder air to the north. In contrast, in the northeastern part, the horizontal transport in the boundary layer is weak.

Figure 11.7b shows the same scenery from a northwestern perspective. The increased transport on the western side that extends upward along the tilted warm front is clearly visible. The figure also displays trajectories that describe the pathways of the observed air parcels in space and time, calculated with the *trajectory model* LAGRANTO (Wernli et al. 1997). Shown are the pathways up to 2.5 days in advance and up to 1.5 days after the lidar observation. The trajectories were calculated using NWP wind data. The strong moisture transport clearly originates from the southwest. Air in the lower troposphere is humidified over the Atlantic and the Mediterranean by evaporation and subsequent vertical transport in the boundary layer. These air masses converge over southeastern Europe. The developing cyclone incorporates and transports the moisture northeastward towards the flight path. Many air parcels slightly ascend towards the lidar cross section and further up until they reach the upper troposphere ($\sim$8 km) 1.5 days after being observed by the lidar. For an in-depth discussion of this case study and a comparison with NWP data the reader is referred to Schäfler et al. (2010). The airborne lidar instruments are able to map the clear-air transport of water vapor into cyclones that is known to play a key role in the structure, intensity, and further evolution of cyclones. The latent heat release is expected to amplify cyclones, and also to affect the upper-level wind fields which control the weather development further downstream. This emphasizes the importance of a correct representation of moisture for the quality of *weather forecasts*. Schäfler et al. (2011) show that the lidar observations are useful for diagnosing model errors which result from either incorrect horizontal or vertical transport of moisture or from insufficiently represented evaporation in moisture source regions.

11.7 Conclusion and Outlook

Accurate water vapor measurements are sparse and challenging because of the high variability of its sources in space and time, and because of the complexity of atmospheric motion and mixing. Advanced airborne water vapor observations are, as shown, valuable for addressing relevant issues of current research in atmospheric water vapor and transport. There is an increasing need for improved measurements, for example of the relative humidity in the upper troposphere in regions with ice *supersaturation* (Gierens et al. 2000). This implies having simultaneous information on temperature from NWP data or auxiliary measurements. A combined water vapor and temperature lidar is a future challenging task, as temperature profiling by lidar is basically possible yet still subject to high detection noise, which necessitates long averaging periods and consequently means coarse spatial resolution. Water vapor in the lower troposphere is particularly relevant for weather (Keil et al. 2008), while water vapor in the upper troposphere and the stratosphere is particularly relevant for climate, where, after vertical transport from the surface sources, it strongly influences Earth's radiation budget (Trenberth et al. 2007). Consequently, understanding the transport and distribution of atmospheric humidity and the processes that control it is of primary importance to both climate research and weather prediction. New instrument developments such as mass spectrometry and multi-wavelength lidar will, on board the new HALO research aircraft, advance the science, building on the knowledge gained in the above-mentioned experiments. The measurements will help to understand significant atmospheric processes and to evaluate state-of-the-art weather and climate models. In addition, future satellite missions to measure wind (ESA *Atmospheric Dynamics Mission*) or water vapor (ESA WALES: Water vapor Lidar Experiment in Space) will be supported and validated.

References

Behrendt, A., Wulfmeyer, V., Kiemle, C., Ehret, G., Flamant, C., Schaberl, T., Bauer, H.-S., Kooi, S., Ismail, S., Ferrare, R., et al.: Intercomparison of water vapor data measured with Lidar during IHOP_2002. Part II: Airborne-to-airborne systems. J. Atmos. Oceanic Technol. **24**, 22–39 (2007)

Busen, R., Buck, A.L.: A high-performance hygrometer for aircraft use: description, installation, and flight data. J. Atmos. Oceanic Technol. **12**, 73–84 (1995)

Chaboureau, J.-P., Cammas, J.-P., Duron, J., Mascart, P.J., Sitnikov, N.M., Voessing, H.-J.: A numerical study of tropical cross-tropopause transport by convective overshoots. Atmos. Chem. Phys. **7**, 1731–1740 (2007)

Couvreux, F., Guichard, F., Redelsperger, J.L., Kiemle, C., Masson, V., Lafore, J.P., Flamant, C.: Water vapour variability within a convective boundary layer assessed by Large Eddy simulations and IHOP_2002 observations. Quart. J. R. Meteorol. Soc. **131**, 2665–2693 (2006)

Davis, K.J., Lenschow, D.H., Oncley, S.P., Kiemle, C., Ehret, G., Giez, A., Mann, J.: Role of entrainment in surface-atmosphere interactions over the boreal forest. BOREAS special issue. J. Geophys. Res. **102**(D24), 29219–29230 (1997)

Davis, K.J., Gamage, N., Hagelberg, C., Kiemle, C., Lenschow, D.H., Sullivan, P.P.: An objective method for deriving atmospheric structure from airborne lidar observations. J. Atmos. Oceanic Technol. **17**, 1455–1468 (2000)

Ehret, G., Kiemle, C., Renger, W., Simmet, G.: Airborne remote sensing of tropospheric water vapor with a near-infrared differential absorption lidar system. Appl. Optics **32**, 4534–4551 (1993). doi:10.1364/AO.32.004534

Ehret, G., Hoinka, K.P., Stein, J., Fix, A., Kiemle, C., Poberaj, G.: Low stratospheric water vapor measured by an airborne DIAL. J. Geophys. Res. **104**(D24), 31351–31359 (1999)

Flentje, H., Dörnbrack, A., Ehret, G., Fix, A., Kiemle, C., Poberaj, G., Wirth, M.: Water vapour heterogeneity related to tropopause folds over the North Atlantic revealed by airborne water vapour differential absorption lidar. J. Geophys. Res. **110**, D03115 (2005). doi: 10.1029/2004JD004957

Fueglistaler, S., Dessler,A.E., Dunkerton, T.J., Folkins, I., Fu, Q., Mote,P.W.: Tropical tropopause layer. Rev. Geophys. **47**, RG1004 (2009). doi:10.1029/2008RG000267

Gettelman, A., Hoor, P., Pan, L.L., Randel, W.J., Hegglin, M.I., Birner, T.: The extratropical upper troposphere and lower stratosphere. Rev. Geophys. **49**, RG3003 (2011). doi: 10.1029/2011RG000355

Gierens, K., Schumann, U., Helten, M., Smit, H., Wang, P.H.: Ice-supersaturated regions and sub visible cirrus in the northern midlatitude upper troposphere. J. Geophys. Res. **105**, 22743–22754 (2000)

Giez, A., Ehret, G., Schwiesow, R.L., Davis, K.J., Lenschow, D.H.: Water vapor flux measurements from groundbased vertically pointed water vapor differential absorption and Doppler lidars. J. Atmos. Oceanic Technol. **16**, 237–250 (1999)

Grund, C.J., Banta, M., George, J., Howell, J., Post, M., Richter, R., Weickmann, A.: High resolution Doppler lidar for boundary layer research. J. Atmos. Oceanic Technol. **18**, 376–393 (2001)

Harnisch, F., Weissmann, M., Cardinali, C., Wirth, M.: Experimental assimilation of DIAL water vapour observations in the ECMWF global model. Q. J. R. Meteorol. Soc. **137**, 1532–1546 (2011)

Hauf, T., Schulte, P., Alheit, R., Schlager, H.: Rapid vertical trace gas transport by an isolated midlatitude thunderstorm. J. Geophys. Res. **100**(D11), 22957–22970 (1995). doi:10.1029/95JD02324

Keil, C., Röpnack, A., Craig, G.C., Schumann, U.: Sensitivity of quantitative precipitation forecast to height dependent changes in humidity. Geophys. Res. Lett. **35**, L09812 (2008). doi:10.1029/2008GL033657

Kiemle, C., Kästner, M., Ehret, G.: The convective boundary layer structure from lidar and radiosonde measurements during the EFEDA′91 campaign. J. Atmos. Oceanic Technol. **12**, 771–782 (1995)

Kiemle, C., Ehret, G., Giez, A., Davis, K.J., Lenschow, D.H., Oncley, S.P.: Estimation of boundary-layer humidity fluxes and statistics from airborne differential absorption lidar (DIAL). BOREAS special issue. J. Geophys. Res. **102**(D24), 29189–29203 (1997)

Kiemle, C., Brewer, W.A., Ehret, G., Hardesty, R.M., Fix, A., Senff, C., Wirth, M., Poberaj, G., LeMone, M.A.: Latent heat flux profiles from collocated airborne water vapor and wind lidars during IHOP_2002. J. Atmos. Oceanic Technol. **24**, 627–639 (2007)

Kiemle, C., Wirth, M., Fix, A., Ehret, G., Schumann, U., Gardiner, T., Schiller, C., Sitnikov, N., Stiller, G.: First airborne water vapor lidar measurements in the tropical upper troposphere and mid-latitudes lower stratosphere: accuracy evaluation and intercomparisons with other instruments. Atmos. Chem. Phys. **8**, 5245–5261 (2008)

Kiemle, C., Wirth, M., Fix, A., Rahm, S., Corsmeier, U., Di Girolamo, P.: Latent heat flux measurements over complex terrain by airborne water vapour and wind lidars. Q. J. R. Meteorol. Soc. **137**, 190–203 (2011)

Kley, D., Russell, J.M., Phillips, C.: SPARC assessment of upper tropospheric and stratospheric water vapour, WCRP report No. 113, WMO/TD Report No. 1043, World Climate Research Programme, Geneva, 325 pp (2000)

Möller, D., Feichter, J., Herrmann, H.: Von Wolken, Nebel und Niederschlag. In: Zellner, R. (ed.) Chemie über den Wolken, p. 237. Wiley, Weinheim (2011)

Poberaj, G., Fix, A., Assion, A., Wirth, M., Kiemle, C., Ehret, G.: Airborne all-solid-state DIAL for water vapour measurements in the tropopause region: system description and assessment of accuracy. Special issue "trends in laser sources, spectroscopic techniques and their applications to trace gas detection". Appl. Phys. B **75**, 165–172 (2002)

Randall, D.A., Wood, R.A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., et al.: Climate models and their evaluation. In: Solomon, S., et al. (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 (2007)

Schäfler, A., Dörnbrack, A., Kiemle, C., Rahm, S., Wirth, M.: Tropospheric water vapor transport as determined from airborne lidar measurements. J. Atmos. Oceanic Technol. **27**, 2017–2030 (2010)

Schäfler, A., Dörnbrack, A., Wernli, H., Kiemle, C., Pfahl, S.: Airborne lidar observations in the inflow region of a warm conveyor belt. Q. J. R. Meteorol. Soc. **137**, 1257–1272 (2011)

Schmidt, H., Schumann, U.: Coherent structure of the convective boundary layer derived from large-eddy simulations. J. Fluid Mech. **200**, 511–562 (1989)

Schumann, U., Weinzierl, B., Reitebuch, O., Schlager, H., Minikin, A., Forster, C., Baumann, R., Sailer, T., Graf, K., Mannstein, H., et al.: Airborne observations of the Eyjafjalla volcano ash cloud over Europe during air space closure in April and May 2010. Atmos. Chem. Phys. **11**, 2245–2279 (2011). doi:10.5194/acp-11-2245-2011

Tollerud, E.I., Caracena, F., Koch, S.E., Jamison, B.D., Hardesty, R.M., McCarty, B.J., Kiemle, C., Collander, R.S., Bartels, D.L., Albers, S., et al.: Mesoscale moisture transport by the low-level jet during the IHOP field experiment. Mon. Weather Rev. **136**, 3781–3795 (2008)

Trenberth, K.E., Dai, A., Rasmussen, R.M., Parsons, D.B.: The changing character of precipitation. Bull. Am. Meteorol. Soc. **84**, 1205–1217 (2003)

Trenberth, K.E., Jones, P.D., Ambenje, P., Bojariu, R., Easterling, D., Klein Tank, A., Parker, D., Rahimzadeh, F., Renwick, J.A., Rusticucci, M., et al.: Observations: surface and atmospheric climate change. In: Solomon, S., et al. (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 (2007)

Voigt, C., Schumann, U., Jurkat, T., Schäuble, D., Schlager, H., Petzold, A., Gayet, J.-F., Krämer, M., Schneider, J., Borrmann, S., et al.: In situ observations of young contrails—overview and selected case studies from the CONCERT campaign. Atmos. Chem. Phys. **10**, 9039–9056 (2010). doi:10.5194/acp-10-9039-2010

Wagner, J.S., Gohm, A., Dörnbrack, A., Schäfler, A.: The mesoscale structure of a polar low: airborne lidar measurements and simulations. Q. J. R. Meteorol. Soc. **137**, 1516–1531 (2011)

Weckwerth, T.M., Parsons, D.B., Koch, S.E., Moore, J.A., Lemone, M.A., Demoz, B.R., Flamant, C., Geerts, B., Wang, J., Feltz, W.: An overview of the international H2O project (IHOP_2002) and some preliminary highlights. Bull. Am. Meteorol. Soc. **85**, 253–277 (2004)

Weissmann, M., Langland, R.H., Cardinali, C., Rahm, S.: Influence of airborne Doppler wind lidar profiles near Typhoon Sinlaku on ECMWF and NOGAPS forecasts. Q. J. R. Meteorol. Soc. **138**, 118–130 (2012)

Wernli, H., Davies, H.C.: A Lagrangian-based analysis of extratropical cyclones. I: the method and some applications. Quart. J. Roy. Meteor. Soc. **123**, 467–489 (1997). doi:10.1002/qj.49712353811

Wirth, M., Fix, A., Mahnke, P., Schwarzer, H., Schrandt, F., Ehret, G.: The airborne multi-wavelength water vapor differential absorption lidar WALES: system design and performance. Appl. Phys. B **96**, 201–213 (2009). doi:10.1007/s00340-009-3365-7

Long-Range Transport of Air Pollutants

12

Anke Roiger, Heidi Huntrieser and Hans Schlager

Abstract

Air pollutants may be transported over several 1 000 km ranges, across continents, oceans (intercontinental), hemispheres, or even globally, depending on lifetime and other properties. Transported emissions include primary pollutants emitted from industry and vehicles (e.g., nitric oxide) and natural pollutants released by forest fires (e.g., soot). Also secondary pollutants (e.g., ozone) that are formed near the source region undergo transport. Recent advances in understanding atmospheric transport pathways, source-receptor relationships, and transformation of pollutants during transport are described.

12.1 Introduction

Early concerns about the harmful effects of imported air pollution resulted from the acidification of surface water in Scandinavia by acid rain in the late 1960s. There are many other examples where the occurrence of long-range transport becomes obvious, for example yellow or red colored snow in the Alps due to Saharan dust transported from Africa, or Arctic haze, a phenomenon recurring each winter/spring

A. Roiger (✉) · H. Huntrieser · H. Schlager
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Anke-Elisabeth.Roiger@dlr.de

H. Huntrieser
e-mail: Heidi.Huntrieser@dlr.de

H. Schlager
e-mail: Hans.Schlager@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_12,

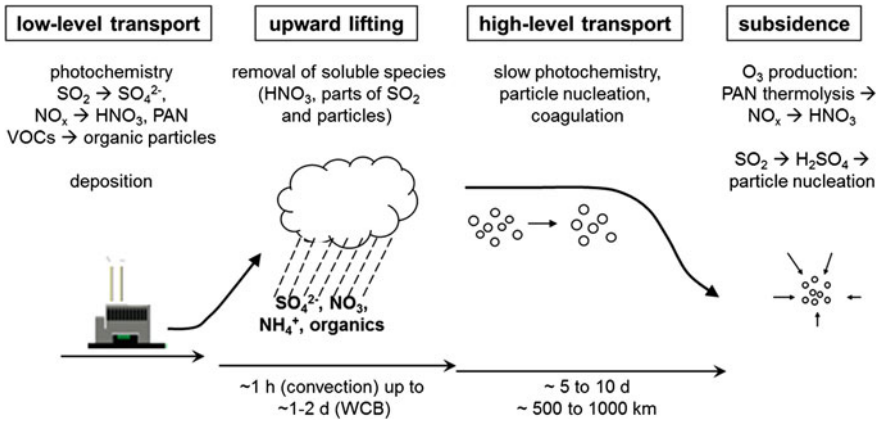

Fig. 12.1 Schematic of processes and time scales during long-range transport of pollutants [adapted from Brock et al. (2004)]. During transport at low levels, chemical and deposition processing dominates the evolution of all pollutants, while during uplift within a warm conveyor belt (WCB) or by deep convection water-soluble gases will be partly removed. At higher altitudes, chemical processing may take place but at slower rates than in the boundary layer. Also, particle formation and coagulation may occur. Finally, subsidence of the pollution plumes to the lower troposphere with higher temperatures can activate further transformation processes, such as PAN ($CH_3C(O)O_2NO_2$) decomposition, ozone production, SO_2 oxidation to H_2SO_4 and subsequent particle formation. Typical time scales involved in long-range transport include about 1 h and 1 day for the vertical uplift to the upper troposphere by deep convection and WCB, respectively, and 5–10 days for intercontinental transport

which can be attributed to imported pollution from Eurasia (Law and Stohl 2007). A recent striking example is the ash cloud from the eruption of Eyjafjallajökull volcano on Iceland, which was transported across large parts of Europe in spring 2010 and resulted in severe disruption of air traffic (Schumann et al. 2011).

The principle links between emissions, atmospheric transport and processing of pollutants are depicted in Fig. 12.1. A precondition for observing long-range transport events are pollutant concentrations which are significantly enhanced relative to atmospheric background conditions. This can be caused by high release rates of the pollutants, an accumulation of emissions during stagnant meteorological conditions near the source region, and/or weak mixing with clean background air during transport. Primary emissions may already be transformed to secondary pollutants near the source and during transport at low levels in the boundary layer. Efficient pollution transport on synoptic and intercontinental scales, however, requires further conditions. The polluted air mass needs to be lifted up from the boundary layer to the upper troposphere by deep convection or frontal systems. Here, atmospheric conditions are favorable for the transport of pollution plumes over large distances in several respects. In the upper troposphere at mid-latitudes, strong winds prevail, in particular near the jet stream, causing fast horizontal advection of the pollution plume. The atmosphere is stably stratified in the vicinity of the tropopause, inducing weak vertical mixing of the pollution

plume. In fact, the vertical diffusivity in the upper troposphere is several orders of magnitude smaller than the horizontal diffusivity (Schumann et al. 1995; Pisso et al. 2009). Therefore, the dilution of pollution plumes in general is weak in the upper troposphere and the mixing of the plumes with background air is mainly dependent on the vertical wind shear and dispersion in the horizontal flow. In addition, the lifetime of many pollutants is longer in the upper troposphere compared to the boundary layer because loss processes like chemical oxidation, thermal decomposition and wet removal are less effective there.

Long-range transport of air pollution can affect the receptor region in various ways including impacts on human health, damage to the ecosystem, and changes in regional climate (Monks et al. 2009). However, understanding the processes involved in long-range transport processes still includes many open questions:

- Are the mass fluxes of pollutants into the upper troposphere larger because of deep convection or synoptic-scale uplift?
- How efficiently are soluble trace compounds removed during vertical uplift by convection and cyclones?
- Are pollutants mixed into the boundary layer in the receptor region after long-range transport and how do they impact chemical composition?
- Can pollutants from Southeast Asia, the major present emission region worldwide, impact Europe?
- Which chemical and aerosol transformations occur during long-range transport along the different pathways and are there interactions among these processes?

Using an integrated analysis of aircraft measurements, trajectory model simulations and satellite observations, DLR's Institute of Atmospheric Physics (IPA) has been involved in many projects in recent years to study the long-range transport of air pollutants. In the following, advances in understanding long-range transport are presented with a focus on investigations by IPA and partners.

12.2 Results on the Long-Range Transport of Air Pollution

Research on the long-range transport of air pollutants started at IPA in the mid-1980s with airborne investigations on the import of typical primary air pollutants like sulfur dioxide (SO_2) and nitrogen oxides (NO_x, the sum of NO and NO_2 species) to West Germany from eastern European countries, e.g., in the project "Measurements of Transboundary Mass Fluxes of Air Pollutants over the Federal Republic of Germany" (Klemm et al. 1995). Thereafter, studies like MEMOSA and POLUMET focused on the transport of ozone precursor species (NO_x, volatile organic compounds) into the Alpine region from major anthropogenic emission source regions in Europe and on the photochemical production of ozone (Krautstrunk et al. 2000). After the unification of West and East Germany in 1990, the transport, mixing and chemistry of exhaust plumes from industrial regions of the former DDR were investigated in the SANA project (Klemm et al. 1995; Krautstrunk et al. 1995; Krautstrunk et al. 2000). All these

Table 12.1 DLR-IPA measurement activities focused on long-range transport of pollution

Field campaign	Main objective	Aircraft involved	References
MEMOSA (1990)	Measurement and modeling of air pollution transport in the Alpine region	DLR Queen Air	Krautstrunk et al. (2000)
SANA (1990/1993)	Origin and distribution of air pollutants over eastern Germany	DLR DO228	Klemm et al. (1995); Krautstrunk et al. (1995); Krautstrunk et al. (2000)
POLINAT (1997)	Aircraft emissions in the North Atlantic	DLR Falcon	Arnold et al. (1997)
EXPORT (2000)	Export pathways of European pollutants	DLR Falcon	Schlager (2003)
CONTRACE (2001)	Influence of convective transport mechanism on the European trace gas budget	DLR Falcon	Stohl (2003a); Huntrieser et al. (2005)
MINOS (2001)	Pollutant transport into the Mediterranean boundary layer	DLR Falcon	Lelieveld et al. (2002); Fischer et al. (2003)
INTEX-A (2004)	Intercontinental pollutant transport in the northern hemisphere—phase A	Nine aircraft, including DLR Falcon, NOAA P3, NASA DC8, BAe-146	Petzold et al. (2007); Real et al. (2007); Real et al. (2008)
INTEX-B (2006)	Intercontinental pollutant transport in the northern hemisphere—phase B	Five aircraft, including DLR Falcon and NASA DC8	Stohl et al. (2007); Fiedler et al. (2009a); Fiedler et al. (2009b)
AMMA (2006)	African Monsoon Multidisciplinary Analyses	Four aircraft, including DLR Falcon, Geophysica, BAe-146	Fiedler et al. (2011); Huntrieser et al. (2011)
ASTAR (2007)	Springtime transport of pollutants into the Arctic	DLR Falcon, DO-228	Engvall et al. (2009)
POLARCAT-GRACE (2008)	Summer-time transport of pollutants into the Arctic	DLR Falcon, ATR-42	Roiger (2011); Roiger et al. (2011)

early investigations concentrated on pollution transport in the boundary layer and lower free troposphere.

One of the first unambiguous airborne measurements of the intercontinental transport of distinct pollution plumes in the upper troposphere was performed during POLINAT (Arnold et al. 1997). A layer with very high SO_2 concentrations was detected in the upper troposphere over the eastern North Atlantic. This pollution plume originated in the Ohio valley south of the Great Lakes in the USA, an industrial region with very high anthropogenic SO_2 emissions. The uplift of the air pollution occurred within a cyclone located off the coast of New England 5 days prior to the Falcon measurements. Table 12.1 summarizes airborne experiments

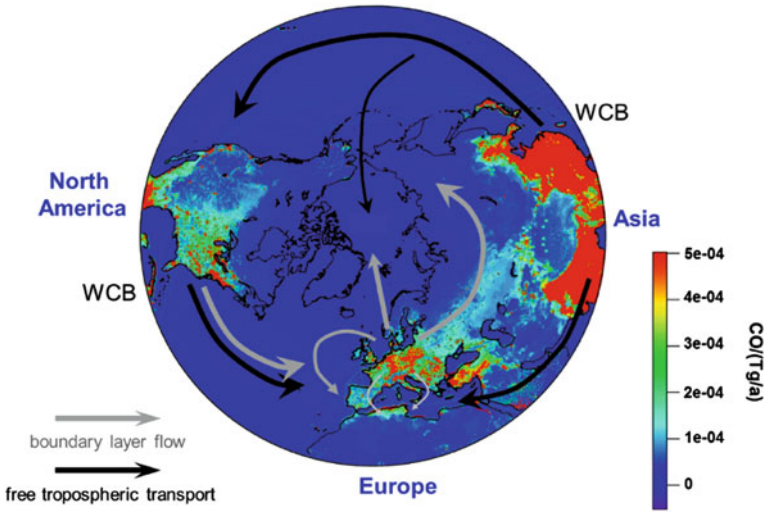

Fig. 12.2 Distribution of anthropogenic surface emissions of CO in the northern hemisphere. Superimposed are major transport pathways of intercontinental pollution transport within the boundary layer (*grey arrows*) and in the free troposphere (*black arrows*). CO emission data taken from http://edgar.jrc.ec.europa.eu/index.php

with IPA coordination and participation on the long-range transport of air pollutants. Selected results from these projects are presented in the following sections.

12.2.1 Transport Pathways and Timescales

The geographical distribution of anthropogenic surface emissions in the northern hemisphere is illustrated in Fig. 12.2. There are several regions with particularly high emissions: Western and central Europe, the North American east coast, and Southeast Asia. The pollutants released in these source regions are transported along several major intercontinental pathways. In general, emissions from Asia are mainly transported to North America within the westerly flow; North American emissions impact Europe, and pollutants from Europe are advected in part to the Arctic, Eastern Europe and the Mediterranean. IPA investigations mainly focused on pollution export and import from and to Europe. The eastward outflow of European pollutants was studied during the EXPORT campaign performed in summer 2000. These pollutants were lifted both by convective and frontal uplift and probed by the Falcon during their transport eastwards to Poland (Schlager et al. 2003). During the MINOS experiment in Crete in summer 2001, the pollution export from Europe into the Mediterranean area was investigated. The Falcon measured high CO values along with elevated ozone over large areas in the boundary layer over Crete, which could be attributed mainly to inflow from Western and Eastern Europe (Lelieveld et al. 2002).

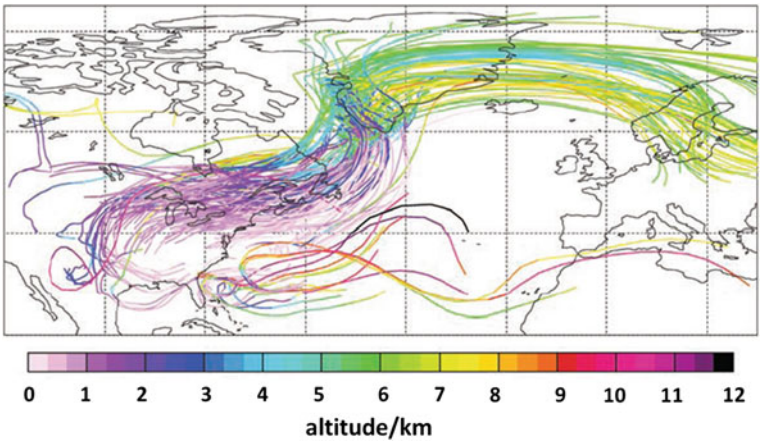

Fig. 12.3 Five-day forward trajectories starting from a regular grid (120°W–60°W, 30°N–60°N) with 1° resolution at an altitude of 500 m above ground level on 12 UTC 14 November 2001. Only trajectories that ascend >5 000 m within 4 days are plotted (WCB trajectories). The trajectories are color-coded according to their actual altitude. The uplifted North American pollution was encountered by the Falcon on 19 November 2001 over Scandinavia. Adapted from Stohl (2003a)

In contrast, the import of pollution to Europe is dominated by emissions originating in North America. Slow low-level transport across the Atlantic could be observed during airborne measurements in summer. For example, during the INTEX campaign the Falcon sampled polluted air masses over the English Channel that originated from the North American east coast. This pollution plume was transported across the North Atlantic at altitudes below ∼2 km (Real et al. 2008).

Fast intercontinental transport is initiated by uplift of pollutants within WCBs, which was studied during the CONTRACE project in 2001. WCBs typically evolve at the beginning of the North Atlantic and North Pacific storm tracks, i.e., on the western sides of the oceans (Eckhardt et al. 2004). For this reason, emissions released at the strong industrialized eastern seaboards of North America and Asia are especially subject to intercontinental transport.

During CONTRACE, for the first time forecasts from a Lagrangian atmospheric transport model (FLEXPART) and the 3D chemistry-transport model MATCH-MPIC were used to predict the North American pollution events transported to Europe, and to direct the Falcon very precisely into these pollution layers above Europe (Lawrence et al. 2003; Stohl et al. 2003a). Figures 12.3 and 12.4 illustrate a typical long-range transport event observed during the CONTRACE field experiment. During the flights on 19 November 2001, North American pollution was sampled by the Falcon over Scandinavia, which was transported in the mid-troposphere to Europe within 4–5 days (Huntrieser et al. 2005). Figure 12.3 shows 5 day forward trajectories started from a regular grid (120–60°W, 30–60°N) at an

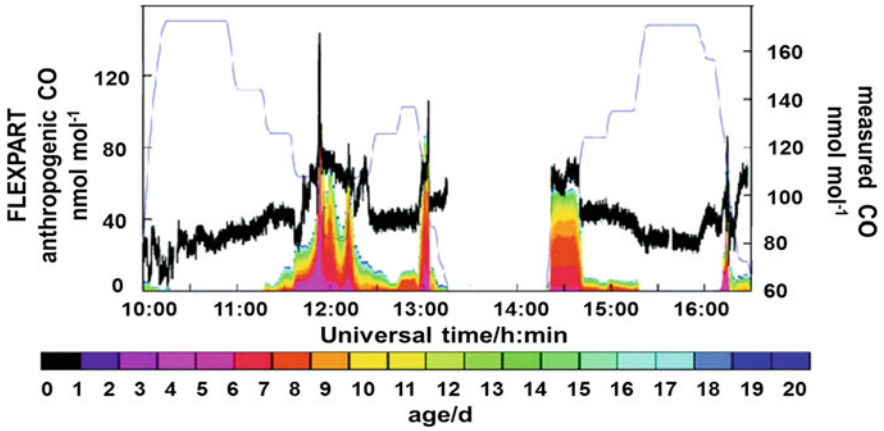

Fig. 12.4 Time series of North American anthropogenic CO tracer mixing ratios obtained from FLEXPART backward model simulations for two flights on 19 November 2001 (*left axis*). Color codes indicate the age of the modeled pollution tracers. Also given is CO measured by the Falcon (*black line*, *right axis*) and aircraft altitude (*dotted line*, relative units, maximum 9400 m). FLEXPART tracers represent only the mixing ratios emitted by pollution sources (and therefore are referred to as "anthropogenic CO"). For a comparison with in situ-measured CO values, mixing ratios of the unpolluted background atmosphere thus have to be added. Adapted from Stohl (2003a)

altitude of 500 m over North America (Stohl et al. 2003a). Only trajectories ascending more than 5 000 m within 4 days are depicted (WCB trajectories). First, the air masses were advected at low altitudes over industrialized regions in eastern North America before the polluted air was uplifted from the boundary layer. Due to the cyclonic motion, air lifted within WCBs generally first ascends polewards until it reaches the jet stream level where it is forced to turn towards the east. During the multiday transport over the Atlantic, the pollution remained in the mid-troposphere.

Figure 12.4 shows carbon monoxide (CO) mixing ratios measured during the corresponding Falcon flights on 19 November 2001 from Oberpfaffenhofen to Stockholm and return. CO originates from incomplete combustion processes and is an excellent pollution tracer on intercontinental scales due to its relatively long lifetime of several weeks in the free troposphere. Also given is the FLEXPART North American anthropogenic CO tracer mixing ratio, which represents only the mixing ratios emitted by pollution sources. For this reason, CO mixing ratios of the unpolluted background atmosphere have to be added for a comparison with in situ measured CO values. The pollution layer was penetrated several times, first at lower levels while obtaining a profile close to Oslo ($\sim$12:00 UTC), and again during the descent close to Stockholm ($\sim$13:00 UTC). After return from Stockholm, the same part of the North American pollution was encountered again ($\sim$14:30 UTC). The plume locations are quite well reproduced by the FLEXPART "anthropogenic CO" tracer for North America, while the contribution from Asian and European

anthropogenic CO tracers was negligible (not shown). Forward calculations initiated along the Falcon flight path indicated that one branch of this pollution plume was advected towards the south and descended to ground levels over the Alpine region. Enhanced levels of CO and ozone were detected at the Zugspitze mountain site and at the Arosa Alpine site station in Switzerland, where the highest daily ozone means of November 2001 were observed during this event (Huntrieser et al. 2005).

Another CONTRACE case study showed that under certain meteorological circumstances transoceanic transport can be much faster. Stohl et al. (2003b) reports on an extraordinary fast pollution transport within a quickly developing cyclone carrying pollution across the North Atlantic within only about one day.

Within the westerly flows, Asian pollution may even reach Europe, a pathway which for the first time was directly observed during the INTEX-B campaign. The Falcon twice intercepted Asian pollution plumes which arrived over Europe after a travelling time of ~6 and 10 days, respectively (Stohl et al. 2007; Fiedler et al. 2009b).

More recently, IPA investigated pollution import into the European sector of the Arctic. Climate change has been observed to proceed fastest at high latitudes and pollutants transported into the Arctic may influence the radiative budget (Law and Stohl 2007). Low-level transport from Eurasia into the Arctic is a well-known phenomenon. However, Falcon measurements during the Arctic field campaigns ASTAR and POLARCAT-GRACE showed that regular transport across the pole occurs also at higher altitudes, thereby importing Asian anthropogenic pollution and Siberian biomass burning emissions into the European sector of the Arctic (Roiger 2011; Sodemann et al. 2011). This represents an alternate pathway for Asian pollution transport to Europe.

An exceptional case of transport of industrial emissions into the Arctic is presented in Fig. 12.5. Depicted is the time series of several trace gases measured during the POLARCAT-GRACE research flight on 10 July 2008, when the Falcon probed an Asian pollution plume over northern Greenland at an altitude of ~11.3 km, well above the tropopause. The analysis revealed that the pollutants uplifted within a strong WCB reached the lowermost stratosphere already within a few days after emission (Roiger et al. 2011). This transport mechanism was already suggested by Lagrangian climatological studies (Stohl et al. 2001; Wernli and Bourqui 2002), but was experimentally proved in this case study.

The anthropogenic pollution plume with a horizontal extension of ~80 km was encountered twice during the northernmost part of the flight at nearly 80°N (~16:08 and 16:37 UTC), at potential vorticity (PV) levels higher than 5 PVU (PV units). When the Falcon entered the Asian plume, CO increased to levels of ~140 nmol/mol (Fig. 12.5b, grey line), while typical CO mixing ratios at this altitude were observed to amount about 40 nmol/mol during the POLARCAT-GRACE campaign. The CO enhancement was accompanied by an increase of the tropospheric smog pollutant peroxyacetyl nitrate (PAN) and CO_2, while trace gases with generally higher mixing ratios in the stratosphere (O_3, NO and NO_y) decreased inside the polluted air mass. This suggests that the Falcon sampled an air mass of both tropospheric and stratospheric origin. According to the FLEXPART

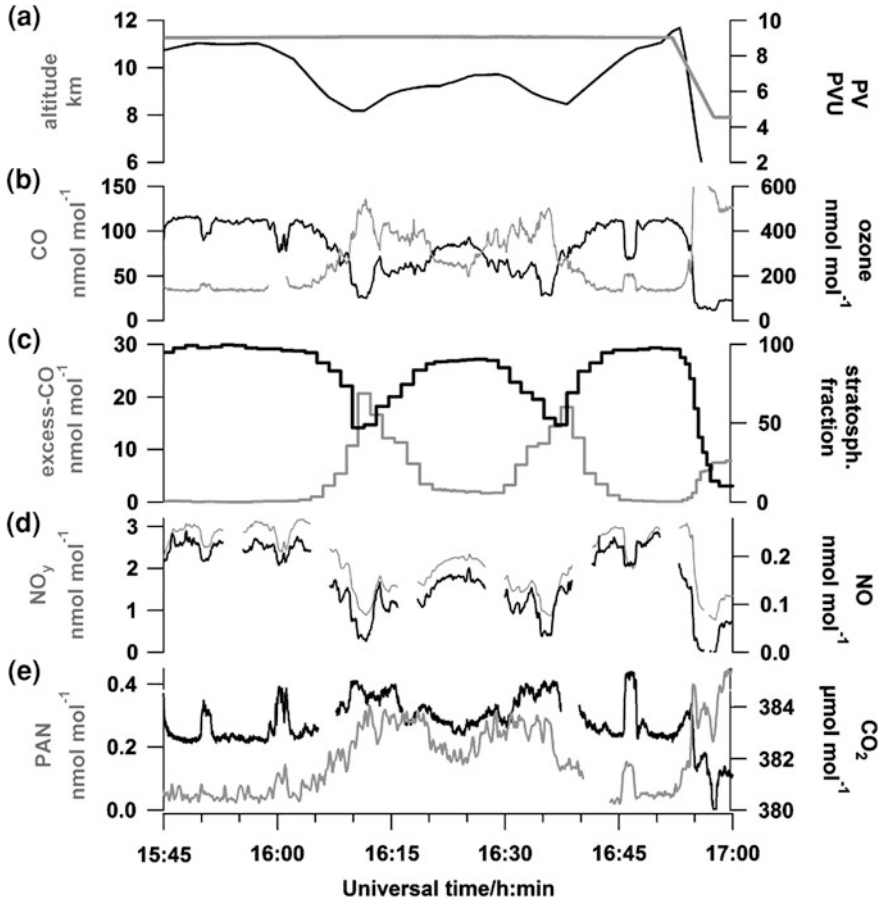

Fig. 12.5 Time series of trace gas mixing ratios measured from the DLR Falcon during the interception of an Asian pollution plume in the polar lowermost stratosphere. **a** Flight altitude (*grey*) and PV (*black*), **b** CO (*grey*) and ozone (*black*), **c** FLEXPART modeled tracers excess-CO (*grey*) and stratospheric fraction (*black*), **d** NO (*black*) and NO_y (*grey*), **e** PAN (*grey*) and CO_2 (*black*). Adapted from Roiger et al. (2011)

model, the tropospheric part contained anthropogenic pollution from East Asia (Fig. 12.5c, grey line). The observed decrease in the FLEXPART stratospheric fraction (Fig. 12.5c, black line) shows that also the mixing between the tropospheric pollution and the stratospheric air is well reproduced by the model.

A detailed backward trajectory analysis of this case revealed that the pollution originated mainly from the North China Plain, one of the most densely populated regions in China. It was then uplifted within a strong WCB associated with a low pressure system over northern Russia. The main part of the lifted pollution was advected at upper levels within the westerlies towards North America, similar to the pathways indicated in Figs. 12.2 and 12.3. However, a small part of the

polluted air was transported directly across the North Pole into the Arctic. Forward trajectory calculations initialized at the location of observation suggest that the pollution was completely mixed into the lowermost stratosphere. This possibly represents an important pathway of pollution import into the lower stratosphere.

It is worth noting that emissions can also be uplifted by deep convection. Here, the export of emissions from the boundary layer into the upper troposphere is much faster (typically within $\sim$1 h), but on smaller scales (typical cross section of a thunderstorm $\sim$30 km). Falcon measurements during MINOS in the outflow of a deep convective event over southern Italy in August 2001 showed that boundary layer air can be directly injected into the lowermost stratosphere (Fischer et al. 2003).

12.2.2 Chemical Processes

The composition of pollution plumes may change during long-range transport due to the production and destruction of species by photochemical reactions and by mixing with background air. As illustrated in Fig. 12.1, different kinds of transformation processes are active within different phases of a long-range transport event, depending on factors like meteorological conditions and type of emissions. The ideal approach to study transformation processes and the role of mixing during long-range transport of a pollution plume is to perform so-called Lagrangian experiments. Here, the same polluted air mass is probed several times along its pathway by one or more research aircraft. During the INTEX campaign in 2004, for the first time several Lagrangian experiments were conducted on intercontinental scales using multiple aircraft, including the DLR Falcon (Methven et al.2006).

One of the most important questions concerning long-range transport is to what extent the transported pollution will affect the ozone concentration in the receptor region. Thermal decomposition of PAN during subsidence into the warmer lower troposphere may initiate photochemical ozone production due to release of the ozone precursor NO_x (Hudmann et al. 2004). This effect was investigated during the INTEX campaign in a Lagrangian experiment. A pollution plume was intercepted three times during its transport over the North Atlantic within 5 days using three different research aircraft, the NASA DC8, NERC BAe146, and DLR Falcon. Figure 12.6 shows the flight tracks of three different aircraft which subsequently probed a biomass burning plume originating in northern Canada and Alaska. The DC8 probed the forest fire plume first over Newfoundland, before the BAe-146 sampled it again 2 days later flying out of the Azores. Again 3 days later, the Falcon intercepted the plume over France. Also given in Fig. 12.6 are the backward trajectories starting along the matching segments of each flight track. The matching segments are the flight sections where the same air mass was sampled by the three aircraft according to an analysis of tracers (Methven et al. 2006).

Ozone changes observed between the different interceptions of the biomass burning plume were analyzed using the photochemical trajectory model CiTTyCAT (Cambridge Tropospheric Trajectory model of Chemistry and Transport).

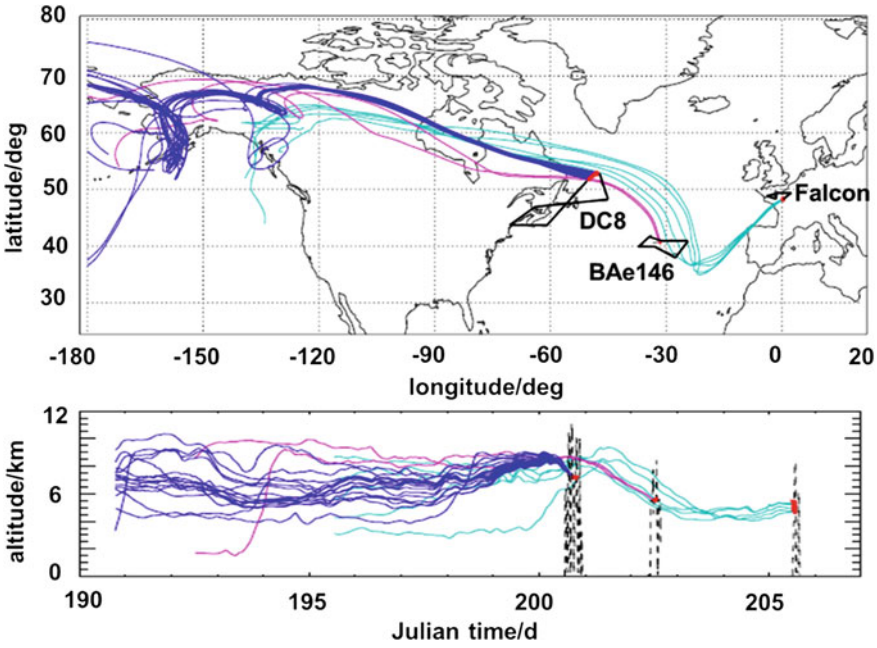

Fig. 12.6 Flight tracks of the DC8 (18 July), BAe-146 (20 July), Falcon (23 July) (*black lines*). Lagrangian matches are highlighted in *red*. Backward trajectories were initialized along these matching segments (*light blue* for the Falcon, *pink* for the BAe-146, and *dark blue* for the DC8). Adapted from Real et al. (2007)

CiTTyCAT is a chemical box model, which follows a 3D Lagrangian air parcel trajectory calculated from large-scale winds (Wild et al. 1996). The model was initialized with the initial plume composition measured by the DC8 and run along a chosen trajectory representative of the plume pathway. Calculated changes in the plume composition were compared with the measurements performed in the matching flight segments of the BAe-146 and Falcon flights, 2 and 5 days after the DC8 observations.

Figure 12.7 shows the temporal evolution of selected trace species along the plume trajectory as simulated by CiTTyCAT. The DC8 probed the forest fire plume 3–4 days after its emission. At this time, CO values in the plume were still high with a mean mixing ratio of about 450 nmol/mol (Fig. 12.7a). The ozone levels were relatively low ($\approx$62 nmol/mol, Fig. 12.7b) compared to values outside of the plume. The plume concentrations of the reactive nitrogen compound PAN were about 3 nmol/mol (Fig. 12.7c). By the time the BAe-146 sampled the plume 2 days later, CO had decreased by about 30 nmol/mol, and when it was sampled again 3 days later by the DLR Falcon southwest of Paris, CO had decreased to half the levels measured by the DC8. In contrast, ozone levels had increased by 17 nmol/mol over 5 days between the different samplings. Figure 12.7d shows

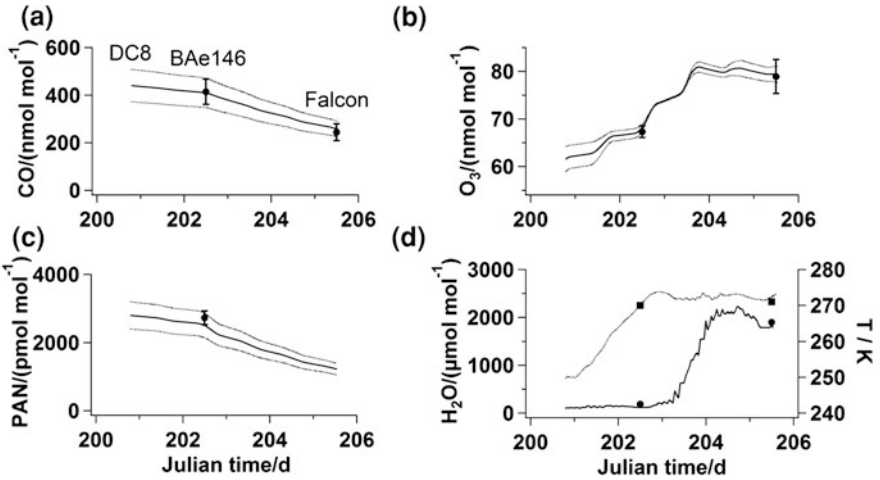

Fig. 12.7 Time evolution of the volume mixing ratio of selected trace gases as derived from the CiTTyCAT simulations. **a** CO, **b** ozone, **c** PAN, **d** water vapor (*solid line*) and temperature (*dotted line*). The simulations were initialized with the trace gas mixing ratios measured by the NASA DC8 (mean ± std), and then compared with the downwind measurements by the BAe-146 and the DLR Falcon. Based on Real et al. (2007)

that the polluted air mass warmed up during subsidence towards Europe, reaching temperatures of more than 270 K. As a consequence, PAN thermally decomposed Fig. 12.7c, thereby releasing NO_x, which act as catalysts for the photochemical production of ozone.

12.2.3 Aerosol Processes

IPA research also contributed much to the present knowledge on aerosol formation and aging processes occurring during long-range transport. Different kind of aerosol transformation processes may occur during transport (see also Fig. 12.1). For example, during the uplift of a pollution plume within a WCB or within deep convection, hygroscopic particles may become incorporated in cloud droplets. Forest fire emissions may initially not experience wet removal. Strong heat development in the fires can directly inject plumes into the free troposphere. Petzold et al. (2007) found export efficiencies for black carbon (BC) of 70–90 % for the bulk of the investigated biomass burning plumes, based on Falcon measurements during the INTEX campaign. BC transported to Europe from North America in such plumes increased the aerosol absorption coefficient by about two orders of magnitude compared to the background level of the free troposphere in Europe.

Aerosol size distributions may change during long-range transport either due to new particle formation or coagulation processes. The most important aerosol

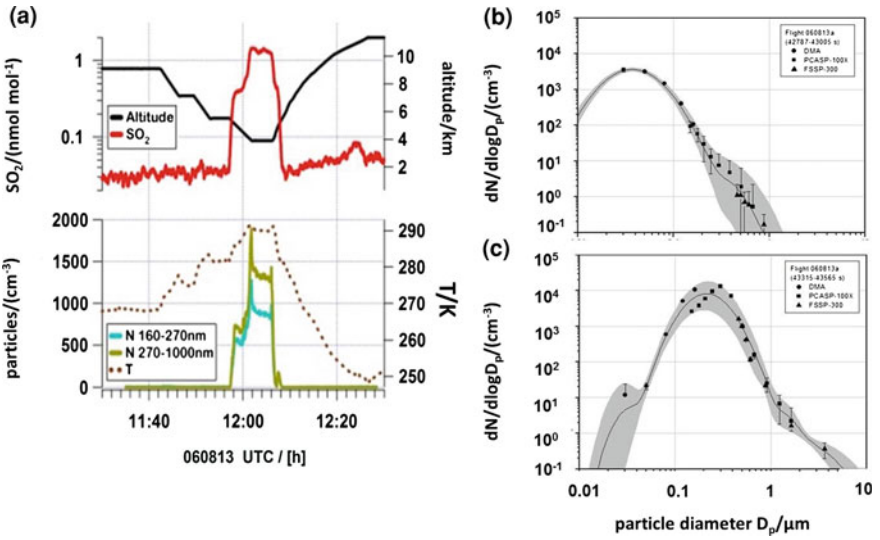

Fig. 12.8 **a** SO_2 mixing ratio, particle concentration, flight altitude, and temperature versus flight time, **b** particle size distribution above and **c** inside a biomass burning plume. Modified from Fiedler et al. (2011)

nucleation precursor is sulfuric acid, which is formed by the OH oxidation of primary emitted SO_2. Cloud processing during the air mass uplift removes only a part of initial SO_2 via chemical conversion to sulfate. For example, an Asian plume sampled by the Falcon over Europe still contained high SO_2 levels of about 900 pmol/mol (Fiedler et al. 2009a). The analysis showed that during its travel to Europe in the upper troposphere, OH-induced gas-phase conversion of SO_2 to sulfuric acid was active. This was followed by condensation and nucleation, leading to sulfuric acid aerosol formation and growth (Fiedler et al. 2009b). In contrast, several other Falcon measurements of aged pollution plumes indicate that no particle formation occurred during transport, either due to the absence of gaseous precursors such as SO_2 (Petzold et al. 2007), or by scavenging of produced sulfuric acid on larger particles also present in the pollution plumes (Fiedler et al. 2011). In this case, enhanced coagulation of smaller particles with larger particles will impact on the particle size distributions, as illustrated in Fig. 12.8.

Figure 12.8a shows time series for SO_2 and ambient number concentrations of aerosol particles as measured during a Falcon flight on 13 August 2006 over the Atlantic in the Gulf of Guinea. As the Falcon descended into the mid-troposphere, aged fire plumes originating from biomass burning in the African savannah belt south of the equator were encountered at $\sim$4 km. Both SO_2 and aerosol particle concentrations increased very steeply along with other measured trace gases (not shown). Figure 12.8b and c presents the corresponding aerosol particle size distributions measured above the plume, i.e., in background conditions (4 500 m), as well as inside the plume (3 900 m). The two particle size distributions differ

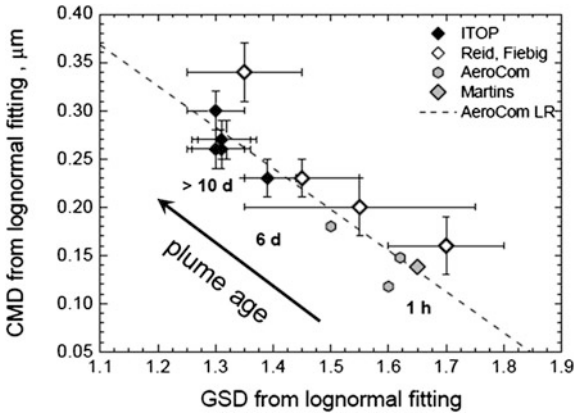

Fig. 12.9 Count median diameter (CMD) of the size distribution of boreal forest fire smoke aerosols as a function of the geometric standard deviation (GSD) of the distribution (Petzold et al. (2007) and references herein)

significantly. In the background case, the size distribution is dominated by smaller particles peaking at diameters around 40 nm. Inside the biomass burning plume, the particle size peaks at much larger diameters around 250 nm. This indicates the presence of many larger particles, most likely aged smoke particles. Smaller particles obviously were removed due to coagulation processes, while new particle formation has been suppressed probably due to scavenging of sulfuric acid (originating from OH induced conversion of SO_2). The net result is an increase of the peak diameter and a narrowing of the size distribution.

Petzold et al. (2007) showed that this modification of particle size distributions is active also for biomass burning plume ages beyond one week. Aerosol properties of aged forest fire plumes were systematically studied using the INTEX Falcon measurements and compared with earlier observations and model simulations of younger biomass burning plumes. Figure 12.9 summarizes the count median diameter (CMD, peak diameter of the size distribution) and the geometric standard deviation (GSD, width of the size distribution) of boreal forest fire aerosols of different ages. A linear relation between the CMD and the GSD is observed for fresh ($\sim$1 h) to aged (>1 week) wildfire aerosols.

12.3 Outlook

Over the past 20 years significant advancements have been made in understanding pollution pathways and transformation processes during long-range transport. In the future, it is expected that the impact of long-range transport of pollutants from North America and Europe will decrease due to efficient pollution control measures there. However, emissions of pollutants in the rapidly growing economies of Asia and in developing countries will further increase in the coming years. In addition, biomass burning emissions are also expected to increase in the future. The number of boreal forest fires has strongly increased in the past decade, especially in North America and Siberia, as a result of increasing temperatures and

dryness. In addition, climate change may modify the transport pathways of pollution due to changes in large-scale wind patterns. Another important scientific issue is the interaction of regional pollution and the climate system. The monsoon region including Asia, India and Indonesia is the most polluted area on Earth. Long-range transport associated with the monsoon circulation may provide an important pathway of pollution into the lower stratosphere with possible impacts on climate on global scales. Several field campaigns with the HALO and Geophysica research aircraft are planned to identify the transport pathways in the Asian monsoon and their importance in pollution transport into the stratosphere.

References

Arnold, F., Schneider, J., Gollinger, K., Schlager, H., Schulte, P., Hagen, D.E., Whitefield, P.D., van Velthofen, P.: Observation of upper tropospheric sulfur dioxide- and acetone-pollution: potential implications for hydroxy radical and aerosol formation. Geophys. Res. Lett. **24**, 1 (1997). doi:10.1029/96GL03693

Brock, C.A., Hudson, P.K., Lovejoy, E.R., Sullivan, A., Nowak, J.B., Huey, L.G., Cooper, O.R., Cziczo, D.J., de Gouw, J., Fehsenfeld, F.C., et al.: Particle characteristics following cloud-modified transport from Asia to North America. J. Geophys. Res. **109**, D23S26 (2004). doi:10.1029/2003jd004198

Eckhardt, S., Stohl, A., Wernli, H., James, P., Forster, C., Spichtinger, N.: A 15-year climatology of warm conveyor belts. J. Clim. **17**, (2004). doi:10.1175/1520-0442(2004)017

Engvall, A.-C., Ström, J., Tunved, P., Krejci, R., Schlager, H., Minikin, A.: The radiative effect of an aged, internally mixed Arctic aerosol originating from lower-latitude biomass burning. Tellus B **61**(4), 677–684 (2009). doi:10.1111/j.1600-0889.2009.00432.x

Fiedler, V., Arnold, F., Schlager, H., Dörnbrack, A., Pirjola, L., Stohl, A.: East Asian SO_2 pollution plume over Europe—Part 2: Evolution and potential impact. Atmos. Chem. Phys. **9**, 4729–4745 (2009a). doi:10.5194/acp-9-4717-2009

Fiedler, V., Nau, R., Ludmann, S., Arnold, F., Schlager, H., Stohl, A.: East Asian SO_2 pollution plume over Europe—Part 1: Airborne trace gas measurements and source identification by particle dispersion model simulations. Atmos. Chem. Phys. **9**, 4717–4728 (2009b). doi:10.5194/acp-9-4717-2009

Fiedler, V., Arnold, F., Ludmann, S., Minikin, A., Hamburger, T., Pirjola, L., Dörnbrack, A., Schlager, H.: African biomass burning plumes over the Atlantic: aircraft based measurements and implications for H_2SO_4 and HNO_3 mediated smoke particle activation. Atmos. Chem. Phys. **11**, 3211–3225 (2011). doi:10.5194/acp-11-3211-2011

Fischer, H., de Reus, M., Traub, M., Williams, J., Lelieveld, J., de Gouw, J., Warneke, C., Schlager, H., Minikin, A., Scheele, R., Siegmund, P.: Deep convective injection of boundary layer air into the lowermost stratosphere at midlatitudes. Atmos. Chem. Phys. **3**, 739–745 (2003). doi:10.5194/acp-3-739-2003

Hudman, R.C., Jacob, D.J., Cooper, O.R., Evans, M.J., Heald, C.L., Park, R.J., Fehsenfeld, F. Flocke, Holloway, J., Huebler, G. et al.: Ozone production in transpacific Asian pollution plumes and implications for ozone air quality in California. J. Geophys. Res., **108**, D23S210 (2004). doi: 10.1029/2004JD004974

Huntrieser, H., Heland, J., Schlager, H., Forster, C., Stohl, A., Aufmhoff, H., Arnold, F., Scheel, H.E., Campana, M., Gilge, S., et al.: Intercontinental air pollution transport from North America to Europe: experimental evidence from airborne measurements and surface observations. J. Geophys. Res. **110**, D01305 (2005). doi:10.1029/2004JD005045

Huntrieser, H., Schlager, H., Lichtenstern, M., Stock, P., Hamburger, T., Höller, H., Schmidt, K., Betz, H.-D., Ulanovsky, A., Ravegnani, F.: Mesoscale convective systems observed during

AMMA and their impact on the NO_x and O_3 budget over West Africa. Atmos. Chem. Phys. **11**, 2503–2536 (2011). doi:10.5194/acp-11-2503-2011

Klemm, O., Wehrhan, J., Schaller, E., Schlager, H., Krautstrunk, M.: Characterization of ozone plumes in eastern Germany. Geophys. Res. Lett. **22**(15), 2021–2024 (1995). doi:10.1029/95GL01322

Krautstrunk, M., Neumann-Hauf, G., Schlager, H., Klemm, O., Schaller, E.: Aircraft-based measurements of the three-dimensional distribution of gaseous pollutants over the highly industrialised area of Lausitz in eastern Germany. Air Pollution III. In: Observation and Simulation of Air Pollution—Results from SANA and EUMAC, Computational Mechanics vol. 4, pp. 69–77 (1995)

Krautstrunk, M., Neumann-Hauf, G., Schlager, H., Klemm, O., Beyrich, F., Corsmeier, U., Kalthoff, N., Kotzian, M.: An experimental study on the planetary boundary layer transport of air pollutants over East Germany. Atmos. Environ. **34**, 1247–1266 (2000). doi: 10.1016/s1352-2310(99)00124-7

Law, K.S., Stohl, A.: Arctic air pollution: origins and impacts. Science. **315**, 1537–1540 (2007). doi: 10.1126/science.1137695

Lawrence, M.G., Rasch, P.J., von Kuhlmann, R., Williams, J., Fischer, H., de Reus, M., Lelieveld, J., Crutzen, P.J., Schultz, M., Stier, P., et al.: Global chemical weather forecasts for field campaign planning: predictions and observations of large-scale features during MINOS, CONTRACE, and INDOEX. Atmos. Chem. Phys. **3**, 267–289 (2003). doi:10.5194/acp-3-267-2003

Lelieveld, J., Berresheim, H., Borrmann, S., Crutzen, P.J., Dentener, F.J., Fischer, H., Feichter, J., Flatau, P.J., Heland, J., Holzinger, R., et al.: Global air pollution crossroads over the Mediterranean. Science **298**, 25 (2002). doi:10.1126/science.1075457

Methven, J., et al.: Establishing Lagrangian connections between observations within air masses crossing the Atlantic during the International Consortium for Atmospheric Research on Transport and Transformation experiment. J. Geophys. Res. **111**, D23S62 (2006). doi: 10.1029/2006JD007540

Monks, P., Granier, C., Fuzzi, S., Stohl, A., Williams, M.L., Akimoto, H., Amann, M., Baklanov, A., Baltensperger, U., Bey, I.N., et al.: Atmospheric composition change—global and regional air quality. Atmos. Environ. **43**, 33 (2009). doi:10.1016/j.atmosenv.2009.08.021

Petzold, A., Weinzierl, B., Huntrieser, H., Stohl, A., Real, E., Cozic, J., Fiebig, M., Hendricks, J., Lauer, A., Law, K., et al.: Perturbation of the European free troposphere aerosol by North American forest fire plumes during the ICARTT-ITOP experiment in Summer 2004. Atmos. Chem. Phys. **7**, 5105–5127 (2007). doi:10.5194/acp-7-5105-2007

Pisso, I., Real, E., Law, K.S., Legras, B., Bousserez, N., Attié, J.L., Schlager, H.: Estimation of mixing in the troposphere from Lagrangian trace gas reconstructions during long-range pollution plume transport. J. Geophys. Res. **114**, D19301 (2009). doi:10.1029/2008JD011289

Real, E., Law, K., Weinzierl, B., Fiebig, M., Petzold, A., Wild, O., Methven, J., Arnold, S., Stohl, A., Huntrieser, H. et al.: Processes influencing ozone levels in Alaskan forest fire plumes during long-range transport over the North Atlantic. J. Geophys. Res. **112**, D10S41 (2007) doi: 10.1029/2006JD007576

Real, E., Law, K.S., Schlager, H., Roiger, A., Huntrieser, H., Methven, J., Cain, M., Holloway, J., Neuman, J.A., Ryerson, T., et al.: Lagrangian analysis of low level anthropogenic plume processing across the North Atlantic. Atmos. Chem. Phys. **8**, 7737–7754 (2008). doi: 10.5194/acp-8-7737-2008

Roiger, A.: Biomass burning pollution in the summer time Arctic atmosphere: development and deployment of a novel CI-ITMS instrument for PAN detection. Dissertation, Universität Heidelberg (2011)

Roiger, A., Schlager, H., Schaefler, A., Huntrieser, H., Scheibe, M., Aufmhoff, H., Cooper, O.R., Sodemann, H., Stohl, A., Burkhart, J.: In situ observation of Asian pollution transported into the Arctic lowermost stratosphere. Atmos. Chem. Phys. **11**, 10975–10994 (2011). doi: 10.5194/acp-11-10975-2011

Schlager, H., Huntrieser, H., Heland, J., Ziereis, H.: Analysis of convective and long-range transport of reactive nitrogen species over Europe using aircraft observations. EXPORT-E2 Final Report, EUROTRAC-2 ISS. In: Penkett, S., Law, K. S., Platt, U., Volz-Thomas, A. (eds.) GSF—National Research Center for Environment and Health, Munich, Germany, pp. 151–155 (2003)

Schumann, U., Konopka, P., Baumann, R., Busen, R., Gerz, T., Schlager, H., Schulte, P., Volkert, H.: Estimate of diffusion parameters of aircraft exhaust plumes near the tropopause from nitric oxide and turbulence measurements. J. Geophys. Res. **100**:(D7), 14, 147–114, 162 (1995) doi: 10.1029/95JD01277

Schumann, U., Weinzierl, B., Reitebuch, O., Schlager, H., Minikin, A., Forster, C., Baumann, R., Sailer, T., Graf, K., Mannstein, H., et al.: Airborne observations of the Eyjafjalla volcano ash cloud over Europe during air space closure in April and May 2010. Atmos. Chem. Phys. **11**, 2245–2279 (2011). doi:10.5194/acp-11-2245-2011

Sodemann, H., Pommier, M., Arnold, S.R., Monks, S.A., Stebel, K., Burkhart, J.F., Hair, J.W., Diskin, G.S., Clerbaux, C., Coheur, P.-F., et al.: Episodes of cross-polartransport in the Arctic troposphere during July 2008 as seenfrom models, satellite, and aircraft observations. Atmos. Chem. Phys. **11**, 3631–3651 (2011). doi:10.5194/acp-11-3631-2011

Stohl, A., Haimberger, L., Scheele, M.P., Wernli, H.: An intercomparison of results from three trajectory models. Meteorol. Appl. **8**, 127–135 (2001). doi:10.1017/S1350482701002018

Stohl, A., Forster, C., Eckhardt, S., Spichtinger, N., Huntrieser, H., Heland, J., Schlager, H., Wilhelm, S., Arnold, F., Cooper, O.: A backward modeling study of intercontinental pollution transport using aircraft measurements. J. Geophys. Res., **108** (D12):4370 (2003a) doi: 10.1029/2002JD002862

Stohl, A., Huntrieser, H., Richter, A., Beirle, S., Cooper, O.R., Eckhardt, S., Forster, C., James, P., Spichtinger, N., Wenig, M., et al.: Rapid intercontinental air pollution transport associated with a meteorological bomb. Atmos. Chem. Phys. **3**, 969–985 (2003b). doi:10.5194/acp-3-969-2003

Stohl, A., Forster, C., Huntrieser, H., Mannstein, H., McMillan, W.W., Petzold, A., Schlager, H., Weinzierl, B.: Aircraft measurements over Europe of an air pollution plume from Southeast Asia—aerosol and chemical characterization. Atmos. Chem. Phys. **7**, 913–937 (2007). doi: 10.5194/acp-7-913-2007

Wernli, H., Bourqui, M.: A Lagrangian "1-year climatology'' of (deep) cross-tropopause exchange in the extratropical Northern Hemisphere. J. Geophys. Res. **107**, (2002). doi: 10.1029/2001jd000812

Wild, O., Law, K.S., McKenna, D.S., Bandy, B.J., Penkett, S.A., Pyle, J.A.: Photochemical trajectory modeling studies of the North Atlantic region during August 1993. J. Geophys. Res. **101**: 29 269–229 288 (1996) doi: 10.1029/96JD00837

Atmospheric Acoustics

13

Dietrich Heimann, Arthur Schady and Joseph Feng

Abstract

This chapter deals with sound propagation in the atmosphere, which is an important link in the functional chain from noise emissions from aircraft, road and rail vehicles, and wind turbines to noise perception. The principle processes in outdoor sound propagation are explained. They include refraction, diffraction, and reflection. Two sound propagation models for scientific applications are briefly outlined. Finally, three illustrative applications and their results are discussed.

13.1 Introduction

The atmosphere is a carrier of various types of internal waves. Among them are sound waves which ensure aural communication between individuals, both people and animals. However, sound waves are also generated by technical devices, in particular those that move or have moving parts. Any kind of motorized means of transportation, such as aircraft, trains, trucks and cars, emits sound waves which

D. Heimann (✉) · A. Schady · J. Feng
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Dietrich.Heimann@dlr.de

A. Schady
e-mail: Arthur.Schady@dlr.de

J. Feng
e-mail: Joseph.Feng@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_13,

propagate through the atmosphere. Such *sound*, once it is perceived by a human being, is often classified as *noise* because it is unwanted, annoying, disturbing, or even pathogenic.

Today, noise and above all traffic noise is seen as a severe environmental nuisance with psychological and physiological effects. Noise reduction is therefore a societal and political goal. This effort includes technical measures at the source to reduce noise *emissions* and measures influencing the propagation path, for instance the construction and installation of noise barriers. A solid knowledge of outdoor sound propagation effects is essential for assessing the noise situation in a given area and planning adequate and effective mitigation measures.

The Institute of Atmospheric Physics has dealt with the field of atmospheric *acoustics* and noise since the late 1990s. First, the activities were incorporated in aeronautics projects on quiet air traffic. Later, the activities were extended to transportation noise in general. A further extension to noise from wind turbines is envisaged.

In the following the principle processes and effects of outdoor *sound propagation* are summarized. A description follows of two model concepts which were implemented at the Institute of Atmospheric Physics. Finally, illustrative results of three applications (aircraft noise, sonic boom, influence of trees) are shown.

13.2 Outdoor Sound Propagation

Sound waves in the atmosphere consist of pressure oscillations associated with longitudinal particle motion, e.g. Embleton (1996), Ostashev (1997). A sound wave is characterized by its frequency f and sound pressure amplitude $\hat{p}$. For healthy young people the audible frequencies range from 16 Hz to 16 kHz with a maximum sensitivity between 1 and 6 kHz. The sound pressure amplitude is commonly expressed by the *sound pressure level* as a logarithmic measure $SPL = 10\lg\left(\hat{p}^2/p_0^2\right)$ dB (decibel) with $p_0 = 2 \cdot 10^{-5}$ Pa, the threshold of hearing at 1 kHz, as reference. The use of the squared sound pressure in the definition of *SPL* is a consequence of the fact that sound energy depends on the squared pressure amplitude. Hence, doubling the sound energy (e.g., two equal sound sources instead of one) raises the sound pressure level by approximately 3 dB because $\lg 2 \approx 0.3$. In practice the sound pressure level is often frequency-weighted to consider the human sensibility to different frequencies. The most common 'A-weighting' reduces the sound level for frequencies below 1 kHz and above 6 kHz according to the human capability of perceiving tones of different frequency.

In the atmosphere the sound speed mainly depends on the temperature T: $c \approx \sqrt{\kappa R T}$ with the ratio of specific heats $\kappa = c_p/c_v = 1.4$, and the gas constant of dry air $R = 287\,\mathrm{J\ kg^{-1}\,K^{-1}}$. Near the ground the sound speed relative to the air varies between 315 m/s and 350 m/s assuming typical mid-latitude temperatures. The *wavelength* is related to the speed of sound c and *frequency* f by $\lambda = c/f$. Hence, the wavelengths of audible frequencies cover a wide range from 2 cm to 20 m.

Outdoor sound propagation is affected by processes which depend on the actual state of the atmosphere, by interactions with the even or uneven ground, and with obstacles on the ground such as buildings, screens or vegetation. Before typical atmospheric propagation effects are illustratively described, the processes involved are first defined separately.

Geometrical spreading. Once a point source emits sound uniformly into the air space (spherical wave fronts), the sound intensity (i.e., the sound energy flux density) decreases proportional to $1/d^2$ where d is the distance from the source. This causes the *SPL* to decrease by approximately 6 dB as the distance doubles. In the case of a line source (cylindrical wave fronts), the *SPL* decreases proportional to $1/d$ or by approximately 3 dB as the distance doubles (because of $10\,\lg(1/2) \approx -3$).

Air absorption. The particle motion associated with sound waves causes friction, which converts a part of the sound energy into heat. Transversal and rotational vibration modes of two-atomic (e.g., O_2 and N_2) and three-atomic (e.g., H_2O) molecules lead to further losses of sound energy. As a consequence sound is attenuated with distance in addition to the effect of geometrical spreading. Air absorption depends on air pressure, temperature, and humidity. Generally, it is the stronger the higher the frequency. For a temperature of 15 °C, a relative humidity of 70 % and an air pressure of 1 000 hPa a 100 Hz wave is attenuated by only 0.25 dB/km, while a 5 kHz wave is attenuated by even 40 dB/km. Therefore, high frequency sound does not play a role beyond some 100 m.

Sound refraction. Refraction occurs if gradients of temperature T (and hence *sound speed* c) or *wind speed* U have a component perpendicular to the direction of propagation. In these cases the wave fronts change their direction, i.e., the sound does not propagate straight ahead (along straight sound rays), but follows a curve (curved sound rays) into the direction of lower temperature T or wind speed U. The curvature (inverse radius) is proportional to the atmospheric velocity gradients. In the near-ground atmosphere, i.e., in the *atmospheric boundary layer*, the gradients of c and U are strongest in the vertical direction. Here, horizontal sound propagation from near-ground sources to near-ground receivers is most affected by refraction. For noise from elevated sources like aircraft flying higher than some 100 m above ground, the effects of refraction occur only far away from the foot of the source. Unless the elevated source is very loud, e.g., sonic booms from supersonic aircraft, refraction effects are therefore of minor importance.

Sound diffraction. Diffraction occurs at convex surfaces or edges of obstacles where a part of the sound is directed into a shielded space. A typical example is diffraction at the top edges of a noise barrier, which allows noise to be diverted to the side of the barrier opposite to the source, i.e., into the protected zone. The longer the wavelength, the more effective is the diffraction.

Sound reflection. Airborne sound is fully or partly reflected at the ground surface or the surfaces of buildings, etc. Specular reflection occurs if the surface is plain with the angle of incidence being equal to the angle of reflection. At rough surfaces the reflection is diffuse. Reflectivity depends on the impedance of the

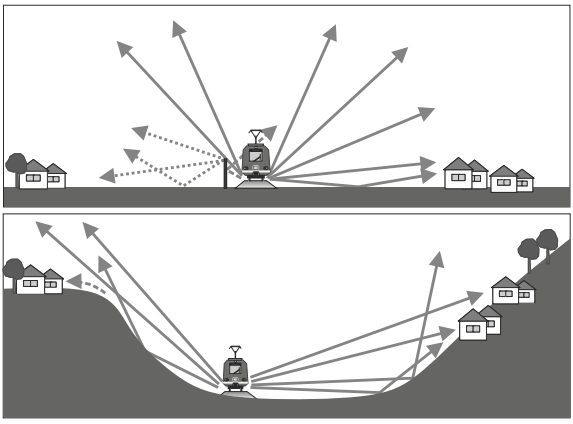

Fig. 13.1 Sound rays showing reflection at plane (*top*) and sloped (*bottom*) ground. Diffracted rays (behind a noise barrier and convex terrain) are broken

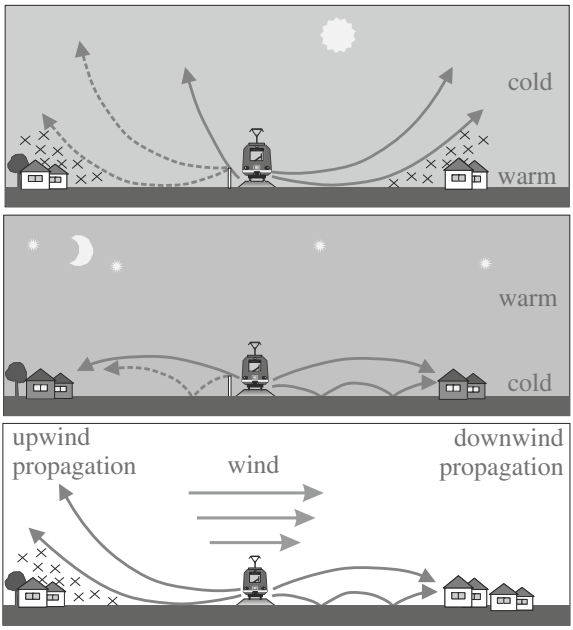

Fig. 13.2 Sound rays with upward refraction during daytime (*top*) and downward refraction during nighttime (*center*). The bottom panel shows upward and downward refraction in upwind and downwind propagation, respectively. Shadow zones are indicated by *crosses*, diffracted rays are drawn *dashed*

reflecting medium (e.g., the ground) relative to the impedance of air: $Z = \rho\, c$ with ρ being the air density. Solid media such as concrete have high impedance, causing strong reflection, while porous media (e.g., the soil) have low impedance, causing only weak reflection. In the case of grazing incidence in near-ground propagation over low impedance, porous ground, the reflected sound is phase shifted by 180°. Hence, the superposition of reflected and direct, non-reflected sound at the receiver causes destructive *interference* and low sound pressure amplitudes. Therefore, acoustically soft, i.e., low-impedance, porous ground or fresh snow leads to a high attenuation of near-ground to near-ground propagation. This damping effect gradually vanishes for steeper incidence.

Sound scattering. Turbulent eddies in the air whose extension is equal to or shorter than the wavelength cause a scattering of sound where forward and backward scattering is more important than sideward scattering. Besides turbulence, scattering is also caused by filigree obstacles like twigs and tree branches.

Typically, noise propagates over irregular ground (solid or porous, plane or hilly) with buildings and vegetation in an inhomogeneous atmosphere with temperature and wind gradients. Therefore, several of the aforementioned processes act in combination. Such combinations are illustrated in Figs. 13.1 and 13.2. They show typical situations of ground-to-ground propagation with a ground-based source (e.g., road or rail traffic, aircraft on runways) and ground-based receivers.

Behind a noise barrier or behind convex terrain an acoustical shadow forms where direct sound incidence is no longer possible. Only diffracted or scattered sound can enter this acoustical shadow.

Downward refraction typically occurs during night in the case of a temperature *inversion* (warm above cold air) with the sound speed increasing with height. Downward refraction also occurs in downwind propagation because the wind speed increases with height. Downward refracted sound is (repeatedly) reflected at ground, which leads to long-range audibility. By contrast, upward refraction occurs during the day when the temperature decreases with height or in upwind propagation. Beyond a certain distance from the source a shadow zone forms where audibility is strongly reduced.

13.3 Sound Propagation Modeling

Sound propagation models are typically used to predict the *SPL* due to given sources and propagation conditions in a selected domain. Various types of such models are published in the literature (Salomons 2001; Bérengier et al. 2003) and differ in the degree of sophistication. Rather simple models are used as noise prediction tools in engineering applications. More sophisticated ones are designed for scientific studies and specific applications. At the Institute of Atmospheric Physics two numerical sound propagation models of the latter type are used. Both models are capable of regarding the influence of meteorology and topography on sound propagation. However, they have specific advantages and disadvantages. The two models are briefly described below.

In the following, the atmospheric parameters are split into a mean part and deviations from that mean. So, we assume that the velocity vector $\mathbf{U} = \bar{\mathbf{U}} + \mathbf{U}' + \mathbf{U}''$ and the pressure $p = \bar{p} + p' + p''$ are each composed of a mean wind speed vector $\bar{\mathbf{U}}$ and a mean air pressure $\bar{p}$, with turbulent deviations from these means $\mathbf{U}'$ and p', and additional acoustical deviations describing the particle velocity vector $\mathbf{U}''$ and the sound pressure p''. Analogously, we may decompose the density ρ and the temperature T. The mean and turbulent parts can be combined to define the meteorological background state ($p_{met} = \bar{p} + p'$, $\mathbf{U}_{met} = \bar{\mathbf{U}} + \mathbf{U}'$, $\rho_{met} = \bar{\rho} + \rho'$, and $T_{met} = \bar{T} + T'$) of the acoustical waves.

Sound Particle Model

This model is based on ray acoustics (Heimann and Gross 1999). At the source a large number of virtual 'sound particles' are emitted into the environment. Each particle is associated with a partial sound energy which is distributed to frequencies according to the acoustic source spectrum. These particles travel with the local sound speed c along sound rays. In the absence of wind or temperature gradients the rays are straight (no refraction). They are curved in the presence of wind or temperature gradients (refraction).

The vector of position $\mathbf{X}_i$ of the ith sound particle and its associated wave front unit vector $\mathbf{N}_i$ are given by

$$\frac{\partial \mathbf{X}_i}{\partial t} = \mathbf{U}_{met} + c\,\mathbf{N}_i \tag{13.1}$$

$$\frac{\partial \mathbf{N}_i}{\partial t} = -\nabla c - \mathbf{N}_i \cdot \nabla \mathbf{U}_{met} \tag{13.2}$$

These equations are integrated in time for each particle until it has left the model domain. Once a particle hits the ground or an obstacle it is reflected (specular reflection at plain surfaces or diffuse reflection at rough surfaces). According to the impedance ratio and the angle of incidence the phase is shifted and a part of its attributed energy is lost.

All particles which are passing receiver points within a defined distance are counted and their respective complex sound pressures (indicating amplitude and phase) are added up to calculate the local sound pressure level *SPL*.

The Lagrangian sound particle model is best suited to high-frequency sound. Diffraction cannot be modeled straightforwardly, but needs special treatments.

Finite-Difference Time-Domain Model

The sound waves that propagate through the atmosphere can be well described by linear acoustics. Nonlinear effects are limited to the near-field of a strong acoustical source or to acoustical shock waves such as sonic booms. A linearization of the Euler equation of motion and the equation of continuity with respect to a given meteorological background, neglecting non-adiabatic processes, Earth's rotation, internal friction and gravitation, leads to a set of prognostic equations for the particle velocity $\mathbf{U}''$ and the sound pressure p'' (Heimann and Blumrich 2002; Salomons et al. 2002; Heimann and Karle 2006). A scale-analysis suggests the negligibility of further terms so the prognostic equations of sound waves read:

$$\frac{\partial p''}{\partial t} = -\mathbf{U}_{met} \nabla p'' - \kappa\, p_{met} \nabla \cdot \mathbf{U}'' \tag{13.3}$$

$$\frac{\partial \mathbf{U}''}{\partial t} = -\mathbf{U}_{met} \nabla \mathbf{U}'' - \frac{1}{\rho_{met}} \nabla p'' \tag{13.4}$$

This set of prognostic equations describes geometrical spreading, refraction, and diffraction. Reflection is considered by appropriate boundary conditions. The equations are numerically integrated in time within a two- or three-dimensional computational domain. Depending on the numerical scheme the cell size of the numerical grid has to be smaller than five to ten times the acoustical wavelength. The sound pressure amplitude is evaluated at the receiver cells and the sound pressure level *SPL* is determined.

The *Eulerian finite-difference time-domain model* is universal and very efficient for low-frequency sound. Long-range simulations of high-frequency sound are often limited by the available computational resources.

13.4 Applications

Sound propagation models can be applied to various problems in atmospheric acoustics. At the Institute of Atmospheric Physics applications included noise propagation in valleys (Heimann and Gross 1999; Heimann et al. 2010) and street canyons (Heimann 2007), interactions with slope (Heimann 2006, 2010), flow and turbulence behind noise barriers (Heimann and Blumrich 2002, 2004), scattering and reflection by trees (Heimann 2003), noise of aircraft during takeoff and landing (Neise 2007), sonic booms from supersonic aircraft (Heimann 2001; Blumrich et al. 2005a, b; Kästner and Heimann 2010), and the role of wake flow behind wind turbines (Heimann et al. 2011). In addition, considerations were made how meteorological conditions can be categorized for sound propagation modeling (Heimann and Salomons 2004; Defrance et al. 2007; Heimann et al. 2007). Three of these applications and their results are discussed in the following.

Noise of Aircraft During Takeoff

Noise in the neighborhood of *airports* is an important issue which often leads to night-flight bans or limitations on the allowed number of operations. Successful efforts to reduce noise by introducing new, less noisy aircraft were often compensated or even overcompensated by the increasing number of flights. Moreover, replacing an aircraft fleet with modern low-noise vehicles takes a rather long time. Therefore, the noise reducing potential of alternative operational procedures is investigated, too. Above all, these are low-noise approach and departure procedures, e.g., continuous descent approaches, steep flight profiles during landing, or reduced power settings during takeoff. In summer 2006 flight tests with different approach and departure procedures were performed in cooperation with Lufthansa

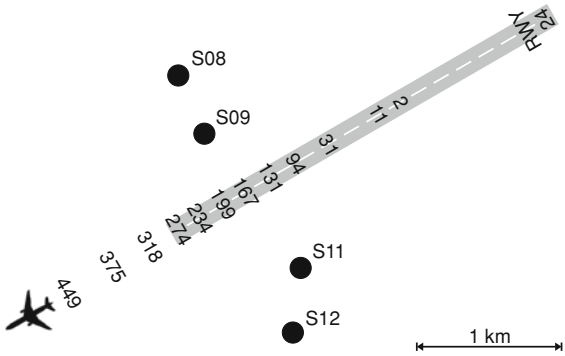

Fig. 13.3 Sketch of runway 24 at Parchim airport and the positions of the four receivers (S08, S09, S11, S12) mentioned in the text. The numbers along the aircraft track indicate the altitude of the departing aircraft in meters above ground

at Parchim airport near Schwerin (Neise 2007). A large number of sound level meters were installed to measure the noise.

Here a part of these data are used to validate sound propagation simulations with the sound particle model described above. Four measurement stations (receivers) were installed perpendicular to the center line of the runway at different distances (Fig. 13.3). The noise was measured during the takeoff of an Airbus A310 twin-jet passenger aircraft at 1.5 m above ground. As the aircraft passed, noise from different angles of sound radiation was encountered by the receivers. Therefore, the noise emission and the emission frequency spectrum were considered as functions of this angle. Figure 13.4 shows how the noise level varies as a function of the azimuth angle under which the departing aircraft is seen from the receiver during the flyby (full-thrust takeoff procedure). While the aircraft climbs and moves against the wind towards the southwest, the outer receivers (S08 and S012) soon come out of the initial acoustical shadow caused by upward refraction during near-ground upwind propagation. Except for the time in the acoustical shadow, the coincidence of measurements and model results is fairly good.

In Fig. 13.5 the observed and simulated spectra are shown for three aircraft positions, given by the azimuth angle under which the aircraft is seen from the receiver during the flyby. Since different aircraft positions also mean different emissions angles, the emission spectra vary. The spectra at the receivers deviate from the emission spectra because of frequency-dependent air absorption and ground reflection which cause partly destructive, partly constructive interference of direct and reflected sound, depending on frequency. Measured and simulated spectra at the receivers are in good agreement.

Sonic-Boom Propagation

Supersonic aircraft flying with a speed V greater than the speed of sound c (i.e., *Mach number* $M = V/c > 1$) generate a shock wave which produces a loud bang near the ground, the "*sonic boom*." The sonic boom is not only very annoying to people, but also a threat to animals, including marine mammals. The discussion to

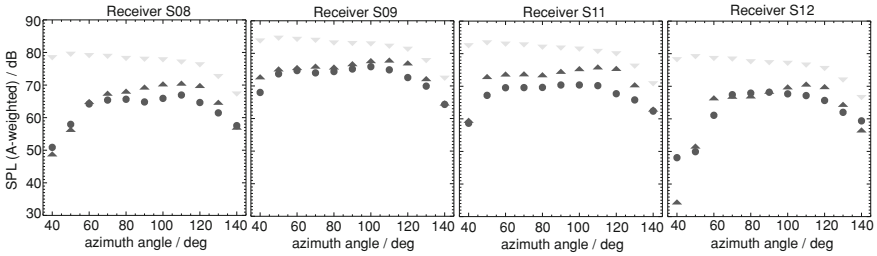

Fig. 13.4 Sound pressure level during flyby at the four receivers shown in Fig. 13.3. The abscissa provides the azimuth angle under which the aircraft is seen from the respective receivers (0° front view, 90° side view, 180° back view). The *bullets* indicate the measurements. The *dark triangles* show the model results. Measurements and model results are for 1.5 m above ground. The *light triangles* provide the emission levels (lateral directivity). All sound levels are A-weighted

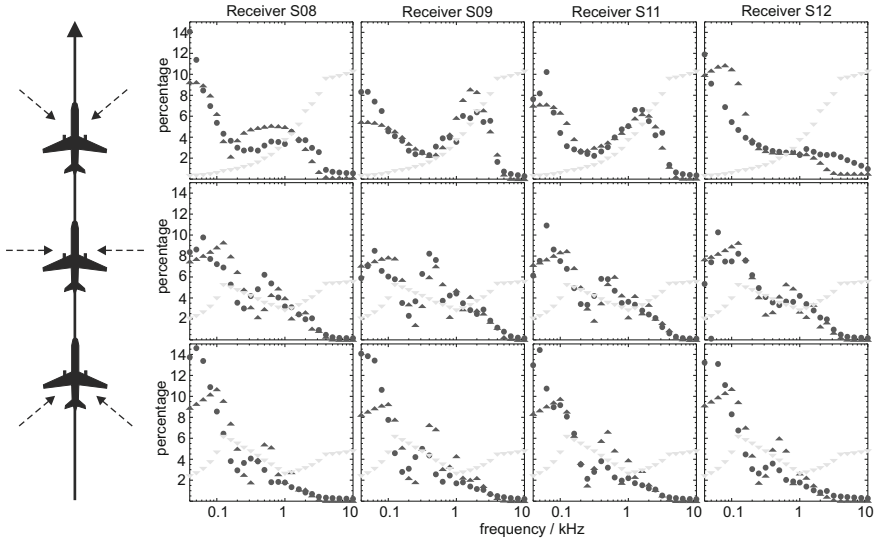

Fig. 13.5 Observed (*bullets*) and simulated (*dark triangles*) spectra at the four receivers shown in Fig. 13.3 (1.5 m above ground). The upper, middle and lower rows are for an azimuth angle of 40°, 90°, and 140°, respectively, under which the aircraft is seen from the receivers (*dashed arrows* on the *left*). The *light triangles* provide the emission spectrum

(re)introduce supersonic passenger or business aircraft brought about several research projects to investigate the potential environmental problems of intensive supersonic air traffic and to explore possible solutions. The sonic-boom problem was especially addressed. One study in this context aimed at the identification of specific refractive atmospheric situations and flight parameters which prevent the sonic boom from reaching the ground (Kästner and Heimann 2010).

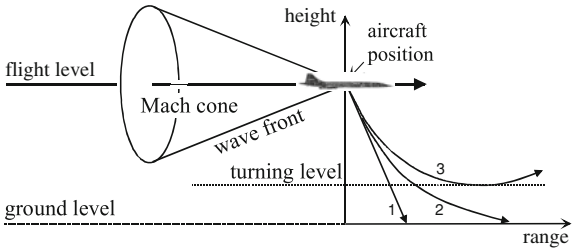

Fig. 13.6 Schematic view of the Mach cone and three sonic-boom sound rays: *1*: non-refracted straight ray reaching the ground, *2*: upward refracted ray reaching the ground, *3*: upward refracted ray not reaching the ground. Taken from Kästner and Heimann (2010)

The sonic-boom shock wave front is defined by the *Mach cone* (Fig. 13.6) and propagates perpendicular to the Mach cone surface. The initial direction of sonic boom propagation of an aircraft flying only marginally faster than 1 Mach is only slightly slanted towards the ground. The faster the aircraft, the steeper the downward direction of sonic-boom propagation. Normally, supersonic vehicles fly at high altitudes between 15 and 25 km, i.e., in the middle stratosphere. During downward propagation the orientation of the wave front changes by refraction depending on the wind and temperature gradients encountered. Important atmospheric features between typical supersonic flight levels and ground are the wind maximum of the jet stream and the tropopause with its characteristic kink in the vertical temperature gradient. Under certain circumstances it may happen that sonic-boom sound rays are refracted upward so that they turn and do not reach the ground. In such cases the sonic boom would not be heard at ground ('no boom') or would at least not be very loud ('low boom').

How frequently a sound boom does not reach the ground depends on many factors, viz. terrain height, regional climatology (frequency distributions of wind and temperature profiles), and relevant flight parameters (altitude, speed or Mach number, heading). An example of calculated sonic-boom sound rays (with Eqs. 13.1 and 13.2) is shown in Fig. 13.7. The chance of a sonic boom not reaching the ground is especially high if the Mach number is low ($M < 1.2$) because in that case the rays start in a nearly horizontal direction and are most affected by refraction due to vertical gradients. The chance is also increased for flight directions against the mean wind direction and for low terrain elevations. Figure 13.8 provides the regional distribution of the frequency of occurrence of low or no boom situations in Europe for a specific set of flight parameters.

Influence of a Tree on Sound Propagation

Single trees, rows of trees or forests influence sound propagation, directly through reflection and scattering of sound wave by trunks, branches and twigs, and/or indirectly through modified refraction due to the influence of the trees on the wind

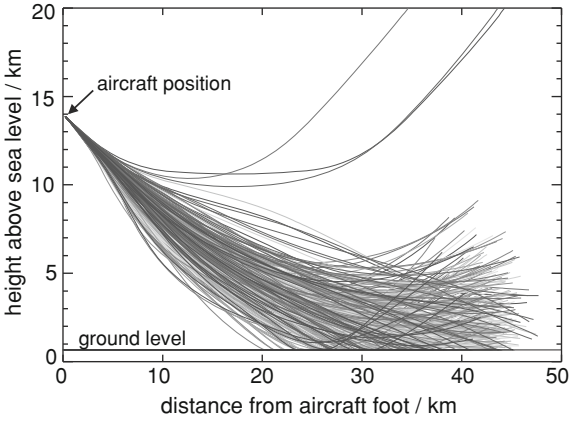

Fig. 13.7 Calculated daily sound rays for the year 2001 at Munich for a westward-heading aircraft flying at 1.2 Mach at 14 km altitude. Taken from Kästner and Heimann (2010)

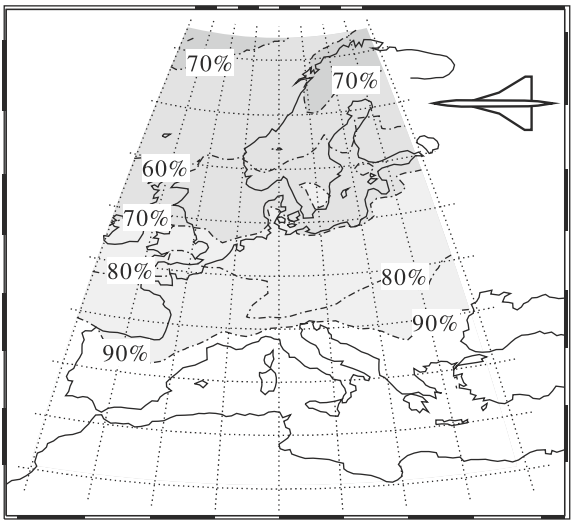

Fig. 13.8 Frequency distribution of low or no boom situations (turning level above ground) for the period 1991–2001 for supersonic flights of 1.15 Mach at 15.2 km altitude and westward headings. Taken from Kästner and Heimann (2010)

field. Many studies have been undertaken to quantify the influence of trees. These include theoretical considerations (e.g., scattering by cylinders), measurements and model simulations. With the finite-difference time-domain model (Eqs. 13.3 and 13.4) it is possible to resolve the tree trunk and major branches in three dimensions. The model is capable of simulating reflection (wavelengths smaller than the trunk/branch diameter) as well as scattering (wavelength equal or larger than the diameter).

To investigate the influence of a row of single trees, a 30 m long, 8 m wide and 24 m high model domain with a numerical resolution of 0.05 m was defined. The resolution allows the evaluation of spectra for $f \leq 1350\,\mathrm{Hz}$. Hence, the domain consists of 46 million numerical grid cells. A 16 m high idealized tree (a vertical straight trunk with straight branches) with a crown diameter of 8 m was placed in the center of the domain. Two arrays of virtual receivers were placed 3 m in front

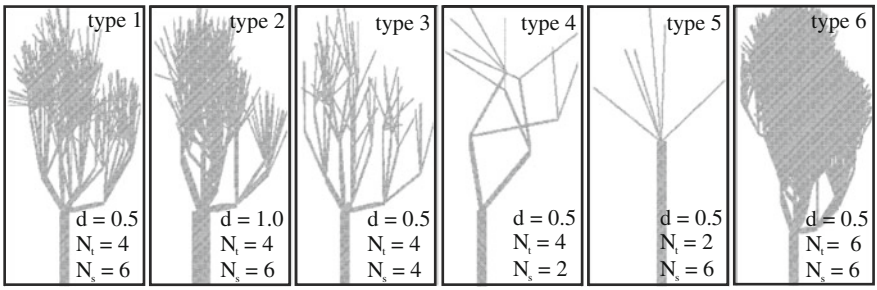

Fig. 13.9 Silhouettes of idealized trees (only trunk and branches) with different parameter settings (d: diameter of the trunks in m, N_t: number of branching orders, N_s: number of branches per branching order). The branch diameter halves per branching sequence. All trees are 16 m high with a crown diameter of 8 m

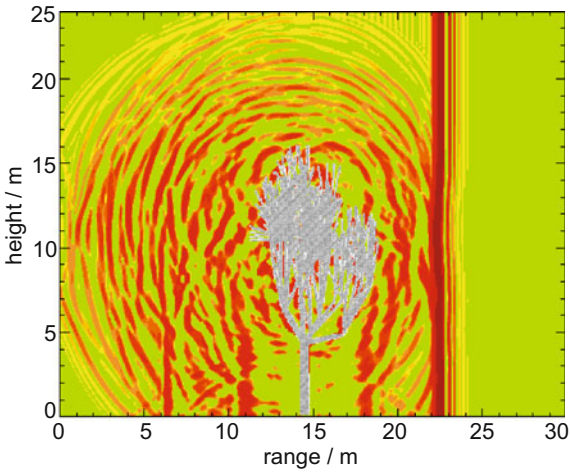

Fig. 13.10 Snapshot of simulated sound pressure (vertical cross-section through the center of the domain) after the wave front has passed an idealized row of trees (type 1; silhouette in *grey*) from *left* to *right*

and 3 m behind the tree. At these arrays the backscattered or backward reflected sound and the forward transmitted sound are recorded. Cyclic boundary conditions in the lateral directions suggest the existence of an infinite row of trees. The chosen simulation time of 0.09 s ensures that besides the signal of the center tree at least the sideward scattered or reflected sound from the adjacent trees is recorded at the arrays. The scattered and reflected signals of farther distant trees do not much contribute anyway.

The simulations were performed for six types of trees (Fig. 13.9). Trunks and branches are assumed to be cylinders. The parameters characterizing the tree type include the trunk diameter d, the number of branching orders N_t, and the number of branches per branching order N_s. The first order branches have a diameter that is half of the trunk diameter, the branch diameter again halves from one order to the subsequent one. For each parameter setting five random realizations were simulated. They differ by the random orientation of the branches.

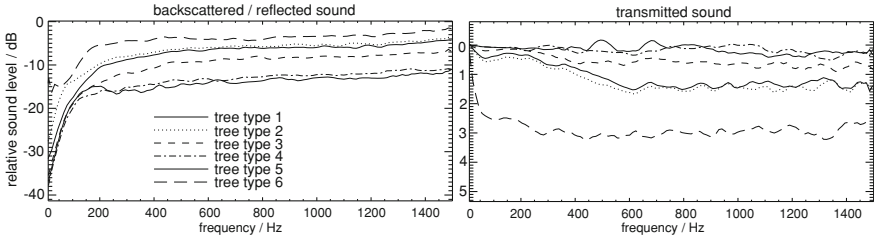

Fig. 13.11 Spectral distribution of backscattered and backward reflected sound (*left*) and of transmitted sound (*right*). The results are shown for six types of idealized trees. The sound levels are relative to the undisturbed sound. Note the different scales at the vertical axes

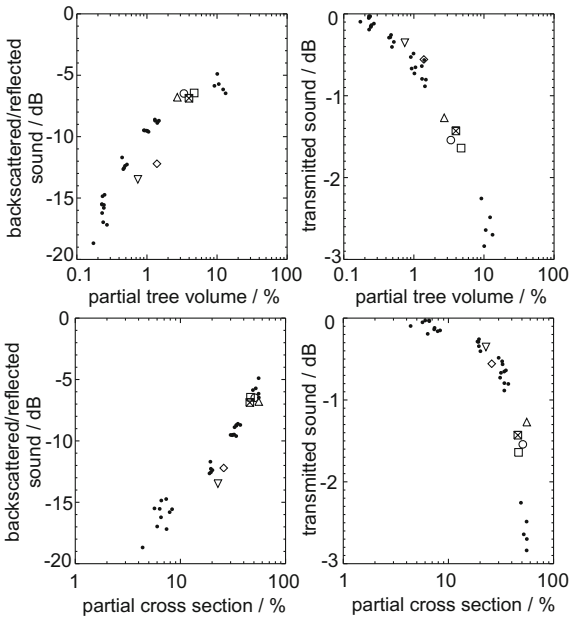

Fig. 13.12 Simulation results of backward directed (*left*) and transmitted (*right*) sound relative to undisturbed sound as a function of partial tree volume (*top*) and partial cross section (*bottom*). The *dots* refer to the six tree types with five realizations per type. Other symbols refer to six different sets of laser-scan tree data (courtesy of A. Bienert, Professur für Photogrammetrie, TU Dresden)

In the model simulations a plane upright sound pressure pulse, oriented parallel to the row of trees, propagates horizontally and penetrates the tree. As an example Fig. 13.10 provides a vertical cross section of the sound pressure field perpendicular to the row of trees after the pulse has passed the row. While the pulse front itself is hardly distorted, irregular patterns around the tree indicate that sound is scattered and reflected in all directions.

Signal evaluation at the receiver arrays in front of and behind the row of trees reveals that low-frequency sound ($f \leq 100\,\mathrm{Hz},\ \lambda \geq d$) is hardly directed backward and passes the row of trees without significant hindrance (Fig. 13.11). The degree of reflection and transmission of medium and high frequency sound depends on the chosen tree parameters.

Figure 13.12 shows the result for all model simulations with various types of trees. Here the trees are characterized by the ratio of wood to air in the enveloping tree volume and in the silhouette as seen from the approaching sound wave. The results suggest a functional relationship between these parameters on the one side and the backward and forward sound energy on the other side. This relationship can be used for the parameterization of trees in simulations which do not permit the resolution of single trees.

13.5 Conclusions and Outlook

Atmospheric acoustics provides knowledge and tools to describe the propagation of sound in the atmosphere. For the solution of outdoor noise problems, in particular noise from aircraft, road vehicles, trains and wind turbines, sound propagation is an important link between source and receiver. It is part of the functional chain between noise emission and noise effects on human beings (e.g., sleep disturbance, annoyance, impairment of heath). While state-of-the-art noise prediction tools are regulated in national and international standards (e.g., ISO 9613-2 (1996)) scientific sound propagation models are much more sophisticated and are capable of describing meteorological and topographical influences in physical detail. However, these models are rather demanding in terms of computational resources, both as to time and storage. Therefore, the use of these models is limited to scientific applications (investigations of processes and relationships, for instance to derive parameterizations) and selected practical problems.

Nevertheless, there is still a high potential for new fields of application and further development of the models. The availability of more powerful computers in future will open applications to larger ranges and higher frequencies. Another extension of applicability is expected from the introduction of refined numerical techniques.

References

Bérengier, M.C., Gauvreau, B., Blanc-Benon, P., Juvé, D.: Outdoor sound propagation: a short review on analytical and numerical approaches. Acta Acust. **89**, 980–991 (2003)

Blumrich, R., Coulouvrat, F., Heimann, D.: Variability of focused sonic booms from accelerating supersonic aircraft in consideration of meteorological effects. J. Acoust. Soc. Am **118**, 696–706 (2005a). doi:10.1121/1.1938547

Blumrich, R., Coulouvrat, F., Heimann, D.: Meteorologically induced variability of sonic-boom characteristics of supersonic aircraft in cruising flight. J. Acoust. Soc. Am. **118**, 707–722 (2005b). doi:10.1121/1.1953208

Defrance, J., Salomons, E., Noordhoek, I., Heimann, D., Plovsing, B., Watts, G., Jonasson, H., Zhang, X., Premat, E., Schmich, I., et al.: Outdoor sound propagation reference model developed in the European Harmonoise project. Acta Acust. **93**, 213–227 (2007)

Embleton, T.F.W.: Tutorial on sound propagation outdoors. J. Acoust. Soc. Am. **100**, 31–48 (1996)

Heimann, D.: Effects of long-term atmospheric variability on the width of a sonic-boom carpet produced by high-flying supersonic aircraft. Acoust. Res. Lett. **2**, 73–78 (2001)
Heimann, D.: Numerical simulations of wind and sound propagation through an idealised stand of trees. Acta Acust. **89**, 779–788 (2003)
Heimann, D.: Sound propagation in a nocturnal slope-wind layer—a numerical model study. Acta Acust. **92**, 362–369 (2006)
Heimann, D.: Three-dimensional linearised Euler model simulations of sound propagation in idealised urban situations with wind effects. Appl. Acoust. **68**, 217–237 (2007)
Heimann, D.: On the efficiency of noise barriers near sloped terrain—a numerical study. Acta Acust. **96**, 1003–1011 (2010). doi:10.3813/AAA.918363
Heimann, D., Blumrich, R.: Anwendungsbeispiele numerischer Schallausbreitungssimulationen mit konsistenter Berücksichtigung der Atmosphäre. Z. Lärmbekämpfung **49**, 91–97 (2002)
Heimann, D., Gross, G.: Coupled simulation of meteorological parameters and sound intensity in a narrow valley. Appl. Acoust. **56**, 73–100 (1999). doi:10.1016/S0003-682X(98)00018-1
Heimann, D., Blumrich, R.: Time-domain simulations of sound propagation through screen-induced turbulence. Appl. Acoust. **65**, 561–582 (2004). doi:10.1016/j.apacoust.2003.09.007
Heimann, D., Karle, R.: A linearized Euler finite-difference time-domain sound propagation model with terrain-following coordinates. J. Acoust. Soc. Am. **119**, 3813–3821 (2006).doi: 10.1121/1.22000139
Heimann, D., Salomons, E.: Testing meteorological classifications for the prediction of long-term average sound levels. Appl. Acoust. **65**, 925–950 (2004). doi:10.1016/j.apacoust.2004.05.001
Heimann, D., Bakermans, M., Defrance, J., Kühner, D.: Vertical sound speed profiles determined from meteorological measurements near the ground. Acta Acust. **93**, 228–240 (2007)
Heimann, D., Schäfer, K., Emeis, S., Suppan, P., Obleitner, F., Uhrner, U.: Combined evaluations of meteorological parameters, traffic noise and air pollution in an Alpine valley. Meteorol. Zeitschrift **19**, 47–61 (2010). doi:10.1127/0941-2948/2010/0426
Heimann, D., Käsler, Y., Gross, G.: The wake of a wind turbine and its influence on sound propagation. Meteorol. Zeitschrift **20**, 449–460 (2011). doi:10.1127/0941-2948/2011/0273
ISO 9613-2: Acoustics—attenuation of sound during propagation outdoors. Part 2: General method of calculation (1996)
Kästner, M., Heimann, D.: Effect of atmospheric variability and aircraft flight parameters on the refraction of sonic booms. Acta Acust. **96**, 425–436 (2010). doi: 10.3813/AAA. 918295
Neise, W. (ed.): Lärmoptimierte An- und Abflugverfahren. Zusammenfassender Schlussbericht (2007)
Ostashev, V.: Acoustics in Moving Inhomogeneous Media. E.&F.N. Spon, London (1997)
Salomons, E.: Computational Atmospheric Acoustics. Kluwer, Norwell (2001)
Salomons, E., Blumrich, R., Heimann, D.: Eulerian time-domain model for sound propagation over a finite-impedance ground surface. Comparison with frequency-domain models. Acta Acust. **88**, 483–492 (2002)

Aircraft Wake Vortices: From Fundamental Research to Operational Application

14

Frank Holzäpfel and Thomas Gerz

Abstract

Aircraft trailing vortices constitute both a kaleidoscope of instructive fluid dynamics phenomena and a challenge for the sustained development of the safety and capacity of the air-transportation system. This section gives an overview of the wake vortex issue commencing at its historical roots, proceeding with a sketch of the nature and characteristics of wake vortices resulting from field measurement and numerical simulation, and concluding with a depiction of the design and performance of wake vortex simulation systems established for the prediction of dynamic aircraft separations in different flight phases and for sensitivity and risk analysis.

14.1 The Wake Vortex Phenomenon

As an unavoidable consequence of lift, *aircraft* generate a pair of counter-rotating vortices, the so-called wing-tip vortices, aircraft trailing vortices, or aircraft wake vortices. Already one century ago, the significance of these vortices was recognized by Lanchester (1907) and Prandtl (1918) during their pioneering search for a theory of human flight. The conceived Lifting-Line Theory recognizes that the

F. Holzäpfel (✉) · T. Gerz
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Frank.Holzaepfel@dlr.de

T. Gerz
e-mail: Thomas.Gerz@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_14,

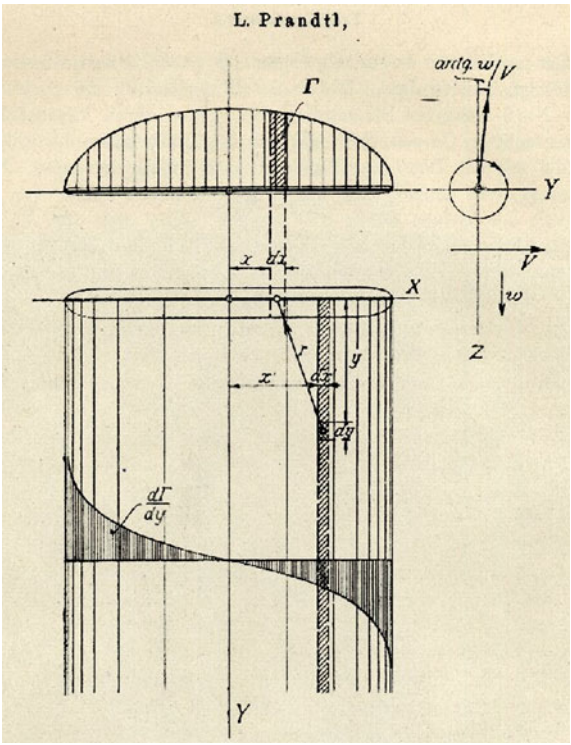

Fig. 14.1 Prandtl's sketch illustrating elliptical circulation distribution along the lifting line (*above*) and the detachment of differential vortex filaments (*below*) (reprinted from Prandtl (1918))

strength of the bound vortex[1] expressed as its *circulation*, Γ, decreases from the wing's root towards its tips (see Fig. 14.1). As a consequence, vorticity detaches from the wing continuously and forms a trailing *vortex* sheet. Driven by self-induction, this vortex sheet rolls up at its ends and generates the trailing vortices. The theory further elucidates the generation of induced drag, which is zero for an airfoil of infinite span. For wings of finite span, the induced drag arises from an effective reduction of the wing's attitude which is caused by the downwash induced ahead of the wing by the trailing vortices. The induced drag roughly equals the kinetic energy of the wake vortices and can be understood in analogy to the *wave* drag experienced by a ship.

Further descriptive explanations for *wake vortex* formation are: (i) The *pressure* difference between suction side and pressure side of the wing accelerates fluid around the wing tips and thus initiates a pair of counter-rotating vortices trailing behind. (ii) Another demonstrative approach uses the notion that the aircraft wings deflect freestream flow downwards, imparting a downwards-directed momentum onto the fluid. At the edges of the generated uniform downwash, a shear layer

[1] The bound vortex is a hypothetical vortex filament located on a lifting line which represents a straight wing. In a uniform flow perpendicular to its axis, the bound vortex experiences a lift force according to the Kutta-Zhukhovski lift theorem.

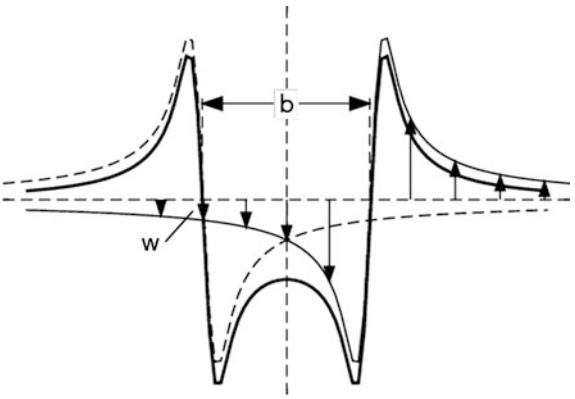

Fig.14.2 Schematic of vertical velocity profiles of two counter-rotating vortices (*fine lines*) separated by b and corresponding vertical velocity envelope (*bold line*)

develops which rolls up into the wake vortex pair. (iii) Finally, the first Helmholtz vortex theorem states that a vortex tube cannot end within the fluid. It must either end at a solid boundary or form a closed loop. For trailing vortices, the closed loop consists of the bound vortex along the wing; it is continued by the trailing vortices and completed by the starting vortex.

The strength of wake vortices is usually expressed by their circulation, Γ. When the forces which act on the aircraft are in balance, the initial circulation corresponds approximately to

$$\Gamma_0 = \frac{W}{(\pi/4)\,\rho\,B\,V}$$

Thus, the strength of the vortices is proportional to the weight of the aircraft, W, and inversely proportional to the air density, ρ, the wing span, B, and the flight velocity, V. Interestingly, the circulation achieves similar magnitudes for both a cruising aircraft at high altitude and an aircraft during approach to an *airport* at low altitude because the inverse variation of flight velocity and air density roughly compensate each other.

As depicted in Fig. 14.2, the vortices mutually induce a descent speed which amounts to

$$w = \frac{\Gamma}{2\pi\,b}$$

where b denotes the vortex separation. Initial descent speeds of trailing vortices of commercial aircraft range from 1 to 2 m/s and under favorable conditions maximum descent distances may potentially reach out more than 600 m (Misaka et al. 2012). During the descent the vortex pair transports fluid within an oval-shaped streamline which connects two stagnation points midway above and below the vortices, respectively. The motion at external points corresponds to the inviscid flow around an equally-shaped rigid body. This quasi-inviscid flow topology

explains the low drag of the descending oval and the sizeable descent distances traveled by wake vortices without environmental disturbances.

Depending on meteorological conditions, wake vortices may persist for several minutes and, therefore, may pose a potential risk to aircraft following behind. Serious problems with wake vortices were first recognized back in the 1970s when the Boeing 747 came into service. To avoid wake-vortex encounters, the Federal Aviation Administration (FAA) and the International Civil Aviation Organization (ICAO) established separation standards for consecutive aircraft. These separation distances vary between 3 and 6 nautical miles for different combinations of a 3×3 weight class matrix for preceding and follower aircraft on approach. Recently, ICAO has initiated a process to harmonize wake turbulence separation standards starting with the introduction of six weight classes instead of three, and with the final aim to establish dynamic pair-wise separations that will vary with *atmospheric* conditions and aircraft performance.

The incentive for today's wake vortex research still rests on the empirically motivated separation standards between consecutive aircraft introduced in the 1970s. These standards may heavily degrade aviation efficiency when *traffic* congestion limits airport *capacity* during landing and take-off. A wake vortex advisory system that allows current aircraft separations to be relaxed under favorable *weather* conditions, whilst keeping *safety* at least at the same level, may provide significant economizations. En route, the reduction of vertical aircraft separations to 1 000 ft in RVSM airspace (Reduced Vertical Separation Minimum) has led to a noticeable but not really known increase of encounters. With the deployment of new operational concepts envisaged by *SESAR* and NextGen, the possibility of encountering wake vortices in the upper *atmosphere* might increase. There is both a need to estimate future encounter frequencies and to develop systems that allow avoidance of encounters based on wake vortex prediction en route. Also, the increasing diversity of aircraft types (e.g., the advent of Very Light Jets) and the introduction of new very large aircraft like the twin-deck A380 or the B747-8 have prompted extensive wake vortex investigations.

The development of sustainable wake vortex advisory systems requires an interdisciplinary effort where disciplines like fluid dynamics, *atmospheric physics*, *aerodynamics*, and flight guidance must cooperate closely. The final acceptance of such systems though, requires early cooperation among a much larger community including aircraft manufacturers, *air traffic* control, airport service providers, *aviation* authorities, airlines, and pilots.

The wake vortex research presented here was conducted in the framework of the DLR projects Wirbelschleppe (1999–2007) and Wetter und Fliegen (2008–2012) and benefited from the EU projects ATC-Wake, AWIATOR, CREDOS, C-Wake, FAR-Wake, FLYSAFE, I-Wake, S-Wake, WakeNet, WakeNet2-Europe, and WakeNet3-Europe.

14.2 Decoding Wake Vortex Physics

A principal objective of wake vortex research is first to understand and then predict, or even accelerate, vortex decay. However, the longevity of wake vortices inhibits investigations of vortex decay in wind tunnels and complicates them in facilities like catapults and large towing tanks. The most effective and indispensable access to real wake vortex behavior is provided by lidar measurement techniques that trace full-scale wake vortices in the *free atmosphere*. Here, environmental conditions that can neither be controlled nor reproduced have a strong impact on vortex evolution. Furthermore, *advection* of the vortices beyond the observation domain and intrinsic difficulties connected with the interpretation of the complex velocity fields of eroded wake vortices constrain detailed insights into vortex decay mechanisms.

Complementary, high-resolution numerical *simulation* constitutes an essential resource for elaborating comprehensive understanding of vortex decay characteristics. The great advantage is that all desired quantities are readily available for analysis. On one hand, numerical simulations of wake vortices in the atmosphere should be conducted in large domains in order to allow for the formation of large atmospheric eddies; on the other hand, the simulations suffer from limited resolution in the vortex core region. Whereas *direct numerical simulation* (DNS) only reaches small *Reynolds numbers*, it is the type of subgrid-scale closure that controls vortex core evolution in *large-eddy simulation* (LES). Neither approach meets the complex interaction of *turbulence* and rotation at high Reynolds-number flows together with the peculiarities of specific aircraft configurations and environmental conditions. Such limitations will persist despite the enormous increase of computational power and the development of smart *numerical methods*. Full-scale *experiments* remain both challenging and mandatory to ensure the validity of the conclusions drawn from numerical simulations.

Flight experiments

The topic of aircraft wake vortices entered the Institute of Atmospheric Physics (IPA) in 1990 when a nationwide project was initiated by IPA to assess the impact of aircraft *emissions* on the atmosphere (Schumann 1990). Originally motivated by measurements of the constituents of exhaust *jets*, numerous flights in the wakes of commercial and *research aircraft* have been performed over the years. From five-hole-probe velocity measurements performed with the DLR research aircraft *Falcon* chasing B737, A340, A310, B727, and VFW614 (*ATTAS*) aircraft at distances ranging from 40 m to 15 km, power density spectra were established (see Fig. 14.3, Holzäpfel et al. (2001)). The in situ observations reveal that the intensity of aircraft-induced turbulence may vary by several orders in magnitude. Maximum turbulence was found in the axial velocity component, u, when the Falcon immerged into the exhaust jet region at small distances. Two spectra obtained from measurements in the undisturbed atmosphere are also included for comparison.

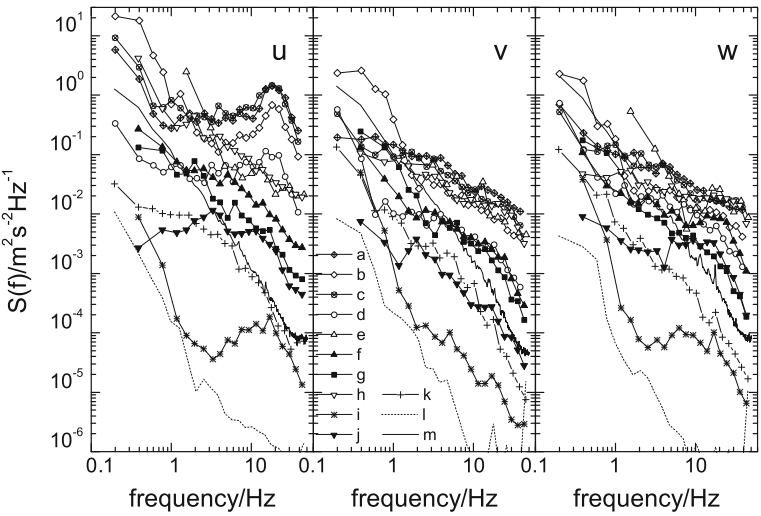

Fig. 14.3 Power density spectra established from in situ measurements behind different cruising aircraft and in free atmosphere. Open symbols denote near-field and closed symbols far-field data. **a** B737-300, $\Delta x = 43–57$ m, jet; **b** B737-300, $\Delta x = 40–60$ m, wake; **c** A340, short distance, jet; **d** A340, short distance, wake; **e** A310-300, $\Delta x = 100–150$ m; **f** A310-300, $\Delta x = 450–600$ m; **g** A310-300, $\Delta x = 800$ m; **h** VFW614, $\Delta x = 100$ m, *center*; **i** VFW614, $\Delta x = 80$ m, 10–20 m *left*; **j** VFW614, $\Delta x = 1.6$ km; **k** B727, $\Delta x = 15$ km; **l** free atmosphere, $z = 8.2$ km; **m** convective boundary layer, $z = 1$ km (reprinted with permission from Holzäpfel et al. (2001))

Lidar

Already in 1983 first *lidar* wake vortex measurements were conducted at Frankfurt airport (Köpp 1994). At that time a 10.6 μm continuous-wave Doppler lidar was applied which originally has been developed for investigations of *wind* and turbulence in the atmospheric *boundary layer*. The results of these pioneering investigations were used for the development of the wake vortex warning system of the Deutsche Flugsicherung (German air traffic control) at Frankfurt airport which, however, never became operational.

Over the years many measurement *campaigns* at airports in Frankfurt, Istres, Munich, Oberpfaffenhofen, Tarbes, and Toulouse were to follow which saw various developments and applications. During the WakeOP campaign in 2001, together with colleagues from ONERA and DERA three continuous-wave lidar systems were deployed at Oberpfaffenhofen airport to simultaneously measure wake vortices generated by the ATTAS aircraft (Köpp et al. 2003). Triangulation methods were applied in order to estimate vortex core locations and vortex separations with high accuracy. Later it was demonstrated that, despite its long pulse length of about 100 m, a 2 μm pulsed Doppler lidar can be modified for *long-range* wake-vortex characterization (Köpp et al. 2004). The pulsed lidar allows for observations over long periods from the moment of vortex generation to a progressed state of vortex decay. High measurement accuracies of vortex trajectories and circulation can be

Fig. 14.4 A380 wake vortex measurement campaign at Oberpfaffenhofen airport

achieved by a time-consuming interactive data processing algorithm. Moreover, the lidar may also provide profiles of wind velocity and turbulence dissipation rate (Smalikho et al. 2005); both quantities that strongly impact wake vortex fate.

With the Airbus A380, an aircraft with a wider span and greater weight than previously operating transport aircraft entered airline service in October 2007. In order to determine appropriate separation standards for the A380, an unprecedented test program taking 388 flight hours was accomplished by Airbus. In the time frame from May 2005 to December 2007 IPA performed over 600 ground-based lidar runs at different airports, establishing a unique high-quality data base of wake-vortex characteristics of the A380 and reference aircraft like the A340 or the B747. On 23 August 2006 hundreds of curious onlookers watched the A380 and a B747 from Lufthansa circling over the Munich area at low altitude when wake vortices shed with different aircraft configurations were measured by lidar at Oberpfaffenhofen airport (see Fig. 14.4).

The comprehensive test program culminated in innovative airborne lidar measurements on board the Falcon research aircraft (see Fig. 14.5; Rahm and Smalikho (2008)). For seeding the wake vortices evolving in clear air, either smoke generators were mounted on the wings or *kerosene* was injected into the exhaust jets. Direct comparison of wake vortices generated by the A380 and the A340 flying wing tip to wing tip enabled the impact of variations in the atmospheric *environment* to be minimized. The analysis of these lidar measurements led to a reduction of A380 separations from the preliminary value of 10 NM for all follower aircraft to 6, 7, and 8 NM for following heavy, medium, and light weight class aircraft, respectively. The 2 NM penalty over heavy leading aircraft includes a rounding up to the nearest mile and an *uncertainty* allowance of 1 NM.

Large-eddy simulation

Highly resolving large-eddy simulations (LES) conducted on *supercomputers* provide valuable insights in the *physics* of wake vortex behavior under various

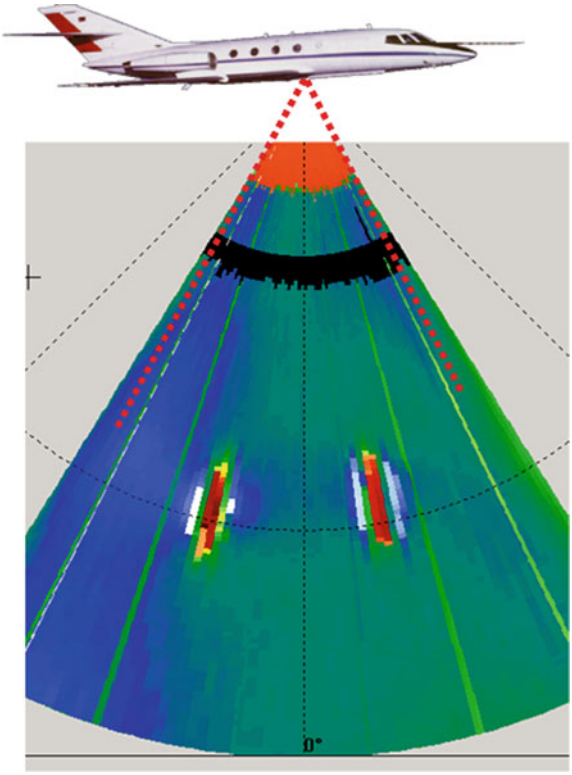

Fig. 14.5 Simultaneous measurement of A380 (*right*) and A340 (*left*) wake vortex pairs with airborne lidar on Falcon aircraft

environmental conditions and as such supply indispensable guidance for the development of simple wake vortex prediction *models* which constitute an essential part of wake vortex advisory systems. For almost two decades the in-house developed codes LESTUF and MESOSCOP have been used to conduct investigations of wake vortex evolution in turbulent, stably stratified, and sheared environments as well as in *convective boundary layers* (Holzäpfel et al. 2000; Hofbauer 2003; Holzäpfel et al. 2003; Hennemann 2010). In the context of assessing the impact of aircraft emissions on the atmosphere in the beginning of the 1990s, the entrainment of aircraft exhaust in the wake vortices, its subsequent detrainment when the vortex pair descends in the atmosphere, and its mixing and final distribution were studied by means of LES using the LESTUF code and comparing the results to in situ *measurements* (Gerz et al. 1998). Exemplarily for other pioneering LES of wake vortices reference is made to Robins and Delisi (1998) and Proctor and Switzer (2000). In 1998 a DLR project named Wirbelschleppe (German for wake vortex) was initiated at the institute to meet a request from Airbus Industries to understand the physics of wake vortices caused by very large transport aircraft (the later A380) and to explore *mitigation* options (Gerz et al. 2002). The LESTUF code played a key role in that project. For

example, the two-phase decay of the vortices was discovered by LES and later confirmed by lidar measurements.

Since 2010 LES of wake vortices has mostly been conducted using MGLET, which is a fourth-order-accurate finite volume solver of the incompressible *Navier–Stokes equations* employing the Lagrangian dynamic *subgrid scale* model (Manhart 2004). The numerical experiments are typically conducted in a Cartesian grid with 1 m resolution and domain sizes of 400 × 512 × 512 m^3. Larger grid sizes and higher resolutions also with adapted grid spacing have been employed in order to better resolve the vortex cores and the surface boundary layer or to realize large turbulent length scales of the ambient flow. For the massive parallel simulations, up to 1 024 processors are used. In the following a few recent LES results are introduced as examples of the comprehensive achievements in this field.

At cruise altitude, winds may be strong but the prevailing turbulence is typically very weak. At these altitudes stable thermal *stratification* typically limits wake vortex *lifetimes* to a maximum of about 3 min. Figure 14.6 depicts wake vortex evolution in a neutrally stratified and weakly turbulent environment where long-lived wake vortex rings may form. The vortex rings feature intriguing phenomena both visible in the photograph (above) and in the LES (below): after linking of the vortex pair, an elongated vortex ring is established and helical instabilities propagate along the vortices. Later, bone-shaped vortex rings transform into the shape of an "8" followed by a double ring phase (Misaka et al. 2012).

In order to establish environmental turbulence with large turbulent eddies, LES have been run in domain sizes up to 1 024 × 1 024 × 1 024 m^3 with over 1 billion grid points. These simulations indicate that with increasing turbulence intensity and increasing turbulence integral length scales the vortex topology is becoming more complex and the classical shape of Crow *instability* and ring formation is getting lost in favor of superimposed random deformations caused by large-scale turbulent eddies (Hennemann and Holzäpfel 2011).

A particular risk prevails during final approach to an airport, where the vortices cannot descend below the flight path, but tend to rebound due to interaction with the ground. Moreover, the possibilities of the pilot to counteract the imposed rolling moment are restricted due to the low height of the aircraft above ground. Figure 14.7 left shows the interaction of the *primary wake* vortices with the turbulent structures generated by some crosswind at the ground surface. At a vortex age of 28 s the vorticity sheet generated by the lee (rear) vortex detaches from the ground and starts rotating around the primary vortex. Triggered by crosswind streaks the secondary vorticity sheet transforms into so-called omega loops wrapping around the primary vortices and initiating vortex decay. Under unfavorable crosswind conditions the rebounding luff vortex may hover over the runway directly in the flight corridor of a landing aircraft.

The introduction of an obstacle at the ground surface (see Fig. 14.7 right) may generate powerful secondary vorticity structures that substantially accelerate vortex decay in the critical area close to the *threshold* where most vortex encounters occur. Such a setup specifically exploits properties of vortex dynamics to accelerate wake vortex decay in ground proximity with the following characteristics: (i) early

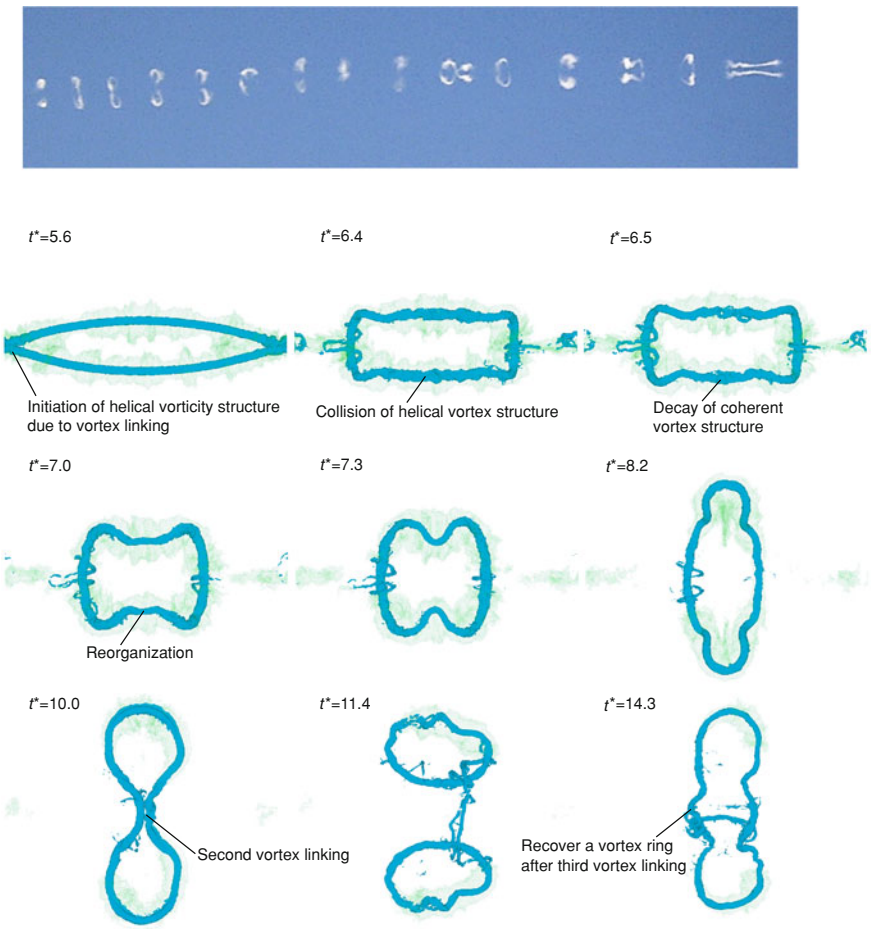

Fig. 14.6 *Above* various stages of vortex rings; flight direction from *left* to *right*, vortex age increasing from *right* to *left* (photo Sven Lüke, 16 Nov. 2006). *Below* LES of vortex ring formation in neutrally stratified and weakly turbulent environment. Reprinted with permission from Misaka et al. (2012). Copyright 2012, American Institute of Physics

detachment of strong omega-shaped secondary vortices, (ii) omega shape causes self-induced fast approach to the primary vortex, (iii) after the secondary vortex has looped around the primary vortex, it separates and travels both ways along the primary vortex, again driven by self-induction, (iv) the artificially generated secondary vortex connects to the regular ground effect vortex and thus obtains a continued supply of energy, (v) the highly intense interaction of primary and secondary vortices leads to rapid wake vortex decay independent of natural external disturbances. Field experiments will have to demonstrate the efficiency of this way

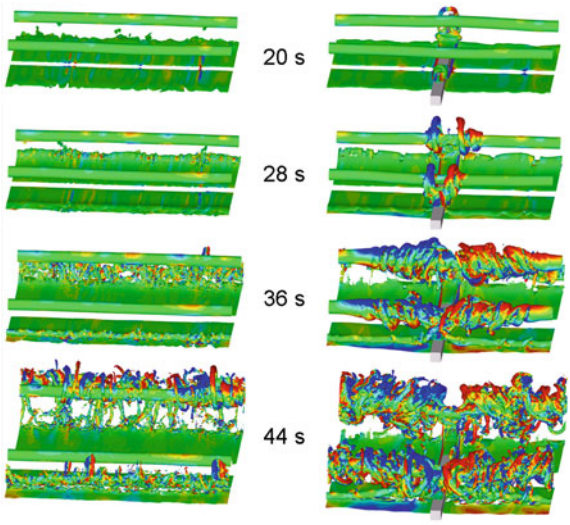

Fig. 14.7 Wall resolving LES of wake vortex evolution in ground proximity with turbulent crosswind (wind direction points into image plane). *Left* without and *right* with obstacle at the ground surface

to provoke premature vortex decay in the most critical flight phase prior to touchdown. Similarly as illustrated by this example for ground effect, most environmental effects like turbulence, stratification, and shear can be understood by analyzing the interaction of the vorticity of the primary vortices and secondary (*baroclinic*) vorticity structures (Holzäpfel et al. 2003). Also, effects like vortex bursting or the formation of vortex funnels can be explained by the interaction of *coherent* vorticity structures (Misaka et al. 2012).

14.3 Simplifying the Complexity

In order to build operational wake vortex advisory systems a number of complex issues (aircraft trajectories, aircraft parameters, meteorological conditions, wake vortex behavior, and safety areas) need adequate simplification and modeling to enable fast and reliable operation of these systems. The challenge is to find the optimal degree of simplification such that the devised models, on one hand, support real-time operation and, on the other hand, cover rare but safety-relevant events. Two ways of modeling aircraft trajectories, aircraft parameters, meteorological conditions, and safety areas for different applications are described in Holzäpfel et al. (2009b) and Holzäpfel and Kladetzke (2011). Exemplarily, the devised Probabilistic Two-Phase wake vortex decay and *transport* model (P2P, Holzäpfel (2003)) is described here in more detail. Other real-time wake vortex models have been developed by Greene (1986), Sarpkaya et al. (2001), and Jackson et al. (2001).

P2P was developed to guide the safe readjustment of aircraft separations during approach and landing. Later it was extended to departures and cruise flight. P2P considers all effects of the leading order impact parameters: aircraft configuration

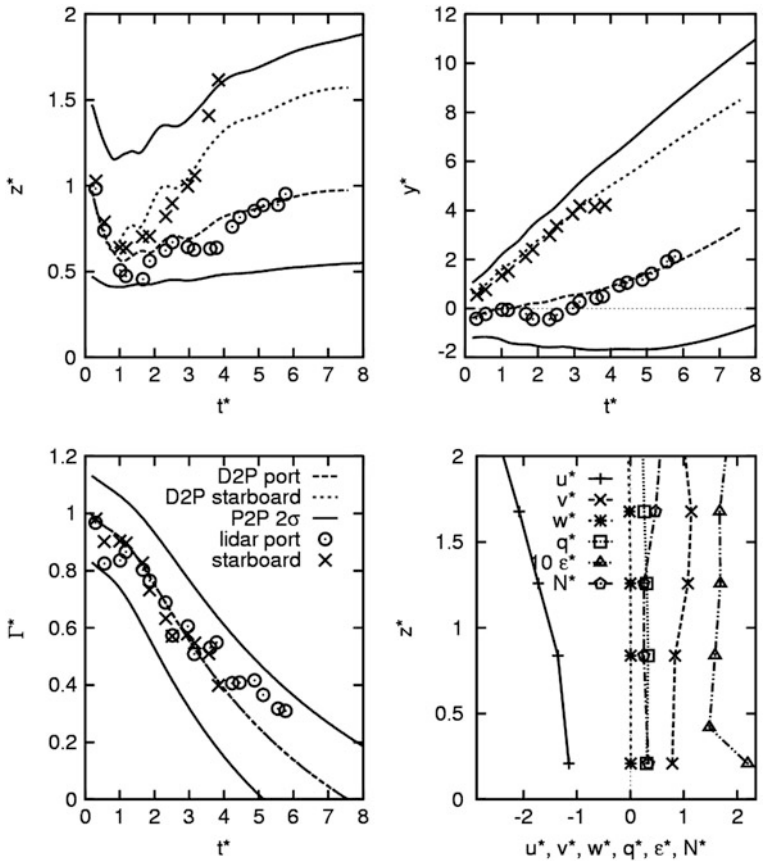

Fig. 14.8 Evolution of normalized vertical (z*) and lateral positions (y*) and circulation (Γ*) in ground proximity. Measurements by lidar (*symbols*) and predictions with P2P wake vortex model (*lines*). *Dashed* and *dotted lines* denote deterministic behavior; *solid lines* are probabilistic envelopes (95.4 %). The *bottom right* panel shows vertical profiles of normalized measured meteorological parameters (reprinted with permission from Holzäpfel et al. (2009b))

(span, weight, speed, location, heading, flight path angle, and bank angle), wind, wind shear, turbulence, *thermal stratification*, air density, and ground effect. The model predicts deterministic (mean) evolution (D2P) together with probabilistic envelopes for vortex trajectories and strength (P2P). Figure 14.8 shows exemplarily the asymmetric rebound of the wake vortices in a crosswind situation close to the runway (port vortex dashed, starboard vortex dotted). Most of the lidar measurements (symbols) are well situated within the predicted probabilistic envelopes (solid lines).

The model design rests on four pillars: (i) The model is formulated in normalized form where the characteristic scales are based on initial vortex separation, b_0, and circulation, Γ_0, leading to the *time scale* $t_0 = 2\pi\, b_0^2/\Gamma_0$, corresponding to

the time that the young vortices need to descend by one vortex separation. This normalization enables the application of the wake vortex model to a large variety of different aircraft types and environmental conditions.

(ii) The equations for circulation decay and vortex descent are built on an exact solution of the Navier-Stokes equations for a laminar decaying potential vortex. This equation has been extended to mimic the behavior of turbulent vortex pairs and (iii) has been adjusted to LES of wake vortex evolution by DLR and NASA. For the prediction of circulation, the concept of two-phase decay is pursued (see Fig. 14.8). The turbulent diffusion phase is followed by a rapid decay phase where the onset time of rapid decay and the respective decay rate depend on ambient turbulence, stratification, and vortex altitude above ground. The descent rate obeys a nonlinear dependence on circulation, which allows for a reduction of circulation without the reduction of the descent rate during the early vortex evolution and stagnating or even rebounding vortices with nonzero circulation in highly stable stratified environments.

Precise deterministic wake vortex predictions are not feasible operationally. Primarily, it is the nature of turbulence that deforms and transports the vortices in a stochastic way and leads to considerable spatio-temporal variations of vortex position and strength (see Fig. 14.6). Moreover, uncertainties about aircraft parameters and the temporal and spatial variability of environmental conditions must be taken into account. P2P is therefore designed to predict wake vortex behavior within defined confidence intervals. For this purpose P2P employs several static and dynamic uncertainty allowances, and (iv) it is calibrated with field measurement data such that the model output can be translated into probability density distributions of vortex position and strength (Holzäpfel 2006). Based on these probability density distributions, the model output can be adjusted to the required degree of probability. The computation time on a standard PC amounts to about 0.01 s.

So far, P2P has been validated against measurement data from two U.S. and more than six European field measurement campaigns employing in total over 10 000 individual cases. The vortex model has been applied to a large variety of aircraft types and flight phases including approach, landing, takeoff, departure, and cruise. A standardized scoring procedure indicates that deviations between measurement and prediction have been reduced continuously over the years, which reflects the progress in simulation, measurement, and data analysis techniques.

14.4 Wake Vortex Simulation Systems

This section describes the ground-based and airborne wake vortex advisory systems WSVBS and WEPS, respectively, and the airspace simulation system WakeScene for sensitivity analysis and risk *assessment*.

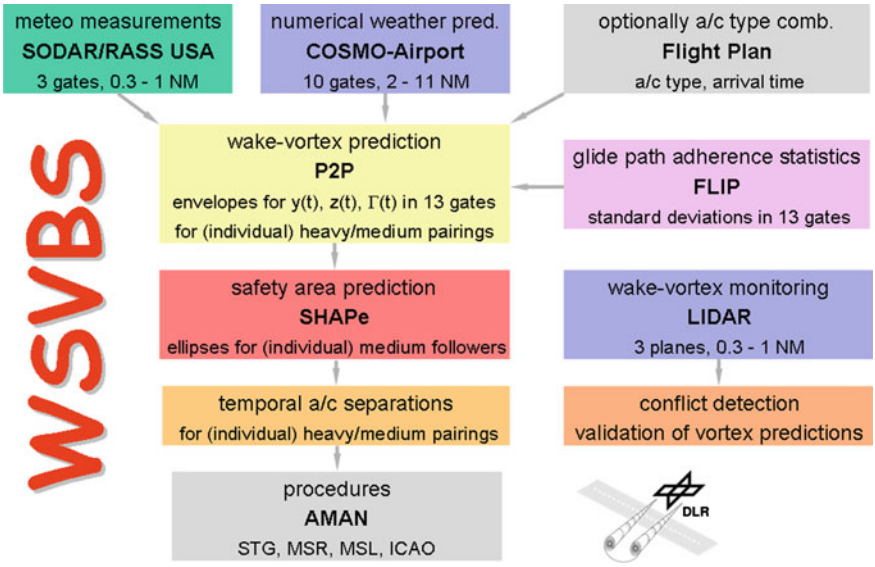

Fig. 14.9 Flowchart of the wake vortex prediction and monitoring system WSVBS

WSVBS

The Wake Vortex Prediction and *Monitoring* System (WSVBS) has been developed to tactically increase airport capacity for approach and landing on closely-spaced parallel runways as well as on single runways (Holzäpfel et al. 2009b). The two wake vortex advisory systems that were most influential for the development of the WSVBS are the Wake Vortex Warning System devised by Deutsche Flugsicherung (Gurke and Lafferton 1997) and the Aircraft Vortex Spacing System (AVOSS) developed by NASA (Hinton 1996). The WSVBS is thought to dynamically adjust aircraft separations dependent on weather conditions and the resulting wake vortex behavior without compromising safety. For this purpose it predicts wake vortex transport and decay and the resulting safety areas along the glide slope from final approach fix to threshold.

Figure 14.9 delineates the components of the WSVBS and their interplay. The bottleneck of runway systems prevails in ground proximity because there stalling or rebounding wake vortices may not descend below the flight corridor. Therefore, in that domain the best wake prediction *skill* is required, which is achieved here based on measurements of meteorological conditions with a SODAR/RASS system and an ultra sonic anemometer (USA). Because it is not possible to cover the whole glide slope with such *instrumentation*, the meteorological conditions in the remaining area are predicted with a *numerical weather prediction* system

(COSMO-Airport), leading to wake predictions with increased uncertainty bounds. Based on glide path adherence statistics (FLIP) the probabilistic wake vortex model P2P predicts upper and lower bounds for the position and strength of vortices generated by heavy aircraft. These bounds are expanded by the safety area around a vortex that must be avoided by follower aircraft for safe and undisturbed flight (SHAPe). Wake vortex and safety area predictions can be conducted optionally based upon either weight class combinations (heavy/medium) or individual aircraft type pairings. The instant when the safety areas do not overlap with the flight corridor defines the temporal aircraft separations. These temporal separations are translated into dynamic separations based on established procedures by the DLR arrival manager (AMAN). The lidar monitors the correctness of WSVBS predictions in the most critical gates at low altitude.

During a performance test for the closely-spaced parallel runway system at Frankfurt airport in winter 2006/2007, capacity-improving wake-vortex separation concepts of operation could have been used 75 % of the time and continuously applied for at least several tens of minutes (Gerz et al. 2009). It was found that the system ran stably and the predicted minimum separation times were totally confirmed by lidar measurements of wake vortex transport. From fast-time simulations the eventual capacity gain for Frankfurt was estimated to be 3 % taking into account the real traffic mix and operational constraints in the period of one month. Performance tests of dynamic pairwise separations were conducted at Munich airport in summer 2010 and spring 2011.

WEPS

An airborne Wake Encounter Prevention System (WEPS) is being developed in SESAR together with Airbus. Airborne prediction of wake vortex behavior shed by neighboring aircraft will allow for the identification of imminent encounters. The required aircraft and meteorological data are transferred by air-to-air data link exchange. A best guess of the meteorological situation will be obtained from the *fusion* of all available data sources. In case of a detected conflict, small-scale maneuvers or speed adjustments will avoid encounters.

WakeScene

The WakeScene (Wake Vortex Scenarios Simulation) package allows assessing the encounter probabilities and the related vortex strengths for different air traffic *scenarios* (Holzäpfel et al. 2009a; Holzäpfel and Kladetzke 2011). For arrivals the simulation domain extends from the final approach fix to the threshold; for departures it ranges from the runway along different departures routes up to heights of about 3 000 ft. Currently, WakeScene is extended to other phases of flight and, in particular, to approaches to closely-spaced parallel runways. WakeScene can be

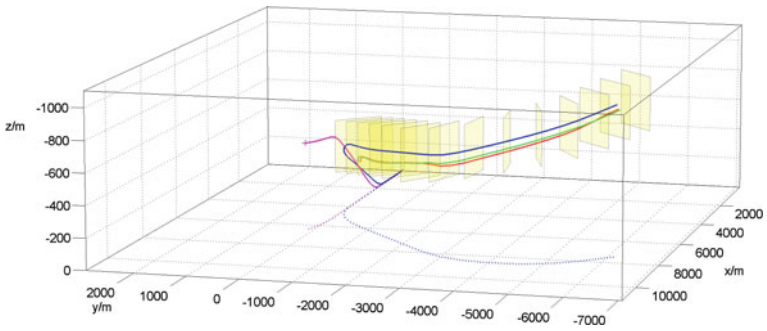

Fig. 14.10 WakeScene departure simulation. Perspective view of trajectories of wake-generating aircraft (*blue*) and follower aircraft (*magenta*) together with wake vortex positions (starboard vortex *green*, port vortex *red*). Projections of aircraft trajectories on vertical and horizontal planes and a number of gates used for wake vortex prediction are displayed for convenience (reprinted with permission from Holzäpfel et al. (2009a))

applied for sensitivity analysis, *optimization*, and risk assessment of new wake vortex advisory systems, air traffic control procedures, and aircraft types or even the elaboration of a new aircraft separation matrix (re-categorization). Figure 14.10 shows a snapshot of a WakeScene departure simulation with diverging flight routes at Frankfurt airport. Because of its late take-off and the steep climb rate, the follower aircraft (magenta trajectory) encounters wake vortices generated by the leading aircraft (blue trajectory) immediately after take-off at low altitude.

The modeling environment supports Monte-Carlo simulation as well as prescribed parameter variations and generates statistical evaluations. The package consists of elements that model traffic mix, aircraft trajectories, meteorological conditions, wake vortex evolution, and potential hazard area. The aircraft-*trajectory model* provides time, speed, attitude, mass, and lift of generator and follower aircraft at different gate positions (simulation planes, see Fig. 14.10), using point-mass aircraft models or the Advanced Flight Management System. A large number of environmental and aircraft specific parameters influence an aircraft trajectory and its deviations from a nominal flight path. The meteorological data base comprises one-year of statistics of realistic meteorological conditions (more than 1.3×10^6 vertical profiles) for the Frankfurt terminal area which were produced with a limited area *weather forecast* model system. Based on vertical *profiles* of environmental conditions and aircraft parameters, the P2P wake vortex model simulates the development of wake vortex trajectories, circulation, vortex core radius, and the attitude of wake vortex axes. The hazard area module defines an area of interest around the wake vortex. When this area of interest is penetrated by the follower aircraft, this is considered to be a "potential encounter." Different options exist for the area of interest definition: a simple circle with 50 m radius around each vortex, or the more differentiated approach of Simplified Hazard Areas (SHA), which are dynamically adjusted according to

vortex strength and aircraft pairing and designed to ensure operationally safe flight outside of the SHA (Schwarz and Hahn 2006). In cases with potential wake encounters, all relevant parameters can be provided to the Airbus tool VESA (Vortex Encounter Severity Assessment), which may subsequently perform detailed investigations of encounter severity (Kauertz et al. 2012).

Comprehensive sensitivity analyses have been conducted to investigate the wake vortex encounter probabilities for crosswind departure scenarios within the EU-project CREDOS (Holzäpfel and Kladetzke 2011). It was found that for crosswinds exceeding 8 kt at 10 m height, aircraft separations could be reduced from 120 to 60 s. Diverging departure routes and veering of the wind with altitude (Ekman spiral) impede the use of lower crosswind thresholds.

14.5 Assessment and Outlook

The steady development of measurement skills with lidar and numerical simulation with LES has brought about detailed understanding of the effects of meteorological conditions and the ground on wake vortex transport, decay, and topology. Meanwhile, even very specific and rare effects like vortex bursting and the generation of vortex funnels can be simulated and explained. These insights went into the development of a fast, robust, and validated wake vortex prediction model (P2P). Employing further models and data bases for aircraft trajectories, aircraft parameters, meteorological conditions, and safety areas, wake vortex advisory systems for landing, departure, and cruise, for single runways and closely-spaced parallel runways, for aircraft weight class combinations and dynamic pairwise separations have been devised (WSVBS, WEPS). Performance tests at Frankfurt and Munich airports have demonstrated the functionality and potential capacity gains of the systems. Software packages for the optimization and for the elaboration of safety cases for these systems are at hand (WakeScene).

Nevertheless, important research needs persist. For example, simulations and field measurements indicate that approaching aircraft frequently fly close to or even through not fully decayed wake vortices in ground proximity. In order to develop the most efficient and safe wake vortex advisory systems for approach and landing, the mechanisms which secure the safety of current operations must be understood and considered appropriately. Possibly, the suggested method to provoke premature vortex decay during the most critical flight phase prior to touchdown will further enhance the reliability and capacity gain of wake vortex advisory systems. Although these systems are basically available they require further refinement and adjustment and the elaboration of safety cases before they can be implemented at airports or in aircraft. Other future activities are estimation of the probability and severity of encounters in the upper atmosphere in the coming SESAR and NextGen airspace. Further, an initialization method for spatial LES is being developed where the highly resolved flow around

detailed aircraft geometries is flown through the simulation domain. This will enable a major step forward in understanding the impact of specific aircraft characteristics or flight maneuvers on wake topology.

References

Gerz, T., Dürbeck, T., Konopka, P.: Transport and effective diffusion of aircraft emissions. J. Geophys. Res. **103**, 25905–25913 (1998). doi:10.1029/98JD02282

Gerz, T., Holzäpfel, F., Darracq, D.: Commercial aircraft wake vortices. Prog. Aerosp. Sci. **38**, 181–208 (2002). doi:10.1016/S0376-0421(02)00004-0

Gerz, T., Holzäpfel, F., Gerling, W., Scharnweber, A., Frech, M., Kober, K., Dengler, K., Rahm, S.: The wake vortex prediction and monitoring system WSVBS Part II: performance and ATC integration at Frankfurt airport. Air Traffic Control Quart. **17**(4), 323–346 (2009)

Greene, G.C.: An approximate model of vortex decay in the atmosphere. J. Aircraft **23**(7), 566–573 (1986)

Gurke, T., Lafferton, H.: The development of the wake vortex warning system for Frankfurt airport: theory and implementation. Air Traffic Control Quart. **5**(1), 3–29 (1997)

Hennemann, I.: Deformation und Zerfall von Flugzeugwirbelschleppen in turbulenter und stabil geschichteter Atmosphäre. Dissertation, DLR-Forschungsbericht 2010–21, pp. 146 (2010)

Hennemann, I., Holzäpfel, F.: Large-eddy simulation of aircraft wake vortex deformation and topology, Proceedings of the Institution of Mechanical Engineers, Part G. J. Aerosp. Eng. **225**(12), 1336–1350 (2011). doi:10.1177/0954410011402257

Hinton, D.A.: An Aircraft Vortex Spacing System (AVOSS) for Dynamical Wake Vortex Spacing Criteria, The Characterization and Modification of Wakes from Lifting Vehicles in Fluids. *AGARD*, CP-584, 23.21–23.11.(1996)

Hofbauer, T.: Numerische Untersuchungen zum Einfluss von Windscherung und Turbulenz auf Flugzeugwirbelschleppen. Dissertation, DLR-Forschungsbericht 2003–01, pp. 115 (2003)

Holzäpfel, F., Gerz, T., Frech, M., Dörnbrack, A.: Wake vortices in a convective boundary layer and their influence on following aircraft. J. Aircraft **37**(6), 1001–1007 (2000). doi:10.2514/2.2727

Holzäpfel, F., Gerz, T., Baumann, R.: The turbulent decay of trailing vortex pairs in stably stratified environments. Aerosp. Sci. Technol. **5**(2), 95–108 (2001). doi:10.1016/S1270-9638(00)01090-7

Holzäpfel, F.: Probabilistic two-phase wake vortex decay and transport model. J. Aircraft **40**(2), 323–331 (2003). doi:10.2514/2.3096

Holzäpfel, F., Hofbauer, T., Darracq, D., Moet, H., Garnier, F., Ferreira Gago, C.: Analysis of wake vortex decay mechanisms in the atmosphere. Aerosp. Sci. Technol. **7**(4), 263–275 (2003). doi:10.1016/S1270-9638(03)00026-9

Holzäpfel, F.: Probabilistic two-phase aircraft wake-vortex model: further development and assessment. J. Aircraft **43**(3), 700–708 (2006). doi:10.2514/1.16798

Holzäpfel, F., Frech, M., Gerz, T., Tafferner, A., Hahn, K.-U., Schwarz, C., Joos, H.-D., Korn, B., Lenz, H., Luckner, R., et al.: Aircraft wake vortex scenarios simulation package—WakeScene. Aerosp. Sci. Technol. **13**(1), 1–11 (2009a). doi:10.1016/j.ast.2007.09.008

Holzäpfel, F., Gerz, T., Frech, M., Tafferner, A., Köpp, F., Smalikho, I., Rahm, S., Hahn, K.-U., Schwarz, C.: The wake vortex prediction and monitoring system WSVBS—Part I: design. Air Traffic Control Quart. **17**(4), 301–322 (2009b)

Holzäpfel, F., Kladetzke, J.: Assessment of wake vortex encounter probabilities for crosswind departure scenarios. J. Aircraft **48**(3), 812–822 (2011). doi:10.2514/1.C000236

Jackson, W., Yaras, M., Harvey, J., Winckelmans, G., Fournier, G., Belotserkovsky, A.: Wake vortex prediction—an overview, transport Canada, Montreal, Rept. TP 13629E (2001)

Kauertz, S., Holzäpfel, F., Kladetzke, J.: Wake vortex encounter risk assessment for crosswind departures. J. Aircraft **49**(1), 281–291 (2012). doi:10.2514/1.C031522

Köpp, F.: Doppler lidar investigation of wake vortex transport between closely spaced parallel runways. AIAA J. **32**(4), 805–810 (1994). doi:10.2514/3.12057

Köpp, F., Smalikho, I., Rahm, S., Dolfi, A., Cariou, J.-P., Harris, M., Young, R.I., Weekes, K., Gordon, N.: Characterization of aircraft wake vortices by multiple-lidar triangulation. AIAA J. **41**(6), 1081–1088 (2003). doi:10.2514/2.2048

Köpp, F., Rahm, S., Smalikho, I.: Characterization of aircraft wake vortices by 2-μm pulsed Doppler lidar. J. Atmos. Ocean. Technol. **21**(2), 194–206 (2004). 10.1175/1520-0426(2004)021<0194:COAWVB>2.0.CO;2

Lanchester, F.W.: Aerodynamics. Constable, London (1907)

Manhart, M.: A zonal grid algorithm for DNS of turbulent boundary layers. Comput. Fluids **33**(3), 435–461 (2004). doi:10.1016/S0045-7930(03)00061-6

Misaka, T., Holzäpfel, F., Gerz, T., Manhart, M., Schwertfirm, F.: Vortex bursting and tracer transport of a counter-rotating vortex pair. Phys. Fluids **24**(2), 25104-25101–25104-25121 (2012). doi: 10.1063/1.3684990

Prandtl, L.: Tragflügeltheorie. I. Mitteilung. *Nachrichten der K. Gesellschaft der Wissenschaften zu Göttingen*, Mathematisch-physikalische Klasse, 451–477 (1918)

Proctor, F.H., Switzer, G.F.: Numerical simulation of aircraft trailing vortices. In: Proceedings of the 9th Conference on Aviation, Range and Aerospace Meteorology, vol. 7.12, pp. 511–516 (2000)

Rahm, S., Smalikho, I.: Aircraft wake vortex measurement with airborne coherent Doppler lidar. J. Aircraft **45**, 1148–1155 (2008). doi:10.2514/1.32896

Robins, R.E., Delisi, D.P.: Numerical simulation of three-dimensional trailing vortex evolution in stratified fluid. AIAA J. **36**(6), 981–985 (1998). doi:10.2514/2.468

Sarpkaya, T., Robins, R.E., Delisi, D.P.: Wake-vortex eddy-dissipation model predictions compared with observations. J. Aircraft **38**(4), 687–692 (2001). doi:10.2514/2.2820

Schumann, U. (ed.): Air traffic and the environment—background, tendencies and potential global atmospheric effects. Lecture Notes in Engineering 60. Springer-V, (1990)

Schwarz, C.W., Hahn, K.-U.: Full-flight simulator study for wake vortex hazard area Investigation. Aerosp. Sci. Technol. **10**(2), 136–143 (2006). doi:10.1016/j.ast.2005.09.005

Smalikho, I., Köpp, F., Rahm, S.: Measurement of atmospheric turbulence by 2-μm Doppler lidar. J. Atmos. Ocean. Technol. **22**(11), 1733–1747 (2005). doi:10.1175/JTECH1815.1

Contrails: Visible Aviation Induced Climate Impact

15

Ulrich Schumann, Kaspar Graf, Hermann Mannstein and Bernhard Mayer

Abstract

Contrails are aircraft induced linear ice particle clouds. In ice-supersaturated air masses, contrails persist and develop into contrail cirrus. Aviation changes cloudiness also by soot emissions and other aerosols. In the global mean, contrail cirrus warms the atmosphere. The local warming induced by a contrail varies strongly relative to the global mean, and may be positive or negative, depending on aircraft, route, solar and Earth-atmosphere parameters. Here we summarize present knowledge on contrails with respect to aviation and atmospheric research.

U. Schumann (✉) · K. Graf · H. Mannstein · B. Mayer
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Ulrich.Schumann@dlr.de

K. Graf
e-mail: Kaspar.Graf@dlr.de

H. Mannstein
e-mail: Hermann.Mannstein@dlr.de

B. Mayer
e-mail: bernhard.mayer@lmu.de

B. Mayer
Ludwig-Maximilians-Universität München (LMU), Meteorological Institute Munich (MIM), Theresienstraße 37, 80333 München, Germany

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_15,

15.1 Introduction

"*Contrails*" (condensation trails, German *Kondensstreifen*) are thin, linear ice particle clouds often visible behind cruising *aircraft*. They form because emitted *water vapor* mixed with cold ambient air leads to local *liquid saturation*, *condensation* of water on *aerosols*, and subsequent *freezing* (see Fig. 15.1). *Linear contrails* can *persist* for hours and may evolve into diffuse *cirrus clouds* (Fig. 15.2). Aviation also changes cloudiness by introducing *soot* and other aerosols.

Contrails represent reproducible prototypes of cirrus clouds which are easier to understand scientifically and offer better chances for experimental investigations than natural cirrus. Contrails occur both behind propeller and jet driven aircraft. Contrails were first studied because of their *visibility*. Since the first observations of contrails in 1915, the investigation of contrail formation has led to important general insights into the atmosphere system, such as the detection of *ice supersaturation*, *homogeneous* and *heterogeneous* ice particle formation, *subvisible* cirrus, and *Brewer–Dobson circulation*.

In 1987, after detection of the ozone hole and growing concern about *anthropogenic* global *climate change*, the impact of aviation on UV radiation, air quality and climate became a topic of public discussion in Europe. A series of scientific research projects started in Europe in 1990 and a little later in the USA, at first extending earlier stratospheric research undertaken around 1970 in the context of supersonic transport. Initially, the main emphasis was on the chemical effects of carbon, nitrogen, sulfur and hydrogen oxides and particles on ozone. Since the 1990s, contrails were increasingly considered because of indications of the climatic importance of *aviation induced cloudiness* (Liou et al. 1990; Fahey et al. 1999; Minnis 2003).

Since 1990, this institute has conducted extensive research on aircraft impact on air composition and climate. Contrails have been investigated using several methods. They include in situ measurements, e.g., with the *Falcon* behind an Airbus A380 (Voigt et al. 2011), *satellite* observations of linear contrail cover (Meyer et al. 2007), satellite observations of cirrus cover in a region with a special *diurnal cycle* of air traffic allowing for identification of aviation induced cloudiness (Graf et al. 2009), *large-eddy simulations* of contrails (Unterstrasser and Sölch 2010), models which deduce contrail cover from the product of *potential contrail cover* and *air traffic* density for given meteorological input (Frömming et al. 2011), a global *climate model* that includes a process-based *subgrid scale* parameterization of contrail cirrus and computes the global coverage of contrail cirrus as well as the changes in natural cloudiness that they induce (Burkhardt and Kärcher 2011), and a contrail cirrus prediction tool (CoCiP) which simulates the properties of a large ensemble of contrails as a function of given air traffic tracks and meteorology, including approximate predictions of contrail cirrus cover and contrail climate impact (Schumann 2012). Here, we summarize present understanding of contrail properties based on these investigations.

Fig. 15.1 Contrail formation at threshold conditions. *Left* Contrail formation for fuels of different sulfur content, FSC (FSC for left/right engine in μg/g, respectively: 1.6/245, Busen and Schumann (1995)]. *Center* Same for 170/5500 (Schumann et al. 1996). *Right* Two aircraft at comparable flight conditions with different overall propulsion efficiency η [*left* < 0.24, *right* > 0.28, Schumann et al. (2000)]. (Photos: R. Busen, P. Thomas, U. Schumann, DLR)

Fig. 15.2 Example of persistent contrail cirrus (Photo Ch. König, Bildverarbeitungs-Büro König, BVBK, 2004; 0948 UTC 24 January 2004, 47.91°N, 11.13°E, SSE direction)

15.2 Contrail Formation

The thermodynamic explanation of *contrail formation* is similar to that of *fog* formation caused by people breathing in cold and humid air. A fog forms when, during mixing of warm and moist *exhaust* air (from a human lung or from an aircraft engine)

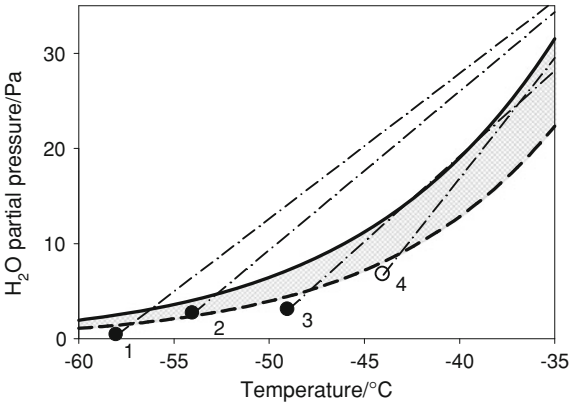

Fig. 15.3 Contrail examples in a diagram of water vapor partial pressure plotted against temperature. The *bold curves* represent the water vapor saturation pressure over liquid (full) and ice water (*dashed*). The dots represent examples of measured ambient humidity-temperature conditions. Details were published previously [SULFUR-2-ATTAS-C2 (Schumann et al. 1996), SULFUR-3-A310 (Petzold et al. 1997), POLINAT-2-A340-MOZAIC (Helten et al. 1999), and SUCCESS-#19 (Jensen et al. 1998b)]. *Full* and *open dots* represent cases where a contrail was definitively visible and not visible, respectively. The *dash-dotted lines* are the mixing lines representing the state of air parcels during mixing from engine exit (high temperature) to ambient conditions; their steepness G differs because of different pressure and propulsion efficiency values. The symbols 1–4 represent different contrail conditions as observed: 1: short-lived, 2: persistent, 3: threshold, 4: no contrail. Contrails form when the mixing line (*dash-dotted*), for given ambient conditions and steepness G, cuts the line of liquid saturation (as for cases 1–3). Contrails are persistent when forming in the *shaded region* between ice and liquid saturation (case 2). No contrail was observed for case 4 despite ice supersaturation during mixing. The contrail observed for case 3 could not be explained without accounting for finite overall propulsion efficiency η

with cooler outside air, the *relative humidity* in the air mixture reaches liquid saturation. The state in an air parcel during mixing follows approximately a straight line in a diagram plotting *partial water vapor pressure* against *temperature*, see Fig. 15.3. [Near engine exit, at temperatures around 600 K, far to the right in this figure, the mixing line deviates to higher humidity values because part of the energy is still to be converted by turbulent dissipation from kinetic jet energy to internal energy and because the specific heat capacity of the hot exhaust products is higher than at ambient conditions, see Schumann (2000).] The saturation pressures for liquid and ice saturation versus temperature follow approximately exponential curves. Hence, the *mixing line*, if steep enough, may be tangent to or cut a saturation curve. If liquid saturation occurs, fog forms. The fog starts evaporating when the mixture returns to subsaturation. At temperatures below about −40 °C, the droplets freeze quickly by *homogeneous nucleation* (Pruppacher 1995) and do not sublimate before reaching *ice saturation*. For an ideal gas, the ratio of partial pressure to absolute pressure equals the mixing ratio of water molecules per air molecules. For a given temperature, the relative humidity is the ratio of partial pressure to *saturation*

pressure. The saturation pressure over ice is lower than that over a liquid surface (less than 60 % below −50 °C).

Thermodynamic contrail formation was first understood by Ernst Schmidt in 1941. The explanation is surprisingly simple and independent of engine, jet and contrail particle details. Such details clouded earlier theories. The theory was soon verified experimentally by *Fritz W. Höhndorf* and Hans-Joachim aufm Kampe, in flight tests under the guidance of *Walter Georgii* at the Deutsche Forschungsanstalt für Segelflug (DFS), one of the forerunner institutions of this institute. Schmidt, later a professor at Munich Technical University, originally assumed that the ice particles in contrails form under equilibrium conditions. For low temperatures, this would imply ice saturation. From in situ measurements in contrails shortly behind aircraft, it was found already in 1941 that contrail formation requires liquid saturation in the exhaust plume (Höhndorf 1941; Schumann 1996). At about the same time, between 1942 and 1946, Alan W. Brewer and others in the United Kingdom and several teams in the United States came to similar conclusions. They also understood that ice supersaturation in ambient air is required to form long-lasting contrails. For humidity above liquid saturation the aircraft would likely fly in liquid water clouds. Hence, if the ambient state is within the shaded area in Fig. 15.3, a persistent contrail should form. Herbert Appleman published in 1953 a thermodynamic explanation similar to that of Schmidt and clearly identified local liquid saturation as necessary for contrail formation. Since then, the so-called *Schmidt–Appleman criterion* (see below) has been used to decide whether an atmosphere is cool and humid enough to let contrails form (Schumann 1996).

Test flights performed behind the DLR research aircraft *ATTAS* in 1994–1998 (Fig. 15.1) confirmed this criterion. The experiments also showed that the *threshold* temperature depends significantly on the fraction η of fuel energy used for aircraft propulsion of the aircraft. Figure 15.1c shows two aircraft flying wing by wing, where the older aircraft has no contrails while the more modern aircraft with higher *propulsion efficiency* η does cause a visible contrail. For a given fuel, this demonstrates that contrails cannot be avoided by improved engines (except by condensing water on a cooler before engine exit), but only by avoiding flights at contrail forming ambient conditions. The tests also showed that the *sulfate* emitted from fuel with typical sulfur content causes a weak increase in the number of ice particles forming and a hardly detectable impact (<0.4 K) on the contrail formation threshold temperature.

Observations of contrails behind aircraft during the Second World War indicated often less or shorter contrails for flights above the *tropopause*. This triggered Brewer to develop and fly an airborne *frost point* hygrometer to measure *humidity*. The air masses above the tropopause in the *lower stratosphere* were often found to have a low frost point, with very dry air. Contrails evaporate quickly if the ambient air is dry. This explained the shortness of high contrails. These measurements led Gordon M.B. Dobson and Brewer to postulate in 1946 what we now call *Brewer–Dobson circulation*, with *freeze-drying* of rising air at the cold tropical tropopause and subsidence of dry air at high extratropical latitudes. Hence, contrail research contributed to the detection of a major component of *atmospheric circulation* (Brewer 2000).

Besides contrails forming from engine exhaust water, contrails may also form aerodynamically by adiabatic cooling of air because of sudden pressure reduction along curved streamlines near curved parts of the aircraft (wings, propellers, etc.) when flying in humid air (Gierens et al. 2009; Kärcher et al. 2009b). Such *aerodynamic contrails* may also occur simultaneously with classical exhaust-water induced contrails and be indistinguishable when the contrails interfere with the wake vortices behind the aircraft.

The dynamics of young contrails interact with the engine jet, the aircraft *wake vortex*, and with ambient *turbulence*, *stratification* and *wind shear* (Lewellen and Lewellen 1996; Unterstrasser et al. 2008). The details are aircraft dependent (larger aircraft have higher fuel consumption, emit more water, more soot particles, and cause deeper wake vortex sinking, hence a thicker contrail, which transforms under shear into a wider contrail, etc.). The fate of the resultant contrail cirrus depends strongly on the number of ice crystals forming from soot and volatile aerosols at high supersaturation locally in the *jet phase* and on the number of ice particles surviving *adiabatic heating* in the initially sinking wake vortex. The more soot, the more ice particles form, which have larger optical depth and hence larger radiative impact, and may stay longer in the atmosphere because of less *sedimentation*.

15.3 The Schmidt–Appleman Criterion

The *Schmidt–Appleman criterion* (SAC) of contrail formation requires liquid saturation to occur locally in the plume of aircraft exhaust gases during *isobaric mixing* (at constant pressure) with cold ambient air. This condition is satisfied when the ambient temperature T is below a *threshold temperature* T_c. This condition is also satisfied when the *relative humidity* U for liquid saturation is above a critical humidity U_c. This humidity condition is computationally more efficient than the temperature criterion (Ponater 2002). To good approximation (Schumann 1996),

$$T_c = 46.46 + 9.43\ln(G-0.053) + 0.72\ [\ln(G-0.053)]^2,\ T_c \text{ is units of } ^\circ\text{C and } G \text{ in Pa/K}, \tag{15.1}$$

$$G = \frac{c_p p EI_{H_2O}}{(M_{H_2O}/M_{air})Q_{fuel}(1-\eta)},\ \eta = \frac{VF}{m_{fuel}Q_{fuel}},\ U_c = \frac{G(T-T_c)+p_{liq}(T_c)}{p_{liq}(T)} \tag{15.2}$$

Here, G is the steepness of the mixing line with specific heat capacity of air $c_p = 1004$ J/(kg K), ratio of molecular masses of water and air $M_{H2O}/M_{air} = 0.622$, ambient pressure p, water vapor emission index EI_{H2O} and combustion heat Q_{fuel} of the fuel used, and overall propulsion efficiency η of the aircraft at cruise. For *kerosene*, $EI_{H2O} = 1.23$ and $Q_{fuel} = 43.2$ MJ/kg. The *overall*

propulsion efficiency η relates the work performed by propulsion of an aircraft with thrust F and speed V relative to the combustion energy provided by a fuel with specific combustion heat Q_{fuel} at flow rate m_{fuel}. Typically, η varies from 0.2 (for older jet aircraft) to 0.35 (modern jet aircraft at cruise). Obviously, increasing efficiency η causes more contrails because of less heat emission into young contrails, with contrail formation at lower altitudes and higher ambient temperatures.

For kerosene driven aircraft with $\eta = 0.3$, the SAC implies contrail formation in the aviation *standard atmosphere* to occur above 8.4 km and below 14 km. For an absolutely dry standard atmosphere, this range starts above 10 km altitude. Contrails may also form near ground in very cold regions of the Earth atmosphere (Alaska, Siberia in winter). For regenerative fuels, which often contain a higher fraction of hydrogen than kerosene, the vertical altitude range of contrail formation is slightly larger. For threshold conditions, contrails become visible about one wing span behind the engines. For lower temperature, contrails can be seen forming already a few meters behind the engines.

The SAC has been verified experimentally (Busen and Schumann 1995; Jensen et al. 1998b; Kärcher et al. 1998; Schumann 2000). For example, Fig. 15.3 shows measured cases. The measurements show that contrails do not form when the mixing line exceeds ice saturation, but not liquid saturation.

15.4 Contrail Cirrus

Contrails are often observed to persist for hours (Schumann and Wendling 1990; Minnis 1998; Haywood et al. 2009). They spread and transform into *contrail cirrus* with considerable *cloud cover* and *optical depth*, e.g., Fig. 15.2. Contrails persist for ambient humidity above ice saturation. In fact, the detection of *ice supersaturated regions* in the atmosphere was triggered by contrail research (Gierens et al. 1999). Measurements show that the relative humidity in the upper troposphere often exceeds ice saturation (Jensen et al. 2001; Ovarlez et al. 2002). Here, the contrail ice particles grow in size and mass due to ambient humidity. During this process the humidity of the air inside the contrails, with high ice particle number concentrations, drops to near ice saturation. The total ice mass content in persistent contrails may be several orders of magnitude larger than the amount of water emitted from the aircraft. Persistent contrails may spread to several kilometers in width, mainly because of vertical wind shear (change of horizontal wind speed with height) (Freudenthaler et al. 1995; Jensen et al. 1998a).

Contrail cirrus form and persist in air that is ice-saturated, whereas natural cirrus often require high ice supersaturation to form. Therefore, contrail cirrus can persist in supersaturated air that is cloud-free, thus increasing high-cloud cover. In addition, persisting contrails may be embedded in thin cirrus clouds. Cirrus with low optical depth (<0.03) can be easily detected by lidar but is often not visible to human observers and therefore classified as subvisible (Kärcher et al. 2009a). Optical depth values of contrail cirrus have been reported in the literature reaching

from 10^{-4} to 2.3. Here the lower bound is observed by *lidar* (Immler et al. 2008; Iwabuchi et al. 2012). The upper bound may be topped in contrail *fallstreaks* (Atlas et al. 2006). Because of skewed *probability density function* the median value ($\sim$0.1) is lower than the mean value ($\sim$0.2). Thin cirrus and ice supersaturation may coexist for low ice particle number concentrations, in particular in rising air masses (Jensen et al. 2001). In air masses with high traffic density and high humidity, many contrails form, persist, intermingle, and overlay with other thin cirrus, lose their linear structure, and develop into amorphous contrail cirrus (Schumann 2012). Such widespread formation of persistent contrails in otherwise apparently clear skies or thin or subvisible cirrus is called a *contrail outbreak* (Duda et al. 2004).

Contrails are generally composed of ice crystals with trace amounts of exhaust products such as soot and sulfate. Most of the contrail particle mass is water from ambient air. Contrail formation and persistence does not require, as sometimes speculated in the internet, artificial spraying of chemicals. Hence, contrail observations do not prove the existence of so-called chemtrails.

The annual and global mean cover of linear contrails has been estimated from the *potential contrail cover* multiplied by a measure of air traffic (such as flight distance) and calibrated to satellite observations (Sausen et al. 1998). The potential contrail cover is the frequency that an air mass is conditioned to form persistent contrails, i.e., is cool enough for contrail formation and ice supersaturation. Global annual mean values of visible and total contrail cover vary over a considerable range of uncertainty, from 0.06 to 0.27 % (Frömming et al. 2011). The cover by contrail cirrus has been estimated from models and satellite data in a few studies to be larger than linear contrail cover by a factor of about 10 (Mannstein and Schumann 2005). Most recently, visible contrail cirrus cover has been assessed to be about 0.23 % (Burkhardt and Kärcher 2011).

15.5 Radiative Forcing by Contrails

The *global warming* induced by contrails from a fleet of aircraft in the global and annual mean can be computed using so-called linear response models for a given radiative forcing (Sausen and Schumann 2000; Ponater et al. 2006; Fuglestvedt et al. 2010). The *radiative forcing* (RF) is the change in net downward energy flux per unit surface for the case with contrails compared to the case without contrails. For the Earth-atmosphere system as a whole, the RF is the change of net energy flux at the top of the atmosphere. However, the change in global mean surface temperature is better related to the *stratosphere-adjusted* RF at the tropopause (Hansen et al. 2005). The climate system eventually responds to a change in RF with a global mean surface temperature increase, ΔT_s, when thermal equilibrium between the atmosphere and ocean is reached after several decades (Hansen et al. 2005; Ponater et al. 2005). Typically the response ΔT_s/RF is of an order of 1 K W^{-1} m^2. However, the global temperature response per unit contrail forcing

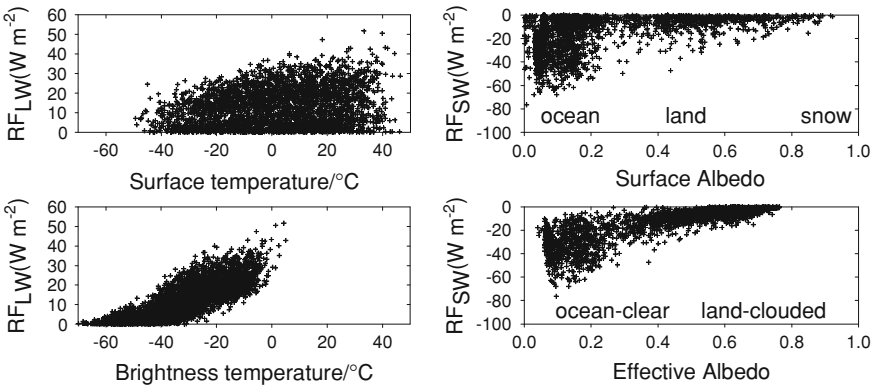

Fig. 15.4 Radiative forcing (RF) for contrail cirrus for unit contrail cover versus surface/brightness temperature and surface/effective albedo, at longwave (LW, *left*) and shortwave (SW, *right*) radiation wavelengths, for a wide set of contrail properties (550-nm-optical depth 0.2 to 0.5, six nonspherical habits), land and ocean surfaces, atmospheres with and without other clouds, and solar zenith angles (0–78°)

relative to the response to CO_2 forcing (the "*efficacy*") is less than one because of the high spatial and temporal variability of contrails compared to the more homogeneously distributed CO_2 disturbance (Ponater et al. 2006).

Contrail RF is composed of a longwave (LW) and a shortwave (SW) part which depend on several contrail, solar, and *Earth–atmosphere system* parameters (Meerkötter et al. 1999). Foremost, the RF depends on the integral product of contrail cover and its effective optical depth τ (typically in the middle of the visible solar spectrum at 550 nm wavelength; at terrestrial radiation wavelengths τ is smaller). For global RF estimates, contrail cirrus may be approximated as horizontally homogeneous *plane-parallel cloud* layers. Three-dimensional photon transport is important for narrow contrails and high solar zenith angles (Forster et al. 2012). Contrail RF has been computed using *libRadtran* (Mayer and Kylling 2005) for a large set of Earth atmosphere contrail cases (Schumann 2012). Figure 15.4 shows the computed RF per unit contrail cover as a function of surface and brightness temperature and surface and system albedo.

Longwave RF is positive (warming), depends on *terrestrial radiation* effective day and night, and is largest over a warm Earth-atmosphere system. As shown in Fig. 15.4, RF_{LW} correlates weakly with the surface temperature but more strongly with the *brightness temperature* of the atmosphere. A cool surface or an atmosphere with cool cloud tops below the contrails would radiate little terrestrial radiation to space anyway so that a contrail cannot much reduce this energy loss. The LW RF causes reduced cooling of the Earth surface compared to clear sky and heats mainly the lower contrail parts.

Shortwave RF is mostly negative (cooling), depends on *solar radiation* and is effective, therefore, during daytime only. Figure 15.4 shows that RF_{SW} correlates weakly with surface albedo but more strongly with the system *albedo*

of the atmosphere-surface system. RF_{SW} is largest for contrails over dark surfaces (e.g., cloud-free oceans). Over white surfaces like snow or over thick clouds, a contrail cannot much increase the albedo and hence cannot increase the reflected solar radiation. The SW RF cools the Earth surface. Hence, contrails may affect the *diurnal temperature range* at the Earth surface, though significant contributions from aviation have not been identified so far (Dietmüller et al. 2008).

Details of RF depend on the *optical properties* of the contrail *ice particles*, mainly their light extinction, absorption and forward scattering properties. Finally, RF depends on absorbing and reflecting clouds and gases above the contrails, in particular on the 550-nm-optical depth τ_c of cirrus above the contrail. All these properties make simple RF estimates difficult. The optical properties of contrail particles depend on their size, *habit*, orientation, and number-size distribution. Ice particles in young contrails are small (1–20 μm) and smaller than typical cirrus particles. Ice particles in contrails are nonspherical. Small crystals are randomly oriented. Many of the ice particles in young contrails are frozen droplets which may be described as *droxtals*. Droxtals have a multifaceted near-spherical shape. With increasing age, the shape and size of contrail particles may approach those of natural cirrus. Nonspherical ice particles scatter light more sideward than spherical ones, and hence cool more effectively, in particular for large solar zenith angles (more backscatter to space during mornings and evenings). Smaller particles interact more strongly with solar radiation than with terrestrial radiation, hence the SW effect (cooling) is more important for smaller particles than for larger ones. Measured *size distributions* of cirrus particles are often approximated by log-normal distributions, but vary a lot. Often size distributions are found in which a small fraction of the number concentration N of particles is very large and carries most of the ice volume V, while the more frequent smaller particles contribute most to the vertically projected particle cross-section area A. Hence, the optically *effective radius* r_{eff} differs from the *volume mean radius* r_{vol},

$$r_{eff} = 3\,V/(4\,A),\; r_{vol} = [3\,V/\,(4\pi N)]^{1/3},\; C = r_{vol}/r_{eff}.$$

C is important because it has linear impact on the optical depth τ for a given ice water volume V and number density N. Depending on the shape of the size distribution, the optically effective radius r_{eff} may be much larger than the volume mean radius r_{vol}, so that $C < 1$. On the other hand, for complex particle habits, the ratio V/A may become small so that $C > 1$ is possible as well. For young contrails, C varies from about 0.6 to 1.2 (Schumann et al. 2011b).

15.6 Approximate Radiative Forcing Dependencies

Despite the complexity of these relationships, the local contrail RF per unit cover can be approximated with reasonable accuracy by simple analytical functions (Schumann et al. 2012). These equations are given here to illustrate the major dependencies, first for the LW part:

$$RF_{LW} = [OLR - k_T(T - T_0)]\left\lfloor 1 - \exp\left(-\left(1 - \exp\left(-\delta_{lr} r_{eff}\right)\right)\delta_\tau \tau\right)\right\rfloor \exp(-\delta_{lc}\tau_c). \tag{15.3}$$

The first factor represents the change in *outgoing LW radiation* flux for opaque contrails, i.e., for $\tau \gg 1$. This term is usually positive, and the larger the lower the contrail temperature T compared to the effective brightness temperature T_B of the atmosphere without contrails. Here T_B does not appear explicitly, but in the outgoing longwave radiation $OLR = \sigma\ T_B^4$ at top of atmosphere, where $\sigma = 5.67 \times 10^{-8}$ W/(m^2 K^4) is the *Stefan-Boltzmann* constant. For typical values of OLR (150–300 W/m^2), T_B is about 235–270 K. The contrail temperature T ($\sim$200–240 K) is mostly below the brightness temperature of the atmosphere without contrails so that $RF_{LW} \geq 0$. The second factor accounts for the effective *emissivity* of the contrail. It depends on the effective optical depth in the thermal infrared (IR). The optical depth in the IR is smaller than in the solar range, in spite of enhanced absorption, because of far smaller *scattering*. This is expressed by a numerical parameter δ_τ and a size dependent correction. Here r_{eff} is the effective particle size in μm. The last factor accounts approximately for cirrus of optical depth τ_c above the contrail. Numerical parameters for the LW part, e.g. for droxtals, are $k_T = 2.304$, $T_0 = 166$ K, $\delta_{lr} = 0.202$, $\delta_\tau = 0.928$, and $\delta_{lc} = 0.067$. The values of these parameters are determined by least square fits to accurate forward calculations of the RF for a large set of test cases.

The shortwave (SW) RF can be modeled by

$$RF_{SW} = -SDR\left(t_A - A_{eff}\right)^2\left[1 - \exp\left(-\Gamma\tau_{eff}\right)\right] \left[C_\mu + A_\mu\left(\frac{(1-\mu)^{B_\mu}}{(1/2)^{B_\mu}} - 1\right)\exp(-\gamma\tau')\right]\exp\left(\delta_{sc}\tau_c - \frac{\delta'_{sc}\tau_c}{\mu}\right) \tag{15.4}$$

It depends on the fraction $(t_A - A_{eff})$ of incoming *solar direct radiation SDR* absorbed by the Earth-atmosphere system without contrails according to the effective albedo $A_{eff} = RSR/SDR$ (typically between 0.1 and 0.8). The coefficient $t_A \cong 0.9$ accounts for the fact that the Earth atmosphere scatters part of the solar radiation back to space by *Rayleigh scattering* even for a cloud-free atmosphere over a black surface. The factor $(t_A - A_{eff})$ occurs squared because multiple scattering between the contrail and the absorbing Earth surface and lower atmosphere reduces the amount of radiation reflected by contrails. The next factor accounts for the contrail albedo, which increases with the effective optical depth

values $\tau' = \tau\ (1-F_r\ [1-exp(-\delta_{sr}\ r_{eff})])$ and $\tau_{eff} = \tau'/\mu$, which are dependent on *solar-zenith angle* and size. Here, $\mu = SDR/S_0$ is the cosine of the solar zenith angle and $S_0 \cong 1365$ W/m^2 is the solar constant. The term with B_μ is constructed empirically to account for the strong dependence of RF_{SW} on the solar zenith angle because of sideward scattering by contrail ice particles. The last factor accounts approximately for cirrus above the contrail. Here two terms are included in the exponential, one accounting for a reduction of reflection due to the optical depth for incoming solar radiation depending on the solar zenith angle, and one accounting for enhanced downward radiation from scattering by the cirrus above the contrail. Numerical parameters of the SW coefficients for droxtals are $\delta_{sr} = 0.052$, $F_r = 0.249$, $t_A = 0.899$, $\Gamma = 0.275$, $\gamma = 0.311$, $A_\mu = 0.343$, $B_\mu = 1.564$, $C_\mu = 0.660$, $\delta_{sc} = 0.172$, $\delta'_{sc} = 0.244$. Again, the values are determined by a least square fit to accurate forward calculations. The numerical values of these coefficients, in particular for SW, vary considerably with particle habit. The approximate RF values correlate with the libRadtran based RF values to better than 98 %.

15.7 Energy Forcing by Contrail Forming Flights

As explained before, the climate impact of contrail cover from a fleet of aircraft is measured by the RF. However, the climate impact of contrails for individual flights depends on the amount of energy induced into the Earth-atmosphere system during the lifetime of the contrail forming along the flight path. This impact is measured by the *energy forcing* (EF, in units of J m^{-1}),

$$EF = \int_{lifetime} RF_{nets}(t', s)W(t',\ s)dt' \tag{15.5}$$

The global and annual mean RF follows from integrating EF along the flight paths of all flights during a year and dividing by the Earth surface and the time of a year (Schumann et al. 2011a). EF is a function of the local net radiative forcing of a flight segment $RF_{nets} = RF_{LW} + RF_{SW}$ with local longwave and shortwave contributions (W/m^2) on average over the contrail width W (m), integrated from time $t' = t$ (s) of contrail formation until time $t = t + t_{age}$ of contrail disappearance. For typical values, $RF_{nets} = 10$ W/m^2, width $W = 1\,000$ m, and contrail lifetime 10 000 s, we expect a mean value of $EF \sim RF \times$ width $\times$ lifetime $\sim$ 100 GJ/km. For individual contrails, the value is often far more than a factor of ten larger or smaller and often negative during daytime. For comparison, the combustion fuel heat for a typical fuel consumption of 6 g/m contributes an EF of $\sim$0.26 GJ/km. The radiative forcing by CO_2 from this fuel consumption contributes a temperature increase equivalent to an energy forcing of $\sim$260 GJ/km. Hence, mitigation of contrails pays in regions with large contrail EF, and with only a few percent of additional fuel consumption.

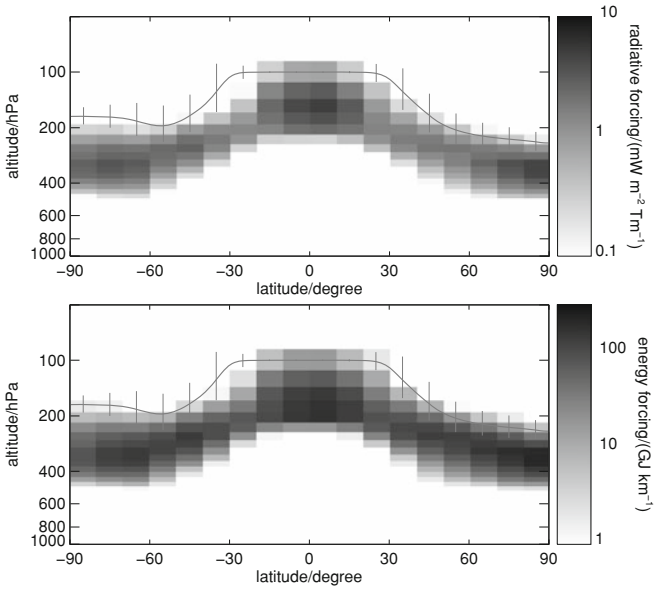

Fig. 15.5 Potential climate impact of contrails over their full life-cycle versus latitude and pressure altitude in terms of **a** global mean net radiative forcing (RF) and **b** global mean energy forcing (EF) from contrails per flight distance as computed with CoCiP. The *white curves* and *error bars* indicate the mean and standard deviations of the tropopause pressure altitude

Figure 15.5 shows a typical distribution of the potential RF and the potential EF from contrails, both including their full life-cycle, versus latitude and altitude. The differences between EF and RF are mainly due to the lifetime dependence of EF. The lifetime is larger than average in the polar mid-troposphere and the upper tropical troposphere. The potential values define the RF or EF values per flight distance in case an aircraft cruises at the given position. It happens that most global air traffic occurs in atmosphere regions which are conditioned to form contrail cirrus with considerable radiative forcing or energy forcing, namely in the upper extratropical troposphere. Air traffic at polar latitudes frequently cruises above the tropopause in often dry air. Air traffic in the tropical atmosphere at the cruise levels of subsonic aviation causes only few contrails because the atmosphere is often too warm and too dry for contrail formation. This would change if supersonic aircraft would be used, since they fly higher and more often reach the cold and humid upper tropical troposphere. RF is small in the tropics because of the low brightness temperature of the cloud tops in the intertropical convergence zone and humid lower atmosphere. It would be high in the subtropics, but this region is usually too dry for contrail formation. The EF is largest in the polar troposphere because of high RF and long contrail lifetimes in these regions. The RF shown is about a factor 10–20 larger than the RF from linear contrails (Fichter 2009). The mean values are positive. The day-to-day variances are comparable to their mean values so that negative RF and EF values occur on many days.

15.8 Estimates of the Climatic Importance of Contrails

Contrails likely contribute the largest fraction to the *non-CO_2 emission* and *short-lived* climate effects of aviation (Lee et al. 2009; Burkhardt and Kärcher 2011). However, the precise value of global and annual mean radiative forcing by contrails is still rather uncertain. There is little doubt that contrail cirrus causes global mean warming. But the sign of the total effect of aviation induced cloud changes (from aerosols and contrails) is still under debate (Penner et al. 2009). Aviation causes cloud changes ("*soot cirrus*") also because of aerosol emissions from aircraft engines, in particular soot particles. The additional particles compete with other particles in forming cirrus clouds by homogeneous and heterogeneous ice nucleation, which leads to either an increase or a decrease in the number of cirrus particles (Hendricks et al. 2011). Aerosol-induced cirrus changes may cause a positive or negative RF. The aerosol effect has longer time scales than contrails because it takes several days before aerosols are removed from the atmosphere.

Best estimates of the global net RF of linear persistent contrails are in the range of 8–20 mW/m^2 for air traffic in the year 2000 (Kärcher et al. 2010). The radiative forcing of the total contrail cirrus has been estimated to amount to about 31 mW/m^2 (Burkhardt and Kärcher 2011). This result is close to estimates derived from decadal cirrus trends of about 30–70 mW/m^2 (Stordal et al. 2005). Some recently published studies (Penner et al. 2009) suggest very large and possibly negative RF from aviation soot induced cirrus changes (possibly exceeding $\pm$100 mW/m^2). However, our global aerosol and cirrus model studies do not support such large values (Hendricks et al. 2011).

Experimental validation of the RF from cloud changes is difficult because contrail cirrus cannot easily be separated from other cirrus. Therefore, a large fraction of contrail research aims at identifying aviation "*fingerprints*" in cirrus properties that provide significant tests for models. Linear contrails served as the first such fingerprints. They were used to calibrate contrail cover models (Sausen et al. 1998; Frömming et al. 2011). Cover by linear contrails can be identified in satellite data (Mannstein et al. 1999; Palikonda 2005). But linear structures represent only part of all contrail cirrus. Thick contrails and their transition to contrail cirrus have been identified from measured radiance patterns and monitored in temporally resolving *Meteosat* pictures for about 1 h on average and 14 h maximum time (Vazquez-Navarro 2009). However, cirrus cover is not a well defined quantity because cirrus is often optically thin and its cover is very sensitive to thresholds and the observation methods used. Direct comparisons between observed and modeled radiances may suffer less from such ambiguity. Moreover, one needs to have information on cirrus properties without aviation to identify the aviation contribution. Observations which provide such information would serve as fingerprints of aviation induced contrail cirrus. We are presently working on various approaches to identify such fingerprints. A particular fingerprint is given by the diurnal cycle of cirrus cover and radiances over the North Atlantic (Graf et al. 2009). Preliminary results obtained from Meteosat-OLR observation data and

CoCiP model studies suggest that the actual RF value of contrail cirrus may be larger than 40 mW/m^2. This is suggested from comparisons of the annual mean *diurnal cycle* of the LW radiative forcing computed for the west and east parts of the *North Atlantic region* using the CoCiP model with the difference in observed OLR between a region in the North Atlantic and a corresponding region in the Southern Atlantic.

Anyway, contrails have far larger regional effects than expected from the global mean values. The local RF for contrail cirrus with optical depth of about 0.1 could vary typically by ± 10 W/m^2, e.g., between day and night. This is at least 100 times larger than the global mean value from all contrail cirrus. This large variability offers *potential for mitigating* the impact of aviation on climate, e.g., by avoiding flights in regions where contrails warm the Earth and by enhanced flying in regions where contrails cool.

15.9 Conclusions

Contrails contribute to the *climate impact* of aviation. In fact, the largest *short-term climate impact* of air traffic may result from *aviation induced cloudiness*. The climate impact of contrails depends on the aircraft, engines, fuels, emissions, and flight routes for given weather conditions. Small changes in aircraft routing may cause large changes in contrail climate impact. The local climate impact of contrail cirrus may be positive or negative and far larger than its global mean value. This variability offers potential for mitigating the climate impact of aviation at small expense of increased carbon dioxide emissions, in particular for increasing air traffic. Using renewable fuels changes the thermodynamic and aerosol properties of the emissions and hence also impacts contrail formation. Hence, contrail research may contribute to keeping aviation *sustainable* in light of growing transport demands. This chapter explained the thermodynamic formation and climatic importance of contrails and contrail cirrus. Much progress has been made in understanding the climate impact of contrails. A strategy was explained for experimental validation of the derived climatic impact of contrails and aviation induced cloud changes.

References

Atlas, D., Wang, Z., Duda, D.P.: Contrails to cirrus—morphology, microphysics, and radiative properties. J. Appl. Meteorol. Clim. **45**, 5–19 (2006)

Brewer, A.W.: The stratospheric circulation: a personnel history. SPARC Newsl. **15**, 28–32 (2000)

Burkhardt, U., Kärcher, B.: Global radiative forcing from contrail cirrus. Nature Clim. Change **1**, 54–58 (2011). doi:10.1038/NCLIMATE1068

Busen, R., Schumann, U.: Visible contrail formation from fuels with different sulfur contents. Geophys. Res. Lett. **22**, 1357–1360 (1995). doi:10.1029/95GL01312

Dietmüller, S., Ponater, M., Sausen, R., Hoinka, K.-P., Pechtl, S.: Contrails, natural clouds, and diurnal temperature range. J. Clim. **21**, 5061–5075 (2008). doi:10.1175/2008JCLI2255.1

Duda, D.P., Minnis, P., Nyuyen, L., Palikonda, R.: A case study of the development of contrail clusters over the Great Lakes. J. Atmos. Sci. **61**, 1132–1146 (2004)

Fahey, D.W., Schumann, U. (Lead-Authors), Ackerman, S., Artaxo, P., Boucher, O., Danilin, M.Y., Kärcher, B., Minnis, P., Nakajima, T., Toon, O.B.: Aviation-produced aerosols and cloudiness. In: Penner, J.E., Lister, D.H., Griggs, D.J., Dokken, D.J., McFarland, M. (eds.) Aviation and the global atmosphere. A special report of IPCC working groups I and III, pp. 65–120. Cambridge University Press, Cambridge (1999)

Fichter, C.: Climate impact of air traffic emissions in dependency of the emission location and altitude, DLR Forschungsbericht 2009-22, ISSN 1434-84543, Oberpfaffenhofen, 152 pp (2009)

Forster, L., Emde, C., Unterstrasser, S., Mayer, B.: Effects of three-dimensional photon transport on the radiative forcing of realistic contrails. J. Atmos. Sci. (2012) (to appear). doi: 10.1175/JAS-D-11-0206.1

Freudenthaler, V., Homburg, F., Jäger, H.: Contrail observations by ground-based scanning lidar: cross-sectional growth. Geophys. Res. Lett. **22**, 3501–3504 (1995). doi:10.1029/95GL03549

Frömming, C., Ponater, M., Burkhardt, U., Stenke, A., Pechtl, S., Sausen, R.: Sensitivity of contrail coverage and contrail radiative forcing to selected key parameters. Atmos. Environ. **45**, 1483–1490 (2011). doi:10.1016/j.atmosenv.2010.11.033

Fuglestvedt, J.S., Shine, K.P., Berntsen, T., Cook, J., Lee, D.S., Stenke, A., Skeie, R.B., Velders, G.J.M., Waitz, I.A.: Transport impacts on atmosphere and climate: metrics. Atmos. Environ. **44**, 4648–4677 (2010). doi:10.1016/j.atmosenv.2009.04.044

Gierens, K., Schumann, U., Helten, M., Smit, H., Marenco, A.: A distribution law for relative humidity in the upper troposphere and lower stratosphere derived from three years of MOZAIC measurements. Ann. Geophys. **17**, 1218–1226 (1999)

Gierens, K., Kärcher, B., Mannstein, H., Mayer, B.: Aerodynamic contrails: phenomenology and flow physics. J. Atmos. Sci. **66**, 217–226 (2009). doi:10.1175/2008JAS2767.1

Graf, K., Mannstein, H., Mayer, B., Schumann, U.: Some evidence of aviation fingerprint in diurnal cycle of cirrus over the North Atlantic. Proceedings of the 2nd International Conference on Transport, Atmosphere and Climate (TAC-2), 22–25 June 2009, DLR-FB 2010-10, ISSN 1434-8454, Aachen and Maastricht, pp. 180–185 (2009)

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G.A., Russell, G., Aleinov, I., Bauer, M., Bauer, S., et al.: Efficacy of climate forcings. J. Geophys. Res. **110**, D18104 (2005). doi:10.1029/2005JD005776

Haywood, J.M., Allan, R.P., Bornemann, J., Forster, P.M., Francis, P.N., Milton, S., Rädel, G., Rap, A., Shine, K.P., Thorpe, R.: A case study of the radiative forcing of persistent contrails evolving into contrail-induced cirrus. J. Geophys. Res. **114**, D24201 (2009). doi: 10.1029/2009JD012650

Helten, M., Smit, H.G.J., Kley, D., Ovarlez, J., Schlager, H., Baumann, R., Schumann, U., Nedelec, P., Marenco, A.: In-flight comparison of MOZAIC and POLINAT water vapor measurements. J. Geophys. Res. **104**, 26087–26096 (1999). doi:10.1029/1999JD900315

Hendricks, J., Kärcher, B., Lohmann, U.: Effects of ice nuclei on cirrus clouds in a global climate model. J. Geophys. Res. **116**, D18206 (2011). doi:10.1029/2010JD015302

Höhndorf, F.: Beitrag zum Problem der Vermeidung von Auspuffwolken hinter Motorflugzeugen, Deutsche Luftfahrtforschung, Aerologisches Institut, Deutsche Forschungsanstalt für Segelflug, FB No. 1371, 15 pp (1941)

Immler, F., Treffeisen, R., Engelbart, D., Krüger, K., Schrems, O.: Cirrus, contrails, and ice supersaturated regions in high pressure systems at northern mid latitudes. Atmos. Chem. Phys. **8**, 1689–1699 (2008). doi:10.5194/acp-8-1689-2008

Iwabuchi, H., Yang, P., Liou, K.N., Minnis, P.: Physical and optical properties of persistent contrails: climatology and interpretation. J. Geophys. Res. **117**, D06215 (2012). doi: 10.1029/2011JD017020

Jensen, E.J., Ackermann, A.S., Stevens, D.E., Toon, O.B., Minnis, P.: Spreading and growth of contrails in a sheared environment. J. Geophys. Res. **103**, 13, 557–513, 567 (1998a). doi:10.1029/98JD02594

Jensen, E.J., Toon, O.B., Kinne, S., Sachse, G.W., Anderson, B.E., Chan, K.R., Twohy, C.H., Gandrud, B., Heymsfield, A., Miake-Lye, R.C.: Environmental conditions required for contrail formation and persistence. J. Geophys. Res. **103**, 3929–3936 (1998b). doi:10.1029/97JD02808

Jensen, E.J., Toon, O.B., Vay, S.A., Ovarlez, J., May, R., Bui, P., Twohy, C.H., Gandrud, B., Pueschel, R.F., Schumann, U.: Prevalence of ice-supersaturated regions in the upper troposphere: implications for optically thin ice cloud formation. J. Geophys. Res. **106**, 17253–17266 (2001). doi:10.1029/2000JD900526

Kärcher, B., Busen, R., Petzold, A., Schröder, F.P., Schumann, U., Jensen, E.J.: Physicochemistry of aircraft-generated liquid aerosols, soot, and ice particles. 2. Comparison with observations and sensitivity studies. J. Geophys. Res. **103**, 17129–17147 (1998). doi:10.1029/98JD01045

Kärcher, B., Burkhardt, U., Unterstrasser, S., Minnis, P.: Factors controlling contrail cirrus optical depth. Atmos. Chem. Phys. **9**, 6229–6254 (2009a). SRef-ID: 1680-7324/acp/2009-9-6229

Kärcher, B., Mayer, B., Gierens, K., Burkhardt, U., Mannstein, H., Chatterjee, R.: Aerodynamic contrails: microphysics and optical properties. J. Atmos. Sci. **66**, 227–243 (2009b). doi:10.1175/2008JAS2768.1

Kärcher, B., Burkhardt, U., Ponater, M., Frömming, C.: Importance of representing optical depth variability for estimates of global line-shaped contrail radiative forcing. Proc. Natl. Acad. Sci. U.S.A., 19181–19184 (2010). doi:10.1073/pnas.1005555107

Lee, D.S., Fahey, D.W., Forster, P.M., Newton, P.J., Wit, R.C.N., Lim, L.L., Owen, B., Sausen, R.: Aviation and global climate change in the 21st century. Atmos. Environ. **43**, 3520–3537 (2009). doi:10.1016/j.atmosenv.2009.04.024

Lewellen, D.C., Lewellen, W.S.: Large-eddy simulations of the vortex-pair breakup in aircraft wakes. AIAA J. **34**, 2337–2345 (1996)

Liou, K.N., Ou, S.C., Koenig, G.: An investigation of the climatic effect of contrail cirrus. In: Schumann, U. (ed.) Air traffic and the environment—background, tendencies and potential global atmospheric effects. Lecture notes in engineering, pp. 154–169. Springer, Berlin (1990)

Mannstein, H., Meyer, R., Wendling, P.: Operational detection of contrails from NOAA-AVHRR data. Int. J. Remote Sens. **20**, 1641–1660 (1999)

Mannstein, H., Schumann, U.: Aircraft induced contrail cirrus over Europe. Meteorol. Z. **14**, 549–554 (2005). doi:10.1127/0941-2948/2005/0058

Mayer, B., Kylling, A.: The libRadtran software package for radiative transfer calculations: description and examples of use. Atmos. Chem. Phys. **5**, 1855–1877 (2005). doi:10.5194/acp-5-1855-2005

Meerkötter, R., Schumann, U., Minnis, P., Doelling, D.R., Nakajima, T., Tsushima, Y.: Radiative forcing by contrails. Ann. Geophys. **17**, 1080–1094 (1999). doi:10.1007/s00585-999-1080-7

Meyer, R., Buell, R., Leiter, C., Mannstein, H., Pechtl, S., Oki, T., Wendling, P.: Contrail observations over Southern and Eastern Asia in NOAA/AVHRR data and comparisons to contrail simulations in a GCM. Int. J. Remote Sens. **28**, 2049–2069 (2007). doi:10.1080/01431160600641707

Minnis, P.: Contrails. In: Holton, J., Pyle, J., Curry, J. (eds.) Encyclopedia of atmospheric sciences. Academic Press, London, pp. 509–520 (2003)

Minnis, P., Young, D.F., Garber, D.P., Nguyen, L., Smith Jr, W.L., Palikonda, R.: Transformation of contrails into cirrus during SUCCESS. Geophys. Res. Lett. **25**, 1157–1160 (1998). doi:10.1029/97GL03314

Ovarlez, J., Gayet, J.F., Gierens, K., Ström, J., Ovarlez, H., Auriol, F., Busen, R., Schumann, U.: Water vapor measurements inside cirrus clouds in Northern and Southern hemispheres during INCA. Geophys. Res. Lett. **29**, 60–61, 60–64 (2002). 10.1029/2001gl014440

Palikonda, R., Minnis, P., Duda, D.P., Mannstein, H.: Contrail coverage derived from 2001 AVHRR data over the continental United States of America and surrounding areas. Meteorol. Z. **14**, 525–536 (2005). doi:10.1127/0941-2948/2005/0051

Penner, J.E., Chen, Y., Wang, M., Liu, X.: Possible influence of anthropogenic aerosols on cirrus clouds and anthropogenic forcing. Atmos. Chem. Phys. **9**, 879–896 (2009). doi:10.5194/acp-9-879-2009

Petzold, A., Busen, R., Schröder, F.P., Baumann, R., Kuhn, M., Ström, J., Hagen, D.E., Whitefield, P.D., Baumgardner, D., Arnold, F., et al.: Near-field measurements on contrail properties from fuels with different sulfur content. J. Geophys. Res. **102**, 29867–29880 (1997). doi:10.1029/97JD02209

Ponater, M., Marquart, S., Sausen, R.: Contrails in a comprehensive global climate model: parameterization and radiative forcing results. J. Geophys. Res. **107**, 4164 (2002). doi:10.1029/2001JD000429

Ponater, M., Marquart, S., Sausen, R., Schumann, U.: On contrail climate sensitivity. Geophys. Res. Lett. **32**, L10706 (2005). doi:10.1029/2005gl022580

Ponater, M., Pechtl, S., Sausen, R., Schumann, U., Hüttig, G.: Potential of the cryoplane technology to reduce aircraft climate impact: a state-of-the-art assessment. Atmos. Environ. **40**, 6928–6944 (2006). doi:10.1016/j.atmosenv.2006.06.036

Pruppacher, H.R.: A new look at homogeneous ice nucleation in supercooled water droplets. J. Atmos. Sci. **52**, 1924–1933 (1995)

Sausen, R., Gierens, K., Ponater, M., Schumann, U.: A diagnostic study of the global distribution of contrails. Part I: present day climate. Theor. Appl. Climatol. **61**, 127–141 (1998)

Sausen, R., Schumann, U.: Estimates of the climate response to aircraft CO_2 and NO_x-emission scenarios. Clim. Change **44**, 27–58 (2000)

Schumann, U.: On conditions for contrail formation from aircraft exhausts. Meteorol. Z. **5**, 4–23 (1996)

Schumann, U.: Influence of propulsion efficiency on contrail formation. Aerosp. Sci. Technol. **4**, 391–401 (2000). doi:10.1016/S1270-9638(00)01062-2

Schumann, U.: A contrail cirrus predicition model. Geosci Model Dev. **5**, 543–580 (2012). doi:10.5194/gmd-5-543-2012

Schumann, U., Wendling,P.: Determination of contrails from satellite data and observational results. In: Schumann, U. (ed.) Air traffic and the environment—background, tendencies and potential global atmospheric effects. Lecture Notes in Engineering, pp. 138–153. Springer, Berlin (1990)

Schumann, U., Ström, J., Busen, R., Baumann, R., Gierens, K., Krautstrunk, M., Schröder, F.P., Stingl, J.: In situ observations of particles in jet aircraft exhausts and contrails for different sulfur-containing fuels. J. Geophys. Res. **101**, 6853–6870 (1996). doi:10.1029/95JD03405

Schumann, U., Busen, R., Plohr, M.: Experimental test of the influence of propulsion efficiency on contrail formation. J. Aircraft **37**, 1083–1087 (2000)

Schumann, U., Graf, K., Mannstein, H.: Potential to reduce the climate impact of aviation by flight level changes. 3rd AIAA Atmospheric and Space Environments Conference No. 1020774, 1020771–1020722 (2011a)

Schumann, U., Mayer, B., Gierens, K., Unterstrasser, S., Jessberger, P., Petzold, A., Voigt, C., Gayet, J.-F.: Effective radius of ice particles in cirrus and contrails. J. Atmos. Sci. **68**, 300–321 (2011b). doi:10.1175/2010JAS3562.1

Schumann, U., Mayer, B., Graf, K., Mannstein, H.: A parametric radiative forcing model for contrail cirrus. J. Appl. Meteorol. Clim. **51** in press (2012) doi: 10.1175/JAMC-D-11-0242.1

Stordal, F., Myhre, G., Stordal, E.J.G., Rossow, W.B., Lee, D.S., Arlander, W., Svendby, T.: Is there a trend in cirrus cloud cover due to aircraft traffic? Atmos. Chem. Phys. **5**, 2155–2162 (2005). doi:10.5194/acp-5-2155-2005

Unterstrasser, S., Sölch, I.: Study of contrail microphysics in the vortex phase with a Lagrangian particle tracking model. Atmos. Chem. Phys. **10**, 10003–10015 (2010). doi:10.5194/acp-10-10003-2010

Unterstrasser, S., Gierens, K., Spichtinger, P.: The evolution of contrail microphysics in the vortex phase. Meteorol. Z. **17**, 145–156 (2008). doi:10.1127/0941-2948/2008/0273

Vazquez-Navarro, M.R.: Life cycle of contrails from a time series of geostationary satellite images, DLR-FB 2010-19, 146 pp (2009)

Voigt, C., Schumann, U., Jessberger, P., Jurkat, T., Petzold, A., Gayet, J.-F., Krämer, M., Thornberry, T., Fahey, D.W.: Extinction and optical depth of contrails. Geophys. Res. Lett. **38**, L11806 (2011). doi:10.1029/2011GL047189

Part II

Methods

Measurements of Nitrogen Oxides and Related Trace Gases

16

Helmut Ziereis, Paul Stock and Hans Schlager

Abstract

Reactive nitrogen species play a key role in the chemistry of the free troposphere and lowermost stratosphere. They have a decisive influence on reaction cycles and on ozone and hydrogen oxide radicals. Therefore, in situ measurements of nitrogen oxides (NO, NO_2, NO_y) and related trace gases (O_3, CO, and CO_2) are of special interest. The Institute of Atmospherics Physics has performed observations of these trace gases aboard different research aircraft for many years. In this chapter instrumental techniques used to measure these species are presented.

16.1 Introduction

Among all trace gases affecting atmospheric *composition* and *chemistry*, *nitrogen oxides* certainly belong to the most prominent ones. It is well known that nitrogen oxides from car exhaust and *combustion* processes are essential for the formation of summer smog in polluted urban areas. Nitrogen oxides, however, also play an

H. Ziereis (✉) · P. Stock · H. Schlager
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Helmut.Ziereis@dlr.de

P. Stock
e-mail: Paul.Stock@dlr.de

H. Schlager
e-mail: Hans.Schlager@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_16,

important role for atmospheric *photochemistry* in the *free troposphere* and the upper troposphere and lowermost stratosphere (UTLS). Their local concentration is controlled by a variety of different sources and processes, chiefly *long-range transport*, lofting from the *boundary layer*, *lightning*, and *air traffic* emissions.

Nitrogen oxide (NO) and nitrogen dioxide (NO_2) are converted into each other within a few minutes during daylight conditions (photostationary state). The sum of these oxides of nitrogen is therefore referred to as NO_x. Further oxidation processes lead to the formation of other oxidized reactive nitrogen species. These includes besides NO_x, nitrate radical (NO_3), dinitrogen pentoxide (N_2O_5), *nitric acid* (HNO_3), *peroxyacetyl nitrate* (PAN), *nitrous acid* (HNO_2), peroxynitric acid (HO_2NO_2), chlorine nitrate ($ClONO_2$), and other organic nitrates. It is useful to combine all these reactive nitrogen species into the "family" of *odd nitrogen* (NO_y) because they are linked together by a variety of reactions. Figure 16.1 shows a simplified scheme of *photochemical reactions* in the UTLS connecting different members of the odd nitrogen family. Odd nitrogen is emitted into the *atmosphere* predominantly as NO. The main sources are anthropogenic combustion processes, *biomass burning*, lightning, and soils. Reactive nitrogen is predominantly removed from the atmosphere by dry and wet *deposition* of nitric acid.

Nitrogen oxides play an essential role in the formation of *ozone* in the troposphere. Ozone is produced by the oxidation of *carbon monoxide*, *methane* and nonmethane *hydrocarbon*s. In these cycles nitrogen oxides are connected to the atmospheric chemistry of the *hydroxyl radical* (OH), one of the most important constituents of the troposphere (see Fig. 16.1). Due to its high reactivity OH determines the *lifetime* of various atmospheric species, for example, methane. Both ozone and methane are important *greenhouse* gases. Therefore nitrogen oxides have a significant indirect impact on the *radiative forcing* of the atmosphere.

Measurements of ozone, carbon monoxide and *carbon dioxide* are helpful in determining the origin of a probed air mass. Carbon monoxide and carbon dioxide are nearly exclusively released on ground, while the main source of ozone in the UTLS region is downward transport from the stratosphere. As a result, these gases serve as tracers for boundary layer and stratospheric air, respectively.

16.2 Challenges of Airborne Trace Gas Measurement

Until the late 1970s very few measurements of nitrogen oxides in the free troposphere were available (Kley et al. 1981; Bradshaw et al. 2000). The Institute of Atmospheric Physics (IPA) has performed measurements aboard different *research aircraft* covering altitudes from the lower troposphere up to the *lower stratosphere* since the early 1990s. Early objectives addressed aircraft emissions (Arnold et al. 1992; Schulte and Schlager 1996), pollution in *air traffic corridors* (Schumann et al. 1995, 2000; Schlager et al. 1997; Ziereis et al. 2000) and lightning NO (Huntrieser et al. 1998; Höller et al. 1999; Schumann and Huntrieser 2007). For observations in the lower and middle troposphere the DLR twin

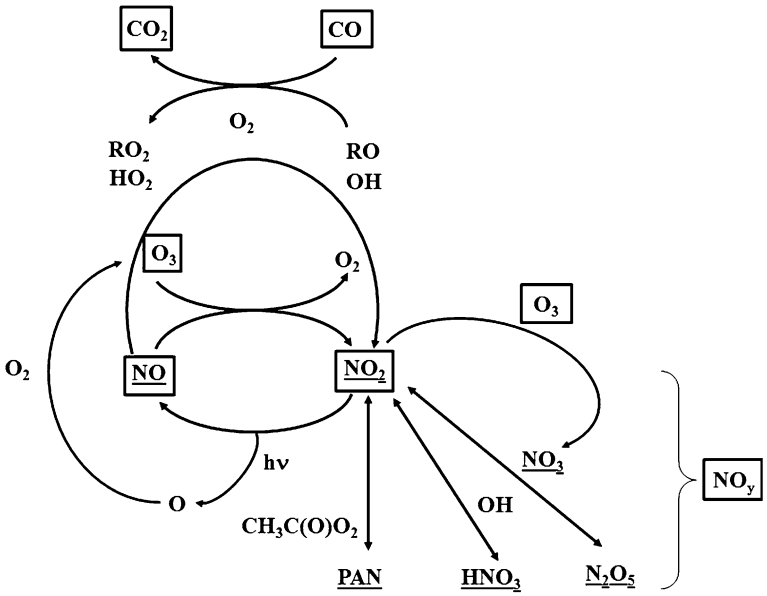

Fig. 16.1 Simplified reaction cycle of nitrogen oxides in the UTLS. Species measured with trace gas instruments described in this chapter are marked by *boxes*. *Underlined* species are members of the odd nitrogen family, NO_y. RO_2: alkylperoxy radical, RO: alkoxy radical, HO_2: hydroxyl radical, $CH_3C(O)O_2$: acetyl peroxy radical, hν: radiation. Other chemical species are described in the text

turboprop aircraft, a Dornier Do 228 may be used. More often, the DLR Dassault *Falcon* 20 E has been used as a platform for *instruments* especially in the free troposphere and lowermost stratosphere for more than 35 years. Measurements at altitudes up to 15 km and with a range of more than 8 000 km will be possible using the newest member of the DLR aircraft fleet for atmospheric research, *HALO*. HALO is based on a Gulfstream G550, a long-range business jet. For altitudes up to about 21 km, the Russian aircraft *Geophysica* M55 is utilized for research. Furthermore, a civil Lufthansa Airbus (A340) is used for measurements within the CARIBIC project (Civil Aircraft for the Regular Investigation of the Atmosphere Based on an Instrument Container).

Trace gases like nitrogen oxides, ozone and other pollutants are quite frequently monitored on ground for environmental purposes. However, the measurement of these species on board a research aircraft poses specific challenges for the instrument techniques used. The concentration of trace gases like nitrogen oxides is usually highly variable in the atmosphere and often shows a strong dependence on altitude. At the ground close to specific sources, the concentrations can be several orders of magnitude higher than in the upper troposphere. In urban areas NO_x *volume mixing ratios* up to 100 nmol/mol can be measured, while in remote regions of the upper troposphere the NO_x volume mixing ratio can be as low as a few pmol/mol. Therefore, instruments used for measurements in the upper

troposphere and lower stratosphere must have significantly lower detection limits than instruments used on ground.

A further challenge for instruments operated from research aircraft is the time response of the applied measurement technique. The concentration of trace gases might be nearly constant at scales covering several hundreds of kilometers in the upper troposphere. However, concentration may be enhanced at small scales of only a few kilometers by lightning or aircraft emissions. Jet research aircraft like the Falcon or HALO cover a distance of about 200 m/s. Therefore, a short response time on the order of 1 s is mandatory for resolving such small-scale enhancements.

Instruments on board research aircraft also have to deal with rapid changes in ambient *temperature*, *pressure* and *humidity*. The respective values may change from about 30 °C, 1013 hPa and a water volume mixing ratio of 1 % at sea level to about −70 °C, 200 hPa and a water mixing ratio of a few parts per million in the UTLS. Also, condensing humidity during ascent due to *adiabatic* cooling may cause instrument problems.

Besides challenges posed by the atmosphere, there are also requirements arising from aircraft *safety* and aircraft operations. Space, payload and available power are limited resources on board most research aircraft. Therefore, there is a strong requirement to construct and build instruments which are as small, light, and economical (in sense of electrical power) as possible. Additionally, the electronics of the instruments must not interfere with the sophisticated avionics of modern aircraft. Further requirements for instruments on aircraft arise from air safety considerations.

To meet these demands, instruments used on research aircraft are often specially designed and constructed for this purpose. Off-the-shelf instruments are rarely used; at least, they have to be adapted for measurements in the atmosphere.

16.3 Atmospheric Trace Gas Measurement Techniques

The measurement of *trace gases* makes use of the typical physical and chemical properties of the species. The technique employed should be sensitive to the respective trace gas and interference with other species should be as low as possible. In the following section, techniques for the measurement of NO, NO_2, NO_y and the associated trace gases O_3, CO, and CO_2 applied by the institute are presented. The main focus will be on measurement techniques for nitrogen oxides.

16.4 Measurement of Nitrogen Oxide

The method predominantly used to measure atmospheric NO is based on *chemiluminescence* (Fontijn et al. 1970; Stedman et al. 1972). In general, chemiluminescence is the emission of light, luminescence, following a chemical reaction. It has been found that the reaction of NO with ozone leads to the emission of light that peaks at *wavelengths* of about 1 200 nm.

$$NO + O_3 \rightarrow NO_2 + O_2 \quad (16.1)$$

$$NO + O_3 \rightarrow NO_2* + O_2 \quad (16.2)$$

$$NO_2* \rightarrow NO_2 + h\nu \quad (16.3)$$

$$NO_2* + M \rightarrow NO_2 + M \quad (16.4)$$

M: neutral *collision* partner like N_2 or O_2 that absorbs excess energy, (*): excited molecule, hν: *radiation*

By reaction of NO with O_3, NO_2 is formed (16.1). A fraction of these molecules is in an electronically excited state (16.2) which returns to the ground state either by emission of radiation (16.3) or by collisional deactivation of the excited NO_2 molecules (16.4). The chemiluminescence technique is widely used for observations in the atmosphere (on ground, on balloons, and on aircraft) (Ridley et al. 1992). At IPA, NO chemiluminescence detectors have been employed on research aircraft since the early 1990s (Schulte and Schlager 1996). A simplified schematic representation of a two-channel chemiluminescence detector is shown in Fig. 16.2. The sample air containing NO is mixed with ozone in a reaction chamber. The ozone is produced from a silent discharge (ozonizer or ozone generator) at concentrations far exceeding the atmospheric background levels (up to 3 % ozone in pure *oxygen*). To optimize the chemiluminescence signal the pressure inside the reaction chamber is reduced to about 10–20 hPa to decrease the collisional deactivation of the excited NO_2 molecules. The chemiluminescence signal is detected by means of a photomultiplier. A photomultiplier is a very sensitive detector of light making use of the photoelectric effect. To reduce the dark current (a signal registered by a photomultiplier even without light irradiation), the photomultiplier is cooled to about −30 °C or less. A red filter is used in front of it to reduce interfering radiation. The signal recorded is not only caused by the chemiluminescence signal from NO but also from the dark current of the photomultiplier itself and the light arising from the chemiluminescence reaction of other species. To subtract these signals most chemiluminescence detectors use a so-called zeroing volume. Upstream of the reaction chamber a pre-reaction chamber (zeroing volume) is installed. The sample air passes both chambers. The ozone flow is switched regularly between the pre-reaction chamber and the reaction chamber by means of a three-way valve. The dimensions of the zeroing volume and the flows are chosen in such a way that the chemiluminescence signal from the NO reaction has already decayed before the air reaches the reaction chamber with the photomultiplier (Drummond et al. 1985). In this mode the signal from the dark current and other species can be determined and finally subtracted from the measurement signal. This method is based on the assumption that the chemiluminescence signal from other reactions does not change significantly between these two modes.

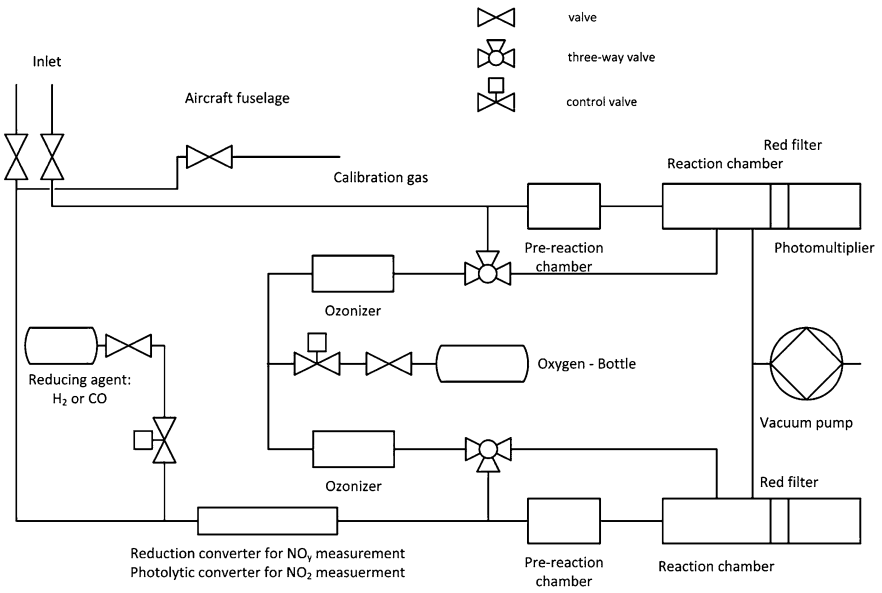

Fig. 16.2 Simplified schematic representation of a two-channel NO chemiluminescence detector

For IPA instruments, a maximum sensitivity of about 20 counts per second and per nmol/mol NO was achieved. The best achievable detection limit for data averaged over 10 s is about 1–2 pmol/mol. The accuracy depends on the measured concentration. It is about 8 % for NO and NO_y levels of about 50 and 450 pmol/mol, respectively. The instruments are usually operated with a 1-s time resolution.

Figure 16.3 gives a view inside the NO, NO_y instrument used on the Geophysica aircraft. Some of the main components of the detector are shown. On the Geophysica, the detector is installed inside a pressure tank and mounted in a wing pod below one of the aircraft wings.

16.5 NO_y and NO_2 Conversion

Odd nitrogen, the family of all reactive nitrogen species, comprises a variety of different compounds. Several specific instruments are necessary to measure each of the main compounds of this chemical family. However, the combination of a special reduction converter with a chemiluminescence detector offers the opportunity for a comparatively simple method to measure the sum of all reactive nitrogen species, NO_y (see Fig. 16.2).

A technique using a gold converter was introduced by Bollinger et al. (1983) and Fahey et al. (1985). Laboratory studies have shown that a gold converter heated to 300 °C converts the most important NO_y species to NO with an

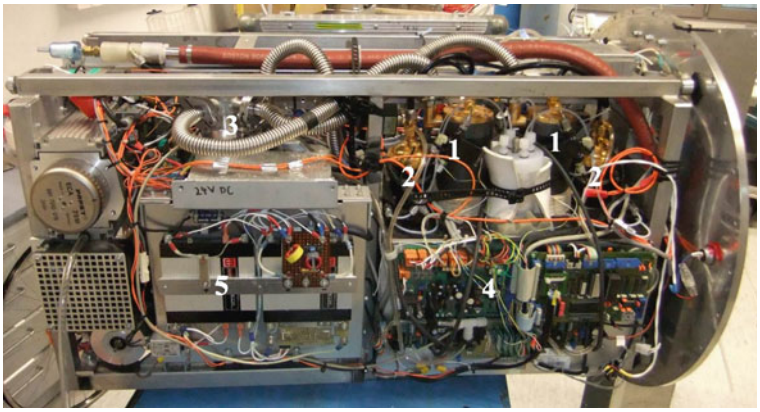

Fig. 16.3 View inside the NO and NO_y detector used aboard Geophysica. *1* Reaction chamber, *2* Pre-reaction chamber, *3* Pump, *4* Control electronics, *5* Voltage converter

efficiency of more than 90 %. As reducing agent either pure carbon monoxide or *hydrogen* (from a metal hydride storage device) can be used. Usually, at least two instruments are employed to allow simultaneous measurements of NO and NO_y. At IPA a converter consisting of a heated gold tube with a length of about 50 cm and a diameter of about 0.6 cm is used.

More specifically, NO_2 can be measured by using a chemiluminescence detector in combination with a photolytic converter (see Fig. 16.2). Radiation from an *ultraviolet* (UV) source is used to convert NO_2 to NO, which is subsequently measured as described above. For earlier measurements, a xenon UV lamp was used for the *photodissociation* of NO_2 (Ziereis et al. 1999). However, interference from the photodissociation or thermal decomposition of other reactive nitrogen species had to be considered. Therefore, a modified commercial converter based on diodes emitting light in the ultra violet range (UV-LEDs) is now used to photolyze NO_2 to NO.

16.6 O_3 Measurement Technique

There are a couple of techniques to measure atmospheric ozone. A widely used technique is based on the fact that ozone molecules absorb light in the UV range. UV *absorption* photometers make use of radiation absorption by ozone molecules at about 254 nm. The air sample is divided into two air streams. Both pass equal absorption chambers with a length of about 40 cm. In one chamber UV absorption is measured with an unchanged ozone concentration. In the other chamber UV absorption is measured in the so-called reference mode. In this chamber ozone has been previously removed by a scrubber. Reference and sample gases are switched regularly between the two chambers. The ozone concentration in the sample air is

calculated from the signal difference between the two modes by using the Lambert–Beer absorption law. The accuracy of the measurement also depends on *calibration* (discussed below) and is about 1 nmol/mol at ground-level.

The measurement of ozone also demonstrates the trade-off between the accuracy of a measurement, its time resolution and its usability on board research aircraft. To improve the accuracy of the ozone measurement the length of the absorbing cell could be enlarged. More ozone molecules would be available to attenuate the UV light. However, a larger absorbing cell would reduce the time resolution of the measurement. Furthermore, more space and stronger pumps with higher power consumption would be needed.

16.7 CO Measurement Technique

The detection of CO relies on the resonance *fluorescence* of CO (Gerbig et al. 1996). The air sample is radiated by light from a CO resonance lamp. The appropriate wavelength of this radiation has been selected by an *optical* filter. Radiation at a wavelength of 150 nm leads to the excitation of the CO in the sample. The excited CO molecules can either be deactivated by collisions with other air molecules or by fluorescence radiation (160–190 nm) that is detected by means of a photomultiplier. For 1-s measurements the accuracy is about 3 nmol/mol.

16.8 CO_2 Measurement Technique

CO_2 is known as one of the most important greenhouse gases and can be measured by a variety of instruments. Measurement of CO_2 usually relies on infrared absorption techniques. Non-dispersive infrared (NDIR) spectroscopy instruments were used during several aircraft *campaigns* (Schulte et al. 1997). The instruments used on board the DLR research aircraft are commercially available instruments that have been adapted for use in the upper troposphere. The instrument consists of two cells. By means of a pump, sample air is sucked through one of these cells. The two chambers are irradiated with infrared light. An optical filter is used to select a certain wavelength of the *infrared radiation* most suitable for absorption by CO_2. The radiation passing through the cell with the sample air is finally detected by a lead selenide solid state device. The reference cell is flushed by a gas with a known amount of CO_2. The difference of the signals registered at the two detectors is used to derive the CO_2 concentration in the sample air. The time response is <1 s; the accuracy is better than 0.3 μmol/mol.

More recently, *cavity ring-down spectroscopy* (CRDS) is used for the detection of atmospheric CO_2. A *laser* pulse trapped between two mirrors is reflected a few thousand times. CO_2 that absorbs at the appropriate wavelength affects the decay (the "ring-down") of the laser pulse. The decay of the laser pulse increases with increasing CO_2 concentration. The advantage of this technique lies in its stability

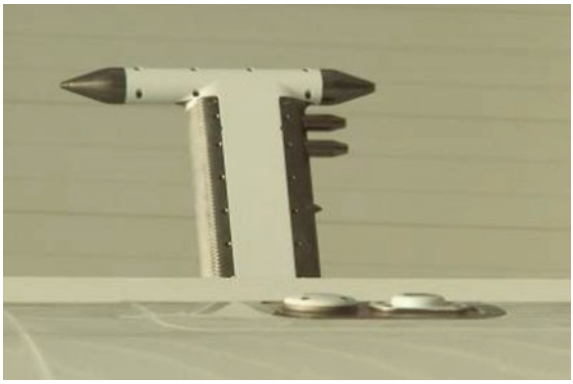

Fig. 16.4 Trace gas inlet of the HALO research jet. The inlet extends about 430 mm outside the fuselage. In total this inlet can house four inlet lines with diameters between about 3 and 13 mm. The inlet lines are usually made of stainless steel or perfluoroalkoxy copolymer (PFA) and may be heated

compared to instruments based on infrared spectroscopy. Furthermore, it allows the simultaneous measurement of methane and *water vapor.*

16.9 Inlets and Sampling Lines

Inlets are designed to allow the sampling of ambient air from outside the aircraft and to transport the probed air to instruments inside the cabin (see Fig. 16.4). They must assure that the sample air is not influenced by contact with the aircraft before the measurement. Therefore, they have to be tall enough to rise above the boundary layer of the aircraft, which is approximately the distance from the aircraft skin to the point where the air velocity around the fuselage again approaches the free stream velocity.

Any kind of in situ *measurement* has to take into account the potential influence of the inlet and the inlet tube on the measured species. In this sense, species like NO and NO_2 are relatively easy to sample because their interactions with the inlet walls are small. On the other hand, trace gases like nitric acid (or ammonia) are very "sticky" and are readily adsorbed to surfaces, especially if the surfaces are wet. As nitric acid is a main component of atmospheric NO_y, these measurements can be perturbed by its reaction with the surface of the inlet and inlet tubing. The measurement is even more complicated because nitric acid can be desorbed again when conditions change in the sample air. The adsorption of nitric acid strongly depends on the surface material and the temperature. Laboratory studies have shown that PFA (perfluoroalkoxy copolymer) tubing heated to temperatures above 10 °C show the best behavior with respect to sampling atmospheric nitric acid (Neuman et al. 1999). To keep interference from the inlet as small as possible, high sample flow rates and short inlets are a prerequisite.

Ozone is also very sensitive to wall loss. It can especially be destroyed on metal surfaces. Therefore, PFA tubing is also used for ozone inlet lines. Furthermore, there are indications that O_3 is formed by electrical discharges on the inlet system close to or inside *thunderstorms* (Zahn et al. 2002; Huntrieser et al. 2007; Schumann and Huntrieser 2007).

16.10 Calibration

None of the instruments described in this chapter relies on an absolute measurement. This means that the output signal of the detector alone does not allow unequivocal determination of the respective trace gas concentration. Therefore, calibration of these instruments is necessary. Calibration can be understood as a comparison between two measurements. The response of the measurement of an unknown atmospheric concentration of a specific trace gas is compared to the response of the instrument when introducing a known concentration of this gas.

Due to their low detection limit and fast response time, chemiluminescence detectors in combination with converters are the optimal choice for aircraft borne measurements of nitrogen oxides. However, to achieve high *accuracy* and *precision* of the measurements regular calibration of the instruments and quality checks are essential.

The detected signal of the chemiluminescence depends on the ozone concentration and therefore on the performance of the ozone generator, as well as on the quality of the photomultiplier. Although the sensitivity of a detector does not change rapidly, it has to be checked regularly by adding a small flow of a calibration gas, usually NO in pure N_2, to NO-free synthetic air. The instrument background is determined by adding NO-free synthetic air to the measuring systems. This instrument background signal or zero air artifact might be caused by wall chemiluminescence or by impurities in the O_2 gas used for the production of ozone (Drummond et al. 1985) and has to be subtracted from the measured NO signal.

The measurement of NO_2 and NO_y using photolytic or gold converters, respectively, requires additional calibration. The conversion efficiency of the converters for NO_2 can be checked by adding known amounts of NO_2 produced by reaction of NO with O_3 (gas phase titration). The conversion efficiency for HNO_3 can also be checked by using specific sources like permeation tubes. The conversion efficiency of both converter types depends on the residence time inside the converter, and thus on the pressure. To avoid pressure changes at varying altitudes the pressure can be kept constant by means of a Teflon valve. However, this may result in additional wall losses at the comparatively high surface of the valve. For this reason, in the applications described, no valve in front of the converter was used for the NO_y measurements. The pressure dependence of NOy conversion efficiency has to be accounted for in the data analysis.

Of special interest is the undesired conversion of species which do not belong to the reactive nitrogen family. For example, HCN is considered to be one major candidate affecting NO_y measurements. Different studies found conversion efficiencies between a few to 100 % depending on temperature, reducing agent (CO or H_2), and O_3 concentration (Weinheimer et al. 1998; Ziereis et al. 2000). Therefore, the measurements have to be adjusted accordingly.

Calibration is also necessary for the CO and CO_2 measurements described above. For these instruments, certified mixtures of CO and CO_2 are introduced to the respective instrument from compressed cylinders.

The sensitivity of the UV-absorption photometer is checked regularly with an ozone standard source. To double-check the accuracy of the ozone measurements the ozone standard source can be compared to other ozone UV-absorption measurements, for example at the global watch station of the German Weather Service at Hohenpeissenberg in Germany.

Time response, sensitivity, detection limit, precision and accuracy of the measurement depend on a variety of parameters and may change with time. Regular intercomparison between instruments operated from different research aircraft is highly desirable and raises confidence in the measurement quality (Brough et al. 2003). Such in-flight intercomparisons yield insights into instrument performance during ambient measurements, which might differ significantly from observations made under controlled laboratory conditions.

16.11 Examples

16.11.1 Measurements of Nitrogen Oxides Emitted by Civil Aircraft at Flight Level

A fast time response of the instruments is necessary for measuring air traffic emission signatures in the UTLS region. Measurements with the Falcon were performed to determine NO_x emission indices of commercial aircraft for cruising conditions (Schulte and Schlager 1996; Schulte et al. 1997). The NO_x *emission index* indicates how much nitrogen oxide (in mass units of NO_2) is produced per kg fuel burnt. This information is necessary to estimate the impact of aircraft on the nitrogen budget of the atmosphere. The measurements were performed in *radar*-controlled zones over southern Germany and the North Atlantic. During the measurements, the Falcon "chased" the source aircraft, a commercial aircraft like an Airbus or a Boeing, at a distance of about ten nautical miles or less. During such flights, the Falcon encountered the exhaust trail of the source aircraft several times. This can be recognized by the simultaneous enhancements in the NO and CO_2 signal. In Fig. 16.5, the time series of CO_2, NO and O_3 measurements performed on 5 July 1995 behind a B747 are shown. While CO_2 and NO concentrations increased significantly, ozone remained nearly constant at about 60 nmol/mol. During flights in the *North Atlantic flight corridor* enhancements in the NO_x mixing ratio of up to 300 nmol/mol were observed (Schulte et al. 1997). Concurrent enhancements in the CO_2 signal of up to 50 μmol/mol were observed. Assuming that nearly all carbon contained in the fuel is transformed to CO_2, the emission index for CO_2 is known with sufficiently high accuracy. This allows calculation of the NO_x emission index from the concurrent measurement of the enhancements of NO_x and CO_2 exceeding the background values.

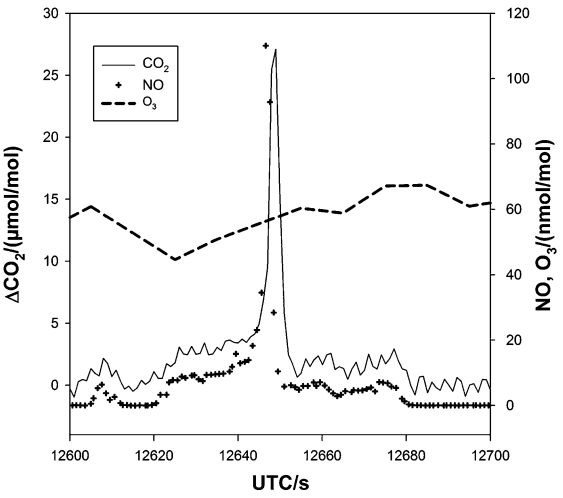

Fig. 16.5 Aircraft "chasing" in the North Atlantic flight corridor. Time series of CO_2, NO, and O_3 obtained during measurements close to a B747 aircraft. The CO_2 signal shown is the difference relative to background. The temporal resolution of the data is 1 s, except for O_3 (10 s)

16.11.2 Nitrogen Oxides Measurements Using Civil Aircraft

Instruments for measuring nitrogen oxides are widely used on research aircraft like Falcon or Geophysica. Another approach to analyze the distribution and budget of atmospheric nitrogen oxides is chosen when using commercial in-service aircraft. Since December 2004, IPA has contributed a NO and NO_y detector to the CARIBIC *instrumentation* (Brenninkmeijer et al. 2007). About once a month the automated measurement container (see Fig. 16.6) is installed in the cargo bay of a Lufthansa Airbus A340-600. Flights are performed between Frankfurt in Germany and airports in South and North America, Asia, and South Africa. During more than 240 flights so far, a unique set of nitrogen oxides data has been obtained for the upper troposphere and lower stratosphere.

While focused aircraft missions are dedicated to specific questions, long-term global measurements of nitrogen oxides allow the analysis of large-scale regional and seasonal variations of the background NO and NO_y burden in the upper troposphere and lowermost stratosphere. However, signatures of air traffic emissions are also regularly observed during measurements with the CARIBIC container. Typically, aircraft plumes can be recognized by sharp enhancements in NO, NO_y, and fine particles (Schlager et al. 1997). The *time scales* of these enhancements are of the order of 1–100 s. An example is given in Fig. 16.7 when the Airbus with the CARIBIC container penetrated aircraft exhaust plumes several times. This flight was performed on 21 June 2007 between Frankfurt and Guangzhou in China. NO and NO_y increased from background values of about 0.2 and 1.8 nmol/mol up to a few nmol/mol. NO exceeded its background levels by a factor of more than 10. The NO/NO_y ratio shows that nearly all NO_y consisted of NO during these events. In the background UTLS, this ratio ranges roughly between 0.1 and 0.25. A high ratio indicates that this particular injection has occurred only recently. Otherwise

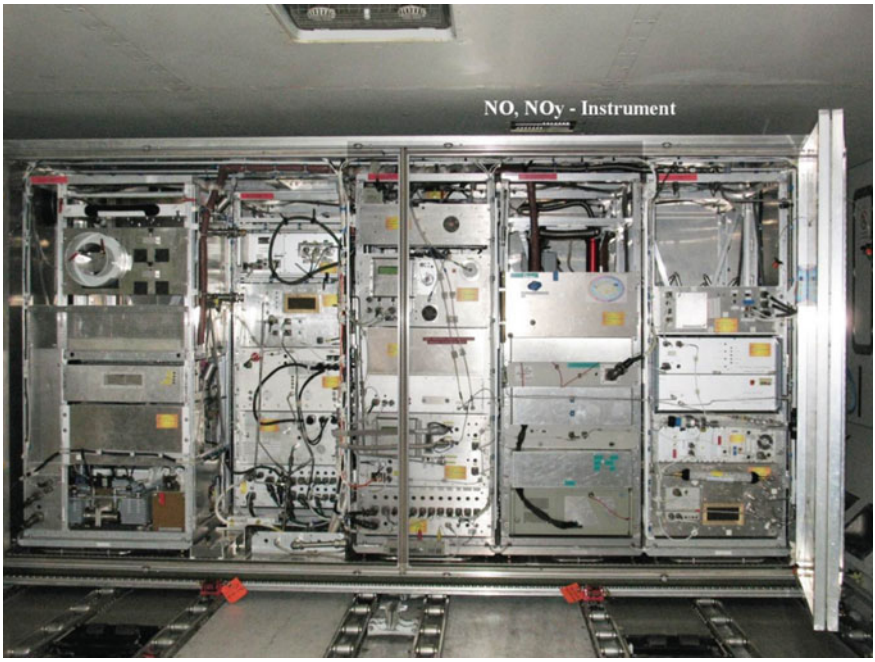

Fig. 16.6 CARIBIC instrumented airfreight container on board a Lufthansa Airbus A340-600. The instrument package comprises a suite of instruments to measure trace gases and aerosol properties. The NO and NO_y detector has been provided as part of this instrumentation since December 2004 (Picture permission granted by C. Brenninkmeijer, MPI-C, Mainz)

NO_x would have been already converted to other NO_y species. High NO/NO_y ratios are therefore an indication of recent aircraft emissions.

16.11.3 Uptake of Reactive Nitrogen on Cirrus Clouds

The technique described in this chapter allows the measurement of gaseous nitrogen oxides. However, it can also be used to study the *uptake* of atmospheric nitric acid in particles. For this kind of measurement two inlets are connected to the two channels of a chemiluminescence detector equipped with gold converters. One inlet faces in the flight direction, the other against it. *Cirrus cloud particles* are sampled with enhanced efficiency relative to the gas phase with the forward-facing inlet. The inlet facing against flight direction discriminates against the collection of cloud particles. The signal measured with this channel is dominated by gas-phase NO_y. Particles sampled with the forward-facing inlet evaporate in the inlet and the released nitric acid is converted into NO. The enhancement of the particles in the forward-facing inlet can be calculated. Subsequently, condensed-phase NO_y expressed as gas-phase equivalent can be inferred from the difference between the NO_y signals obtained from measurements with the forward- and backward-facing inlets.

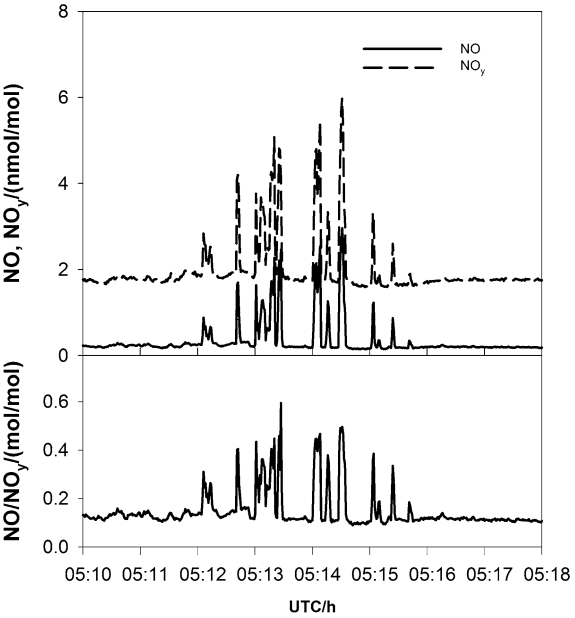

Fig. 16.7 Time series of NO, NO_y, and NO/NO_y for the CARIBIC flight from Frankfurt to Guangzhou over northern China, 21 June 2007

The uptake of nitrogen species on *ice crystals* in cirrus clouds was studied in 2000 during the INCA (*Interhemispheric* differences in cirrus properties from anthropogenic emissions) *experiment* out of Punta Arenas (Chile) (Ziereis et al. 2004). It was found that the uptake of nitric acid on ice particles depends on atmospheric temperature, the nitric acid *partial pressure*, and the available surface area of the cirrus cloud particles. For the conditions encountered during INCA, less than 1 % of the *total reactive nitrogen* was found as particulate nitrate. Further measurements also studied the uptake of nitric acid on persistent contrails (Schauble et al. 2009). It was found that the mean fraction of nitric acid in *contrails* is higher than in natural *cirrus clouds*.

16.12 Conclusions and Outlook

Knowledge of distribution and sources of nitrogen oxides is essential for understanding the photochemistry of the UTLS and their impact on radiative forcing. Several examples were presented. Further examples concerning lightning-induced NO_x concentrations are described elsewhere (Huntrieser et al. 1998; Schumann and Huntrieser 2007). In this context, the measured trace gas concentrations have been used for comparison with *satellite* observations (Heland et al. 2002; Emmons et al. 2004) and modeled results (Schlager et al. 1999; Sovde et al. 2011). The observation of nitrogen oxides and related trace gases like carbon monoxide, ozone and carbon dioxide will therefore remain a major task for measurements in

the UTLS. The new DLR research aircraft HALO will significantly extend the possibilities for measurements from aircraft. The trace gas instrumentation described is already scheduled for several HALO missions in the near future. While measurements on aircraft like the Falcon or HALO are focused on specific scientific questions, measurements from civil in-service aircraft like CARIBIC contribute to establishing an observational NO and NO_y climatology.

References

Arnold, F., Scheid, J., Stilp, T., Schlager, H., Reinhardt, M.E.: Measurements of jet aircraft emissions at cruise altitude. I: The odd-nitrogen gases NO, NO_2, HNO_2 and HNO_3. Geophys. Res. Lett. **19**, 2421–2424 (1992)

Bollinger, M.J., Sievers, R.E., Fahey, D.W., Fehsenfeld, F.C.: Conversion of nitrogen-dioxide, nitric-acid, and normal-propyl nitrate to nitric-oxide by gold-catalyzed reduction with carbon-monoxide. Anal. Chem. **55**, 1980–1986 (1983)

Bradshaw, J., Davis, D., Grodzinsky, G., Smyth, S., Newell, R., Sandholm, S., Liu, S.: Observed distributions of nitrogen oxides in the remote free troposphere from the NASA global tropospheric experiment programs. Rev. Geophys. **38**, 61–116 (2000)

Brenninkmeijer, C.A.M., Crutzen, P., Boumard, F., Dauer, T., Dix, B., Ebinghaus, R., Filippi, D., Fischer, H., Franke, H., Friess, U., et al.: Civil aircraft for the regular investigation of the atmosphere based on an instrumented container: the new CARIBIC system. Atmos. Chem. Phys. **7**, 4953–4976 (2007)

Brough, N., Reeves, C.E., Penkett, S.A., Stewart, D.J., Dewey, K., Kent, J., Barjat, H., Monks, P.S., Ziereis, H., Stock, P., et al.: Intercomparison of aircraft instruments on board the C-130 and Falcon 20 over southern Germany during EXPORT 2000. Atmos. Chem. Phys. **3**, 2127–2138 (2003)

Drummond, J.W., Volz, A., Ehhalt, D.H.: An optimized chemi-luminescence detector for tropospheric NO measurements. J. Atmos. Chem. **2**, 287–306 (1985)

Emmons, L.K., Deeter, M.N., Gille, J.C., Edwards, D.P., Attie, J.L., Warner, J., Ziskin, D., Francis, G., Khattatov, B., Yudin, V., et al.: Validation of measurements of pollution in the troposphere (MOPITT) CO retrievals with aircraft in situ profiles. J. Geophys. Res. **109**, (2004). doi:10.1029/2003jd004101

Fahey, D.W., Eubank, C.S., Hubler, G., Fehsenfeld, F.C.: Evaluation of a catalytic reduction technique for the measurement of total reactive odd-nitrogen NOy in the atmosphere. J. Atmos. Chem. **3**, 435–468 (1985)

Fontijn, A., Sabadell, A.J., Ronco, R.J.: Homogeneous chemiluminescent measurement of nitric oxide with ozone—implications for continuous selective monitoring of gaseous air pollutants. Anal. Chem. **42**, 575 (1970)

Gerbig, C., Kley, D., VolzThomas, A., Kent, J., Dewey, K., McKenna, D.S.: Fast response resonance fluorescence CO measurements aboard the C-130: instrument characterization and measurements made during North Atlantic Regional Experiment 1993. J. Geophys. Res. **101**, 29229–29238 (1996)

Heland, J., Schlager, H., Richter, A., Burrows, J.P.: First comparison of tropospheric NO2 column densities retrieved from GOME measurements and in situ aircraft profile measurements. Geophys. Res. Lett. **29**, (2002). doi:10.1029/2002gl015528

Höller, H., Finke, U., Huntrieser, H., Hagen, M., Feigl, C.: Lightning-produced NOx (LINOX): experimental design and case study results. J. Geophys. Res. **104**, 13911–13922 (1999)

Huntrieser, H., Schlager, H., Feigl, C., Höller, H.: Transport and production of NOx in electrified thunderstorms: survey of previous studies and new observations at midlatitudes. J. Geophys. Res. **103**, 28247–28264 (1998)

Huntrieser, H., Schlager, H., Roiger, A., Lichtenstern, M., Schumann, U., Kurz, C., Brunner, D., Schwierz, C., Richter, A., Stohl, A.: Lightning-produced NOx over Brazil during TROCCINOX: airborne measurements in tropical and subtropical thunderstorms and the importance of mesoscale convective systems. Atmos. Chem. Phys. **7**, 2987–3013 (2007)

Kley, D., Drummond, J.W., Mcfarland, M., Liu, S.C.: Tropospheric profiles of NOx. J. Geophys. Res.-Oceans Atmos. **86**, 3153–3161 (1981)

Neuman, J.A., Huey, L.G., Ryerson, T.B., Fahey, D.W.: Study of inlet materials for sampling atmospheric nitric acid. Environ. Sci. Technol. **33**, 1133–1136 (1999)

Ridley, B.A., Grahek, F.E., Walega, J.G.: A small, high-sensitivity, medium-response ozone detector suitable for measurements from light aircraft. J. Atmos. Ocean. Technol. **9**, 142–148 (1992)

Schauble, D., Voigt, C., Karcher, B., Stock, P., Schlager, H., Kramer, M., Schiller, C., Bauer, R., Spelten, N., de Reus, M., et al.: Airborne measurements of the nitric acid partitioning in persistent contrails. Atmos. Chem. Phys. **9**, 8189–8197 (2009)

Schlager, H., Konopka, P., Schulte, P., Schumann, U., Ziereis, H., Arnold, F., Klemm, M., Hagen, D.E., Whitefield, P.D., Ovarlez, J.: In situ observations of air traffic emission signatures in the North Atlantic flight corridor. J. Geophys. Res. **102**, 10739–10750 (1997)

Schlager, H., Schulte, P., Flatoy, F., Slemr, F., van Velthoven, P., Ziereis, H., Schumann, U.: Regional nitric oxide enhancements in the North Atlantic flight corridor observed and modeled during POLINAT 2—a case study. Geophys. Res. Lett. **26**, 3061–3064 (1999)

Schulte, P., Schlager, H.: In-flight measurements of cruise altitude nitric oxide emission indices of commercial jet aircraft. Geophys. Res. Lett. **23**, 165–168 (1996)

Schulte, P., Schlager, H., Ziereis, H., Schumann, U., Baughcum, S.L., Deidewig, F.: NOx emission indices of subsonic long-range jet aircraft at cruise altitude: in situ measurements and predictions. J. Geophys. Res. **102**, 21431–21442 (1997)

Schumann, U., Konopka, P., Baumann, R., Busen, R., Gerz, T., Schlager, H., Schulte, P., Volkert, H.: Estimate of diffusion parameters of aircraft exhaust plumes near the tropopause from nitric-oxide and turbulence measurements. J. Geophys. Res. **100**, 14147–14162 (1995)

Schumann, U., Schlager, H., Arnold, F., Ovarlez, J., Kelder, H., Hov, O., Hayman, G., Isaksen, I.S.A., Staehelin, J., Whitefield, P.D.: Pollution from aircraft emissions in the North Atlantic flight corridor: overview on the POLINAT projects. J. Geophys. Res. **105**, 3605–3631 (2000)

Schumann, U., Huntrieser, H.: The global lightning-induced nitrogen oxides source. Atmos. Chem. Phys. **7**, 3823–3907 (2007)

Sovde, O.A., Hoyle, C.R., Myhre, G., Isaksen, I.S.A.: The HNO3 forming branch of the HO2 + NO reaction: pre-industrial-to-present trends in atmospheric species and radiative forcings. Atmos. Chem. Phys. **11**, 8929–8943 (2011). doi:10.5194/acp-11-8929-2011

Stedman, D.H., Stuhl, F., Daby, E.E., Niki, H.: Analysis of ozone and nitric-oxide by a chemiluminescent method in laboratory and atmospheric studies of photochemical smog. J. Air Pollut. Control Assoc. **22**, 260 (1972)

Weinheimer, A.J., Campos, T.L., Ridley, B.A.: The in-flight sensitivity of gold-tube NOy converters to HCN. Geophys. Res. Lett. **25**, 3943–3946 (1998)

Zahn, A., Brenninkmeijer, C.A.M., Crutzen, P.J., Parrish, D.D., Sueper, D., Heinrich, G., Gusten, H., Fischer, H., Hermann, M., Heintzenberg, J.: Electrical discharge source for tropospheric "ozone-rich transients". J. Geophys. Res. **107**, (2002). doi:10.1029/2002jd002345

Ziereis, H., Schlager, H., Schulte, P., Kohler, I., Marquardt, R., Feigl, C.: In situ measurements of the NOx distribution and variability over the eastern North Atlantic. J. Geophys. Res. **104**, 16021–16032 (1999)

Ziereis, H., Schlager, H., Schulte, P., van Velthoven, P.F.J., Slemr, F.: Distributions of NO, NOx, and NOy in the upper troposphere and lower stratosphere between 28° and 61°N during POLINAT 2. J. Geophys. Res. **105**, 3653–3664 (2000)

Ziereis, H., Minikin, A., Schlager, H., Gayet, J.F., Auriol, F., Stock, P., Baehr, J., Petzold, A., Schumann, U., Weinheimer, A., et al.: Uptake of reactive nitrogen on cirrus cloud particles during INCA. Geophys. Res. Lett. **31**, (2004). doi:10.1029/2003gl018794

Chemical Ionization Mass Spectrometric Measurements of Atmospheric Trace Gases

17

Heinfried Aufmhoff, Dominik Schäuble, Anke Roiger, Frank Arnold, Tina Jurkat, Christiane Voigt and Hans Schlager

Abstract

Chemical Ionization Mass Spectrometry (CIMS) is a versatile, fast and sensitive method for the detection of atmospheric trace gases, which play key roles in atmospheric chemistry and climate. A combination of different types of mass spectrometers, ion sources and inlet configurations are deployed by IPA. Sophisticated calibration techniques and in-flight calibration systems are developed to ensure high data quality. The CIMS instruments are employed preferably on research aircraft; in addition, observations are performed on ships, ground based sites and in laboratory scenarios including atmospheric chambers.

H. Aufmhoff (✉) · D. Schäuble · A. Roiger · F. Arnold · T. Jurkat · C. Voigt · H. Schlager
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20, 82234 Oberpfaffenhofen, Germany
e-mail: Heinfried.Aufmhoff@dlr.de

D. Schäuble
e-mail: Dominik.Schaeuble@dlr.de

A. Roiger
e-mail: Anke-Elisabeth.Roiger@dlr.de

F. Arnold
Max-Planck -Institute for Nuclear Physics, Saupfercheckweg 1, 69117 Heidelberg, Germany
e-mail: frank.arnold@mpi-hd.mpg.de

T. Jurkat
e-mail: Tina.Jurkat@dlr.de

C. Voigt
e-mail: Christiane.Voigt@dlr.de

H. Schlager
e-mail: Hans.Schlager@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_17,

17.1 Introduction

Several atmospheric *trace gases* detectable by *Chemical Ionization Mass Spectrometry* (CIMS) play a key role in particle formation processes and *ozone* chemistry. A major *precursor* gas for atmospheric *nucleation* leading to new *aerosol* particles is gaseous H_2SO_4 (*sulfuric acid*). H_2SO_4 is produced via the OH (*hydroxyl* radical) oxidation of SO_2 (*sulfur dioxide*) and subsequent reactions with oxygen and H_2O (*water*). All key trace compounds such as SO_2, OH, H_2O and H_2SO_4 in this important reaction mechanism have been measured by CIMS. In addition, the atmospheric ozone budget and the *oxidation capacity* of the atmosphere are influenced by the abundance of *reactive nitrogen* species, of which HONO (*nitrous acid*) and the reservoir gases HNO_3 (*nitric acid*) and PAN (*peroxyacetyl nitrate*) are detected with CIMS. CIMS offers the ability to measure atmospheric trace gases in the pmol/mol-range (volume *mixing ratio* of 10^{-12}) and for ultra trace gases and long integration times even down to a few fmol/mol (volume mixing ratio of 10^{-15}). It is a very sensitive, fast and flexible in situ measurement technique, thus well suited especially for airborne measurements. The time resolution is typically about a second. For ultra-sensitive measurements an averaging time of about a minute is needed. For the detection of the trace gas of interest, suitable reagent ions are used generated by an ion source (IS). These *ions* react with the trace gas molecules via *ion molecule reactions* in the *ion flow reactor* and specific *product ions* are generated. *Reagent* and product ions are detected by mass filtration or by trapping and subsequent reading out in the mass spectrometer. For small ratios of product and reagent ions, the measured ratio is proportional to the number of molecules of the trace gas per one cm^3 air inside the flow reactor or per the corresponding number of molecules of air. It is also proportional to the reaction rate coefficient and the ion residence time inside the flow reactor.

In this section an overview of present CIMS techniques deployed at IPA is given. This description of CIMS techniques deals with different kinds of mass spectrometers for trapping or mass filtration, ion sources for the production of appropriate reagent ions, and *calibration* sources and setups for atmospheric measurements. Airborne CIMS measurements of, e.g., SO_2, PAN, HNO_3, HONO, and H_2SO_4 have been made by IPA or by collaborations involving DLR (Arnold et al. 1992; Möhler and Arnold 1992; Curtius et al. 1998; Speidel et al. 2007; Fiedler et al. 2009; Voigt et al. 2010; Jurkat et al. 2011; Roiger et al. 2011; Schlager et al. 2012). Further references to airborne and ground based atmospheric CIMS measurements of these and other trace species are given in Table 17.1. The actual CIMS measurements at IPA described here are built on close collaboration with the Atmospheric Physics research group at the Max-Planck Institute of Nuclear Physics in Heidelberg (MPI-K HD) headed by Frank Arnold, which introduced airborne CIMS in atmospheric research and has more than 30 years of experience in mass spectrometry. This group has designed and built various types of airborne CIMS instruments and deployed them on different types of flying platforms, including *rockets*, *dropsondes*, stratospheric research *balloons*, and

Table 17.1 Atmospheric IPA (/MPI-K HD) CIMS measurements of trace species (selection) and details like ion molecule reactions, mass spectrometer types and measurement platforms

Trace gases(exemplary)	Reagent ion $E^{\pm}$	IS gas	Type	Platform	Exemplary reaction[a]	References (exemplary)
SO_2, HNO_3, HONO, HCN (and other organic acids)	CO_3^-	O_2 (purity ≥99.998 %)	ITMS, LQMS	Falcon, HALO	$CO_3^- + SO_2 \rightarrow SO_3^- + CO_2$, $SO_3^- + O_2 + M \rightarrow SO_5^- + M$	(Arnold et al. 1992; Möhler and Arnold 1992; Speidel et al. 2007; Fiedler et al. 2009; Schlager et al. 2012)
PAN, PPN, MPAN	I^-	N_2/CH_3I[b]	ITMS	Falcon, HALO	$CH_3C(O)O_2$ (fragment of PAN) + $I^- \rightarrow CH_3C(O)O^-$ (potential reagent) + IO	(Roiger et al. 2011)
N_2O_5, halogen compounds		N_2/CH_3I, N_2/CF_3I	AIMS	Falcon, ship	$I^- + Cl_2 \rightarrow ICl_2^-$	
HNO_3, HCl, HONO, SO_2	SF_5^-	N_2/SCF_8[b]	ITMS, AIMS	Falcon, EUPHORE, ship	$SF_5^- + HONO \rightarrow (FHONO)^- + SF_4$	(Jurkat et al. 2010; Voigt et al. 2010; Jurkat et al. 2011)
HO_2 (+ RO_2), OH[c], H_2SO_4,[d] MSA	NO_3^-		LQMS ITMS[b] LQMS[d]	Dornier 228, UFS, ship Falcon	$NO_3^- + H_2SO_4 \rightarrow HSO_4^- + HNO_3$	(Reiner et al. 1999) (Aufmhoff et al. 2011) (Möhler and Arnold 1992; Curtius et al. 1998; Curtius et al. 2002)
H_2O VOCs (e.g. acetone)	H_3O^+[e]	Air	AIMS LQMS ITMS	Falcon Falcon, Learjet	H_3O^+**+$(CH_3)_2CO$** $\rightarrow$ $H^+(CH_3)_2CO + H_2O$	(Arnold et al. 1997; Kiendler and Arnold 2003)

[a] For simplification, additional ligands like H_2O are not written down
[b] Alternative ion source, see reference (right column)
[c] HO_2, RO_2 and OH are detected via chemical conversion to H_2SO_4
[d] H_2SO_4 was also measured by a volatile aerosol component CIMS analyzer and by passive CIMS with ambient ions (Möhler and Arnold 1992)
[e] Product (for H_2O detection) and potential reagent (for VOC detection)

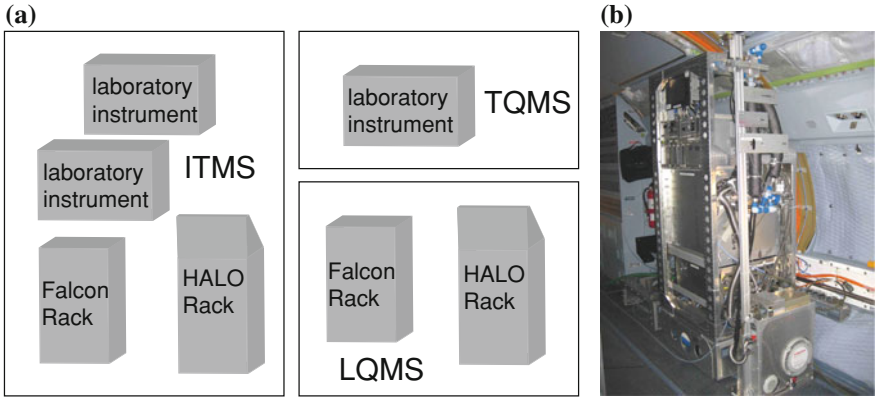

Fig 17.1 **a** Mass spectrometers in operation at IPA and in field experiments (selection). **b** HALO CI-ITMS instrument installed inside the aircraft cabin

research *aircraft*. In addition, international relationships to institutes in the USA developing and operating CIMS instruments have been established, including NCAR, NOAA and Georgia Institute of Technology.

17.2 Overview of Present CIMS Techniques at IPA

Several CIMS techniques are in use at IPA and have recently been deployed in the field, especially for airborne measurements. The individual technique is generally dependent on the trace gas to be measured and the measurement platform. Critical technical parameters of an airborne CIMS instrument that have to fulfill certain standards are weight, size, electrical power requirement, mechanical stability, and degree of automation. The main parts of a CIMS instrument are a *mass spectrometer* and an ion flow reactor. The different types of mass spectrometers used have specific advantages, e.g., high ion transmission or very high mass resolution, and are deployed and optimized for the specific measurement tasks. A single mass filter is a linear quadrupole (LQMS), another mass analyzer is an ion trap (ITMS) and a third mass spectrometer type is a triple quadrupole mass filter with special hyperbolic rods (TQMS). A selection of available instruments in use is given in Fig. 17.1a.

An extensive task is integrating CI (*chemical ionization*) mass spectrometers with peripheral devices into aircraft racks and obtaining air worthiness certification. CI-ITMS have been successfully certified for three different aircraft: *HALO*, *Falcon*, *Learjet*. A photograph of the HALO CI-ITMS instrument is shown in Fig. 17.1b.

The mass spectrometer is the detection unit of a CIMS instrument. The atmospheric air sample with trace gas molecules A is sucked into the *inlet* and the *sampling line* leading to the ion flow reactor (see Fig. 17.2). The ion flow reactor contains an ion source producing reagent ions $E^{\pm}$ that react with trace gas A to

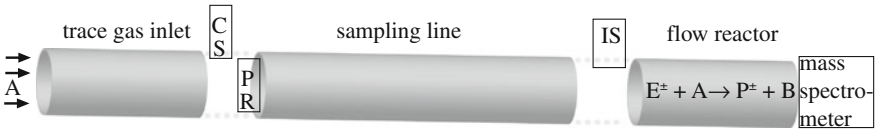

Fig. 17.2 Schematics of CIMS flow tubes and key components—trace gas molecules A, calibration source CS, pressure regulator PR, ion source IS, reagent ion $E^{\pm}$, product ion $P^{\pm}$, byproduct B

product ions $P^{\pm}$. The molecules A are detected by measuring the reagent ions $E^{\pm}$ and generated charged products $P^{\pm}$ with the mass spectrometer. A selection of trace gases and appropriate reagent ions is given in Table 17.1. In principle, the atmospheric trace gas concentration can be calculated from the ratio of product and reagent ions, the ion residence time inside the flow reactor and a reaction rate coefficient. Due to potential wall losses in the sampling line, inaccuracy of the *reaction rate* coefficient, interferences with *water vapor*, and other uncertainties, calibration with a gas standard introduced at the beginning of the sampling line is necessary to reduce measurement uncertainty. Details of sampling issues concerning the optimization of the probe air flow, different measurement platforms used by IPA, and examples of atmospheric measurements including calibrations are described.

17.2.1 Mass Spectrometers

The principle purpose of a mass spectrometer is to separate ions with different mass–to-charge ratios in *atomic mass units* (amu[1]) per z unit charges. The relative abundance of the different ions is reflected in a mass spectrum. Predominant single charged ions with $z = 1$ from the ion flow reactor enter the vacuum chambers of the mass spectrometer. They are accelerated by pressure differences and attractive electrical potentials, focused by an ion guide, separated by mass via a mass analyzer and finally detected by a device that gives an amplified electrical signal. The pressure inside the mass analyzer must be kept very low by a very efficient pumping system to avoid collisions of ions with undesired gas molecules and subsequent perturbations of the ion motion inside the analyzer.

17.2.1.1 Ion Trap Mass Spectrometer

The deployed CI-ITMS instruments contain a modified commercial ion trap mass analyzer. For conversion to a flight instrument several modifications have been made in collaboration with MPI-K HD including an adaption to a CIMS ion source, change of power supplies, assembling of peripheral devices, and integration into aircraft instrument racks. The developed CI-ITMS aircraft instruments have proven to be very robust: they have been successfully deployed during aircraft campaigns in different climate zones all over the world (*arctic*, *tropics*)

[1] 1 amu = $1.66054 \cdot 10^{-27}$ kg.

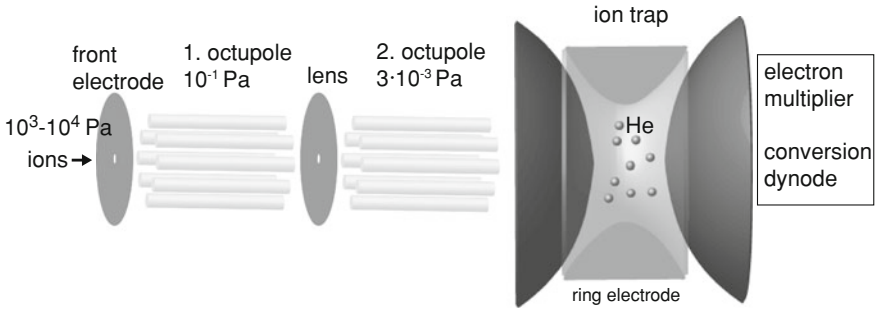

Fig. 17.3 ITMS. The vacuum in the last two pressure chambers is achieved with the help of a two-split turbomolecular pump and a forepump. The chambers are separated by aperture plates. The pressure before the front electrode (*left*) is reduced by a scroll-pump or a rotary vane pump

and first CIMS measurements aboard HALO have recently been successfully carried out (Schlager et al. 2012). The CI-ITMS instruments were also used by the IPA/MPI-K-HD collaboration not only on aircraft but also at the earth's surface (on high-altitude mountain sites, on a research ship), and in laboratories (to investigate aircraft gas turbine exhaust, *lightning* processes, aerosol formation, and modern *diesel car engine emissions*).

The ion guide of the ITMS consists of a front electrode, two octupoles and an interoctupole lens; the mass analyzer is a Paul ion trap and the detection is done by an electron multiplier via a conversion dynode (see Fig. 17.3). The lens between the two octupoles serves as a "gate" that can be opened or closed by applying an appropriate DC voltage. The ions are trapped by losing most of their kinetic energy with the help of helium acting as a damping gas. The stabilizing of the ions inside the trap's electric field can be visualized by the following mechanical analog which was originally developed by Nobelist W. Paul: A ball falls down from a saddle. But if the saddle is rotating, the ball can be stabilized. After trapping the ions, which can be optimized by tuning the ion guide, a mass separation in time is recorded. The ion trap is a three dimensional *quadrupole* allowing ions to be ejected by changing the RF (radio frequency) voltage applied to the ring electrode (see Fig. 17.3), starting at low masses and ending at high masses. A wide mass range is possible (up to 4 000 amu/z) at high resolution (ideally 0.3 amu/z) and importantly a large sensitivity is achieved also for ions with large masses. The identification of ions via $(n-1)$-fold mass-selected ion fragmentation (MS^n) by n-stage mass spectrometry offers the possibility to identify unknown product ions. An example of a MS^2 fragmentation study is described in Sect. 17.2.2. A unique advantage of CI-ITMS is a quasi-simultaneous detection of ions with different masses, yielding a snapshot of the actual situation. This is especially important for conditions with highly variable atmospheric mole fractions of the trace gases in space or time. Hydration of ions, meaning the attachment of water molecules to the

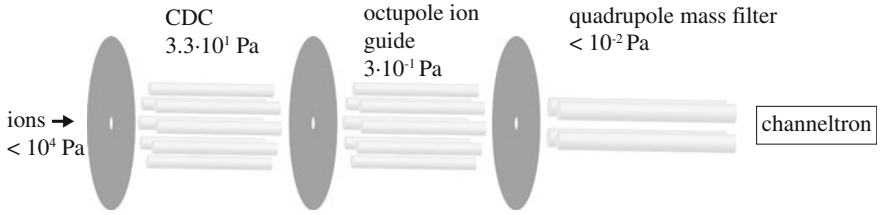

Fig. 17.4 AIMS (LQMS). The vacuum in the last three pressure chambers is achieved with the help of two turbomolecular and one hybride turbomolecular pump (and a forepump). The chambers are separated by aperture plates. The pressure before the front electrode (*left*) is reduced by a scroll-pump

ions in the flow reactor, can unintentionally perturb the measurement. In an ion trap spectrum almost only the core ions of hydrated ions are visible as the H_2O ligands are detached inside the ion trap by collisions with helium atoms. Thereby the mass spectrum becomes easier to interpret due to fewer interferences.

17.2.1.2 Linear Quadrupole Mass Spectrometer

Linear quadrupole mass filters have been used frequently not only on aircraft but also on a dropsonde separated from a rocket for measurements in the mesosphere and stratosphere by MPI-K HD in collaboration with DLR. Here, the focus is on two Atmospheric chemical Ionization linear quadrupole Mass Spectrometers (AIMS), which were constructed by the Georgia Institute of Technology. The instruments are similar to the one described in Huey et al. (1998) with an additional collision dissociation chamber (CDC). The ion flow reactor, the ion source, calibration devices and the inlet system are being further developed at IPA. The instruments are operated on research aircraft, e.g., during the CONCERT2011 (Contrail and Cirrus Experiment) campaign with the Falcon, on a *research vessel*, and in an atmospheric chamber.

AIMS consists of a CDC, an octupole ion guide, a quadrupole mass filter and a channel electron multiplier (channeltron) detector (see Fig. 17.4). The CDC contains an octupole at relatively high pressure (33 Pa) that accelerates and collimates the ion beam and allows fragmentation of cluster ions. In this way, hydrated ions can be declustered to get a clearer mass spectrum. Alternatively, fragmentation can be minimized to allow detection of cluster ions. The ions are focused by the electric field of the second octupole. The linear quadrupole in the last chamber is the mass analyzer; a DC voltage with superimposed RF voltage is applied to the rods for mass separation. Only ions of one mass (to charge ratio) are stable for a given pair of amplitudes of these voltages. To get a full mass spectrum from about 10 to 300 amu the DC and RF voltages are increased at a constant DC-to-RF ratio. Depending on the trace gas to be measured, a trade-off between ion transmissivity and mass resolution has to be made. A LQMS is particularly sensitive to lower masses.

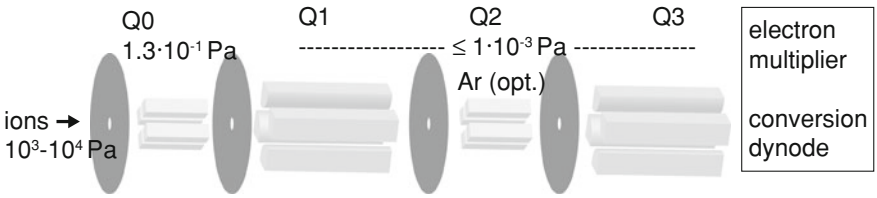

Fig. 17.5 TQMS—the different chambers are divided by aperture plates/lenses

17.2.1.3 Triple Quadrupole Mass Spectrometer

Presently, the development of a new CI-TQMS (triple quadrupole mass spectrometer) has started at IPA, building on the experience with a previous system built by MPI-K HD. This previous airborne CI-TQMS was designed and built by MPI-K HD and, in collaboration with DLR, deployed on DLR-aircraft. This instrument allowed unambiguous identification of numerous ambient positive and negative ion species as well as numerous trace gas species. The core of the new instrument is a triple stage quadrupole from Thermo Scientific, and a CIMS ion source has already been adapted. A TQMS represents a mass analyzer containing three quadrupoles (see Fig. 17.5, Q1–Q3). An ideal quadrupole mass filter uses hyperbolic rods instead of circular ones like Q1 and Q3 of the new TQMS. This ensures a very high mass resolution without signal loss. Therefore even mass defects[2] can be resolved, which can be of additional help for the identification of product ions. The combination of three quadrupoles Q1-Q3 and the addition of argon gas offer the possibility of MS^2 fragmentation studies. With an additional collision-induced dissociation by Q0 (quadrupole ion guide with square rods) (primitive) MS^3 analysis is also possible.

17.2.2 Ion Flow Reactor

In the ion flow reactor, located just before the front electrode of the mass spectrometer, the trace gas molecules of interest can be converted by appropriate reagent ions to measurable product ions (see Fig. 17.2). Reagent ions are produced by a gas discharge ion source IS (a high AC or DC voltage up to 15 kV applied to a noble metal needle as an electrode). The ion source gas and the sample air are generally separated until ionization of the source gas, then the ionized source gas is injected via a tube or just an aperture into the sampled air stream. An exception is the direct ionization of the probed air yielding potential reagent ions. An overview of most of the reagent ions currently used at IPA including the ion source gas and the measured trace gases is shown in Table 17.1. The choice of a suitable reagent ion depends on how selectively this ion reacts with a particular trace gas (to avoid too many other products) and how sensitively this gas can be measured. Also,

[2] (sum of masses of protons and neutrons) − (mass of nucleus) = $\Delta m \neq 0$.

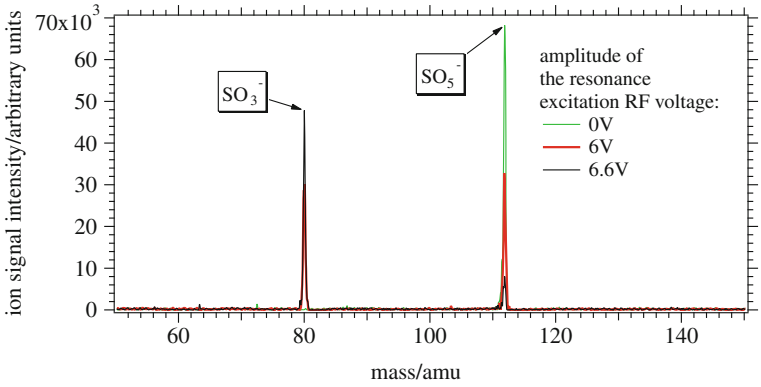

Fig. 17.6 MS^2 (fragmentation) mass spectrum obtained with ITMS, including the resonance excitation RF voltages applied to the ring electrode of the ion trap

interference with other trace gases like H_2O has to be considered. For qualitative analysis of the probed air it is sometimes very interesting to detect and identify additional product ions. Especially ITMS and TQMS are well suited to fragmentation studies that can help to identify a product ion. An example of a fragmentation study by CI-ITMS and CO_3^- as reagent ion is shown in Fig. 17.6.

The product or parent ion with a mass of 112 amu is isolated inside the ion trap and fragmented to a daughter ion with a mass of 80 amu by applying a resonance excitation RF voltage to the ring electrode of the trap. This fragment or daughter ion reduces to a minimum the number of possible candidates for the parent ion. This study was performed to verify the reaction chain of SO_2 detection by CO_3^- via SO_3^- (80 amu) and the subsequent reaction to SO_5^- (112 amu) shown in Table 17.1. SO_2 measurements with CO_3^- have been introduced by Möhler and Arnold (1992) and deployed by IPA already for several years, including online in-flight calibration (see Sect. 17.2.3.1). CO_3^- is also used to detect HNO_3, HONO (Arnold et al. 1992), HCN and other organic acids.

SO_2 is a precursor of H_2SO_4 in the atmosphere. H_2SO_4 is detectable via reactions with the reagent ion NO_3^- (see Table 17.1). The corresponding ion molecule reaction is a fast proton exchange reaction with a reaction rate coefficient close to the collision rate; that means that almost each collision of an ion with a H_2SO_4 molecule leads to a reaction. SO_2, OH and H_2O are essential for the production of H_2SO_4. The reaction of SO_2 with OH can also be used to detect atmospheric OH. Chemical conversion takes place close to the trace gas inlet by *isotopically labeled* SO_2 and ambient water vapor. The method of adding isotopically labeled trace gases to the sampling line is described in more detail in Isotopically Labeled Calibration Gas. A special feature of the ion source deployed for H_2SO_4/OH detection is that an additional quench gas is added to prevent the production of artificial H_2SO_4 from OH produced by the ion source itself; see Aufmhoff et al. (2011).

Besides negative ions, an ion source also produces positive ions. If water vapor is present as in an open ion source, meaning that no special ion source gas but rather atmospheric air is ionized directly, H_3O^+ ions are produced. The amount of these ions correlates directly with the atmospheric water vapor concentration. To prevent hydration of these hydronium ions the pressure inside the ion flow reactor is reduced to a few hPa as the hydration is correlated to the water vapor partial pressure in the flow reactor. The low pressure and short ion residence time also minimize ion–molecule reactions between hydronium ions and trace gases that have a higher proton affinity than H_2O (like many volatile organic compounds). As the H_3O^+ signal is not linear to the water vapor concentration, a calibration source (CS) operating over the whole H_2O range is essential. This CS and an example of airborne H_2O measurements with AIMS building on NOAA developments are described in Sect. 17.2.5.1.

For PAN measurements, the reagent ion I^- can be used. A special feature of PAN ($CH_3C(O)O_2NO_2$) measurements with I^- is that PAN is thermally decomposed by deploying a heated PFA tube at 180 °C and the fragment $CH_3C(O)O_2$ (PA radical) reacts with the reagent ion. PAN detection with I^- was introduced by Slusher et al. (2004) and deployed by DLR-IPA (Roiger et al. (2011) and Sect. 17.2.5.1.). Attachment reactions of I^- with HNO_3 and halogen compounds are also possible. The product cluster ions can be detected sensitively if the ligands are bound strongly (CI-ITMS) or if the product ion mass is sufficiently small (CI-LQMS).

The relative new SF_5^- reagent ions introduced by Marcy et al. (2005) have been used for measurements of HONO and SO_2 in aircraft exhaust plumes and HNO_3 and HCl in the *Upper Troposphere and Lower Stratosphere* (UTLS) (Jurkat et al. 2010, 2011). Using this reactive reagent ion, the four product ions are formed through a fluoride transfer (e.g., $FHONO^-$) and detected on small masses. Fragmentation of these HF-ion clusters inside the mass spectrometer is usually minimized to achieve high sensitivity. For boundary layer conditions, water vapor dependent calibration and background measurement is mandatory. An example of a HONO measurement using SF_5^- reagent ions during a HONO intercomparison in a photochemistry simulation chamber will be described in Sect. 17.2.5.4.

17.2.3 Calibration Sources

A number of different *calibration* sources are used at IPA for the CIMS measurements. With these sources, a known mole fraction of the trace gas is added to the sampling line. For relatively stable trace gases like SO_2, this calibration standard can be stored in a gas cylinder. If the substance is sticky like HNO_3 it has to be stored in a suitable reservoir like a permeation device (see textbox). Thermally unstable gases like PAN have to be produced by a special reactor. The process of generating gases with a reactor is complex, therefore we explain the principles in detail for PAN, H_2SO_4 and H_2O reactors in textboxes. Besides calibration standard gases with natural isotopic abundances, isotopically labeled standards are also used in some cases allowing continuous online calibrations.

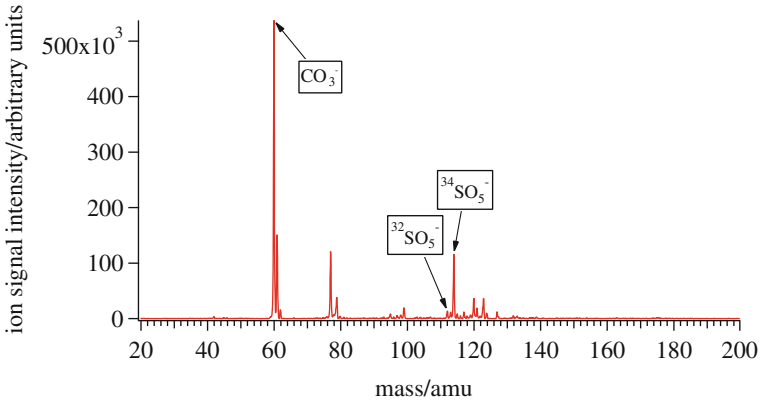

Fig. 17.7 Mass spectrum taken by CI-ITMS with a SO_2 online calibration source during airborne measurements

Intercomparison with other measurement devices under realistic conditions is very helpful in understanding the overall error of an instrument. Additionally to laboratory and in-flight calibrations many intercomparisons are conducted by IPA and other institutes on different platforms. Examples are described in Sect. 17.2.5.

Permeation devices

A permeation device is a small casing that emits a temperature and in some cases a pressure-dependent amount of a standard gas (e.g. $H^{15}NO_3$). Different permeation ovens for these devices are deployed at IPA. As an extraordinary example, a compact latent heat accumulator has been developed at IPA and MPI-K HD. It contains a heating control and is operated at constant temperature and pressure. During power-off periods inside a research aircraft, the permeation device can still be flushed with dry nitrogen at constant temperature (for about 45 min). The pressure is kept constant by a mechanical pressure regulator.

17.2.3.1 Isotopically Labeled Calibration Gas

This calibration method is used during the atmospheric measurement of an appropriate trace gas (e.g. SO_2). A reference mass line is produced online with the help of an isotopically labeled standard of slightly different mass (e.g., $^{34}SO_2$, Speidel et al. (2007)). The height of the measured mass line of the atmospheric trace gas species can be compared directly with the height of the mass line of the standard (with small corrections) because the chemical behavior of the different isotopes is the same. Calibration takes place under realistic conditions parallel to the atmospheric measurement. A typical mass spectrum taken during airborne CI-ITMS SO_2 measurements via CO_3^- (mass 60 amu) is shown in Fig. 17.7. The reference mass

line is at 114 amu ($^{34}SO_5^-$). The atmospheric signal of SO_2 is at 112 amu ($^{32}SO_5^-$). The isotopic fraction of the most abundant molecule mass of the standard should be high (ideally >90 %), as should the fraction of the alternative mass (e.g. $^{32}SO_2$) of the molecules in the atmosphere. Otherwise, the undesired background on the corresponding mass lines would be too high to be correctable.

17.2.4 Sampling

Much optimization work has to be done for the individual mass spectrometers, ion sources, sampling setups and their combinations, aiming at sensitive measurements with high accuracy. The choice of a suitable material for the sampling line and trace gas inlet is very critical. PFA (perfluoroalkoxy, a fluoropolymer) is one of the most suitable materials used as a sampling line, e.g., for HNO_3. It is not well-suited to very sensitive H_2O measurements because it adsorbs/desorbs more H_2O than, e.g., stainless steel. This adsorption/desorption process is a general problem, especially for sticky trace gases. So a plane surface with a small surface-to-volume ratio must be used for the sampling line to minimize this problem. Heating can also help to minimize the adsorption. For H_2O measurements a special inlet, a coated flexible polyethylene/aluminum construction, is used. Background measurements with air that is free of the particular trace gas are necessary to distinguish between artificial and true atmospheric signals. The atmospheric air can be filtered in situ (e.g., with nylon wool to remove HNO_3) or heated (e.g., to decompose PAN). The addition of purified air from a container serves as a dry background measurement. The background in general is subtracted for each product mass. The noise of the artificial signal determines the detection limit, i.e., the value above which the atmospheric signal becomes significant relative to this noise. The trace gas inlet is also very critical in the sampling process. The sample air should be unaffected by the walls of the measurement platform (e.g., aircraft, ship container, building). Guided inlets for ground-based or ship measurements have been deployed using a high air flow produced by a blower. From the central part of this flow an air column is sampled by the probe air inlet. This way air with minimized contact to the sampling platform surface, such as a container roof, is sampled (see Fig. 17.11). For airborne measurements the pressure inside the sampling line and ion flow tube reactor has to be regulated because of variable pressure conditions inside the inlet. With the help of a compact and strong industrial servo (Volz Servos, Germany) PFA or stainless steel valves are regulated automatically to ensure a constant pressure in combination with a pump that sucks a constant flow with a critical orifice in front of the pump. Compact pumps (rebuilt scroll pumps with a robust brushless three-phase 115 V/400 Hz motor) have been developed and already deployed at DLR-IPA for HALO.

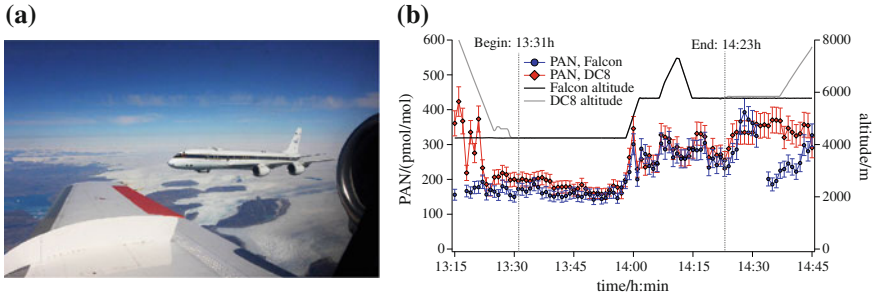

Fig. 17.8 **a** Formation flight with the DLR Falcon and the NASA DC8 for measurement intercomparison (wing of Falcon at *left*, DC-8 aircraft at *right*). **b** Time series of PAN measurements aboard Falcon and DC8 during the wing-to-wing intercomparison. The intercomparison period was between 13:31 and 14:23 UTC. PAN data (DC-8) provided by L. Gregory Huey of Georgia Institute of Technology

17.2.5 Platforms for Atmospheric CIMS Measurements

17.2.5.1 Research Aircraft

Several successful measurement campaigns have been performed using CIMS, mostly with the Falcon aircraft. Examples of atmospheric measurements and the accompanying *intercomparisons* are shown below and listed in Table 17.1. First atmospheric CI-ITMS measurements aboard HALO were also successfully made by IPA (Schlager et al. 2012). The first example is a wing-by-wing intercomparison flight between the DLR Falcon and the NASA DC8 in the framework of the POLARCAT-GRACE[3] campaign in Greenland in summer 2008. An IPA CI-ITMS instrument was aboard the Falcon measuring PAN (Roiger et al. 2011) with an online calibration (see textbox). Figure 17.8a shows a photograph of the NASA DC8 taken from the Falcon cabin. Figure 17.8b gives the PAN mole fraction over time measured by the DLR-IPA CI-ITMS instrument aboard the Falcon, and by a similar CIMS instrument using a LQMS aboard the NASA DC8 operated by the Georgia Institute of Technology. During the intercomparison measurement period, the aircraft separation was less than $\sim$10 m in the horizontal, and less than $\sim$200 m in the vertical. Throughout the whole intercomparison period, the PAN measurements of the two instruments showed very good agreement within the reported measurement uncertainties (10 %). The PAN time-series before and after the intercomparison when the aircraft were still separated illustrate the high variability of ambient PAN at different locations and altitudes.

[3] Polar Study using Aircraft, Remote Sensing, Surface Measurements and Models, of Climate, Chemistry, Aerosols, and Transport–Greenland Aerosol and Chemistry Experiment.

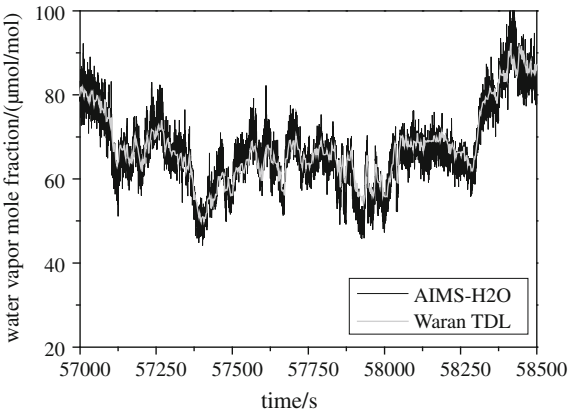

Fig. 17.9 Water vapor mole fractions versus measurement time detected on 16 September 2011 with the Falcon at 11 km altitude above Germany. An intercomparison of H_2O data from the AIMS instrument and a tunable diode laser hygrometer is shown (figure provided by S. Kaufmann)

PAN calibration source (isotopically labeled)
PAN is thermally very labile so it has to be produced in situ. It is photochemically produced via UV photolysis of isotopically labeled acetone ($(^{13}CH_3)_2^{13}CO$). NO is added and oxidized by the product (PA-radical, $CH_3C(O)O_2$) of the reaction of CH_3CO and O_2.These reactions take place in a temperature- and pressure-controlled reaction chamber. For details and online in-flight calibrations see (Roiger et al. 2011a).

Atmospheric water vapor is an important *greenhouse gas*. Low levels of water vapor in the *lower stratosphere* influence the *radiative budget* of the atmosphere and *temperatures* at the earth's surface. Large uncertainties exist in the accurate determination of low water vapor concentrations prevailing in the UTLS region and the lower stratosphere. To this end in cooperation with NOAA (D. Fahey, T. Thornberry), the atmospheric water vapor mass spectrometry originally developed at NOAA has been further developed at IPA and first successful flights were performed with the Falcon during the CONCERT2011 campaign above Germany. AIMS-H2O has a high time resolution of 0.2 s enabling fast water vapor measurements, for example in turbulent or cloudy regions where small scale fluctuations and strong H_2O gradients are expected. The high sensitivity and an in-flight calibration (see textbox) make water vapor detection with AIMS-H2O particularly interesting in the UTLS region at low water vapor mole fractions. A H_2O intercomparison of water vapor detected with AIMS-H2O and a tunable diode laser hygrometer on the Falcon aircraft on 16 September 2011 in *cirrus* clouds above Germany is shown in Fig. 17.9. Good agreement between the instruments is achieved for H_2O mole fractions above 50 μmol/mol; lower water vapor mole fractions were accurately measured with AIMS-H2O during other flights. The *accuracy* of AIMS-H2O is about 8 %.

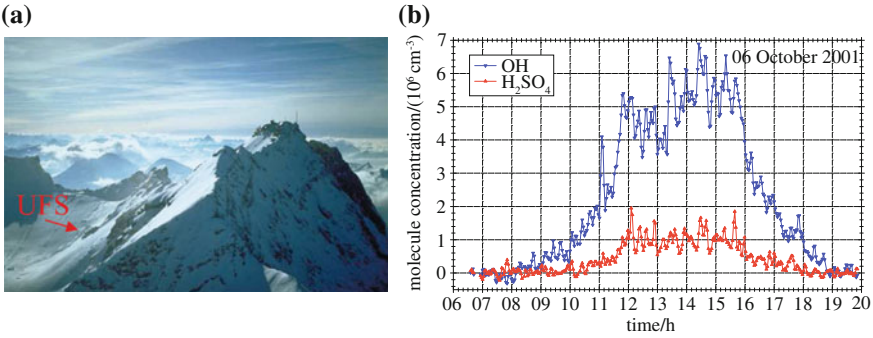

Fig. 17.10 (**a**, at *left*) Zugspitze: summit and Environmental Research Station Schneefernerhaus (UFS). (**b**, at *right*) Diurnal atmospheric variations of OH and gaseous H_2SO_4 measured with CIMS during a sunny day at Schneefernerhaus (UFS). Adapted from Aufmhoff et al. (2011) with permission from Elsevier

H_2O calibration source
In the H_2O calibration source originally developed at NOAA (Rollins et al. 2011), specific mole fractions of H_2 in zero air are converted to H_2O on a platinum foil heated to 200 °C. H_2 is not affected by wall effects like H_2O, so it can be handled much easier. Low water vapor mole fractions can be produced with this catalytic H_2O calibration source with high accuracy. Well defined H_2O mole fractions are directly delivered to the inlet system of the mass spectrometer during in-flight calibration sequences.

17.2.5.2 Schneefernerhaus Mountain Station at the Zugspitze

Gaseous H_2SO_4 and OH were measured at the *Schneefernerhaus* UFS (2650 m above sea level, below the summit of Mount Zugspitze in the Bavarian Alps, Germany). The UFS has the status of a global station within the *Global Atmosphere Watch Programme* of *World Meteorological Organization* (WMO). It is located ≈310 m below the summit on the southern slope of the mountain (see Fig. 17.10a). The air masses came mostly out of the west from the Schneefernerscharte between the summits of Schneefernerkopf (at the left edge of Fig. 17.10a) and Zugspitze.

H_2SO_4/OH calibration source
The H_2SO_4/OH source originally developed at MPI-K HD is based on water vapor photolysis. Purified air is humidified to a certain amount and irradiated with UV light in a well-defined zone. The trace gas inlet (the same as the one used during atmospheric measurements) sucks a fraction of this air containing OH radicals. To produce sulfuric acid, an inlet for a SO_2/N_2 gas mixture is

positioned just after the trace gas inlet in an opposite flow direction to ensure fast mixing with the probe air. Water vapor photolysis takes place in a homogeneously illuminated well defined zone under preferably well-defined laminar gas flow conditions. For details see Aufmhoff et al. (2011).

Simultaneous measurements of gaseous H_2SO_4 and OH by CI-ITMS showed a diurnal variation of both trace species due to photochemically produced OH radicals (see Fig. 17.10b). The very low detection limit of OH-produced H_2SO_4 with a corresponding mole fraction of about 6 fmol/mol at a time resolution of about 2 min is remarkable. A detailed description of these measurements is given in Aufmhoff et al. (2011).

17.2.5.3 Ships

Two ship containers were each equipped with two CIMS instruments and additional measurement devices at DLR Oberpfaffenhofen and MPI-K HD. Air conditioning, guided (blower) inlets, several mounting parts, etc. were installed. One ship container with two CI-ITMS instruments was installed aboard the research vessel Celtic Explorer in the framework of the EU project MAP (Marine Aerosol Production from Natural Sources, 2006) (see Fig. 17.11). In this project *marine aerosol* production from natural sources was investigated. Besides other trace gases, sulfuric acid (showing diurnal variations) was observed over the North Atlantic Ocean by CI-ITMS. The other ship container with two CI-LQMS (AIMS) instruments was installed aboard the research vessel SONNE in the framework of the EU project SHIVA (Stratospheric Ozone: Halogen Impacts in a Varying Atmosphere). The cruise took place in the South China Sea and the Sulu Sea near the coast of Malaysia and the Philippines 2011 with the aim of measuring halogen compounds.

17.2.5.4 Simulation Chambers

In May 2010 the HONO intercomparison campaign FIONA (Formal Intercomparisons of Observations of Nitrous Acid) took place at the European *Photoreactor* (EUPHORE) in Valencia, Spain. At the heart of EUPHORE are two identical Teflon domes (volume $\sim$200 m^3 each) in which tropospheric *photochemistry* is simulated with natural sunlight (Fig. 17.12b). Nighttime conditions can be reproduced by closing the solid domes surrounding the Teflon chambers. The objectives of FIONA were twofold: to improve the understanding of the formation of nitrous acid (HONO) and its role as a source of OH radicals and to evaluate the various measurement techniques under appropriate controlled conditions (semi- to heavily-urban) by means of open and blind experiments. There were 16 groups using nine different measurement techniques participating in FIONA, including the DLR-LQMS AIMS. AIMS was operated with SF_5^- reagent ions. It was calibrated by the addition of known amounts of HONO before and after the individual FIONA experiments (textbox). The detection of HONO with SF_5^- ions is water vapor dependent. To reach high accuracy (about 20 %) calibration was done

Fig. 17.11 CI-ITMS ship container with two blower inlets on top of the roof aboard the research vessel celtic explorer

for a wide range of HONO (0.3–5 nmol/mol) and water vapor concentrations (1 500–11 000 μmol/mol). AIMS measured reasonable HONO fractions between 0.3–15 nmol/mol (the lower limit is defined by the calibrated humidity range rather than the detection limit). In general, the results were in good agreement with the measurements performed by the other participants. Figure 17.12a shows the HONO mole fractions measured with the different instruments during the second open experiment on 18 May 2010. The instruments were tested for interference by the addition of methyl nitrite and t-butyl nitrite at 11:38 UTC and 13:01 UTC, respectively. While some instruments reacted to the addition of the nitrites, no interference was observed for AIMS within the measurement precision.

HONO calibration was performed with a HONO source operated by the University of Wuppertal. Gaseous HONO was produced via the liquid phase reaction of sodium nitrite with sulfuric acid and subsequent HONO release to the gas phase. HONO mole fractions were simultaneously measured with a LOPAP instrument (LOng Path Absorption Photometer, University of Wuppertal) with an uncertainty of 8 % (Kleffmann et al. 2006).

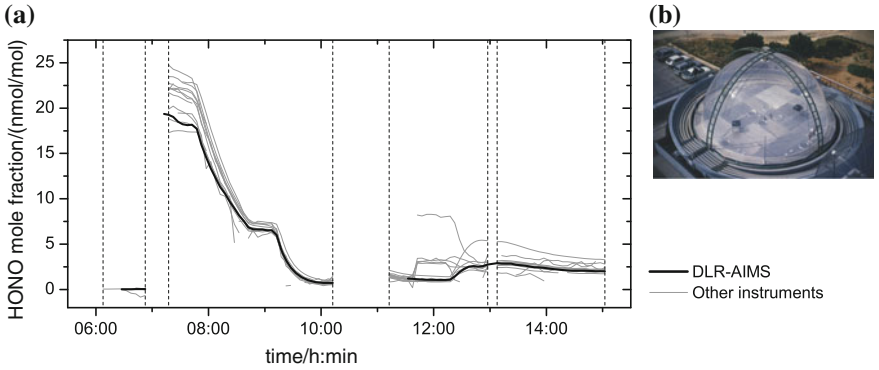

Fig. 17.12 **a** HONO mole fractions measured by the different groups on 18 May 2010 during the FIONA campaign. The experiment protocol: around 07:00 UTC HONO addition, chamber flushed with synthetic air from 07:46 to 08:43, at 09:11 UTC solid dome opened followed by photolysis of HONO, solid dome closed at 10:01 UTC, addition of methyl nitrite and t-butyl nitrite at 11:38 UTC and 13:01 UTC, respectively. **b** One of the two domes of EUPHORE (http://euphore.es/)

17.3 Outlook

Besides the continuation and extension of current CIMS measurements (e.g., SO_2 and H_2O in the lower stratosphere, HONO *flux measurements*, many other trace gases like HNO_3 and HCl in the UTLS), new CIMS analyzers will be built and used by IPA (two examples are mentioned here).

A new *volatile aerosol* component CIMS analyzer (via vaporization of aerosol particles) will be deployed. Further identifications of aerosol components will help to gain new insights in new *particle formation* and growth.

There is a need to identify condensable atmospheric *hydrocarbons* which play an important role in the formation of aerosols. CI-TQMS will be deployed as an aircraft instrument for measuring and identifying these compounds. Many hydrocarbons can be detected with CIMS, e.g., by using H_3O^+ ions and proton transfer reactions (see Ion Flow Reactor. The CI-TQMS instrument is ideal for this task because of its high mass resolution (mass defects as a help for identification), sensitivity, and fragmentation possibilities.

References

Arnold, F., Scheid, J., Stilp, T., Schlager, H., Reinhardt, M.E.: Measurements of jet aircraft emissions at cruise altitude I: the odd-nitrogen gases NO, NO_2, HNO_2 and HNO_3. Geophys. Res. Lett. **12**, 2421–2424 (1992). doi:10.1029/92GL02926

Arnold, F., Schneider, J., Gollinger, K., Schlager, H., Schulte, P., Hagen, D.E., Whitefield, P.D., van Velthofen, P.: Observation of upper tropospheric sulfur dioxide- and acetone-pollution:

potential implications for hydroxy radical and aerosol formation. Geophys. Res. Lett. **24**, 1 (1997). doi:10.1029/96GL03693

Aufmhoff, H., Hanke, M., Uecker, J., Schlager, H., Arnold, F.: An ion trap CIMS instrument for combined measurements of atmospheric OH and H_2SO_4: First test measurements above and inside the planetary boundary layer. Int. J. Mass Spectrom. **308**, 26–34 (2011). doi:10.1016/j.ijms.2011.07.016

Curtius, J., Sierau, B., Arnold, F., Baumann, R., Busen, R., Schulte, P., Schumann, U.: First direct sulfuric acid detection in the exhaust plume of a jet aircraft in flight. Geophys. Res. Lett. **25**, 923–926 (1998). doi:10.1029/98GL00512

Curtius, J., Arnold, F., Schulte, P.: Sulfuric acid measurements in the exhaust plume of a jet aircraft in flight: implications for the sulfuric acid formation efficiency. Geophys. Res. Lett. **29**, 1113 (2002). doi:10.1029/2001GL013813

Fiedler, V., Nau, R., Ludmann, S., Arnold, F., Schlager, H., Stohl, A.: East Asian SO_2 pollution plume over Europe—Part 1: airborne trace gas measurements and source identification by particle dispersion model simulations. Atmos. Chem. Phys. **9**, 4717–4728 (2009). doi:10.5194/acp-9-4717-2009

Huey, L.G., Dunlea, E.J., Lovejoy, E.R., Hanson, D.R., Norton, R.B., Fehsenfeld, F.C., Howard, C.J.: Fast time response measurements of HNO_3 in air with a chemical ionization mass spectrometer. J. Geophys. Res. **103**(D3), 3355–3360 (1998). doi:10.1029/97JD02214

Jurkat, T., Voigt, C., Arnold, F., Schlager, H., Aufmhoff, H., Schmale, J., Schneider, J., Lichtenstern, M., Dörnbrack, A.: Airborne stratospheric ITCIMS measurements of SO_2, HCl, and HNO_3 in the aged plume of volcano Kasatochi. J. Geophys. Res. **115**, D00L17 (2010). doi:10.1029/2010JD013890

Jurkat, T., Voigt, C., Arnold, F., Schlager, H., Kleffmann, J., Aufmhoff, H., Schäuble, D., Schaefer, M., Schumann, U.: Measurements of HONO, NO, NO_y and SO_2 in aircraft exhaust plumes at cruise. Geophys. Res. Lett. **38**(L10807), 1–5 (2011). doi:10.1029/2011GL046884

Kiendler, A., Arnold, F.: Detection of gaseous oxygenated hydrocarbons in upper tropospheric and lower stratospheric aircraft borne experiments. Int. J. Mass Spectrum. **223**, 733–741 (2003). doi:10.1016/S1387-3806(02)00969-7

Kleffmann, J., Lörzer, J., Wiesen, P., Kern, C., Trick, S., Volkamer, R., Rodenas, M., Wirtz, K.: Intercomparison of the DOAS and LOPAP techniques for the detection of nitrous acid (HONO). Atmos. Environ. **40**, 3640–3652 (2006). doi:10.1016/j.atmosenv.2006.03.027

Marcy, T.P., Gao, R.S., Northway, M.J., Popp, P.J., Stark, H., Fahey, D.W.: Using chemical ionization mass spectrometry for detection of HNO_3, HCl, and $ClONO_2$ in the atmosphere. Int. J. Mass Spectrom. **243**, 63–70 (2005). doi:10.1016/j.ijms.2004.11.012

Möhler, O., Arnold, F.: Gaseous sulfuric acid and sulfur dioxide measurements in the artic troposphere and lower stratosphere: implications for hydroxyl radical abundances. Geophys. Res. Lett. **19**, 1763–1766 (1992). doi:10.1029/92GL01807

Reiner, T., Hanke, M., Arnold, F., Ziereis, H., Schlager, H., Junkermann, W.: Aircraft-borne measurements of peroxy radicals by chemical conversion/ion molecule reaction mass spectrometry: calibration, diagnostics, and results. J. Geophys. Res. **104**, 18647–18659 (1999). doi:10.1029/1999JD900312

Roiger, A., Aufmhoff, H., Stock, P., Arnold, F., Schlager, H.: An aircraft-borne chemical ionization—ion trap mass spectrometer (CI-ITMS) for fast PAN and PPN measurements. Atmos. Meas. Tech. **4**, 173–188 (2011). doi:10.5194/amt-4-173-2011

Rollins, A.W., Thornberry, T.D., Gao, R.-S., Hall, B.D., Fahey, D.W.: Catalytic oxidation of H_2 on platinum: a robust method for generating low mixing ratio H_2O standards. Atmos. Meas. Tech. **4**, 2059–2064 (2011). doi:10.5194/amt-4-2059-2011

Schlager, H., Arnold, F., Aufmhoff, H., Baumann, R., Pirjola, L., Roiger, A., Sailer, T., Wirth, M., Schumann, U.: First detection of East-Asian anthropogenic SO_2 lifted to the lower stratosphere by a warm conveyor belt: HALO-aircraft measurements above Germany. Geophys. Res. Lett. (2012) (submitted)

Slusher, D.L., Huey, L.G., Tanner, D.J.: A thermal dissociation—chemical ionization mass spectrometry (TD-CIMS) technique for the simultaneous measurement of peroxyacyl nitrates and dinitrogen pentoxide. J. Geophys. Res. **109**, D19315 (2004). doi:10.1029/2004JD004670

Speidel, M., Nau, R., Arnold, F., Schlager, H., Stohl, A.: Sulfur dioxide measurements in the lower, middle and upper troposphere: Deployment of an aircraft-based chemical ionization mass spectrometer with permanent in-flight calibration. Atmos. Environ. **41**, 2427–2437 (2007). doi:10.1016/j.atmosenv.2006.07.047

Voigt, C., et al.: In situ observations of young contrails—overview and selected case studies from the CONCERT campaign. Atmos. Chem. Phys. **10**, 9039–9056 (2010). doi:10.5194/acp-10-9039-2010

In Situ Measurement Methods for Atmospheric Aerosol Particles and Cloud Elements

18

Andreas Minikin, Andreas Petzold, Bernadett Weinzierl and Jean-François Gayet

Abstract

A multitude of measurement techniques has been developed in atmospheric science in order to study atmospheric aerosols. In situ measurements can address many characteristics of aerosol particles such as concentration, size, shape, composition, and optical properties. Measuring aerosol particles as well as cloud droplets and ice crystals throughout the troposphere and lower stratosphere from fast flying research aircraft poses specific challenges with respect to suitable instrumentation.

This chapter is dedicated to the memory of Hans Rüba (*1953, †2011), who worked little wonders as a technician at DLR with almost all the instruments mentioned here. He was also a dear friend.

A. Minikin (✉) · A. Petzold · B. Weinzierl
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Andreas.Minikin@dlr.de

A. Petzold
e-mail: a.petzold@fz-juelich.de

B. Weinzierl
e-mail: Bernadett.Weinzierl@dlr.de

J.-F. Gayet
Laboratoire de Météorologie Physique, Université Blaise Pascal, 24 av. des Landais,
BP80026 Clermont-Ferrand, 63171 Aubière Cedex, France
e-mail: j.f.gayet@opgc.univ-bpclermont.fr

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_18,

18.1 Introduction

Atmospheric *aerosol particles* are ubiquitous in Earth's atmosphere. Their occurrence in terms of total number concentration is highly variable in both the spatial and the temporal domain. However, unlike gaseous species, a particle population in the atmosphere is not adequately described in terms of a concentration quantity only, because airborne aerosol particles occur over a very wide spectrum of particle sizes (from 10^{-9} to 10^{-4} m scale, not even considering *cloud droplets* and *ice crystals*, which are even larger). Moreover, particles may have varying shapes, varying *chemical compositions* and varying internal structures, which in turn results in varying optical or physical properties (refractive index, hygroscopicity, etc.). A particle population may be formed of a mixture of particles which individually do not have the same characteristics. Furthermore, the chemical and physical characteristics of aerosol particles are not conservative. They change with time and during transport through the atmosphere due to interaction with other particles and gases. This complexity resulted in the development of a multitude of measurement techniques in atmospheric science in order to study atmospheric aerosols.

If aerosol properties in the atmosphere are determined with suitable techniques, either ground-based or airborne, directly at their location, one commonly refers to an "in situ" measurement. In contrast, "*remote sensing*" measurements are those in which observations are taken from a distance. Remote sensing techniques of atmospheric aerosols can be either passive, based on the interference of aerosol particles with natural *radiation*, or active, using a *lidar* technique. This chapter will highlight the in situ measurement methods established for aerosol characterization at the Institute of Atmospheric Physics at DLR Oberpfaffenhofen for a broad range of research topics. The focus will be on discussion of the techniques applied and developed for use on research aircraft, primarily on the DLR *Falcon* aircraft. It should be mentioned, however, that the DLR aerosol *instrumentation* was also often successfully deployed for ground-based or laboratory research, most notably for test rig studies of particle properties in fresh *emissions* of aircraft turbines or ship engines (Petzold et al. 2005, 2008).

18.2 Aerosol Microphysical In Situ Measurement Techniques

Condensation particle counters. For decades the *condensation particle counter* (CPC; also known as condensation nuclei counter, CNC) technique has been well established for the measurement of the integral number of particles down to very small sizes of a few nm in diameter. Particles below roughly 0.1 μm diameter cannot be easily detected directly by optical means with visible light. CPCs therefore adopt the principle of exposing the sample flow to an environment of a strong *supersaturation* of a working substance, which condenses rapidly onto the pre-existing particles and grows them into much larger sizes suitable for optical

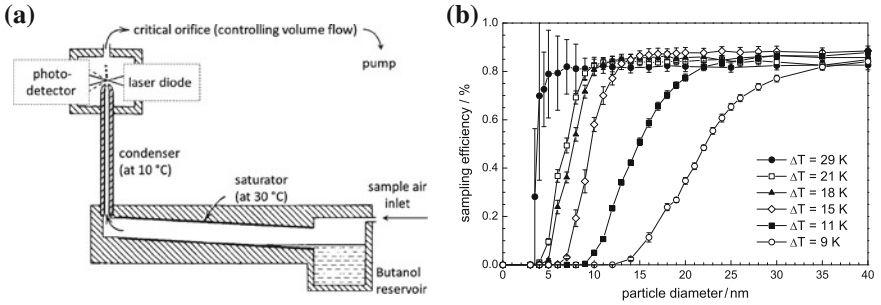

Fig. 18.1 **a** Schematic of a condensation particle counter (CPC); adapted from Hinds (1999). **b** Laboratory *calibration* of CPC module with different supersaturation (ΔT) settings showing *sampling efficiency* (against reference CPC) versus particle size

detection. An overview of different types of CPCs is given by Kulkarni et al. (2011). Figure 18.1a illustrates a common, relatively simple layout of a CPC using an alcohol (1-Butanol) as a working substance. The sample air passes through a warm part, the saturator, where the alcohol evaporates and mixes with sample air. Thereafter, the sample flow is directed through a cold section, the condenser, causing the sample air to become supersaturated and all particles to increase in size due to uptake of the alcohol. Subsequently, particles are counted by detecting the light scattered when they pass through a laser beam. The CPC measurement is an integral measurement of the total number concentration of particles above the smallest detectable particle size, the so-called lower cut-off size. The CPC measurement typically works very well in a particle concentration range of 1–10 000 cm^{-3}. The cut-off diameter is primarily dependent on the degree of supersaturation achieved inside the instrument. In the CPCs used so far on the Falcon aircraft (mostly modified TSI 3010 or 3760A models), the cut-off diameter is controlled by the temperature difference between the saturator and the condenser and it can be operationally varied between 4 and 20 nm (Stein et al. 2001; Fig. 18.1b). The upper limit in the size range of a CPC is typically $\sim$1 μm. It is determined by the particle transmission characteristics of the *inlet* system, but it is usually of no relevance to the integral measurement, because in virtually all atmospheric environments the larger particles in the *coarse mode* are by far outnumbered by the smaller particles.

CPC measurements are fairly fast (typically 1 Hz resolution). Combining CPCs with different lower cut-off sizes allows the particle number *size distribution* to be assessed in the ultrafine size range down to a few nm (see below, Sect. 18.4.1). Adding diffusion screen separators in front of the CPC inlet (Feldpausch et al. 2006) allows the lower cut-off diameter to be increased up to $\sim$80 nm. Compared to the *differential mobility analyzer* (DMA) method (discussed below) this is an important technique to allow size segregated measurements in the *Aitken mode* size range (10–100 nm) if fast instrument response is required. Results of CPC measurements in the *troposphere* have been reported in many DLR studies

(Schröder et al. 2002; Minikin et al. 2003; Hamburger et al. 2011). In environments with very high particle concentrations ($>10^4$ cm^{-3}), like in in-flight jet engine emission studies (Brock et al. 2000; Schröder et al. 2000), a dilution system is needed upstream of the CPC system.

CPCs are usually not selective in detecting particular types of aerosol particles. However, if a *thermal denuder* (Fierz et al. 2007) is used upstream of a CPC, information on the thermal properties of the aerosol particles can be obtained, which in turn are related to their chemical nature. In a thermal denuder, the sample air is heated. Particulate components which are *volatile* at the set temperature evaporate. A temperature setting of 250 °C, for instance, results in the removal of volatile components like *sulfuric acid* or *semivolatile organic aerosol* components, whereas *nonvolatile* particles like *soot*, *mineral dust* or *sea salt* remain in the sample flow. An application of this technique will be discussed in Sect. 18.4.3 further below.

Differential mobility analyzers. The DMA technique allows the size segregated measurement of particles typically in the size range of Aitken mode and *accumulation mode* aerosol (10−500 nm particle diameter). The method only works for charged particles. It is based on the effect that charged particles move in an electrical field on prescribed trajectories defined by their electrical mobility, a quantity which is dependent on the ratio of the particle diameter and the elemental charge the particle is carrying. DMA theory is explained in more detail for instance in Kulkarni et al. (2011). A commonly used set-up is that of a cylindrical DMA, where sample air is drawn through two coaxial electrodes in laminar flow. According to the voltage set between the two electrodes, all charged particles follow trajectories bending towards either of the electrodes. The geometrical design of the DMA, however, allows only particles of a certain electrical mobility (for a given voltage) to exit the DMA for subsequent detection by a CPC or electrometer. All particles having other electrical mobilities are removed from the sample air stream because they attach to the walls (electrodes) inside the DMA.

In order to relate the particle's electrical mobility to its size, the charge of the particle needs to be known. In DMA applications this is achieved by passing the aerosol particles prior to entering the DMA through a so-called neutralizer, which usually contains a weak radioactive source providing ionizing radiation. The neutralizer brings the aerosol into charge equilibrium. The probability of particles to take none, one, or more charges can be calculated in this case. It is dependent on particle size and thermodynamic conditions and needs to be taken into account in the data inversion, as well as the geometric design and flow settings of the DMA. (The term "neutralizer" may seem a bit odd in this context, but the same technique is also used to reduce the particles' charges, i.e., to nearly "neutralize" the aerosol, if these were highly charged before.) The size distribution of the aerosol sample can be finally inferred by applying consecutively a range of fixed voltages (differential mobility particle sizer, DMPS) or scanning continuously through a voltage range (scanning mobility particle sizer, SMPS). For use in aircraft the major shortcoming of the technique is the limited time resolution for determining a

size distribution (typically 1–2 min). Results of the DMA system operated on board the Falcon are reported in, e.g., Weinzierl et al. (2009).

The Neutral Cluster/Air Ion Spectrometer (NAIS) is a new technique based on the DMA principle developed in aerosol science to determine particle sizes down to the smallest possible size of 1.5 nm (such as a cluster of only two molecules; Sipilä et al. 2010). The first successful airborne measurements with this technique were conducted on the Falcon in collaboration with the University of Tartu (Mirme et al. 2010).

Optical particle counters. Many variants of *optical particle counters* (OPCs) have been designed to determine the concentration and size of aerosol particles which are optically active (at least approximately 0.1 μm in size), many of them being commercially available and tailored towards use in atmospheric science and even specifically for aircraft measurements. The aircraft instruments are also known under the term "aerosol spectrometer probes." The measurement principle of these instruments is to size single particles based on the detection of the amount of light scattered if the particles pass through a laser beam. The instruments convert the scattering signal into an electrical signal, which corresponds to size information and is resolved into an array of "size" channels according to its amplitude. The instruments differ in optical geometry and electronic signal processing, and, as a result, cover different particle size ranges. For a comprehensive overview see again for instance Kulkarni et al. (2011).

A classical instrument of this type specifically designed for aircraft measurements and which has been in service on the Falcon for many years is the Forward Scattering Spectrometer Probe (FSSP, see Baumgardner et al. 1992; Gayet et al. 2002; Weinzierl et al. 2009) of the former company PMS Inc., now serviced by DMT Inc. (The terminology "*PMS probes*" has survived until today to describe all kind of aircraft instruments of other manufactures as well which can be fitted into standardized, cylindrical containers outside the fuselage or under the wing of an aircraft, and which are therefore independent of a sample air inlet system for the cabin.) The FSSP-300 version, for instance, has a nominal size range of 0.3–20 μm. Primary aerosol particles in this size range (soil and mineral dust, sea salt, *volcanic ash* and biogenic particles) may in general be few in comparison to smaller particles, which are always present in the troposphere, but they may dominate the total mass concentration as well as the aerosol optical effects. Sizing as accurately as possible in this size range is therefore a key requirement for many studies. An instrument of FSSP type is also essential for the measurement of small airborne hydrometeors, that is, small ice crystals and cloud droplets, of which the smallest relevant sizes fall into the FSSP size range. Figure 18.2 illustrates the working principal and optical pathways of the FSSP probe and how it looks in reality. The amount of light scattered by a single particle into the instrument's collection optics depends on particle size, but also on the complex *refractive index* and the *shape* of the particle. The latter two properties may be a priori not known and can pose a serious limit on the accuracy of size distribution measurements with OPCs. Size calibration is further discussed in Sect. 18.4.2 below. On the *HALO* aircraft a DMT Cloud and Aerosol Spectrometer (CAS-DPOL) will come

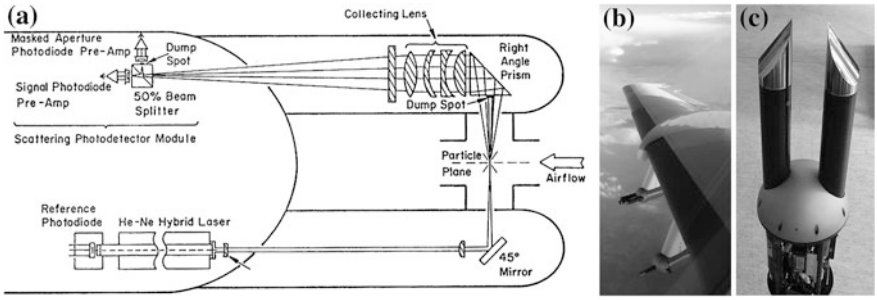

Fig. 18.2 **a** Operation principle of a PMS Forward Scattering Spectrometer Probe (FSSP) adapted from the manufacturer's manual, **b** FSSP-300 and PCASP-100X (foreground) probes in the Falcon wing pods, **c** Photograph of the modernized FSSP-300 probe with replaced *Korolev tips* which aim to reduce measurement artifacts in clouds due to shattering of cloud droplets or ice crystals

into service, which compared to the FSSP offers much improved capabilities (size range 0.5–50 μm with concurrent measurement of *backscatter* and *depolarization*, providing additional information on particle type and phase state).

Another work horse for airborne aerosol research at DLR has proven to be the PCASP-100X probe (Passive Cavity Aerosol Spectrometer Probe) as documented for instance by Fiebig et al. (2002). This is an OPC which can measure particles down to 0.12 μm in size by creating a more intense illumination of the particles and utilizing a much wider collection angle of the light scattered mainly sidewards, in contrast to the FSSP. Thereby the size range covered by this instrument includes largely the very important accumulation mode size range, in which essentially all aged atmospheric aerosol particles end up. The instrument is currently being replaced with the new Ultra-high Sensitivity Aerosol Spectrometer (UHSAS) of DMT Inc., which is available both in a rack-mounted and an airborne "PMS type" version and is considerably improved in offering size segregated measurements in a size range extended further downwards (60 nm to 1 μm).

OPCs like the Grimm SKY-OPC model 1.129 (nominal size range 0.25–30 μm) have proven to be useful because they are built to be small and lightweight, and like CPCs can also be operated downstream of a thermal denuder inside the aircraft cabin. However, their upper size range can be limited by the inlet system (see below).

Imaging and other cloud probes. The size of cloud droplets and ice crystals in the atmosphere covers several orders of magnitudes from 1 μm to 1 mm. OPC techniques as described in the previous section are only sufficient for covering the lower end of this size range. Instruments like the PMS 2D-C probe, the DMT *Cloud Imaging Probe* (CIP) or the SPEC *Cloud Particle Imager* (CPI) actually take images or shadow images of particles limited to upper sizes of 800, 1 500 and 2 500 μm, respectively, based on the principles of geometric optics. Modern instruments achieve a fine resolution on the order of a few μm based on pixel size and gray scale options. 2D-C and CPI results from Falcon measurements were

reported by, e.g., Gayet et al. (2002), Mioche et al. (2010), and Gayet et al. (2011a). The *Polar Nephelometer* is an instrument determining the phase function of single cloud particles, i.e., it measures the angular distribution of scattered light. Other techniques are under continuous development in the research community: The *Small Ice Detector* (SID) is an instrument detecting the refraction pattern of single ice particles, also allowing assessment of particle morphology (Cotton et al. 2010). Finally, holographic techniques are being applied to allow simultaneous measurements of an ensemble of particles inside a larger sample volume.

Optical properties. Atmospheric particles interact with radiation due to light scattering and *absorption*. Integrating nephelometers are commonly used to measure the bulk scattering coefficient of an air sample, often at multiple *wavelengths* (Kulkarni et al. 2011). The method, however, is not well applicable on aircraft platforms, because the instrument must operate downstream of an inlet system which might not have sufficient transmission for coarse mode particles (see Sect. 18.3). Based on so-called closure studies between in situ and lidar measurements, Petzold et al. (2002) demonstrated that the scattering contributed by coarse mode aerosol can be essential. On aircraft, scattering should rather be deduced for instance with *Mie theory* algorithms from size distribution measurements using OPCs mounted outside the aircraft cabin.

In contrast, the most particle absorption is usually caused by soot containing aerosol particles which are present at smaller sizes not so much affected by inlet effects. Measuring the light absorbing properties of aerosol particles is a key requirement for many atmospheric studies, because the absorption of a particle is directly related to its *climate impact*. The Particle Soot Absorption Photometer (PSAP) is an instrument which has been used in many DLR studies (e.g. Fiebig et al. 2002; Petzold et al. 2009, 2011). The measurement principle is to infer the particle absorption coefficient from the differential measurement of the transmission of light through a particle filter, through which sample air is continuously drawn. The concurrent measurement at three different wavelengths ($\lambda = 467$, 530 and 660 nm) of the 3λ-PSAP version (Virkkula 2010) was used by Petzold et al. (2009) in combination with size distribution measurements by optical probes to narrow down the complex refractive index of Saharan dust aerosols.

The Single Particle Soot Photometer [SP2; e.g. Schwarz et al. (2006)] is a modern instrument now being introduced for measurements with the Falcon and HALO aircraft to measure the *black carbon* (soot) in single particles based on detecting the incandescence radiation emitted by absorbing particles if they get heated and evaporated in a strong laser light beam.

Chemical composition. Ground-based analysis of the chemical composition of near-surface aerosol is often based on bulk chemical analysis (e.g., by ion chromatography) of filter samples, through which a volume of $\sim$100–10 000 m^3 of sample air has been drawn. This is not possible on an aircraft because there are restrictions on volume flow rates for a given inlet and on sampling times due to the aircraft's fast motion. Methods based on single-particle off-line analysis do however work well. A two-stage mini-impactor system segregating the aerosol sample into a particle fraction below and above 0.5 μm particle size was used

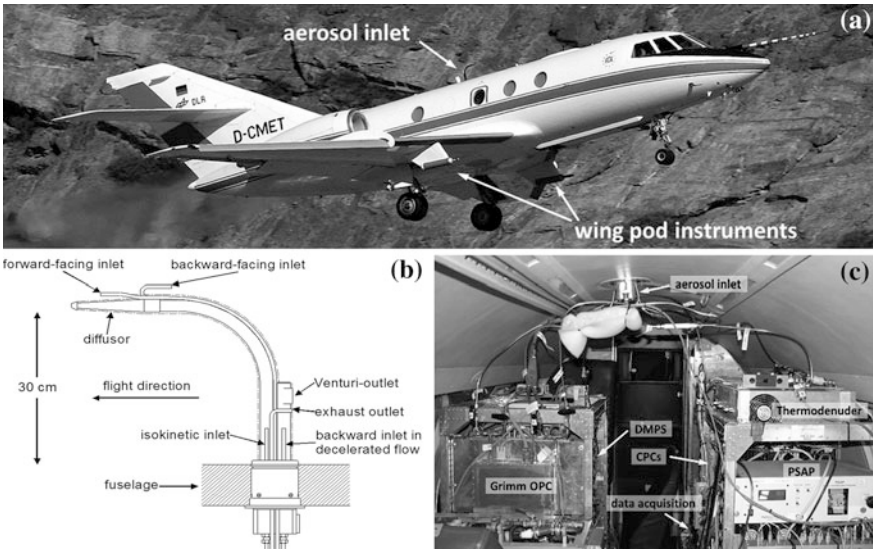

Fig. 18.3 **a** DLR Falcon shortly after take-off. The aerosol inlet is the large forward-facing, "banana"-shaped inlet on the top. In this photograph, the Falcon carries one wing pod for aerosol spectrometer probes under each wing, **b** schematic drawing of the Falcon aerosol inlet, **c** view inside the Falcon cabin with rack-mounted aerosol instruments in 2006, showing also the aerosol inlet and stainless steel tubing between inlets and instruments

during several studies with the Falcon for analysis by scanning electron microscope techniques (Kandler et al. 2009). The upper size limit for particle sampling techniques of this type is again defined by the inlet characteristics (see Sect. 18.3).

A different approach to measuring aerosol composition is the application of aerosol *mass spectrometers* on board aircraft. These can be targeted to analyze either individual particles or ensembles of particles. Schneider et al. (2006) and Schmale et al. (2010), for instance, describe results of an Aerodyne Mass Spectrometer (AMS) flown on the Falcon on tropospheric and volcanic aerosols.

18.3 Challenges of Aircraft Measurements

Research aircraft like the Falcon (Fig. 18.3a) or HALO fly at cruising speeds on the order of 200 m s^{-1}. Within 1 min the aircraft easily covers a distance of more than 10 km. The investigation of clouds or pollution plumes and other atmospheric phenomena often requires looking at shorter scales. Therefore, aerosol in situ measurement methods with fast response times and high time resolution on the order of seconds are generally preferred. As a result there is, broadly speaking, a general preference in airborne research to use microphysical methods able to count or sample single particles. As a consequence, *particle size distribution*s are usually based on the measurement of number concentrations (in different size bins),

whereas surface or volume/mass size distributions in these cases are derived properties.

If aerosol in situ instrumentation is installed inside the (pressurized) cabin of a research aircraft (Fig. 18.3c), one needs to ensure the transport of outside air into the cabin to the instruments with minimal impact on particle properties. Special aircraft aerosol inlets have been developed to ensure this (Wendisch et al. 2004), but there are physical limitations which cannot be evaded. Sample air is warmed before entering the instruments due to the deceleration of the air flow and heat transfer through sample lines inside the cabin. This means that aerosol particles will be exposed to lower *relative humidity* before the actual measurement by a cabin instrument. Even more important is the fact that aerosol particles have inertia: Larger particles cannot follow the streamlines of a gas flow. If the flow is bent, which is inevitably the case when taking sample air through inlet and tubing towards cabin instruments, large particles (typically of several μm size) will be removed due to impact on walls. Moreover, sampling of the aerosol inlet should be isokinetic, which means that a jump of flow velocity should not occur when sample air enters a tube, because otherwise larger supermicron particles can get enriched or depleted with respect to the ambient outside size distribution. Similarly, particle enrichment or depletion effects will occur inside the boundary layer around the aircraft. To overcome this, aerosol inlets designed for the Falcon (Fig. 18.3b) or HALO are mounted on the top of the fuselage and sample the air about 300 and 400 mm, respectively, away from the aircraft skin.

Because of the aforementioned shortcomings of the aerosol inlet system it becomes evident that the measurement of cloud elements as well as coarse mode aerosol particles of a few μm size requires probes specifically designed for use on aircraft external stores (like some of the OPCs and all cloud microphysics probes discussed in the previous section).

For the Falcon aerosol inlet, the sampling efficiency falls off rigorously for large particles. The cut-off diameter, at which 50 % of the particles pass the isokinetic inlet, depends on outside pressure and air speed. For typical Falcon speed, the cut-off is near 2.5 μm at ground level and decreases to about 1.5 μm at 10 km altitude (Wendisch et al. 2004). This is also nicely demonstrated in volcanic ash measurements, discussed by Schumann et al. (2011) in their Fig. 7 when comparing the cabin mounted OPC with the outside PCASP/FSSP-300 instruments.

Another aspect relevant to the operation of many aircraft-based aerosol instruments is the low ambient pressure at maximum flight altitude (e.g., 140 hPa in the case of HALO) and the fact that pressure continuously changes during climbs and descents. Many instruments like commercially available CPCs need to be modified to sustain these conditions, including sealing towards cabin pressure. In DMPS or SMPS systems the high voltage must be limited to avoid sparking between the electrodes, which otherwise is likely to occur at low pressures. Some instruments do not perform well during pressure changes in climbs and descents, and care must be taken in data evaluation also with sequences of steep turns of the aircraft or very high attack angles (nose up against air stream), because the flow may not be well enough aligned with the axis of the aerosol inlet. The low pressure

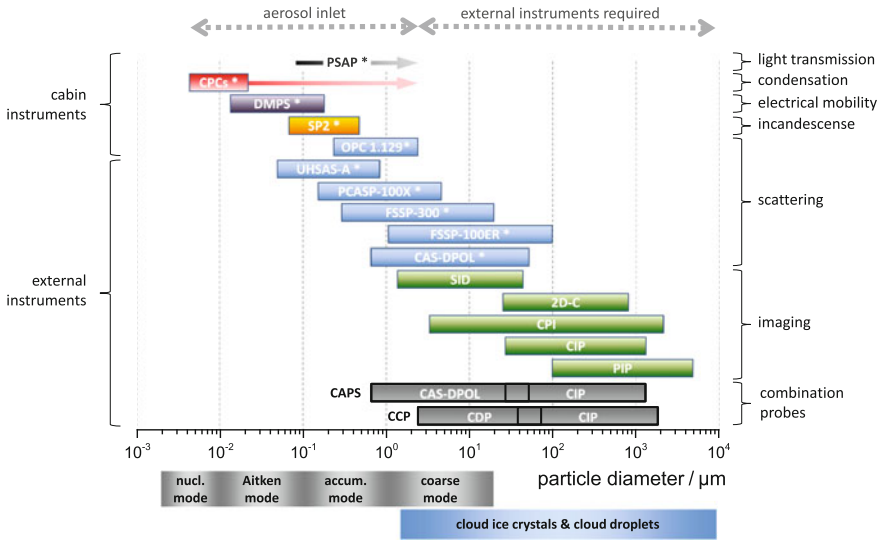

Fig. 18.4 Size ranges of aerosol and cloud in situ instruments used on the Falcon or planned for HALO. * denotes instruments operated currently by DLR. *Arrows* indicate integral measurements, whereas *boxes* represent size-segregated measurements. The basic measurement principles are listed on the right hand side of the figure

also affects the sampling efficiency and lower cut-off diameter of CPCs. This requires careful low-pressure calibration of CPCs in the laboratory.

Inside the cabin, the distribution of sample air to the different aerosol instruments needs to follow a couple of requirements to minimize particle losses. Sharp bends and T-pieces in the tubing should be avoided; they further reduce the transmission efficiency of particles in the 1–2 µm size range. Stainless steel tubing is in general preferred because it is conductive and minimizes particle losses due to build-up of static charges. It is important to select low enough flow rates according to the inner diameter of tubing so that laminar flow is ensured (*Reynolds number* <2000) wherever possible to avoid particle losses due to turbulence and wall contact. On the other hand, if CPC or DMA measurements are targeted to measure the smallest particle sizes below approximately 15 nm, diffusion loss to the inner walls become relevant, more so for the smallest particles due to their higher diffusivity. To optimize this, one should keep the distance short between inlet and instrument. Flow rates should be kept in the laminar flow regime but as high as possible (ensuring short residence times inside the tubing).

Figure 18.4 summarizes the size ranges covered by the instruments which have been deployed by DLR or collaborating groups in the past on the Falcon or other aircraft, or are planned for use on HALO. It should be noted that research aircraft like the Falcon or HALO in general offer the possibility for an onboard science crew to interact with the instruments. Other developments of aerosol instrument packages are designed to work autonomously, for instance on commercial airliners

within the framework of the IAGOS project (http://www.iagos.org) or on special aircraft like the *Geophysica*.

Finally, it should be mentioned that the deployment of any instruments on board larger research aircraft like the Falcon or HALO is subject to rigorous procedures in terms of air worthiness certification. While this is essential to ensure flight safety, it nowadays draws considerably on financial and manpower resources and poses limits to the natural interest of every experimental scientist to continuously change and improve the instrument hardware. For further reading on aircraft measurements see also the EUFAR (http://www.eufar.net/) initiated text book by Wendisch and Brenguier (2012).

18.4 Application of Particle Measurement Methods: Exemplary IPA Research Results

18.4.1 Ultrafine Particle Observations with Condensation Particle Counters

So-called secondary particles in the atmosphere are formed by the gas-to-particle conversion of natural or *anthropogenic precursor* gases like SO_2 or volatile organic compounds. Secondary particles form as atmospheric clusters which grow under suitable thermodynamic conditions into particles of only a few nm size, which makes them detectable by CPC instruments (Sect. 18.2). Combining concurrent measurements of multiple CPCs set to different lower cut-off sizes allows a size-segregated measurement to be achieved in the size range below $\sim$20 nm, often denoted as the *nucleation mode* size range. An application is depicted in Fig. 18.5 showing recordings of different CPCs when the Falcon was flying through the outflow of a *mesoscale convective system* over West Africa during the 2006 monsoon season (AMMA project). The Falcon flies away from the system, thereby sampling outflow air of increasing outflow age ($\sim$1–10 h). Initially (period 1) aerosol number concentrations are very low due to wet removal inside the storm. Subsequently, the particle concentrations in the nucleation mode (>4 nm) first pick up due to new *particle formation* setting in (period 2). Further growth with time leads subsequently also to an increase in Aitken mode particles >10 nm (period 3). Further implications are discussed by Reeves et al. (2010). New particle formation events were monitored with multichannel CPCs also during other aerosol field studies involving the Falcon, e.g. Khosrawi et al. (2010), Mirme et al. (2010), and Weigel et al. (2011).

18.4.2 Determining Size Distributions with Optical Particle Counters

In Sect. 18.2 OPC methods were introduced which have been deployed by DLR in the past in a multitude of aircraft studies on a variety of research topics. Accurate

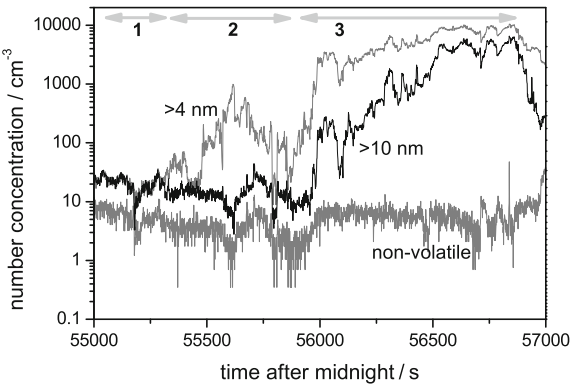

Fig. 18.5 Time series of number concentrations of condensation particles (>4 nm; >10 nm; nonvolatile particles) measured with the Falcon during a traverse of outflow of a mesoscale convective system at 11 km altitude. Flight time is given on the x axis. See text fur further explanation

data analysis in this context requires in-depth understanding of the method and careful size calibration often beyond the manufacturers' specifications. From physical principles the scattering signal detected by the OPC type instrument depends not only on particle size, which is the key target of the measurement, but also on the complex refractive index, and the shape of the particle. Therefore, a size calibration of the instrument (establishing a relationship between particle size and channel, or electrical response) is actually specific for a particle type and cannot generally hold. As a consequence, the following scheme for size calibration of OPCs has been established at DLR:

(a) Laboratory calibration of the instrument with known particle size standards, usually certified spherical polysterene latex particles (PSL) or glass beads.
(b) This establishes a relationship between the electrical signal (or channel number) and the particle size for a known refractive index (that of the calibration spheres).
(c) Using Mie theory for spherical particles for the known refractive index (of the calibration spheres) and the optical geometry of the instrument, the fraction of the scattering cross section seen by the instrument (called the "instrument response") can be theoretically calculated.
(d) This establishes a relationship of instrument response with the actual electrical signal (or channel number). This should be an instrument characteristic which does not change as long as the optical system stays clean.
(e) Finally, for any other particle type (with given refractive index and shape) the size-dependent instrument response in terms of the scattering cross section fraction needs to be calculated (e.g., based again on Mie theory), to establish with the result of step d) a new adequate size calibration of the instrument for the particular particle type in question.

In many cases, the data analysis of OPCs assumes spherical particles and applies Mie theory for the calculation of particle optical properties. However, as shown in Fig. 18.6a for the FSSP-300 probe, the relationship between particle size and particle scattering cross-section shows significant ambiguities in the size range

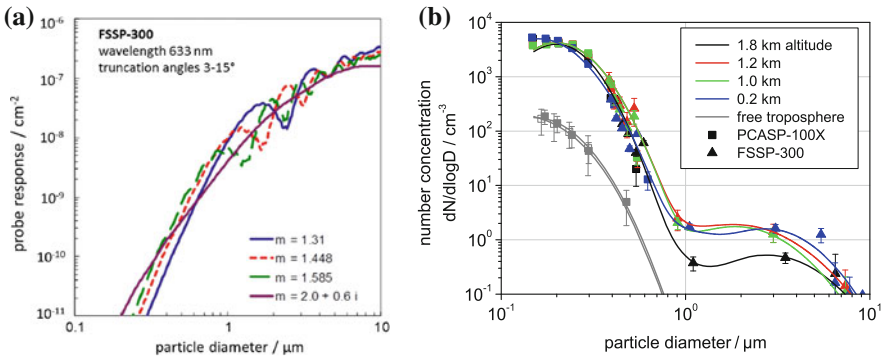

Fig. 18.6 **a** Probe response (fraction of scattering cross section) of FSSP-300 instrument as a function of particle diameter calculated with Mie theory for different refractive indices. **b** Size distribution composite of PCASP-100X and FSSP-300 data measured west of Ireland on 14 May 2008, in different altitudes

1–10 µm where the typical Mie interference patterns occur. Moreover, there is a significant dependence on the refractive index. Therefore, many error sources for the determination of particle size from light scattering have to be investigated: complex refractive index, particle mixing state, particle shape, and possibly inaccurately known optical geometry of the instrument. Fiebig et al. (2002) describe in detail how the various uncertainties and ambiguities affected the data analysis of OPCs during the Lindenberg Aerosol Characterization Experiment (LACE). Schumann et al. (2011) address the problem of measuring the mass concentration of airborne volcanic ash by means of OPC methods on the Falcon in the Eyjafjalla volcanic ash plume over Europe in 2010. It was found that in this context the uncertainties in the refractive index for the supermicron aerosol fraction dominated the fairly large uncertainty of the final mass concentration determination.

Figure 18.6b shows an actual example of aerosol size distributions derived during a flight of the Falcon near Ireland addressing the transport of central European pollution within the EUCAARI project (Kulmala et al. 2011). The graph illustrates how PCASP and FSSP-300 data combine. In this case, pollution aerosol prevails in the lower troposphere below 2 km altitude with a well pronounced coarse mode (>1 µm) which is entirely absent in the clean free troposphere.

18.4.3 Particle Mixing State: Using Thermal Fractionation Techniques

The size distribution of the nonvolatile aerosol fraction can be measured if CPC, DMPS and OPC (see Sect. 18.2) measurements are performed concurrently for the total ambient aerosol as well as behind a thermal denuder heating the air sample, for instance to 250 °C. This is a powerful method allowing conclusions to be

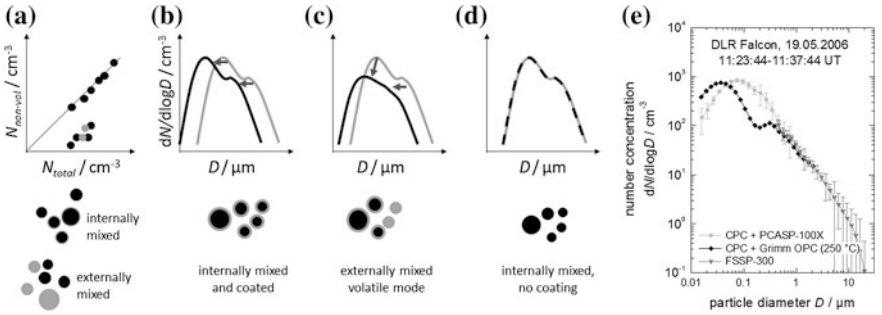

Fig. 18.7 Concept of using measurements of total (volatile) and nonvolatile aerosol particles to determine particle mixing state. **a** Conceptual scatter plot of integral number concentrations (total and nonvolatile). **b–d** Conceptual particle size distributions (*grey curve*: total; *black curve*: nonvolatile). **e** Example of an actual size distribution measurement of Saharan dust with and without employment of a thermal denuder

drawn about the aerosol *mixing state*, as illustrated in Fig. 18.7 and also discussed by Weinzierl et al. (2009). Panel a) depicts a conceptual scatter plot of the integral nonvolatile number concentration and the total particle number concentration. If one assumes all particles to be composed of a nonvolatile core surrounded by a volatile coating (a so-called *internal mixture*), data points should group along the 1:1 line, because even if a volatile coating is evaporated inside the thermal denuder, the nonvolatile core remains for detection by the CPC. However, if data points appear below the 1:1 line, the measurements represent an externally mixed aerosol, meaning that there was in addition to nonvolatile (possibly coated) particles also a population of entirely volatile particles which evaporated and would be missing in the particle count.

To investigate in more detail possible coatings on particles, the actual size distributions of the untreated fraction and the nonvolatile fraction have to be compared. Panels (b–d) in Fig. 18.7 show the relation of the total aerosol size distribution (grey line) and the nonvolatile size distribution (black line). In an internally mixed and coated aerosol (Panel b), the total and the nonvolatile size distribution would have the same shape, but the maxima of the nonvolatile size distribution would be shifted to smaller particle sizes. One would typically expect such a mixing state for an aged aerosol. In an externally mixed aerosol that contains a number of completely volatile particles, the shape and height of the total and the nonvolatile size distribution would differ (Panel c). Such behavior can be observed for example in fresh ship emissions. In an internally mixed aerosol without coatings, like in freshly emitted dust, the total and the nonvolatile size distribution would be identical (Panel d). Panel (e) finally depicts a real measurement taken over Morocco: Two size regimes can be distinguished in this case: one below 0.5 µm resembling the case in panel (b) (possibly anthropogenic background aerosol) and one above 0.5 µm resembling the case in panel (d) (Saharan dust).

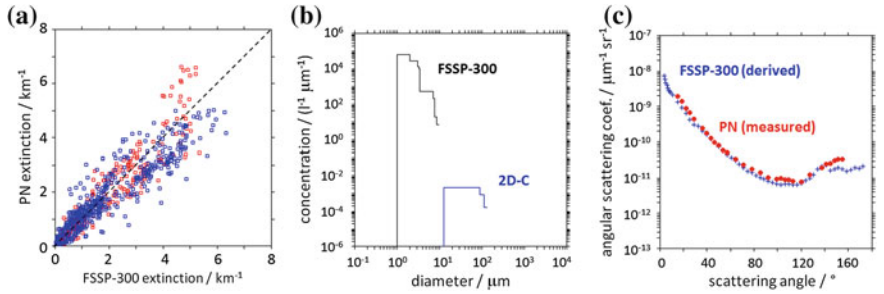

Fig. 18.8 Measurements inside a young A380 contrail. **a** Scatter plot of extinctions derived from PN and FSSP-300, for FSSP-300 calibration assuming spherical (*red symbols*) and nonspherical (*blue symbols*) ice particles, **b** Number size distribution of ice crystals measured by FSSP-300 and 2D-C, **c** Phase function measured by PN and inferred from FSSP-300 data

18.4.4 Combining Data of Wing Pod Instruments: Contrail Microphysical Properties

For more than 15 years the Falcon has also been intensively used during experiments devoted to *cirrus cloud* and *contrail* observations, for instance: SULFUR/ Aerocontrail (Schröder et al. 2000; Schumann et al. 2002); INCA (Gayet et al. 2004); PAZI/CIRCLE-1 (Febvre et al. 2009); PAZI/CIRCLE-2 (Gayet et al. 2011a); CONCERT (Voigt et al. 2011). Based on the data of these experiments further important work addressed the *validation* of *satellite* observations, e.g., of CALIPSO/CALIOP (Mioche et al. 2010), or estimates of the optical *extinction* and climate effect of line-shaped contrails (Voigt et al. 2011). During most of the experiments the aircraft was equipped with several in-situ instruments to infer microphysical and optical ice particle properties: e.g., the FSSP-300, 2D-C, CPI and the Polar Nephelometer (PN). It is only the combination of these types of instruments which provides a comprehensive description of ice crystal properties within a diameter range from 1 µm to some mm. As an example, Fig. 18.8 presents results from measurements in a contrail emitted by an A380 aircraft during the CONCERT experiment in 2008 (Gayet et al. 2011b). Panel (a) displays a comparison between the extinction coefficients derived from the FSSP-300 and the PN inside the contrail, demonstrating good agreement between the two independent measurements in this case. Panel (b) shows the ice particle size distribution measured in the contrail as a composite of FSSP-300 and 2D-C measurements. The *scattering phase function* as measured by the PN (circles) and as derived by calculation from the FSSP (*plus symbols*) is shown in Panel (c). It is an important conclusion that in this case the two independent measurements give consistent results. This is not necessarily the case because it has been found in the past that in some cirrus studies particularly the measurements of the small fraction of the ice particles (<20 µm) with probes like the FSSP-300 were affected by the shattering of large ice crystals at the frontal parts of the probe, generating an artificially high number of small crystals. There are different concepts to technically tackle this in

modern instruments: measurements of inter-arrival times of particles (shattering large crystals create a burst of small crystals), construction of optimized probe tips to prevent shattered particles from entering the sampling volume (Fig. 18.2c), or the design of open path instruments. Shcherbakov et al. (2004) and Gayet et al. (2006) have analyzed possible errors in determining cloud properties due to possible measurement artefacts and have argued that in cases where large ice particles are not present, shattering should not be a critical effect.

18.5 Conclusions and Outlook

The investigation of climate and other effects of atmospheric aerosol requires measuring microphysical, optical and chemical properties of aerosol particles in high quality. In situ measurements by aircraft are needed in particular for the quantification of key properties of the many different atmospheric aerosol types encountered away from Earth's surface in the troposphere and *lower stratosphere*. Such measurements also help to improve the process understanding of the sources, sinks and transformations of atmospheric aerosols, a prerequisite to including such processes correctly in *model simulations*. Moreover, in situ measurements are also essential for testing and validating the *retrieval algorithms* developed for remote sensing observations from satellites. At the Institute of Atmospheric Physics, considerable experience in the measurement of microphysical aerosol properties on research aircraft has been built up. In general, meaningful results can only be obtained if different measurement techniques are combined. Many of the instrument types discussed in this chapter are now being prepared for deployment on the new research aircraft HALO.

References

Baumgardner, D., Dye, J.E., Gandrup, B.W., Knollenberg, R.G.: Interpretation of measurements made by the forward scattering spectrometer probe (FSSP-300) during the airborne arctic stratosphere expedition. J. Geophys. Res. **97**, 8035–8046 (1992)

Brock, C.A., Schröder, F., Kärcher, B., Petzold, A., Busen, R., Fiebig, M.: Ultrafine particle size distributions measured in aircraft exhaust plumes. J. Geophys. Res. **105**, 26555–26568 (2000)

Cotton, R., Osborne, S., Ulanowski, Z., Hirst, E., Kaye, P.H., Greenaway, R.S.: The ability of the small ice detector (SID-2) to characterize cloud particle and aerosol morphologies obtained during flights of the FAAM BAe-146 research aircraft. J. Atmos. Oceanic Technol. **27**, 290–303 (2010). doi:10.1175/2009JTECHA1282.1

Febvre, G., Gayet, J.-F., Minikin, A., Schlager, H., Shcherbakov, V., Jourdan, O., Busen, R., Fiebig, M., Kärcher, B., Schumann, U.: On optical and microphysical characteristics of contrails and cirrus. J. Geophys. Res. **114**, D02204 (2009). doi:10.1029/2008JD010184

Feldpausch, P., Fiebig, M., Fritzsche, L., Petzold, A.: Measurement of ultrafine aerosol size distributions by a combination of diffusion screen separators and condensation particle counters. J. Aerosol Sci. **37**(5), 577–597 (2006). doi:10.1016/j.jaerosci.2005.04.009

Fiebig, M., Petzold, A., Wandinger, U., Wendisch, M., Kiemle, C., Stifter, A., Ebert, M., Rother, T., Leiterer, U.: Optical closure for an aerosol column: method, accuracy and inferable properties

applied to a biomass-burning aerosol and its radiative forcing. J. Geophys. Res. **107**, D21 (2002). doi:10.1029/2000JD000192

Fierz, M., Vernooij, M.G.C., Burtscher, H.: An improved low-flow thermodenuder. J. Aerosol. Sci. **38**, 1163–1168 (2007)

Gayet, J.-F., Auriol, F., Minikin, A., Ström, J., Seifert, M., Krejci, R., Petzold, A., Febvre, G., Schumann, U.: Quantitative measurement of the microphysical and optical properties of cirrus clouds with four different in situ probes: evidence of small ice crystals. Geophys. Res. Lett. **29**(24), 2230 (2002). doi:10.1029/2001GL014342

Gayet, J.-F., Ovarlez, J., Shcherbakov, V., Ström, J., Schumann, U., Minikin, A., Auriol, F., Petzold, A., Monier, M.: Cirrus cloud microphysical and optical properties at southern and northern midlatitudes during the INCA experiment. J. Geophys. Res. **109**, D20206 (2004). doi:10.1029/2004JD004803

Gayet, J.-F., Shcherbakov, V., Mannstein, H., Minikin, A., Schumann, U., Ström, J., Petzold, A., Ovarlez, J., Immler, F.: Microphysical and optical properties of mid-latitude cirrus clouds observed in the southern hemisphere during INCA. Quart. J. R. Meteorol. Soc. **132**, 2719–2748 (2006)

Gayet, J.-F., Mioche, G., Shcherbakov, V., Gourbeyre, C., Busen, R., Minikin, A.: Optical properties of pristine ice crystals in mid-latitude cirrus clouds: a case study during CIRCLE-2 experiment. Atmos. Chem. Phys. **11**, 2537–2544 (2011a). doi:10.5194/acp-11-2537-2011

Gayet, J.-F., Shcherbakov, V., Voigt, C., Schumann, U., Schäuble, D., Jessberger, P., Petzold, A., Minikin, A., Schlager, H., Dubovik, O., et al.: The evolution of microphysical and optical properties of an A380 contrail in the vortex phase. Atmos. Chem. Phys. Discuss. **11**, 26867–26895 (2011b)

Hamburger, T., McMeeking, G., Minikin, A., Birmili, W., Dall'Osto, M., O'Dowd, C., Flentje, H., Henzing, B., Junninen, H., Kristensson, A., et al.: Overview of the synoptic and pollution situation over Europe during the EUCAARI-LONGREX field campaign. Atmos. Chem. Phys. **11**, 1065–1082 (2011). doi:10.5194/acp-11-1065-2011

Hinds, W.C.: Aerosol Technology: Properties Behaviour and Measurement of Airborne Particles. John Wiley & Sons, Inc., Hoboken, USA (1999), 483

Kandler, K., Schütz, L., Deutscher, C., Hofmann, H., Jäckel, S., Knippert, P., Lieke, K., Massling, A., Schladitz, A., Weinzierl, B., et al.: Size distribution, mass concentration, chemical and mineralogical composition, and derived optical parameters of the boundary layer aerosol at Tinfou, Morocco, during SAMUM 2006. Tellus B **61**, 32–50 (2009). doi: 10.1111/j.1600-0889.2008.00385.x

Khosrawi, F., Ström, J., Minikin, A., Krejci, R.: Particle formation in the Arctic free troposphere during the ASTAR 2004 campaign: a case study on the influence of vertical motion on the binary homogeneous nucleation of H_2SO_4/H_2O. Atmos. Chem. Phys. **10**, 1105–1120 (2010)

Kulkarni, P., Baron, P.A., Willeke, K., (eds.): Aerosol Measurement: Principles, Techniques, and Applications. Wiley, Hoboken, USA (2011)

Kulmala, M., Asmi, A., Lappalainen, H.K., Baltensperger, U., Brenguier, J.-L., Facchini, M.C., Hansson, H.-C., Hov, Ø., O'Dowd, C.D., Pöschl, U., et al.: General overview: European integrated project on aerosol cloud climate and air quality interactions (EUCAARI)—integrating aerosol research from nano to global scales. Atmos. Chem. Phys. **11**, 13061–13143 (2011). doi:10.5194/acp-11-13061-2011

Minikin, A., Petzold, A., Ström, J., Krejci, R., Seifert, M., Schlager, H., van Velthoven, P., Schumann, U.: Aircraft observations of the upper tropospheric fine particle aerosol in the northern and southern hemispheres at midlatitudes. Geophys. Res. Lett. **30** (10), 1503 (2003). doi:10.1029/2002GL016458

Mioche, G., Josset, D., Gayet, J.-F., Pelon, J., Garnier, A., Minikin, A., Schwarzenboeck, A.: Validation of the CALIPSO/CALIOP extinction coefficients from in situ observations in mid-latitude cirrus clouds during CIRCLE-2 experiment. J. Geophys. Res. **115**, (1–7), D00H25, (2010). doi:10.1029/2009JD012376

Mirme, S., Mirme, A., Minikin, A., Petzold, A., Hörrak, U., Kerminen, V.-M., Kulmala, M.: Atmospheric sub-3 nm particles at high altitudes. Atmos. Chem. Phys. **10**, 437–451 (2010). doi:10.5194/acp-10-437-2010

Petzold, A., Fiebig, M., Flentje, H., Keil, A., Leiterer, U., Schröder, F., Stifter, A., Wendisch, M., Wendling, P.: Vertical variability of aerosol properties observed at a continental site during the lindenberg aerosol characterization experiment (LACE 98). J. Geophys. Res. **107**, (2002). doi:10.1029/2001JD001043

Petzold, A., Fiebig, M., Fritzsche, L., Stein, C., Schumann, U., Wilson, C.W., Hurley, C.D., Arnold, F., Katragkou, E., Baltensperger, U., et al.: Particle emissions from aircraft engines—a survey of the European project PartEmis. Meteorol. Z. **14**, 465–476 (2005)

Petzold, A., Hasselbach, J., Lauer, P., Baumann, R., Franke, K., Gurk, C., Schlager, H., Weingartner, E.: Experimental studies on particle emissions from cruising ship, their characteristic properties, transformation and atmospheric lifetime in the marine boundary layer. Atmos. Chem. Phys. **8**, 2387–2403 (2008). doi:10.5194/acp-8-2387-2008

Petzold, A., Rasp, K., Weinzierl, B., Esselborn, M., Hamburger, T., Dörnbrack, A., Kandler, K., Schütz, L., Knippertz, P., Fiebig, M., et al.: Saharan dust absorption and refractive index from aircraft-based observations during SAMUM 2006. Tellus B **61**, 118–130 (2009). doi:10.1111/j.1600-0889.2008.00383.x

Petzold, A., Veira, A., Mund, S., Esselborn, M., Kiemle, C., Weinzierl, B., Hamburger, T., Ehret, G., Lieke, K., Kandler, K.: Mixing of mineral dust with urban pollution aerosol over Dakar (Senegal): impact on dust physico-chemical and radiative properties. Tellus B **63**, 619–634 (2011)

Reeves, C.E., Formenti, P., Afif, C., Ancellet, G., Attié, J.-L., Bechara, J., Borbon, A., Cairo, F., Coe, H., Crumeyrolle, S., et al.: Chemical and aerosol characterisation of the troposphere over West Africa during the monsoon period as part of AMMA. Atmos. Chem. Phys. **10**, 7575–7601 (2010). doi:10.5194/acp-10-7575-2010

Schmale, J., Schneider, J., Jurkat, T., Voigt, C., Kalesse, H., Rautenhaus, V., Lichtenstern, M., Schlager, H., Ancellet, G., Arnold, V., et al.: Aerosol layers from the 2008 eruptions of Mount Okmok and Mount Kasatochi: in situ upper troposphere and lower stratosphere measurements of sulfate and organics over Europe. J. Geophys. Res. **115** (D00L07), 1–18 (2010). AGU, doi:10.1029/2009JD013628

Schneider, J., Hings, S.S., Hock, B.N., Weimer, S., Borrmann, S., Fiebig, M., Petzold, A., Busen, R., Kärcher, B.: Aircraft-based operation of an aerosol mass spectrometer: Measurements of tropospheric aerosol composition. J. Aerosol Sci. **37**, 839–857 (2006). doi:10.1016/j.jaerosci.2005.07.002

Schröder, F., Brock, C.A., Baumann, R., Petzold, A., Busen, R., Schulte, P., Fiebig, M.: In situ studies on volatile jet exhaust particle emissions: impacts of fuel sulfur content and environmental conditions on nuclei mode aerosols. J. Geophys. Res., **105** (D15), 19941–19954 (2000)

Schröder, F., Kärcher, B., Petzold, A., Fiebig, M.: Aerosol states in the free troposphere at northern midlatitudes. J. Geophys. Res., **107**, LAC 8-1–LAC 8-8 (2002). doi:10.1029/2000JD000194

Schumann, U., Arnold, F., Busen, R., Curtius, J., Kärcher, B., Kiendler, A., Petzold, A., Schlager, H., Schröder, F., Wohlfrom, K.H.: Influence of fuel sulfur on the composition of aircraft exhaust plumes: the experiments SULFUR 1-7. J. Geophys. Res. **107**, 4247 (2002). doi:10.1029/2001JD000813

Schumann, U., Weinzierl, B., Reitebuch, O., Schlager, H., Minikin, A., Forster, C., Baumann, R., Sailer, T., Graf, K., Mannstein, H., et al.: Airborne observations of the Eyjafjalla volcano ash cloud over Europe during air space closure in April and May 2010. Atmos. Chem. Phys. **11**, 2245–2279 (2011). doi:10.5194/acp-11-2245-2011

Schwarz, J.P., Gao, R.S., Fahey, D.W., Thomson, D.S., Watts, L.A., Wilson, J.C., Reeves, J.M., Darbeheshti, M., Baumgardner, D.G., Kok, G.L., et al.: Single-particle measurements of

midlatitude black carbon and light-scattering aerosols from the boundary layer to the lower stratosphere. J. Geophys. Res. **110**, D16207 (2006)

Shcherbakov, V., Gayet, J.-F., Jourdan, O., Minikin, A., Ström, J., Petzold, A.: Assessment of cirrus cloud optical and microphysical data reliability by applying statistical procedures. J. Atmos. Oceanic Technol. **22**, 409–420 (2004)

Sipilä, M., Berndt, T., Petäjä, T., Brus, D., Vanhanen, J., Stratmann, F., Patokoski, J., Mauldin III, R.L., Hyvärinen, A.-P., Lihavainen, H., et al.: The role of sulfuric acid in atmospheric nucleation. Science **327**(5970), 1243–1246 (2010). doi:10.1126/science.1180315

Stein, C., Schröder, F., Petzold, A.: The condensation particle size analyzer: a new instrument for the measurement of ultrafine aerosol size distributions. J. Aerosol Sci. 381–382 (2001)

Virkkula, A.: Correction of the calibration of the 3-wavelength particle soot absorption photometer (3λ-PSAP). Aerosol. Sci. Technol. **44**, 706–712 (2010)

Voigt, C., Schumann, U., Jessberger, P., Jurkat, T., Petzold, A., Gayet, J.-F., Krämer, M., Thornberry, T., Fahey, D.: Extinction and optical depth of contrails. Geophys. Res. Lett. **38**, L11806 (2011). doi:10.1029/2011GL047189

Weigel, R., Borrmann, S., Kazil, J., Minikin, A., Stohl, A., Wilson, J.C., Reeves, J.M., Kunkel, D., de Reus, M., Frey, W., et al.: In situ observations of new particle formation in the tropical upper troposphere: the role of clouds and the nucleation mechanism. Atmos. Chem. Phys. **11**, 9983–10010 (2011). doi:10.5194/acp-11-9983-2011

Weinzierl, B., Petzold, A., Esselborn, M., Wirth, M., Rasp, K., Kandler, K., Schütz, L., Koepke, P., Fiebig, M.: Airborne measurements of dust layer properties, particle size distribution and mixing state of Saharan dust during SAMUM 2006. Tellus B **61**, 96–117 (2009). doi: 10.1111/j.1600-0889.2008.00392.x

Wendisch, M., Coe, H., Baumgardner, D., Brenguier, J.L., Dreiling, V., Fiebig, M., Formenti, P., Hermann, M., Krämer, M., Levin, Z., et al.: Aircraft particle inlets: state-of-the-art and future needs. Bull. Am. Meteor. Soc. **85**, 92 (2004)

Wendisch, M., Brenguier, J.-L. (eds.): Airborne Measurements: Methods and Instruments, to be published in 2012.

Dropsondes and Radiosondes for Atmospheric Measurements

19

Reinhold Busen

Abstract

Knowing the actual state of the atmosphere is the basis for any kind of weather forecast. The overall atmospheric condition is analyzed periodically using numerous ground station observations and radiosondes rising on a balloon and gathering the vertical profiles of pressure, temperature, humidity, wind speed and wind direction. These data are transmitted directly to a ground station. Dropsondes perform in a very similar way. They are dropped from an aircraft or a balloon mainly for research purposes over any location, in narrow temporal or spatial intervals and in inaccessible areas. The data gathered by dropsondes are transmitted to and recorded inside the aircraft.

19.1 Introduction

Numerical weather prediction models result in prognoses for different time frames extending from several hours to more than a week. They all need to be provided with the actual state of the atmosphere for successful model initiation. Several data sources can be utilized for this purpose.

Ground station data are collected by observers or automatic stations in intervals of 3 or 6 h and are transmitted to a national data center, for example the German Weather Service (DWD). About 230 ground stations are spaced regularly over

R. Busen (✉)
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Reinhold.Busen@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_19,

Germany; more than 10 000 land-based stations are registered worldwide. Beyond standard parameters (*temperature*, *pressure*, *humidity*, *wind speed* and *direction*), these ground stations provide additional data such as visibility, cloud appearance and cloud base, radiation, precipitation and sunshine duration.

Further ground based data are collected from about 100 moored and 600 drifting buoys as well as from 7 300 ships. There are no dedicated weather ships on the oceans any longer, but weather observations from ships continue from a fleet of voluntary merchant vessels in routine commercial operation.

Commercial *aircraft* collect weather data within the AMDAR network (Aircraft Meteorological Data Relay). More than 30 000 reports per day are transmitted from the climb and descent phases as well as from the flight altitude level by either radio or *satellite* communication. These reports contain flight number, date and time, position and flight altitude and meteorological information on temperature as well as on wind speed and direction.

Radar networks observe cloud cover, precipitation and *lightning* activity over the continents; these data can be used to exactly localize frontal systems, *thunderstorms* or other meteorological phenomena connected with cloud formation.

Satellites provide a view of the atmosphere from above. They can be positioned at a *geostationary* point over the equator at an altitude of about 36 000 km. This type of satellite can provide frequent images of the same area, but the resolution is limited and higher geographical latitudes are not well resolved. Satellites orbiting the earth on a polar track take about 100 min for each circuit, fly distinctly lower (about 800 km) with a better resolution at the ground, but cover only a smaller section of the earth surface and are therefore not suitable for a regular view of a specific area. For both categories, special satellites carry instruments for meteorological needs, gathering information on clouds, specific *trace gases*, *aerosols* and more, with coarse information on the vertical distributions of these species.

For measuring the vertical fine structure of the atmosphere, sounding devices like radiosondes (also called upsondes) and *dropsondes* are used. About 650 radiosonde stations launch sondes sometimes several times a day, whereas dropsondes basically serve as a research tool for specific atmospheric investigations. Both sonde types are discussed in detail in the next sections.

19.2 Radiosondes

The earliest document on engagement in *radiosonde* research at the Institute of Atmospheric Physics is a proposal on the construction of a new type of radiosonde written by the later director Prof. Hans Gerhard Müller in 1948. After critically reviewing the latest technical sensor developments he suggests a radiosonde containing sensors for pressure, temperature and humidity and as a really new feature a turning device switching through the sensor outputs and leading them subsequently to the radio transmitter. In contrast to earlier sondes this would allow a continuous data stream and therefore an uninterrupted *vertical profile* of

meteorological parameters. This technique, quite new at that time, has become a standard feature in all modern sondes.

It took years to realize this vision. In 1962, Prof. Müller became the first director of the newly founded "Institut für Physik der Atmosphäre," which was formed by combining former institutions. The first radiosondes specifically for radiation research were built and thermistors were tested and modified especially to avoid solar radiation influence and resulting misreadings at high altitudes. In 1967, sensors for temperature and pressure were installed on a small remotely controlled aircraft and the first radiosondes continuously transmitting data were built.

In 1969, the today still existing balloon hall located beside the institute building was completed and used for preparing the further advanced sondes and filling the balloons. In 1971 for the first time, dropsondes were built and used for investigations on power plant cooling towers.

During those years, several developments of the institute's co-workers went into operational use at the national weather service. This is especially true for ground receiving stations and data acquisition units, which were always constructed using the latest technical and electronic knowledge.

In 1975 and 1976, work on radiosondes slowly came to an end. At that time a fleet of three motor gliders was set up and one of the last activities was dropping sondes from those gliders, which provided on-board data acquisition, a procedure still in use today with up-to-date dropsonde systems. In 1976, the research aircraft Falcon 20 went into service and the institute concentrated on instrumenting and operating this new research tool.

Nowadays, the start of radiosondes is internationally synchronized four times a day at 00, 06, 12 and 18 UTC. There are 12 stations located in Germany and about 700 operating in the northern hemisphere (Figs. 19.1 and 19.2). All stations maintain at least the midnight and noon starts, and some of the stations work fully automatically. Special data formats (TEMP message, WMO code FM35) are used to distribute the large amount of radiosonde data to users. This format contains only numerical data on meteorological standard pressure levels and additional special noteworthy points on the profile. Transmitting the *sounding* this way, only little data can be sent without losing much information, typically 10 kB when transmitting per electronic mail.

19.3 Dropsondes

The development of a new dropsonde system started in 1994 as a joint effort between NCAR (National Center of Atmospheric Research, USA), NOAA (National Oceanic and Atmospheric Administration, USA) and the DLR Institute of Atmospheric Physics. All three research organizations are operating or having access to research aircraft capable of dropping sondes. The system is described in (Hock and Franklin 1999). This development was based on earlier dropsonde types, but was the first to use GPS for wind finding. Earlier sondes had to be

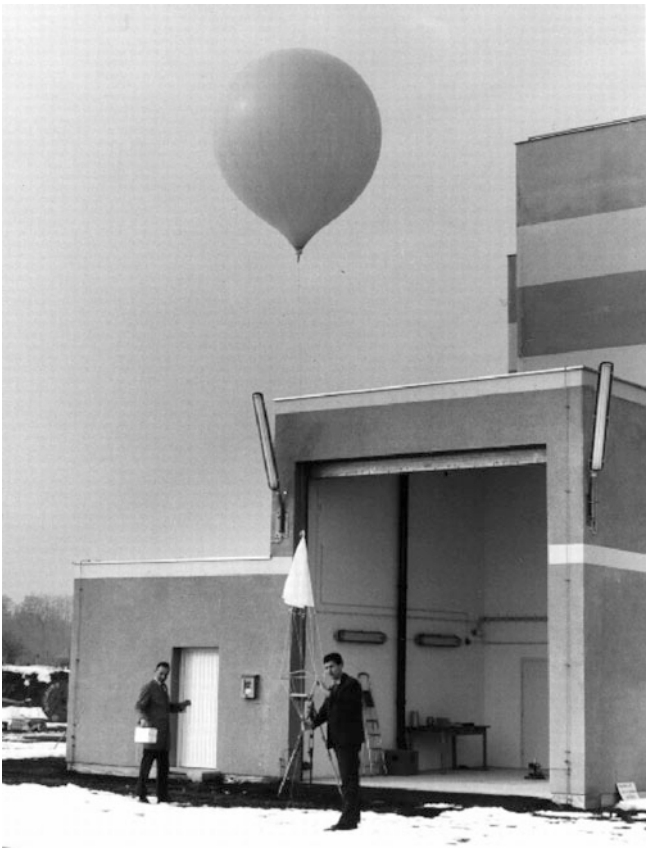

Fig. 19.1 Preparation of a radiosonde launch in front of the balloon hall of the institute (picture taken at IPA about 1974)

tracked by radar or used Omega or Loran-C systems. The latter was the first type with digital electronics and got into the weight range of about 400 g for the first time, whereas previous sondes used to be much heavier.

Launching a Radiosonde

A rubber balloon made from special latex material with a homogeneous wall thickness filled with helium gas carries a radiosonde up to an altitude of typically 30 km. The main advantage of probing the atmosphere in this way is the high vertical resolution and accuracy of the data. After launch the radiosonde system rises at a speed of about 300 m/min transmitting its data once a second, thus providing a vertical resolution of about 5 m. As it reaches higher altitudes, air pressure decreases and the balloon accordingly increase its volume. The higher vertical speed induced by the lower pressure

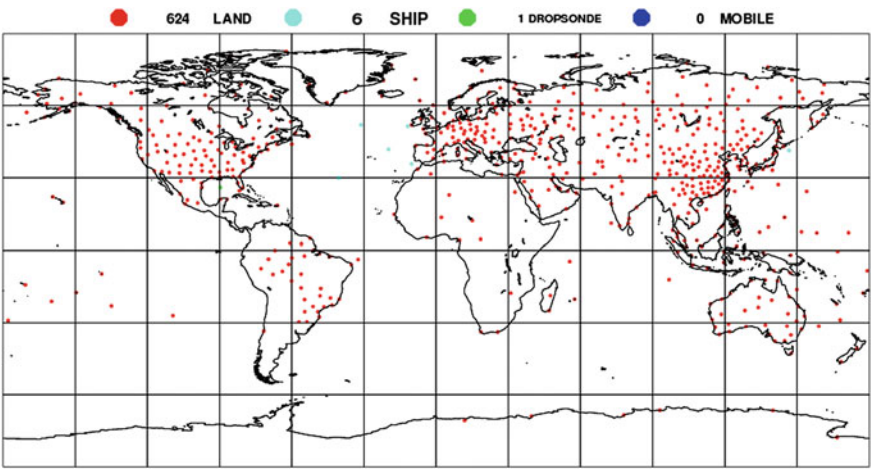

Fig. 19.2 Total number of worldwide upper air observations assimilated and used for numerical weather prediction by ECMWF at 00 UTC 14 March 2012. There are 624 soundings over land, 6 radiosondes are launched from ships, and 1 dropsonde in the Gulf of Mexico. The data network is sparse over Africa and in the southern hemisphere generally. The European Centre for Medium-Range Weather Forecasts (ECMWF) is gratefully acknowledged for providing this figure

is just about compensated by the higher drag of the expanding balloon, which finally results in an approximately constant ascent speed.

The system basically consists of a transmitting section (radio-) and a part containing the sensors (sonde). The radiosonde continuously measures air pressure, air temperature and humidity. The wind vector (speed and direction) can be derived from subsequent positions of the radiosonde as it is horizontally displaced by the prevailing wind. In earlier times a lightweight cardboard structure with a reflecting surface was used to track the system with a ground-based radar and to obtain the sonde positions for the wind calculation. In the meantime radiosondes are equipped with small GPS receivers and broadcast their position together with the meteorological data.

After the radiosonde system has reached an altitude of about 30 km after 100 min flight time, the balloon bursts and the sonde sinks to the surface on a parachute.

The requirements of NCAR; NOAA and DLR, based on existing aircraft or those to be expected within a short time, led to the following design specifications for the new GPS dropsonde and the aircraft data system:

1. Global operation, at altitudes up to 24 km
2. Deployment at indicated air speeds up to 130 m s^{-1}
3. Simultaneous operation of up to 4 sondes per aircraft

4. RF transmission bandwith of < 20 kHz
5. Telemetry range of 325 km
6. Sonde descent time of about 12 min when released from 12 km
7. Sensor measurement rate of 2 Hz, and
8. Sonde shelf life of at least 3 years.

Launching a dropsonde from an aircraft

The aircraft data system of the AVAPS (Airborne Vertical Atmospheric Profiling System) consists of a telemetry unit receiving the signals from the sondes, a system computer visualizing and storing the data, and a computer screen and keyboard. AVAPS is connected to two external antennas, one at the bottom of the aircraft to receive the 400 MHz sonde signal and the other being a GPS antenna placed on top of the aircraft. The on-board operator prepares a sonde before launch, which means transfer of calibration coefficients, frequency setting, and providing the sonde with the current configuration of GPS satellites. The sonde is launched via a launch tube, and the data broadcasted by the sonde are received and processed inside the aircraft. As many as four sondes can be in the air simultaneously with permanent display of the data on the computer screen. Finally, there is an option to analyze any vertical profiles including quality control and to produce a TEMP message as used for distributing radiosonde data or to display the data in a thermodynamic diagram. Depending on the communication capabilities of the aircraft, the TEMP message can be broadcasted to a ground station or immediately to a national weather center for further use in weather forecast models.

The sonde carries three meteorological sensors, which are contained in a PTH (Pressure, Temperature, Humidity) module manufactured by Vaisala, Finland. This company is the supplier of most of the many thousand radiosondes released per year, all containing an identical PTH module. Further, the sonde consists of a GPS receiver module, a digital microprocessor unit controlling and digitizing all sensor outputs, and a 400 MHz transmitter for sending the data telemetrically to the aircraft. Near the ground the descent speed is about 12 m/s.

In the meantime further development steps led to the latest RD-94 dropsonde (Fig. 19.3) with advanced electronics and a better and smaller GPS receiver, and to a lightweight system with only one sonde at a time in the air which was built for small aircraft. NCAR took the next step in building miniature dropsondes containing the same technology as standard-size sondes. These minisondes are used on stratospheric balloons drifting over long distances and releasing sondes automatically. The latest application is an automated launching system on the unmanned Global Hawk aircraft, and there are plans to install a semiautomatic system in the NCAR Gulfstream V aircraft as well.

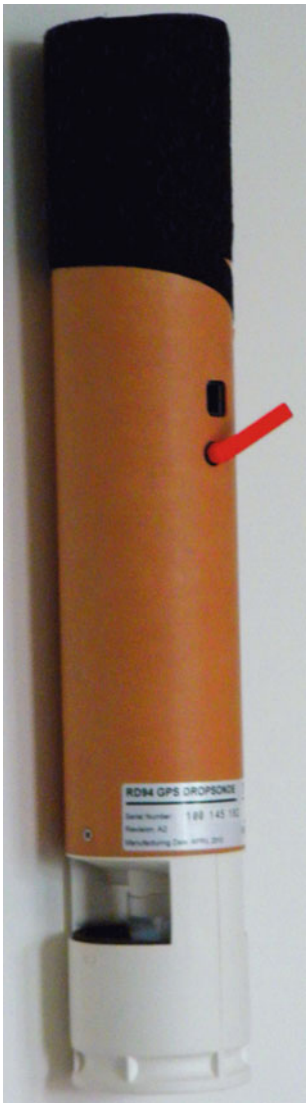

Fig. 19.3 RD94 dropsonde measuring temperature (±0.2 °C), *relative humidity* (±2 %), pressure (±0.4 hPa) and wind. Numbers in parentheses indicate sensor accuracy. Sonde weight 350 g, size is 7 cm in diameter and 41 cm in length. The parachute contained in the top section inflates when the black ribbon winds off after launch

At DLR a special device to measure *ozone* was designed and patented (Stäger et al. 1999). A small optical sensing device can be added to a radiosonde, and probably to a dropsonde as well, resulting in atmospheric ozone profiles. Regular soundings of the atmospheric ozone profile are performed by DWD at its Hohenpeissenberg observatory. A technique based on liquid chemistry reactions is used in combination with a standard radiosonde.

Table 19.1 List of experimental flight campaigns with deployment of dropsondes. Given are the names of the campaign in chronological order together with the experiment area, the number of dropsondes launched, and the main results obtained using the dropsonde data together with a corresponding reference

Campaign/year/location	No. of sondes	Scientific results and reference
AVAPS system test flights/1997/ Oberpfaffenhofen	7	Proof of functionality and data quality, AVAPS ready for scientific usage
EULINOX (European Lightning Nitrogen Oxides)/1998/Oberpfaffenhofen	7	Input data for the initialization of a mesoscale atmospheric model (MM5), (Fehr et al. 2004)
MAP (Mesoscale Alpine Program)/1999/ Swiss Alps	59	Analysis of dynamic processes over the Alps and verification of lidar measurements; (Hoinka et al. 2003)
IHOP (International H_2O Project)/2002/ Oklahoma, USA	91	Vertical profiles for evaluating lidar measurements of water vapor fluxes in the boundary layer; (Tollerud et al. 2008)
VERTIKATOR (Vertical Transport and Orography Project)/2002/ Oberpfaffenhofen	5	Comparison with wind lidar for analyzing local wind systems between the Alps and the Bavarian foreland; (Weissmann et al. 2005a)
A-TReC (Atlantic THORPEX Regional Campaign)/2003/Iceland	33	Comparison and quality control of wind lidar measurements to be assimilated in the ECMWF-model. (Weissmann et al. 2005b)
COPS (Convective and Orographically-induced Precipitation Study)/2007/Spain	11	Vertical profiles for analyzing air masses expected to initialize precipitation in the COPS intensive observation area. (Wulfmeyer et al. 2011)
SAMUM-2 (Saharan Mineral Dust Experiment)/2008/Cape Verde Islands	25	Atmospheric vertical profiles for a model calculating the solar radiative forcing of aerosols; (Bauer et al. 2011)
THORPEX-IPY (The Observing System Research and Predictability Experiment-International Polar Year)/2008/Norway	153	High density vertical profiles for investigating the development, intensity and track direction of polar lows; (Wagner et al. 2011)
T-PARC (THORPEX Pacific Asian Regional Campaign)/2008/Japan	325	High density vertical profiles for improving the track forecast of typhoons in the western North Pacific; (Weissmann et al. 2011); (Harnisch and Weissmann 2010)

Fig. 19.4 Dropsonde release from the Falcon aircraft for measurements during the T-PARC project over the western North Pacific in September 2008. Shortly after launch the parachute is just being inflated, as seen in the enlarged cut-out. Picture with courtesy of *The Yomiuri Shimbun*

19.4 Scientific Results from Dropsonde Deployments

On board the Falcon aircraft the AVAPS system was utilized in various campaigns listed in Table 19.1. A visualization of a dropsonde launch from T-PARC is shown in Fig. 19.4. Presently, efforts are under way to install the AVAPS system on the new Gulfstream 505 research aircraft HALO.

Figure 19.5 shows a typical example of a drop sounding with the view an observer on board the aircraft would obtain once the sounding is completed and the sonde has reached the ground. The temperature and relative humidity as well as the wind speed and direction are plotted from quality-controlled data, meaning that data disturbed by bad transmission conditions and periods when the sonde loses GPS contact are omitted or corrected. An example for lost data can be found in the wind speed profile, which shows gaps in the profile curve.

Another view of the same data can be found in the Skew T-log p diagram (Fig. 19.6) showing temperature and dew point profiles. The different appearance of the temperature and dew point slopes in this diagram is a result of the skewed temperature lines. This specific diagram has its advantage when graphically analyzing energetic interpretations.

The same data can be displayed as a TEMP message, as described in the section on radiosondes. These messages can be sent from the aircraft as emails to a ground station or to a national weather service using a satellite link or other communication links available on the aircraft.

19.5 Operational Constraints for Dropsondes and Radiosondes

Sondes ascending or descending through the atmosphere can collide with aircraft at a probability definitely higher than zero, and every object hitting the ground could cause injuries to people or buildings.

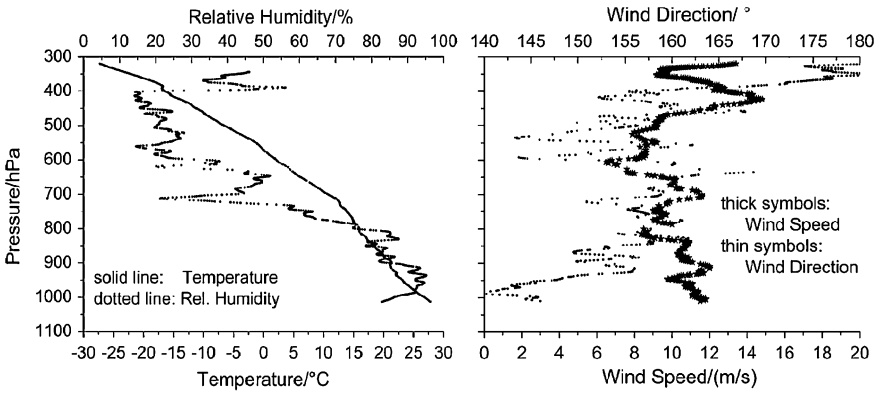

Fig. 19.5 Dropsonde data from 00 UTC 30 August 2008 during T-PARC. The height-scale on the y-axis shows atmospheric pressure from 1 100 hPa at the bottom to 300 hPa at the top. *Left panel*: temperature and relative humidity. *Right panel*: wind speed and direction

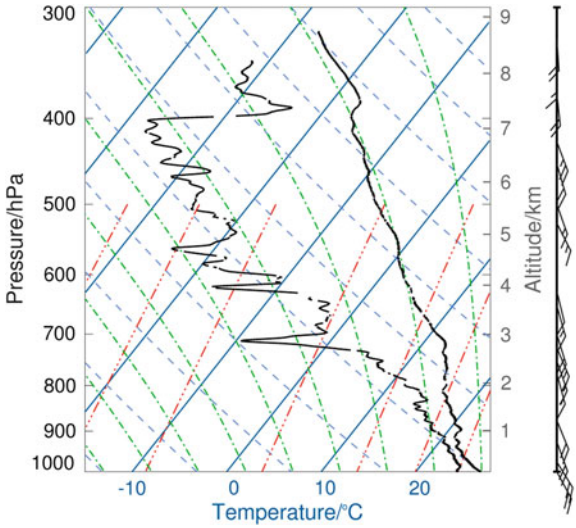

Fig. 19.6 Skew T-log p diagram showing temperature (*right curve*) and dew point temperature (*left curve*) for the same dropsonde. The wind barbs on the right indicate wind speed and direction. The *dew point* is defined as the temperature at which the prevailing humidity would form dew or ice on a cooled mirror surface

To estimate these risks in a statistical manner the *air traffic* over Germany was analyzed for the probability of hitting a lower flying aircraft with a randomly dropped sonde (Busen 2000).

Any aircraft in the air can be hit by a sonde descending on a parachute by means of two different mechanisms:

1. The sonde may hit the upper surface of the aircraft. However, as the aircraft is flying very fast compared to the falling speed of the sonde, the probability for this mechanism is low and any damage is expected to be minor.

2. Whenever the sonde falls through the air space in front of an aircraft, the aircraft itself may hit the sonde. In this case a *collision* of both may be more probable and much more dangerous, considering the relationship of the aircraft's horizontal speed (about 200 ms^{-1}) to the sonde's vertical speed (about 20 ms^{-1} at the aircraft cruising altitude). To calculate the probability of a collision the vertical extensions of both the aircraft and the sonde have to be considered. Depending on the vertical extension of the different aircraft sections, different collision areas can be calculated, adding up to a total collision area of the aircraft. This area is different for various aircraft models and their individual cruising speeds.

Calculating the total collision area of all aircraft over Germany for a certain moment and relating it to the overall area of Germany results in a probability of about 2.6×10^{-6} for hitting another aircraft with a randomly dropped sonde. In other words, one of about 386 000 randomly dropped sondes is statistically to be expected to collide with an aircraft. For this estimation the daily peak air traffic density over Germany was taken into account.

In reality, ATC (Air Traffic Control) supervises any dropsonde release. While so far sondes have been treated as randomly dropped, in reality every flight mission with dropsonde release has to be announced to ATC well in advance. During the mission, for every single drop the airspace below is checked by the ground controller to avoid any danger to other aircraft. Final clearance for dropping the sonde is then given. Following this procedure a collision is very unlikely in a controlled airspace.

Radiosondes, or better upsondes, are released from a ground station and ascend on a balloon with typical vertical speeds of about 5 ms^{-1}. The upsonde body size and weight is quite comparable to that of a dropsonde. Following the same argumentation and mathematical calculation, the probability to collide with an aircraft should be even higher, as the lower vertical speed of the sonde should increase the collision area for each aircraft type by at least a factor of three.

Furthermore, every upsonde passes through the airspace twice, ascending on the balloon and descending on a *parachute* after the burst of the *balloon*. The time and the geographical coordinates of the launch are known and published. However, upsondes strongly drift horizontally with the wind. Therefore, they may pass the altitude band covered by air traffic some distance away from their launch position. When the balloon finally bursts about 90 min after launch, the area where the sonde descends is completely unpredictable and random.

The overall collision probability for a single upsonde is therefore estimated to be about 10^{-5}. With a minimum of about 30 upsondes released in Germany every day, this means that one collision could be expected over a time period of about 9 years.

The chance of injury on the ground cannot be estimated at all. A dropsonde with a mass of about 400 g reaching the ground at a speed of 10 ms^{-1}contains a kinetic energy of 20 J (kg m^2s^{-2}). In comparison, a tile weighing 1 kg and falling from a roof of 10 meters height would contain a kinetic energy of about 100 J,

about 5 times that of a dropsonde. The real injury of people or buildings depends on many circumstances, such as the angle of impact, the exact position of the object, where the sonde hits, and physiological aspects when hitting a person. To avoid any injuries as far as possible, it is quite difficult to get the permission for dropping over land, and in any case this needs to happen over remote areas with a very low population density. Most dropsondes are released over marine areas anyway, as those locations are not accessible to radiosondes with the exception of a very few cases when they are started from research vessels. For these applications the probability of causing any damage at the sea surface is close to zero. US. military and research aircraft drop hundreds of sondes every year during the hurricane season to significantly improve the forecasting of hurricane tracks and the prospective position of landfall. So far, no damage has been reported from this application despite the large number of sondes dropped.

19.6 Summary and Outlook

In conclusion, both radiosondes and dropsondes will certainly stay important, the former for routine use and for *monitoring* the vertical structure of the atmosphere on a regular basis at many sites around the world. Dropsondes, on the other hand, are a valuable research tool for specific scientific investigations. Their special advantage is that, apart from highly populated areas, they can be dropped exactly where interesting weather phenomena is to be explored, at least at locations and atmospheric conditions where an aircraft can fly.

Dropsondes are undergoing further development in two different ways: (1) Sondes are being miniaturized as electronic parts become smaller. These small sondes are used for any kind of automated systems where a large number of them needs to be stored for automatic deployment. Examples are driftsondes (Wang et al. 2010), dropsonde deployment systems on unmanned aircraft like the Global Hawk (Hock et al. 2012), or on other aircraft without a dedicated dropsonde operator on board. (2) There are attempts to add additional sensing elements to existing RD94-dropsondes using the empty space inside the sonde housing. Carrying additional sensors, e.g., for SO_2 or atmospheric aerosols, the sonde are expected to be valuable in emergency cases like *volcanic ash* clouds, nuclear disasters, chemical accidents or other occasions where aircraft are not able to fly directly into the affected area. Dropsondes would then be used to monitor the atmosphere, being dropped from an aircraft flying sufficiently high over the affected atmospheric layers.

References

Bauer, S., Bierwirth, E., Esselborn, M., Petzold, A., Macke, A.: Airborne spectral radiation measurements to derive solar radiative forcing of Saharan dust mixed with biomass burning smoke particles. Tellus B **638**, 742–750 (2011). doi:10.1111/j.1600-0889.2011.00567.x

Busen, R.: The release of dropsondes: a hazard for commercial air traffic? Air Traffic Control Quart. **8**(2), 155–171 (2000)

Fehr, T., Höller, H., Huntrieser, H.: Model study on production and transport of lightning-produced NOx in an EULINOX supercell storm. J. Geophys. Res. **109**, D09102 (2004). doi: 10.1029/2003JD003935

Harnisch, F., Weissmann, M.: Sensitivity of typhoon forecast to different subsets of targeted dropsonde observations. Mon. Weather Rev. **138**, 2664–2680 (2010). doi: 10.1175/2010MWR3309.1

Hock, T.F., Franklin, J.L.: The NCAR GPS Dropwindsonde. Bull. Am. Meteorol. Soc. **80**, 407–420 (1999)

Hock, T.F., Lauritsen, D., Martin, C., Potts, N., Bradley, C.: Overview of NCAR dropsonde technologies for various applications. 92nd American Meteorological Society Annual Meeting, 22–26 Jan 2012

Hoinka, K.P., Richard, E., Poberaj, G., Busen, R., Caccia, J.-L., Fix, A., Mannstein, H.: Analysis of a potential vorticity streamer crossing the Alps during MAP IOP-15 on 6 Nov. 1999. Q. J. Royal. Meteorol. Soc. **129**, 609–632 (2003)

Stäger, R., Delany, A., Uhlemann, G., Busen, R., Güsten, H., Nolle, M.: Vorrichtung zur Bestimmen der Ozonkonzentration in Luft. Patent Nr. *1,964,718,5* des Deutschen Patent- und Markenamtes, erteilt am, 18 Feb 1999

Tollerud, E., Caracena, F., Koch, S., Jamison, B., Hardesty, M., McCarty, B.J., Kiemle, C., Collander, R., Bartels, D., Albers, S., et al.: Mesoscale moisture transport by the low-level-jet during the IHOP field experiment. Mon. Weather Rev. **136**, 3781–3795 (2008). doi: 10.1175/2008MWR2421.1

Wagner, J.S., Gohm, A., Dörnbrack, A., Schäfler, A.: The mesoscale structure of a polar low: airborne lidar measurements and simulations. Q. J. Royal. Meteorol. Soc. **137**, 1516–1531 (2011)

Wang, J., Zhang, L., Lin, P.-H., Bradford, M., Cole, H., Fox, J., Hock, T., Lauritsen, D., Loehrer, S., Martin, C., et al.: Water vapor variability and comparisons in the subtropical pacific from the observing system research and predictability experiment-pacific asian regional campaign (T-PARC) Driftsonde, constellation observing system for meteorology, ionosphere, and climate (COSMIC), and reanalyses. J. Geophys. Res. **115**, D21108 (2010). doi: 10.1029/2010JD014494

Weissmann, M., Braun, F.J., Gantner, L., Mayr, G., Rahm, S., Reitebuch, O.: The Alpine mountain-plain circulation: airborne doppler lidar measurements and numerical simulations. Mon. Weather Rev. **133**, 3095–3109 (2005a). doi:10.1175/MWR3012.1

Weissmann, M., Busen, R., Dörnbrack, A., Rahm, S., Reitebuch, O.: Targetted observations with an airborne wind lidar. J. At. Ocean. Technol. **22**, 1706–1719 (2005b)

Weissmann, M., Harnisch, F., Wu, C., Lin, P., Otha, Y., Yamashita, K., Kim, Y., Jeon, E., Nakazawa, T., Aberson, S.: The influence of assimilating dropsonde data on typhoon tracks and mid-latitude forecasts. Mon. Weather Rev. **139**, 908–920 (2011). doi: 10.1175/2010MWR3377.1

Wulfmeyer, V., Behrendt, A., Kottmeier, C., Corsmeier, U., Barthlott, C., Craig, G., Hagen, M., Althausen, D., Aoshima, F., Arpagaus, M., et al.: The Convective and Orographically-induced Precipitation Study (COPS): the scientific strategy, the field phase, and research highlights. Q. J. Royal Meteor. Soc. **137**, 3–30 (2011). doi:10.1002/qj.752

Lightning Detection

20

Hartmut Höller, Hans-Dieter Betz, Ullrich Finke
and Kersten Schmidt

Abstract

Thunderstorm electrification is mainly due to ice particle interaction processes leading to a typical main dipole structure. Intra-cloud and cloud-to-ground discharges can be detected at low and very low radio frequencies by the lightning detection network LINET. At high radio-frequencies one can detect the step-like flash structure. The optical emissions can be observed from space. The operational LINET network over Europe and networks deployed during field experiments enabled a comparison of lightning characteristics from mid-latitudes and the tropics. Important applications are nowcasting and operational weather forecasting, climate research, and the planning and support of new satellite lightning observing systems.

H. Höller (✉) · K. Schmidt
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: hartmut.hoeller@dlr.de

K. Schmidt
e-mail: Kersten.Schmidt@dlr.de

H.-D. Betz
Faculty of Physics, nowcast GmbH und LMU, 81377 München, Germany
e-mail: hans-dieter.betz@physik.uni-muenchen.de

U. Finke
Hochschule Hannover, Fakultät I - Elektro- und Informationstechnik,
Ricklinger Stadtweg 120, 30459 Hannover, Germany
e-mail: ullrich.finke@fh-hannover.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_20, © Springer-Verlag Berlin Heidelberg 2012

20.1 Introduction

Thunderstorms are very complex clouds. Driven by buoyancy effects they develop strong vertical motions in both upward and downward directions. The *storms* mostly extend throughout the whole troposphere from the boundary layer to the upper troposphere and even penetrate into the stratosphere with their overshooting tops. Water vapor condenses to form cloud *droplets* or *ice crystals*. *Collisions* among these particles lead to larger *precipitation* particles like raindrops, snowflakes, *graupel*, or hailstones. Such cloud microphysical interaction processes cause particle charging. The subsequent size-sorting according to the different fall velocities of particles of opposite charge is the origin of large electric fields within the thundercloud. Once the electrical field is strong enough, a discharge develops either between the different charge centers within the cloud or between the cloud and the ground. For basic information on *electrification* and *lightning* the reader is referred to textbooks like Rakov and Uman (2003) and (Betz et al. 2009b).

When the hot, luminous flash channels cool down, *nitrogen oxides* are released. Due to their climatic impact as ozone precursors they have been in the focus of research at IPA for many years. Comparing ground based lightning observations with airborne chemical measurements in different climatic regions and using these experimental data in atmospheric *models* has allowed uncertainties in knowledge of the source strength of lightning-produced NO_x to be narrowed down (Schumann and Huntrieser 2007). Another focus of the application of lightning data at IPA is severe weather *nowcasting* and *forecasting* based on *radar* and *satellite* observations of thunderstorms. Including lightning data in nowcasting applications can significantly improve local short range forecasting of thunderstorms and is expected to lead to improvements in large scale and long term forecasts, especially with the lightning detection capabilities of future satellites like *Meteosat* Third Generation (MTG).

20.2 Thunderstorm Electrification and Lightning

Even though there are numerous possible electrification processes in clouds, it is now generally accepted that the main causes of thundercloud electrification are noninductive charging processes during collisions of different types of *ice particles* in the presence of cloud water. The ice particles involved are ice crystals and graupel (rimed ice crystals) which grow by *sublimation* at *ice supersaturation*, e.g., at near water saturation in a mixed-phase cloud containing cloud droplets in an updraft. As the different kinds of ice particles have different surface properties like surface roughness and impurities, their depositional growth rates are different. This causes the quasi-liquid layer at the surface to be of different thickness and the mobility of the OH^- and H^+ ions to be different. Thus during collisional contact a net charge transfer can occur between the particles due to exchange of ions in the quasi-liquid surface layers. The sign of the charge flow critically depends on the

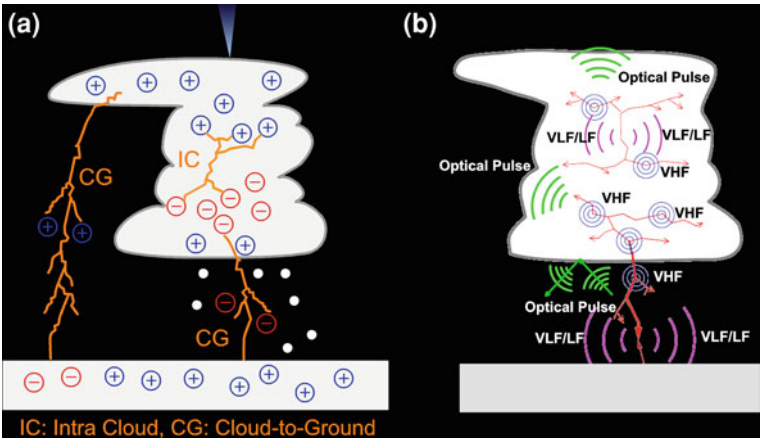

Fig. 20.1 Schematic of **a** the main thunderstorm charge regions and flash types, and **b** electromagnetic emissions of the different flash components. *VLF* very low frequency, *LF* low frequency, *VHF* very high frequency

ambient temperature and the liquid water content which determines the riming rate and, thus, the temperature (heat of *freezing*) of the graupel particles. Under 'normal' thunderstorm conditions the graupel have a negative charge and the smaller ice crystals a positive one. The differential fall velocities of graupel and ice crystals cause a net charge separation in a thunderstorm, leading to the well-known dipole structure with a negative charge center (graupel) at mid-levels and a positive charge center in the upper levels (ice crystals). This idealized main dipole structure is often accompanied by additional minor charge layers. Lower positive charge centers often form close to the cloud base either in the updraft region or in more extended stratiform regions at the melting level. On the other hand, a screening layer of negative charge can be found at cloud tops. Figure 20.1a schematically illustrates the main charge structure and the different *flash* types.

When the amount of charging is strong enough, a discharge occurs between two or more of the in-cloud charge centers (intra-cloud flash, IC), between the cloud and the ground (cloud-to-ground flash, CG), or between the cloud and the surrounding air. As discovered only recently, there are also numerous types of discharge processes (transient luminous events, TLE) between the cloud and the ionosphere. The discharge process is accompanied by a breakdown of the electric field. The initial phase of these processes is called the 'initial breakdown'. Recent measurements using different kinds of lightning detection systems indicate that a rather large portion of these events and associated strokes could be triggered by energetic particles originating from cosmic radiation, which can originate electron avalanches upon interaction with the *atmosphere*. This can lead to the formation of a relatively long ionized channel early in the discharge development, leading to the *emission* of long wave VLF (very low frequency; 3–30 kHz; *wavelength*s from 10 to 100 km) and LF (low frequency; 30–300 kHz; wavelengths from 1 to 10 km)

electromagnetic radiation originating from the lightning pulse. Observations at higher electromagnetic frequencies (VHF—very high frequency; 30–300 MHz; wavelength from 1 to 10 m) indicate that often the ionized lightning channels proceed from a source region in between two main charge centers in the cloud into both directions along the electric field. Negative breakdown into the positively charged region is manifested by a negative leader channel. The positive breakdown occurs by a positive leader propagating into the negatively charged region. This type of leader process is called a bipolar leader. The leader propagates in discrete steps of a few tens or hundreds of meters and tends to radiate in VHF. Once a leader contacts a region of enhanced cloud charge an impulsive current (IC stroke) can result in the already existing channel in the cloud and a VLF pulse is generated. This process is referred to as a recoil streamer or K-process (for 'Kleiner' change, or for Kitagawa and M. Kobayashi).

Flash initiation cannot only take place between the main upper positive and negative charge centers in the cloud, but also between the main negative and a lower positive charge center. This type of discharge can initiate a subsequent ground flash. The negative leader progresses further down to ground and finally connects to the ground, initiating a return stroke progressing in the opposite direction. Upon ground contact, the return stroke produces a strong VLF/LF pulse, which is generally the most prominent emitting process during a flash. Following the first return stroke the flash can develop subsequent discharge processes like in-cloud processes which extend the return stroke into the cloud (junction or J-processes), K-processes, which are pulses superimposed on the slower J-processes or initiate one or more subsequent strokes to ground starting with so-called dart-leaders in existing ionized channels. All these processes are luminous and give rise to *optical* light emission of different intensity and duration. Some ground flashes develop long lasting continuing currents which are well visible for some hundred milliseconds but do not produce *sferics* during that time. The whole flash might last for up to 2 s and typically extend horizontally for a few tens of kilometers. The main emission sources are illustrated schematically in Fig. 20.1b.

20.3 Principles of Lightning Detection

The different components of a flash emit *radiation* in a broad range of the electromagnetic spectrum. The most commonly used radio frequency ranges for ground-based lightning detection are the VLF/LF band and the VHF band. The waves propagate as ground waves along the Earth's surface or as sky waves which can be reflected at the ionosphere. The long waves can be detected at longer distances than the short waves as they are less attenuated. Also, the peak intensity is connected to frequencies around 10 kHz. Thus, networks of VLF sensors can be used even for global lightning detection with sensor baselines in the order of 1 000 km or more. The VLF/LF range is used by country- or continental-scale networks with sensor baselines of some tens to some hundreds of km. Either

electric dipole antennas or magnetic loop antennas are in use. The short VHF waves require a short baseline sensor network in the order of tens of km as they can be detected only *line-of-sight*. Their advantage is that they can detect the details of the flash branch structure, but they miss the higher energy currents produced by the return strokes, IC-strokes and K-processes. Other lightning detection systems use interferometer methods or optical lightning detection. The following discussion is restricted to research around the VLF/LF *lightning detection network* LINET used by the DLR's Institute of Atmospheric Physics (IPA) in connection with projects on lightning-NO_x research, severe weather nowcasting and spaceborne optical lightning detection, the latter becoming increasingly important with future satellite instruments.

The lightning detection network LINET was developed at the University of Munich (LMU) primarily for scientific use but also for commercial operation (Betz et al. 2004). Since 2005 an operational network has been set up in Germany (Betz et al. 2007) and subsequently extended over neighboring European countries. At present the European network operates slightly more than 120 stations (Betz et al. 2009a). IPA has operated a six-sensor network since 2005 in different *field campaigns* in Europe as well as in tropical regions (Huntrieser et al. 2008, 2009, 2011; Höller et al. 2009; Labrador et al. 2009). Each LINET sensor consists of two magnetic loop antennas oriented perpendicular to each other preferably in north-south (N–S) and east-west (E–W) directions. A schematic of the LINET measuring principle is shown in Fig. 20.2. For exact time stamping of the signals a GPS antenna is used at each station. Primarily, this time information is applied for flash geolocalization. It is generally assumed that the emitting flash channel is vertically orientated. Thus, the magnetic field varies horizontally coaxial from the line source. Therefore, a vertical orientation of the field antennas gives maximum response. With nonvertical alignment of the lightning channels the *detection efficiency* of the system decreases. The electromagnetic pulse (radio atmospheric signal or 'sferic', sometimes 'spheric') measured by each station is analyzed for intensity peaks and a time of arrival (TOA) is determined (see Fig. 20.2b). Electric current amplitude information is also used for inferring the *peak current* of the lightning stroke. Together with the calculated distance to the stroke location (D), the propagation law $\sim 1/D$ can be used to infer the original current with an accuracy of around 20 %. Moreover, the polarity of the current flow can be inferred. The time stamps of at least four different stations are used for inferring the horizontal and vertical source location of the stroke (Fig. 20.2b). It is a unique feature of LINET that a height can be attributed to each event and thus a discrimination of intra-cloud and cloud-to-ground strokes becomes possible. The source height H is determined by an *optimization* procedure assuming the increased signal propagation time dT from the source point P to the station S as compared to the horizontal projection of the source location to the ground (see Fig. 20.2c). Other prominent features of LINET are the low detection threshold for both IC and CG strokes and the high location accuracy. The most crucial advantage of LINET is its unusually high temporal resolution, a necessary

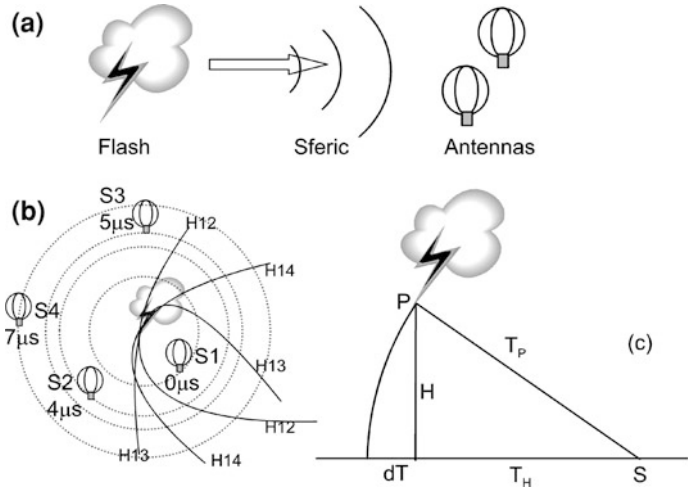

Fig. 20.2 Schematic of LINET lightning detection using magnetic loop antennas. **a** Schematic of stroke propagation, **b** Time of arrival (TOA) method using at least four different stations S1–S4 for flash geolocation on hyperbolas H12–H14 constructed with the closest station S1, and **c** height determination H using the time difference $dT = T_P - T_H$ for a stroke at location P. (**b** from Schmidt (2007), **c** from Betz et al. (2004))

condition for both location precision down to 100 m and the capability to derive IC parameters (Betz et al. 2007).

An example of the peak current distribution measured by LINET over Europe on 3 July 2008 is shown in Fig. 20.3. The maximum values for all types of strokes are found at low electric current amplitudes: at around 3 kA for the IC components and at about 4 to 5 kA for the CGs. These values are dependent on the typical baseline between the stations and would be even lower for smaller baselines. Note that the CG strokes dominate the distributions over a wide range. This is indicative of ground flashes lowering the negative charge from the main negative (graupel) charge center in the cloud down to the ground. For very large peak currents, positive CGs (larger than about 100 kA, not shown in Fig. 20.3) can play a prominent role. This is indicative of positive CGs from the storm *anvil* down to the ground. The IC components become dominating only at very low peak currents but they can outnumber the CGs. Thus, in total the IC stroke numbers can be higher than the number of CGs. At present, the formation of IC strokes is not well understood, but when the discharge occurs in a predominantly vertical channel, the polarity is defined just as for CG strokes; in horizontal channels the polarity of IC strokes does not seem to be meaningful.

A Meteosat snapshot at one instant and the LINET stroke history over central Europe on 3 July 2008 including this time are shown in Fig. 20.4a, b, respectively. Note that thunderstorms were active during night and morning hours over Southern France. The activity was connected to a frontal system moving over central Europe from west to east. Thunderstorm activity continued and intensified in the late noon

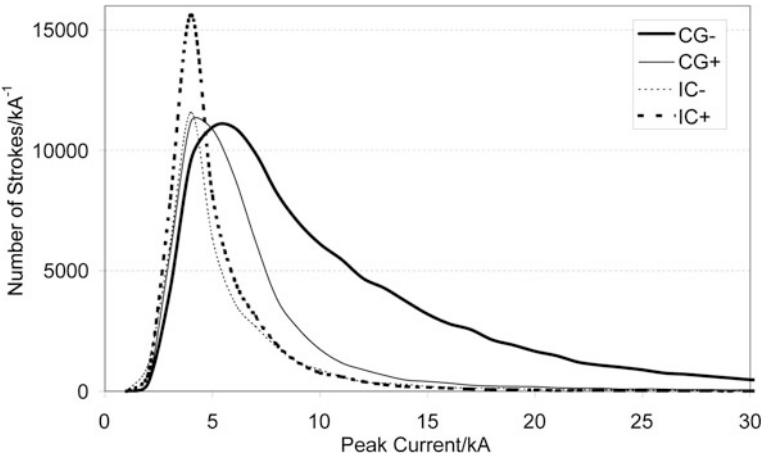

Fig. 20.3 Distribution of LINET peak currents of both polarities for intra-cloud (IC) and cloud-to-ground (CG) strokes over Europe on 3 July 2008

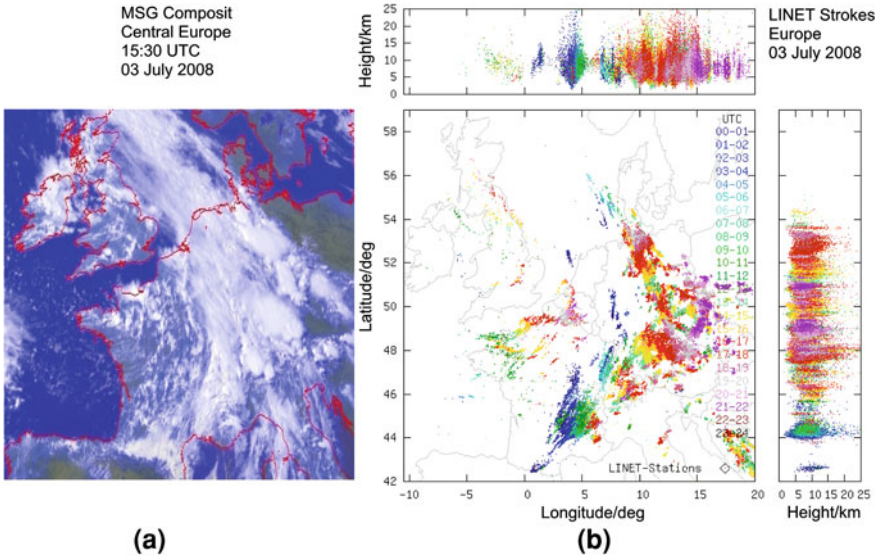

Fig. 20.4 **a** Meteosat image at 15:30 UTC and **b** history of LINET stroke evolution on 3 July 2008 over Central Europe. *Color coding* indicates Universal Time Coordinated (UTC) time/h

and early afternoon hours over mountainous terrain like the northern Alps. General storm propagation from southwest to northeast can be noted, as is typical for a mid-European thunderstorm situation under unstable conditions in a southwesterly flow ahead of an approaching trough. The IC stroke components are additionally shown in the vertical projections. In general, the vertical location accuracy is not as

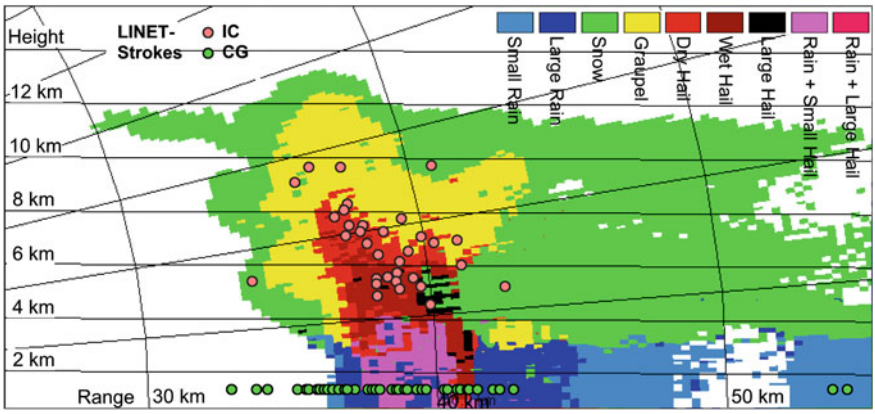

Fig. 20.5 POLDIRAD hydrometeor classification and LINET lightning stroke composite at 13:57 UTC 29 June 2005 and 289° azimuth. LINET strokes from a ± 2 min interval and a ± 2 km slice around the radar scan. *Red* IC—intra cloud; *Green* CG—cloud-to-ground stroke

good as the horizontal accuracy, which explains the scatter of the data at cloud top. Averaged peak IC height was at 9 km on this day.

20.4 Lightning Characteristics of Mid-Latitude and Tropical Thunderstorms

IPA has performed a series of lightning-related international field *experiments* in Europe as well as in tropical areas. In these campaigns the six-sensor DLR-LINET system was deployed and offered the unique possibility to obtain comparable data sets of lightning in different climatic regions. Moreover, with IPA's *C-band* weather radar POLDIRAD (Polarimetric Diversity Doppler Radar) at Oberpfaffenhofen it became possible to study lightning and the microphysical precipitation processes in concert.

An instructive example of a *mid-latitude* thunderstorm is shown in Fig. 20.5. Using the POLDIRAD differential reflectivity (ZDR) of horizontally and vertically transmitted and received radiation as well as the *linear depolarization* ratio (LDR) and the radar *reflectivity* Z one can determine the *hydrometeor* type. Large positive values of ZDR are indicative of horizontally aligned raindrops whereas tumbling or spherical graupel or *hail* exhibits small ZDR vales. Additionally, large LDR values are an indicator of wet hail or melting particles. The storm is approaching the radar from the west (right hand side in Fig. 20.5). New cell growth is to be seen on the eastern flank of the storm in the direction of propagation. The main updraft can be identified by the overshooting top at 13 km height. Older and decaying cells are to be noted on the down-shear side (west). Hail is reaching the ground at about 40 km distance from the radar. The IC strokes are to be found in the area of graupel and small, dry hail aloft from 5 to 10 km height. This is in accordance with

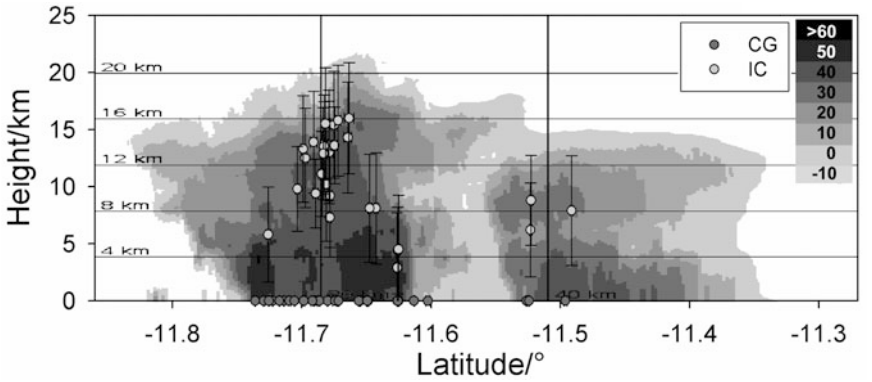

Fig. 20.6 Vertical section (131.1° longitude) of C-POL (C-band *polarimetric* radar) radar reflectivity through major Hector storm cell at 14:15 UTC 19 November 2005 and position of IC and CG strokes with error bars during a 4-min time interval around the radar scan (strokes in the vertical section from a 4-km wide slice). The C-POL radar was operated by the Bureau of Meteorology Research Center at Darwin, North Australia; from Höller et al. (2009)

the graupel charging mechanism as described above. The ground strokes are also roughly coincident with the high precipitation region. For better visibility at far ranges (Earth curvature) the CGs are plotted at 1.5 km height.

During the tropical campaigns, the DLR-LINET was deployed to Brazil in 2005, to northern Australia in 2005/2006 and to Benin/West Africa during summer 2006. One of the prominent differences between mid-latitude and tropical lightning is the height of IC occurrence. Due to the larger height of the tropical tropopause the storm tops often extend up to 18–21 km. As a prominent example Fig. 20.6 shows a vertical section of radar reflectivity and LINET lightning through '*Hector*,' a regularly occurring thunderstorm over the Tiwi Islands north of Darwin. Due to the special forcing conditions of interacting sea breeze *fronts* and old thunderstorm outflows, the Hector cells can grow rather vigorously. We note from Fig. 20.6 that graupel, indicated by radar reflectivity up to 30 dBZ, are lifted up to around 17 km. This implies that also IC strokes are detected up to that height. Intense storm phases are generally characterized by large IC stroke fractions. Due to the large vertical extension of the graupel cores the IC contribution to total lightning is generally also very high. A similar feature is often observed in mid-latitude storms during phases of storm intensification. It can be used as a severe weather warning indicator.

The basic difference in the vertical IC lightning profiles is illustrated in Fig. 20.7. Due to the higher vertical extent of *convection* connected to the higher tropopause height peak frequencies are at higher altitude in the tropics. The characteristics for Brazil and Benin are rather similar to each other with respect to the IC stroke profiles, resulting in mean heights close to 11 km. The Australian strokes peak a little higher up at 12.2 km altitude. The mid-latitude region in southern Germany clearly shows the lowest IC maximum height at 8.5 km.

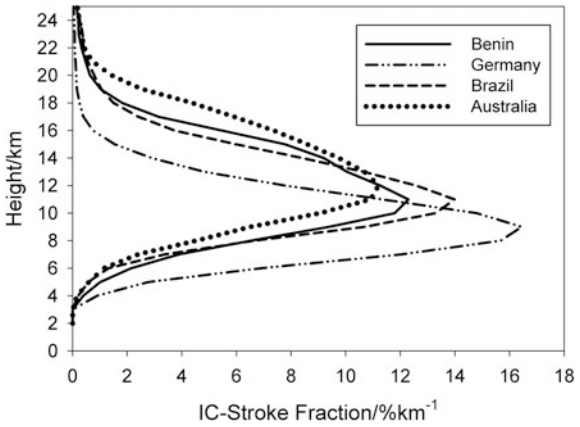

Fig. 20.7 Vertical profiles of IC (intra cloud) stroke fraction of the total IC strokes per km height interval. (From Höller et al. (2009))

20.5 Nowcasting, Modeling and Forecasting of Lightning

The short (nowcasting) and medium range forecasting of severe weather is an important application of lightning detection. Nowcasting techniques try to extrapolate existing trends in lightning structures into the future. Thus a thunderstorm already has to exist in order to represent useful nowcasting information. The quality of the nowcast decreases rapidly with increasing forecasting time and thus is useful in most cases only up to 0.5–1 h. In some cases of persistent storms or storm structures (e.g., long-lived *supercells*, convective complexes or fronts) the forecast might be useful for a longer period of time and longer warning times can be feasible.

In extension of a satellite and radar data tracking algorithm, a combined lightning and radar tracking and nowcasting system has been developed at IPA (Meyer 2010). Radar and lightning information is merged into a combined ECTRAM product (Electrically-charged Cells TRacking And Monitoring). Thus, different properties can be followed for a specific cell and provide more information and increased warning potential. An example is shown in Fig. 20.8. From the cell track (white line) a persistent supercell thunderstorm moving from west to east is noted over the domain shown. It had affected the city of Augsburg (A) and is approaching Munich international *airport* (MUC) at 16:30 UTC. The lightning cell is plotted as a red contour enveloping the individual lightning strokes from the previous 3 min (green crosses). The black and grey contours are the 10- and 20-min forecasts for the radar cell position, respectively. From these data dedicated warnings could be issued for the airport.

There is increasing effort to incorporate lightning information into numerical weather prediction models. Forecasting convection and convective precipitation is still a big challenge and has great potential for further improvement. Such

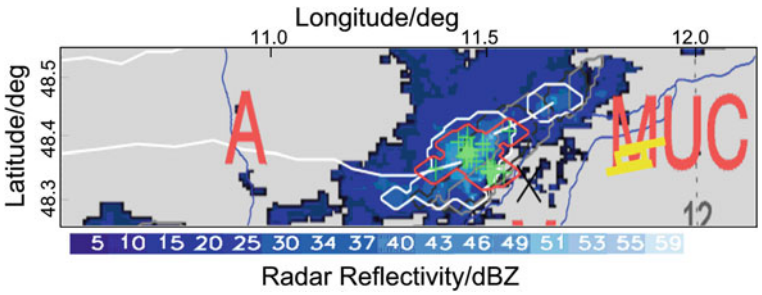

Fig. 20.8 EC-TRAM nowcasting at 16:30 UTC 26 June 2008. Radar reflectivity in *blue*, actual radar cell (33 dBZ) as *white contour*, radar cell track as *white line*, lightning cell as *red contour*, strokes from the previous 3 min as *green crosses*, radar cell forecasts as *black* (10 min) and *grey* (20 min) contours. *MUC* Munich airport, *A* the city of Augsburg. (From Meyer (2010))

developments have to keep up with improvements in the observing systems planned for the near future. With the next generation of European operational *geostationary* meteorological satellites MTG as well as the American GOES-R (Geostationary Operational Environmental Satellite, R-series) and Chinese satellites it will be possible to cover the globe with continuously recording optical lightning detection systems. The capabilities of regional, continental or global ground based lightning detection systems will also continuously improve and, thus, will add a further data source for lightning occurrence.

In numerical modeling the connection between the *microphysics* of precipitation and lightning can be used for incorporating a parameterization of lightning in the models. Recently, this approach has been successfully implemented at IPA in the COSMO-DE (Consortium for Small-Scale Modeling, Germany) model, the operational model used by the German Weather Service (Deutscher Wetterdienst, DWD) (Dahl et al. 2011a, b). Based on POLDIRAD graupel and ice particle concentrations and LINET flash frequencies, a parameterization scheme was developed which relates the model-derived graupel and ice contents to the experimental values. The parameterization uses a simple two-plate capacitor model for the main negative (graupel) and main positive (ice crystal) charge center and assumes equilibrium between the charging current due to microphysical particle interactions and the discharging due to the lightning current. A key parameter for the flash rate of a thunderstorm was found to be the graupel and ice volume and their mass concentration. One limitation for each *simulation* of lightning in a *numerical model* is the ability of the model to represent the convection itself at the right place, the right time and in the right intensity. Thus it is not surprising that when comparing LINET observations to the COSMO modeling shown in Fig. 20.9 differences are noted in the location of flash occurrence and the number of thunderstorm cells. For the specific case shown and under the chosen initial conditions the model produced too many thunderstorm cells. On the other hand, the intense supercell storm which passed over southern Germany on that day was also simulated in the model, except for the exact location. The model track

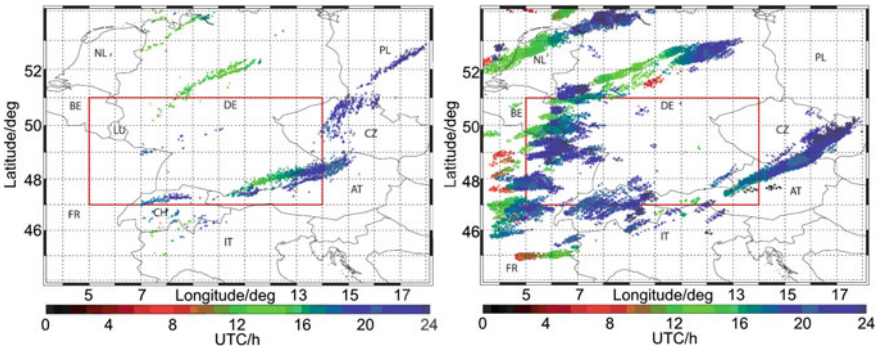

Fig. 20.9 Flashes on 22 August 2008 in central Europe. *Left* LINET observations. *Right* COSMO-DE with parameterized flash rate. (From Dahl et al. (2011b))

was initiated too far east over Austria, but the length, cell duration and lightning intensity were reasonably well simulated.

20.6 Lightning Detection From Space

In the past, optical lightning detection has successfully been used for lightning identification from space. Since the mid-1990s two optical *instruments* have been operated by NASA on board low orbit satellites. Both instruments are starring wide-angle cameras detecting the optical lightning signal. The Optical Transient Detector (OTD) was operated on board the MicroLab-1 satellite, which had an orbit inclination of 70° at an altitude of 735 km above ground and an orbital period of about 100 min. Lightning was detected with a 128 × 128 pixel CCD (charge-coupled device) matrix. The Lightning Imaging Sensor (LIS) is an integral part of the Tropical Rainfall Measuring Mission (TRMM) satellite, which was launched in 1997. In line with the objectives of TRMM, the satellite's orbit limited the observation area to the latitudes ± 35°. The altitude of 350 km and orbital period of 92 min were changed in August 2001 during the TRMM boost to 402 km and 93 min. The construction of the LIS instrument resembles the OTD with several improvements in sensitivity, design and electronics. The instantaneous field of view for the LIS is about 670 × 670 km^2; the local view time is 93 s. LIS also uses a 128 × 128 pixel CCD matrix leading to a 5 km pixel side length. The time resolution is 2 ms (single frame integration time). The OTD and LIS instruments are designed to detect the short transient optical signal from lightning. The typical optical signal from lightning is a light pulse containing characteristic spectral lines with durations shorter than 500 μs and a spatial size of about 10 km. The sensors make use of the emissions within the atomic *oxygen* line at about 777 nm.

In general, the optical data obtained by the pixel matrix are clustered into areas, flashes, groups, and events, based on their proximity in time and space. It is intended that this clustering reflect as much as possible the conventional

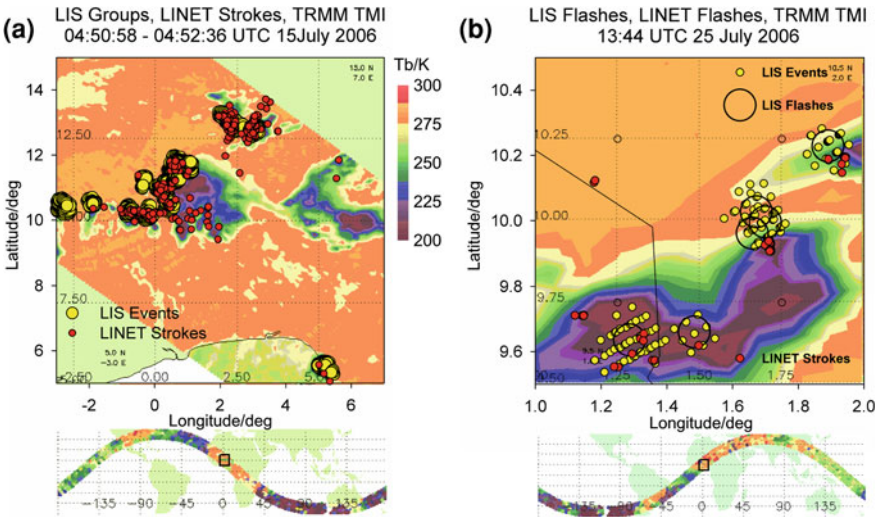

Fig. 20.10 **a** *Mesoscale* convective complexes over West Africa on 15 July 2006. TRMM-TMI brightness temperatures Tb indicate centers of convective elements. These are often co-located with flash locations from LIS and LINET. **b** Thunderstorm cells over Benin on 25 July 2006 as viewed by TRMM-TMI, LIS and LINET. The TRMM Science Data were obtained from the NASA EOSDIS Global Hydrology Resource Center (GHRC) DAAC, Huntsville, AL. http://thunder.nsstc.nasa.gov/

understanding of the lightning and thunderstorm processes. This implies that a LIS group is somewhat comparable to a LINET stroke. The LIS events composing a group do not have a direct counterpart in the LINET data as their number is determined by additional discharge processes not seen in VLF/LF and by the *scattering* processes of light within the cloud. The LINET strokes can also be grouped into flashes in order to compare them with LIS flashes. IPA uses a simple grouping algorithm clustering all strokes within 1 s and 10 km in time and space, respectively, starting from an initial stroke component.

Optical observation detects the optical radiation from both intra-cloud and cloud-to-ground lightning. This total lightning reflects more completely the electrical activity of the storm cloud. On the other hand, the separation of the cloud-to-ground lightning by optical means alone is difficult and may require additional data and new algorithms. Generally, it is thought that intra-cloud lightning radiates lower optical power, but contains more pulses per flash and has longer duration than cloud-to-ground lightning.

During the LINET field experiments in the *tropics* a large data base suitable for comparison of both systems was obtained. Two examples are shown in Fig. 20.10. Part (a) is a large-scale view of a convective complex over West Africa propagating from east to west on 15 July 2006. The color overlay from the TRMM *microwave* imager (TMI) shows the low *brightness temperatures* (T_b) at the leading edge (west) of the system. This is also the active cell region with imbedded

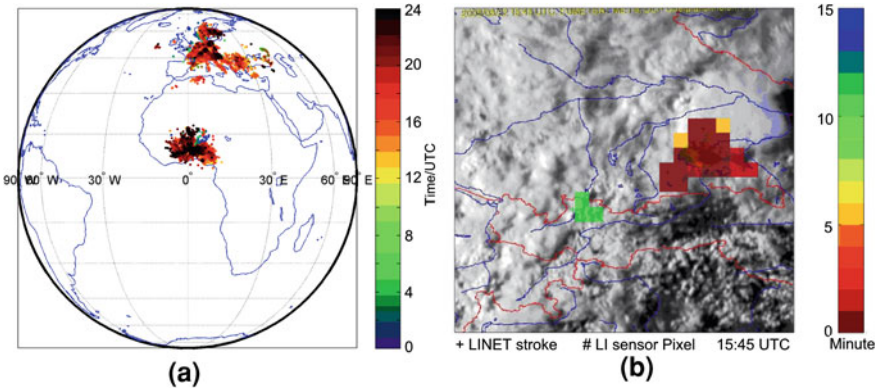

Fig. 20.11 Simulation of lightning optical event pixels from a future Meteosat instrument using LINET data from (**a**) simultaneous observations of the European and African networks on 28 July 2006, and (**b**) from the supercell on 22 August 2008 over southern Germany (*crosses* indicate LINET strokes). The 15:45 UTC Meteosat image is used as underlay and time reference for color coding (time/min before reference time)

lightning. The LINET network is centered in the middle of the picture, so that the stroke detection efficiency decreases towards the edges. This effect is visible by the many LINET strokes in the center without a corresponding optical signal and the opposite effect far away from the center. Part (b) shows smaller cells on 25 July 2006 close to the LINET network. The optical event matrix of the illuminated instrument pixels becomes visible (scaled down for better illustration). All flashes at that time are seen by both systems. The lightning area observed by LIS is somewhat larger than the stroke area observed by LINET.

While the VLF/LF radio emissions from lightning are caused by relatively large currents flowing in rather long channels, *optical emissions* may be caused by a large variety of discharge processes leading to luminous events. This explains the general finding from the LIS-LINET comparison that, on average, there were seven LIS groups observed for one LINET stroke. For most of the LINET strokes a directly coincident LIS group could be detected. As the optical detectors alone cannot discriminate CG flashes from IC flashes, those optical properties can be inferred by a comparison with LINET. As a result it was found that the detection efficiency for CG flashes was quite high even for optical observations from the cloud top. This might be explained by the in-cloud components of CGs, whose lower parts otherwise would have suffered more from scattering and would be expected to be less visible at cloud top. Also, the corresponding optical characteristics of strokes at different emission heights are often obscured and subject to a large amount of statistical variability with only weak correlations.

The statistical properties found from LIS and LINET comparisons can be used to assess the future performance of geostationary lightning detection. Such systems are under development and will be part of the next generation of operational weather satellites like the European MTG or the U.S. GOES-R. The transition to

lightning detection from a geostationary orbit is challenging due to the much larger distance to Earth. The entrance optics has to realize a 100-fold higher angular resolution in order to achieve the same spatial resolution as with the low orbit instruments. Moreover, the location accuracy will depend on the location on Earth due to the geometric increase of the pixel footprint with the satellite zenith angle and parallax effects. The advantage of geostationary satellite-based lightning detection is the continuous coverage of huge areas of the globe with similar detection efficiency. It also offers a large potential for new synergetic products in combination with other cloud observing satellite sensors and ground based lightning observing systems, and as a new source of information for weather prediction and severe weather warning.

As an example of the expected capabilities of a future lightning sensor on MTG (MTG-LI), Fig. 20.11 shows proxy lightning data calculated by a statistical simulation of optical events using LINET data as input. The assumed resolution of the optical array is about 10 km in the tropics. The time resolution is assumed to correspond to the present LIS resolution of 2 ms. Each pixel would be illuminated several times by each flash.

References

Betz, H.-D., Schmidt, K., Oettinger, P., Wirz, M.: Lightning detection with 3-D discrimination of intracloud and cloud-to-ground discharges. Geophys. Res. Lett. **31**, L11108 (2004). doi:10.1029/2004GL019821

Betz, H.-D., Schmidt, K., Laroche, P., Blanchet, P., Oettinger, W.P., Defer, E., Dziewit, Z., Konarski, J.: LINET—an international lightning detection network in Europe. Atmos. Res. **91**, 564–573 (2009a). doi:10.1016/j.atmosres.2008.06.012

Betz, H.-D., Schumann, U., Laroche, P. (eds.): Lightning: Principles, Instruments and Applications, 641 pp. Springer, New York (2009b)

Betz, H.-D., Schmidt, K., Fuchs, B., Oettinger, W.P., Höller, H.: Cloud Lightning: detection and utilization for total lightning measured in the VLF/LF regime. J. Lightning Res. **2**, 1–17 (2007)

Dahl, J.M.L., Höller, H., Schumann, U.: Modeling the flash rate of thunderstorms. Part I: framework. Mon. Weather Rev. **139**, 3093–3111 (2011a). doi:10.1175/MWR-D-10-05031.1

Dahl, J.M.L., Höller, H., Schumann, U.: Modeling the flash rate of thunderstorms. Part II: implementation. Mon. Weather Rev. **139**, 3112–3124 (2011b). doi:10.1175/MWR-D-10-05032.1

Höller, H., Betz, H.-D., Schmidt, K., Calheiros, R.V., May, P., Houngninou, E., Scialom, G.: Lightning characteristics observed by a VLF/LF lightning detection network (LINET) in Brazil, Australia, Africa and Germany. Atmos. Chem. Phys. **9**, 7795–7824 (2009). doi:10.5194/acp-9-7795-2009

Huntrieser, H., Schumann, U., Schlager, H., Höller, H., Giez, A., Betz, H.-D., Brunner, D., Forster, C., Pinto Jr, O., Calheiros, R.: Lightning activity in Brazilian thunderstorms during TROCCINOX: implications for NOx production. Atmos. Chem. Phys. **8**, 921–953 (2008). doi:10.5194/acp-8-921-2008

Huntrieser, H., Schlager, H., Lichtenstern, M., Roiger, A., Stock, P., Minikin, A., Höller, H., Schmidt, K., Betz, H.-D., Allen, G., et al.: NOx production by lightning in Hector: first airborne measurements during SCOUT-O3/ACTIVE. Atmos. Chem. Phys. **9**, 8377–8412 (2009). doi:10.5194/acp-9-8377-2009

Huntrieser, H., Schlager, H., Lichtenstern, M., Stock, P., Hamburger, T., Höller, H., Schmidt, K., Betz, H.-D., Ulanovsky, A., Ravegnani, F.: Mesoscale convective systems observed during AMMA and their impact on the NOx and O3 budget over West Africa. Atmos. Chem. Phys. **11**, 2503–2536 (2011). doi:10.5194/acp-11-2503-2011

Labrador, L., Vaughan, G., Heyes, W., Waddicor, D., Volz-Thomas, A., Pätz, H.-W., Höller, H.: Lightning-produced NOx during the Northern Australian monsoon; results from the ACTIVE campaign. Atmos. Chem. Phys. **9**, 7419–7429 (2009). doi:10.5194/acp-9-7419-2009

Meyer, V.: Thunderstorm tracking and monitoring on the basis of three dimensional lightning data and conventional and polarimetric radar data. Dissertation, Faculty of Physics, LMU München (2010)

Rakov, V.A., Uman, M.A.: Lightning, p. 687. Cambridge University Press, New York (2003)

Schmidt, K.: Ortung und Analyse von Blitzentladungen mittels Registrierung von VLF-Atmospherics innerhalb eines Messnetzes. Dissertation, Faculty of Physics, LMU München (2007)

Schumann, U., Huntrieser, H.: The global lightning-induced nitrogen oxides source. Atmos. Chem. Phys. **7**, 3823–3907 (2007). doi:10.5194/acp-7-3823-2007

Cloud and Precipitation Radar

21

Martin Hagen, Hartmut Höller and Kersten Schmidt

Abstract

Precipitation or weather radar is an essential tool for research, diagnosis, and nowcasting of precipitation events like fronts or thunderstorms. Only with weather radar is it possible to gain insights into the three-dimensional structure of thunderstorms and to investigate processes like hail formation or tornado genesis. A number of different radar products are available to analyze the structure, dynamics and microphysics of precipitation systems. Cloud radars use short wavelengths to enable detection of small ice particles or cloud droplets. Their applications differ from weather radar as they are mostly orientated vertically, where different retrieval techniques can be applied.

21.1 Introduction

First attempts to use electromagnetic waves for the detection of distant objects date back to Christian Hülsmeyer in 1904. During World War II this principle was reinvented for military purpose. The fast development of *RADAR* (an acronym for

M. Hagen (✉) · H. Höller · K. Schmidt
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Martin.Hagen@dlr.de

H. Höller
e-mail: Hartmut.Hoeller@dlr.de

K. Schmidt
e-mail: Kersten.Schmidt@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_21,

radio detection and ranging) enabled the detection of ships and aircraft over long distances. About the year 1942 spurious radar echoes were detected and related to precipitating clouds. After World War II a rapid development of radar for meteorological purposes took place. Meteorological radar systems now differ in many respects from air traffic control radars or navigational radars used on ships.

Meteorological radar—normally termed weather radar—is the only system able to observe the three-dimensional structure of cloud and *precipitation* systems with high spatial (100–1 000 m) and temporal (2–10 min) resolution. Weather radars are used for a wide spectrum of operational applications: identifying precipitation systems (like fronts or thunderstorms) for synoptic purpose; detailed analysis of dynamic and microphysical structures in thunderstorms; monitoring thunderstorm motion for nowcasting; estimating precipitation amount for hydrological applications like flood forecasting for river catchments; identification of thunderstorm or hail swaths; establishing long term observations for precipitation climatologies. A total of about 200 weather radars are operated by weather services in Europe today. Research using weather radar covers three aspects: (i) development of technologies (ii) development of data evaluation products and algorithms, and (iii) investigation of meteorological phenomena.

The DLR Institute of Atmospheric Physics has been using weather radar for research since the 1960s. This work began with a conventional weather radar, first at the former München-Riem airport and since 1968 at Oberpfaffenhofen (Müller and Singler 1972). Since 1986 the Polarization Diversity Doppler Radar POLDIRAD has crowned the institute's building (Schroth et al. 1988; Meischner et al. 1991).

21.2 Radar Principles

A radar transmits a short pulse of high power electromagnetic waves through a directional antenna. This pulse is reflected by an object and received again by the radar receiver. The direction the antenna is pointing and the round-trip time of the pulse are used to locate the object *scattering* the pulse. The principal components of a weather radar are the transmitter, the antenna, the receiver, the signal processor, and the product and image generator. Radar operation control, signal processing, and image generation is accomplished with powerful standard computers. On receive the signal is digitized at an early stage in order to apply sophisticated digital filters and to avoid calibration problems due to aging or temperature-sensitive hardware components. The characteristics of a typical weather radar are listed in Table 21.1.

Electromagnetic waves of different wavelengths are used for meteorological radars. The *wavelength* is selected according to the target of interest, but also with regard to logistics in terms of antenna size and costs (both increase with increasing wavelength). Since radiation with short wavelength experiences stronger attenuation in cases of precipitation, radars using long wavelengths are preferred in

Table 21.1 Characteristics of a typical C-band weather radar and for POLDIRAD

Parameter	Typical C-band radar	POLDIRAD
Transmitted pulse peak power	200–400 kW	200 kW
Frequency/wavelength	5 600–5 650 MHz/5.4−5.3 cm	5 504 MHz/5.4 cm
Pulse duration	0.6–1.5 μs	0.5, 1.0, 2.0 μs
Depth of measurement volume	90–225 m	75, 150, 300 m
Pulse prepetition frequency	600–1 200 Hz	400–2 400 Hz
Minimum detectable signal	10^{-11} mW	10^{-11} mW
Maximum range	120–250 km	60–300 km
Antenna diameter	4 m	4.5 m
Antenna rotation speed	2–6 rpm	2 rpm
Beam width	1°	1°
Polarization	Conventional radar: linear horizontal; dual-polarization radar: linear 45° transmit, linear horizontal and vertical receive	Any for transmit, any for receive

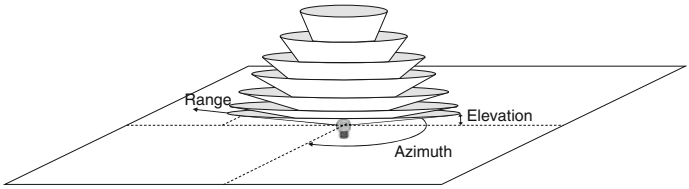

Fig. 21.1 Schematic of a volume scan consisting of a number of conical scans along azimuth at fixed elevation

regions with strong precipitation. Most weather radars in Europe are *C-band* radars (frequency 5.6 GHz, wavelength 5.4 cm); only some radars in the Mediterranean area are S-band radars (frequency 2.8 GHz, wavelength 10.7 cm). X-band radars (9.4 GHz, wavelength 3.2 cm) are used for short ranges like cities or river catchments, but also in mountainous regions to cover valleys which are shielded from the long-range radars located outside the mountains or on mountain peaks.

The radar normally takes measurements with a fixed elevation angle and a complete revolution in azimuth. This type of scan and the related graphical representation of the data is termed a PPI (plan position indicator). To explore the three-dimensional structure of the precipitation field a number of PPIs at different elevations (typically 10–20 PPIs at elevations between 0.5 and 40°) are produced (Fig. 21.1). The volume scan takes 5–15 min and allows for the evaluation of a number of different radar products like pseudo three-dimensional display, arbitrary

vertical cross sections through the volume, or echo top display. Scans with a fixed azimuth angle and varying elevation are called RHI (range height indicators) and are used for detailed exploration of the vertical structure of precipitation fields.

This section describes some fundamental radar principles. More details can be found in textbooks like Sauvageot (1992), or for advanced applications, Meischner (2004).

21.2.1 The DLR C-Band Polarization Diversity Doppler Radar

POLDIRAD was one of the first fully polarimetric weather radar in the world and is—after several upgrades—still unique in aspects like the polarization switch network and the offset antenna. A *polarization* or *polarimetric* radar uses different polarizations for the transmitted and received electromagnetic waves. The radar is a joint development of the DLR Microwaves and Radar Institute, the DLR Institute of Atmospheric Physics, and the manufacturing companies Enterprise Electronics Corporation and Sigmet, both from the USA. The unique high-power ferrite polarization network allows the setting of different polarizations, changing from radar pulse to pulse and also different for transmit and receive. This makes the radar ideally suited for the investigation and comparison of different polarization modes and different polarization bases. The prominent antenna with the offset-feed has better performance than standard parabolic antennas with a central feed in terms of side-lobe suppression and channel isolation. It was specially designed for a C-band polarimetric weather radar.

A large number of projects were carried out with POLDIRAD over the last 25 years. The main focus was thunderstorm research including various aspects like hail initiation, hail suppression, transport of trace gases, lightning initiation, hydrometeor identification, thunderstorm *nowcasting*, and orographic effects. But other topics like electromagnetic wave propagation, effects of precipitation on satellite links, aircraft icing, wind field estimation for aviation, and snow identification for aviation were also covered with POLDIRAD. A highlight was the deployment of POLDIRAD during COPS (Convective and Orographically-induced Precipitation Study field campaign) for three month in summer 2007 at Waltenheim-sur-Zorn in Alsace, France. One goal of COPS was to identify the effects of *orography*on the initiation of convection. It was shown that the wind profile plays an important role, among other factors (Hagen et al. 2011).

21.2.2 Radar Parameters

Whereas other types of radar only aim to detect targets, the scope of weather radar is to quantify the target. A number of different radar parameters are retrieved from a polarimetric Doppler weather radar. In addition to a conventional weather radar measuring only the intensity of the returned signal, a *Doppler radar* additionally

estimates the motion of the target. Since a polarimetric radar uses electromagnetic waves with two different polarizations, the derived radar parameters allow a characterization of the shape and orientation of the target.

21.2.2.1 Radar Reflectivity Factor

Radar *reflectivity* is a synonym for the magnitude of the reflected radar pulse. The radar equation for volume targets gives the relationship between the received signal P_r and the scattering *cross section* σ of the target

$$P_r = \frac{P_t g^2 \lambda^2}{64\pi^3 r^4} \frac{\pi r^2 \theta_0^2 h}{8} \frac{1}{2\ln(2)} \frac{1}{V} \sum_{i=1}^{n} \sigma_i$$

with P_t the transmitted power, g the antenna gain, λ the wavelength, θ_0 the beam width, h the pulse length, r the distance to the targets, and σ_i the scattering cross sections of the targets within the unit volume V. It is assumed that the measurement volume is uniformly filled with scattering targets. The first term describes the radar equation for a single point target, the second the measurement volume, and the third term accounts for the fact that only part of the transmitted power is within the measurement volume bounded by the beam width.

For particles which are large compared to the wavelength (diameter $D > 10\ \lambda$), the scattering cross section σ of a target is given by its geometric cross section. For particles much smaller than the wavelength ($D < 1/10\ \lambda$), the scattering cross section can be approximated by *Rayleigh scattering*

$$\sigma = \frac{\pi^5}{\lambda^4} |\mathrm{K}|^2 D^6$$

with $K = (m^2-1)/(m^2 + 2)$ representing the complex refractive index m of the scattering particle ($|\mathrm{K}|^2 \approx 0.93$ for water and $|\mathrm{K}|^2 \approx 0.2$ for ice). With the assumption that the targets are water particles and much smaller than the wavelength, the radar reflectivity factor[1] z is defined as the sum of D^6 of all particles within the unit volume V

$$z = \frac{1}{V} \sum_{i=1}^{n} D_i^6$$

The unit of z is used in the form $\mathrm{mm}^6/\mathrm{m}^3$. Electrical engineers usually express a ratio R of two powers (p_1 and p_2) in a logarithmic notation $R = 10\log(p_1/p_2)$. The unit of such a ratio is—in honor of Alexander Graham Bell—the *decibel* (dB). It is common in radar meteorology to express the reflectivity factor as a logarithmic ratio with respect to the reflectivity factor of a raindrop with a diameter of 1 mm.

[1] Informally, but also in a large number of publications, the term "reflectivity" is used instead of "reflectivity factor."

$$Z = 10\log\left(\frac{z}{1\,\mathrm{mm}^6\,\mathrm{m}^3}\right)$$

The unit of the logarithmic quantity Z is termed a dBZ (decibel Z).

21.2.2.2 Doppler Velocity

If there is relative motion between the radar and the target, a frequency shift of the reflected wave is observed by the radar. This effect was noted by the Austrian astronomer Christian Doppler in 1842 and the difference between the transmitted frequency and the received frequency is defined as the Doppler frequency. Even though a radar cannot measure the frequency shift because of principal technical limitations, in radar meteorology the term Doppler velocity is used to indicate the measurement of the radial motion of the target relative to the radar.

Weather radar estimates the Doppler frequency from a time series of phase measurements using a number ($\sim$20–60) of radar pulses. With this technique the maximum Doppler frequency or Doppler velocity is ambiguous if the target moves more than half a wavelength within two pulses. The interval where the Doppler velocity can be estimated unambiguously is also called the "*Nyquist interval*." For a typical C-band weather radar, the maximum unambiguous Doppler velocity is in the order of $\pm$ 8 to $\pm$ 16 m/s. Higher velocities will be folded into the Nyquist interval. A number of techniques exist to overcome this limitation, like varying the time between pulses, so called dual-PRF, dual-PRT, or staggered PRT techniques (PRF = pulse repetition frequency; PRT = pulse repetition time). With these techniques the Nyquist interval can be doubled or tripled.

Only a motion towards or away from the radar (the radial or *line-of-sight*[2] component) can be detected by the Doppler effect. If the target moves perpendicular to the radar beam, no Doppler effect is visible. In general, it is assumed that precipitation particles move with the wind; only their fall speed has to be considered. As long as the elevation angle is low, the contribution of the fall speed to the radial velocity is negligible.

21.2.2.3 Differential Reflectivity

Common dual-polarization weather radars use linear polarization with horizontal and vertical polarization planes. Z_{DR} is the ratio between the reflectivity factor measured at horizontal and vertical polarization

$$Z_{DR} = 10\log\left(\frac{z_H}{z_V}\right)$$

with z_H the reflectivity factor at horizontal polarization and z_V the reflectivity factor at vertical polarization. The unit of Z_{DR} is dB. Z_{DR} is the classic radar

[2] Both terms are equivalent and describe Doppler velocity. In radar meteorology the term "radial" component or velocity is preferred; scientists working with lidar prefer the term "line-of-sight" component or velocity.

parameter, which can be measured with dual-polarization radar. Positive Z_{DR} is observed when particles are flat and fall horizontally aligned. This is the case for raindrops larger than 1 mm in diameter and for ice needles or plates. Z_{DR} is used to improve the rainfall rate estimation and to identify hydrometeors. The range of Z_{DR} in rain is 0–5 dB. Particles like graupel or hail have an irregular shape and tumble and rotate during descent. Their Z_{DR} is typically around zero.

21.2.2.4 Linear Depolarization Ratio

The linear *depolarization* ratio (*LDR*) describes how much energy of the transmitted horizontal polarized wave is scattered back vertically polarized

$$LDR = 10 \log\left(\frac{z_{VH}}{z_H}\right)$$

with z_{VH} the reflectivity factor (in mm^6/m^3) received with vertical polarization while transmitted with horizontal polarization. The unit of *LDR* is dB. *LDR* is caused by particles which are canted during falling or tumble and rotate heavily while falling. These are particles which have an irregular shape and/or are water coated like melting snowflakes or wet graupel or hail. *LDR* is about -35 dB in weak rain, -25 dB for graupel and -15 dB or higher for melting hail. *LDR* is used to identify hydrometeors.

21.2.2.5 Co-Polar Correlation Coefficient

The co-polar correlation coefficient $\rho_{HV}(0)$ at time lag zero describes how the time series of the received complex signal for a number of pulses at horizontal and vertical polarization are correlated to each other. The correlation is high for spherical particles which fall uniformly and low for particles with irregular shapes and unsteady falling behavior. The observed $\rho_{HV}(0)$ in rain is 0.995–0.999. Only in strong rain does $\rho_{HV}(0)$ drop down to 0.98 or 0.97. Much lower values (down to 0.75) are observed in water-coated hail or melting snowflakes. $\rho_{HV}(0)$ is used to identify hydrometeors.

21.2.2.6 Differential Propagation Phase

While all parameters described above depend on the scattering process in the measurement volume, the differential propagation phase describes the propagation medium. When an electromagnetic wave passes through rain with oblate and horizontally aligned raindrops, the wave with horizontal polarization is delayed more than the wave with vertical polarization. This causes a phase shift between the two waves. The waves will be scattered at some distance and on the way back to the radar another differential phase shift occurs. The differential phase shift on propagation ϕ_{DP} increases with distance from the radar and is normally caused by oblate raindrops. The unit is degrees. The specific differential phase K_{DP} describes the slope of ϕ_{DP}, its unit is degrees/km. K_{DP} is in the order of 0–6°/km in rain. K_{DP} is normally used for rain rate estimation and to correct attenuation.

21.2.3 Radar Products

The measured radar parameters discussed so far do not represent meteorological quantities. A number of different meteorological quantities like rain rate, wind field or hydrometeor classification can be retrieved from the measurements.

21.2.3.1 Rain Rate Estimation Using Radar

Since *rain* rate and reflectivity factor are different moments of a raindrop *size distribution*, no universal relation between rain rate (R) and reflectivity factor is possible. Common are empirical relations (z-R relations) in the form $z = a{\cdot}R^b$ which were retrieved from long-term drop size distribution measurements. With these relations the rain rate can be estimated with an accuracy of about 20–50 %. Observations in different climatological regions show that the coefficient a has large variations from region to region. For example, the German Weather Service (DWD) uses $z = 256{\cdot}R^{1.42}$.

With polarimetric radar, additional parameters like Z_{DR} or K_{DP} are available. Both parameters can be used to improve the quantitative rain rate estimation. However, an uncertainty remains due to the unknown processes between the radar measurements at some height above ground in a large measurement volume over a short time (app. 0.1 s) every 5–10 min and the continuous measurement of a rain gauge at ground with a standardized sampling size of 200 cm^2. A number of techniques exist to combine radar and rain gauge measurements to optimize the areal estimation of rainfall.

21.2.3.2 Retrieval of Wind Field

As explained above, radar measurements of Doppler velocity only represent the radial component of the *wind* vector. For meteorological interpretation, all three components of the wind vector are of interest. Several methods exist to estimate the remaining components, or at least the horizontal wind vector. Exemplarily, the implementation of single-Doppler and multiple-Doppler wind-field *retrieval* techniques with POLDIRAD are shown below.

Single-Doppler Uniform Wind Technique

With the assumption of a constant wind field in a certain region, it is possible to estimate the horizontal wind vector field from measurements of the radial velocity by a single radar. The radial component v_r of the wind vector field is known; the missing component is the tangential component v_t

$$v_r = u_0 \sin\phi \cos\theta + v_0 \cos\phi \cos\theta$$
$$v_t = -u_0 \cos\phi \cos\theta + v_0 \sin\phi \cos\theta$$

with u_0 and v_0 the horizontal wind components west–east and south–north, θ elevation angle and ϕ azimuth angle. It follows that $v_t = -\partial v_r / \partial\phi$, i.e., the slope of the radial velocity along an arc. To achieve sufficient accuracy it is common to

average the radial velocity over a range of about 10 km and to estimate the slope over a sector of 10–20°. The assumption is that the wind is constant within this sector element of 10 km by 10–20°. This assumption is normally valid for stratiform precipitation systems (Hagen 1992), but in convective situations it is certainly violated.

Multiple-Doppler Wind-Field Retrieval

If more than one Doppler radar observes the same target from different directions, the individual Doppler velocities can be combined to retrieve the complete wind vector. One constraint is that the individual measurements are taken simultaneously (within a few minutes), and that the intersection angle between the radar beams is in the range of 30–150°. The optimal intersection angle would be 90°, but, in the indicated range the accuracy of the retrieved wind vector is below ca. 3 m/s. Multiple- or dual-Doppler techniques are often used for field campaigns where a number of radars are located close together. A typical distance between the radars is about 50 km. In operational radar networks the distance between the radars is in the order of 100–200 km, but even then multiple-Doppler wind-field retrieval is still possible. MeteoFrance operationally retrieves the wind vector field in the Paris airports area.

To limit effort of operating several Doppler radars it is also possible to install a *bistatic* Doppler radar network. It consists of one active Doppler radar and a number of passive Doppler receivers. The receivers measure the Doppler velocity from the sideward scattered energy. The receivers have a small fixed antenna. The principle of retrieval is the same for a bistatic Doppler radar network as for a multiple-Doppler radar network; the major differences are the different geometry (ellipsoidal coordinates instead of spherical coordinates) for the intersection angle and the fact that the size of the measurement volume depends on the intersection angle (Friedrich and Hagen 2004). The advantage of a bistatic over a conventional Doppler radar network is its simplicity for the receivers. All measurements are from the same measurement volume; there is no time lag between the observations. The disadvantage is the lower sensitivity of the bistatic receivers and the smaller area in which the wind field can be retrieved. However, since the costs of a bistatic receiver are less than 1/20 of an active transmitting radar, additional receivers can be installed easily to improve the areal coverage.

21.2.3.3 Dual-Polarization Hydrometeor Classification

Dual-polarization radar parameters are sensitive to the shape and phase of the scattering particles. From observations, numerical simulations, and theoretical considerations it is possible to define the expected range of the different radar parameters (Höller et al. 1994). Additional information about the height of the 0 °C isotherm is included to distinguish between light rainfall and *snow*. The basis for the classification is Z_{DR} and *LDR,* which give information about the shape and falling behavior. Graupel and hail are spherical on average, thus Z_{DR} is around zero. Wet tumbling particles like melting hail give high *LDR* because of the

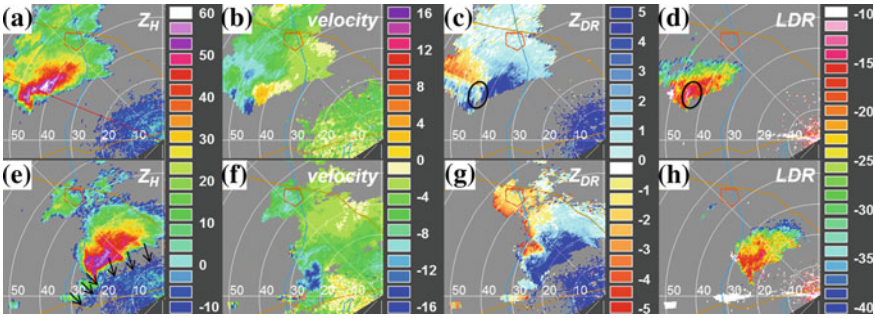

Fig. 21.2 PPI of POLDIRAD measurements at 1° elevation. *Top row* at 12:28 UTC 23 June 2008, *bottom row* at 12:56 UTC. (**a**, **e**) reflectivity factor Z_H, (**b**, **f**) Doppler velocity, (**c**, **g**) differential reflectivity Z_{DR}, (**d**, **h**) linear depolarization ratio *LDR*. Range rings every 10 km

irregular falling behavior and the fact that the refractive index of water is seven times that of ice.

With additional radar parameters and more advanced decision techniques like fuzzy logic, it is possible to improve the classification and to introduce further *hydrometeor* types or nonmeteorological targets like ground echoes. Insects and birds may also be detected.

21.3 Radar Observations of Thunderstorms

21.3.1 POLDIRAD Observations of A Supercell Storm

The observation of a super cell thunderstorm with POLDIRAD will show exemplarily the radar products obtained and their use for *thunderstorms* investigations.

Figures 21.2 and 21.3 show detailed measurements with POLDIRAD of a super-cell thunderstorm which lasted for about 7 h, moving across southern Germany on 22 June 2008. The PPIs in Fig. 21.2 show the typical features known from supercell *storms*. The reflectivity factor (Fig. 21.2a, e) shows a strong precipitation core with values up to 60 dBZ and a hook-like structure at the southern end, termed a "hook echo." As can be seen from the Doppler velocity (Fig. 21.2b), the hook echo is associated with strong rotation; the velocity band of 3 m/s away (orange) from the radar appears next to the band with 16 m/s towards (blue) the radar. This structure is typical for the occurrence of *tornados*; however there are no reports of a tornado for this particular event. Z_{DR} (Fig. 21.2c) shows large values in the eastern part of the storm indicating the presence of large oblate raindrops which have an axis ratio in the order of 0.6. The streaks with negative Z_{DR} are an effect of differential attenuation which occurs if there are large oblate raindrops or melting hailstones. High *LDR* indicates the presence of hail and *graupel*. Strong *hail* is expected in the small region about 45 km from the radar (black circle in

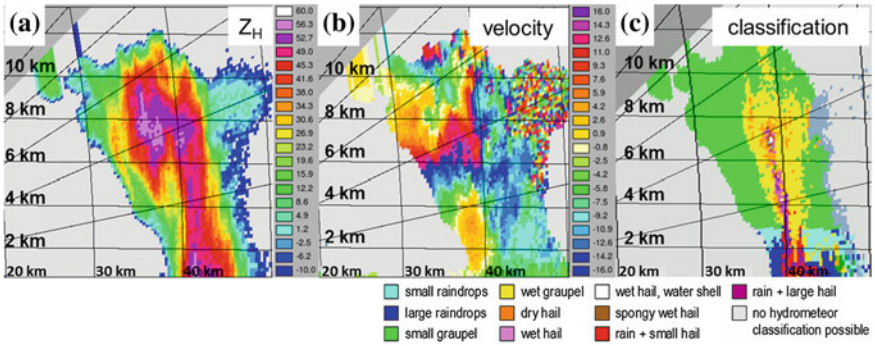

Fig. 21.3 POLDIRAD measurements range height indicator RHI versus distance from radar and altitude at 292° azimuth, 12:35 UTC 23 June 2008. **a** Reflectivity factor Z_H in dBZ, **b** Doppler velocity in m/s, **c** hydrometeor classification

Fig. 21.2c, d). Z_{DR} is around 0 dB and *LDR* is large with values around −11 dB. Hailstones 3–5 cm in diameter have been reported for this storm. The echoes close to the radar (range <20 km) are caused by insects in the convective *boundary layer*. This kind of echo is called a clear-air echo because it is unrelated to precipitation. As their reflectivity factor is weak, these echoes are seen only up to 20 km in range. The high Z_{DR} is an indication for insects with elongated bodies. The structure of the clear-air echoes (Fig. 21.2e, f) has some indications for a cold-air outflow out of the storm. The thin line with weak reflectivities leading from the hook echo to the south and near the echo free region east of the hook echo (marked with black arrows in Fig. 21.2e) are the leading edge of the outflow. The cold outflow rushes beneath the warmer air. *LDR* shows high values close to the radar, caused by ground echoes. The radar echoes associated with *LDR* values of −10 dB and higher towards the west are echoes from storms beyond the maximum range of the PPI measurement (125 km). These so-called "second-trip" echoes result from distant cells along the Danube River and Swabian Jura.

Figure 21.3 shows a vertical cross section (RHI scan) at 12:35 UTC towards 292° (note the red line in Fig. 21.2a). The reflectivity factor shows a vertical column of high values (40–45 km range) and an overhanging region (35–42 km range, 6–10 km height). The overhang region is the graupel and hail growth region. The region beneath the overhang is termed a weak echo or bounded weak echo region; both are an indication of strong updrafts. The high reflectivity factor core consists of heavy precipitation or graupel and hail. The radial Doppler velocity cannot be used to measure the vertical velocity since the elevation angles are low and vertical velocity includes the fall speed of the particles. However, there are typical signatures indicating strong vertical motion. At low levels, convergent flow (40 km range) leads to updrafts. More dominant is the divergent flow in the *anvil* region (above 8 km height). Here, the vertical motion is suppressed by stratification at the tropopause and diverges in horizontal directions.

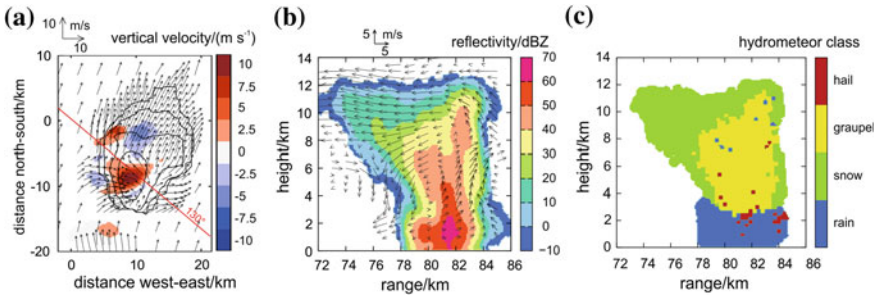

Fig. 21.4 **a** Multiple-Doppler horizontal wind vector field at 5 km above sea level at 13:35 UTC 15 July 2007: The vectors represent the horizontal velocity, colors the vertical velocity, *red* updraft, *blue* downdraft. Contours represent reflectivity factor starting at 10 dBZ with interval 10 dB. **b** Range height indicator RHI towards 130° (*red line* in **a**) from POLDIRAD at 14:43 UTC, *colors* reflectivity factor, *vectors* horizontal and vertical multiple-Doppler wind field components. **c** as **b** for hydrometeor classification

The Doppler velocity is folded twice into the Nyquist interval (±16.4 m/s); the maximum flow is 50 m/s towards the radar (red/purple area at 35 km range and 11 km height), whereas the flow beyond 40 km is about 5 m/s towards the radar. The noisy part beyond 44 km range and above 7 km height is a fake echo, also called a hail-spike or flare echo (also visible in the reflectivity factor). This is caused by waves scattered from the hail shaft towards the ground, back to the hail shaft and then back to the radar. This causes a delay of the echo and therefore it is shown at a farther distance. Figure 21.3c) depicts automatic hydrometeor classification results. The main hail core extends up to about 8 km in height. The inner core is classified as wet growing hail, caused by high liquid water content in the main updraft of the storm. The hail shaft extends all the way down to the ground where hail mixed with rain is analyzed.

21.3.2 Multiple-Doppler Observations of a Single Cell Storm

During COPS, a fast developing *single cell storm* was observed on 15 July 2007 by POLDIRAD and other Doppler radars at Karlsruhe, Feldberg and Türkheim (Schmidt et al. 2012). This example illustrates the added information obtained with additional radars, in particular the wind field that can be retrieved using a multiple-Doppler algorithm. Figure 21.4a) shows the horizontal wind field from a multiple-Doppler analysis, with vertical motion and reflectivity of the thunderstorm at 14:35 UTC in its growing stage. A strong updraft is next to the core with the highest reflectivity, and only a weak downdraft is observed. A RHI scan (Fig. 21.4b) by POLDIRAD along 130° (red line in Fig. 21.4a) shows the reflectivity core at 14:43 UTC. At that time, the storm was already decaying. The high reflectivity core is subsiding to lower levels and updraft prevails only on

the right edge of the storm. A hail spike is visible at 84–86 km range and 2 km height. The classification (Fig. 21.4c) shows some hail, and hail was observed at the ground.

21.4 Cloud Radar

The name "cloud radar" implies observations of small *cloud particles* with high resolution. Cloud radars are typically high-frequency radars (at Ka- or W- band, frequency 35 or 95 GHz, wavelength 8 or 3 mm, respectively) which can detect small-scale particle complexes (down to 10 μm) in contrast to low frequency radars (at S-, C-, or X-band) which are more suited for precipitation particles. Similar to lidar, cloud radars are mostly applied pointing vertically, but in contrast to lidar, the beam is much less attenuated by cloud or precipitation particles. At IPA no cloud radar is operated, but the institute has proposed and has access to the cloud radar MIRA-36 (constructed by METEK Inc.) at the Schneefernerhaus Research Station just beneath the peak of Germany's highest mountain, the Zugspitze, as well as the MIRA-36 cloud radar on board the German research aircraft *HALO* within HAMP (HALO microwave package). MIRA-36 has Doppler capability, operates at a wavelength of 8.6 mm (35.2 GHz) and achieves a maximum pulse power of 30 kW. The radar is polarimetric and can measure the linear depolarization ratio *LDR* in addition to reflectivity.

The radar is used for cloud parameter investigations. Reflectivity is related to the size and amount of particles. *LDR* is low for spherical liquid *droplet*s, but higher for irregular particles like large ice aggregates. Doppler velocity, as discussed above, is only one product—the first moment—of the full Doppler power spectrum. Microphysical cloud properties such as the characteristic particle size or the liquid and ice water contents are conventionally retrieved with vertically pointing radars using only the first three radar moments (power, recalculated as reflectivity, mean radial velocity, and Doppler spectral width). Powerful signal processing and data recording now allow for the recording and analysis of the full Doppler spectrum. The Doppler spectrum shows the reflectivity-weighted velocity contributions to the total reflectivity and mean velocity. Different hydrometeor types (cloud droplet, drizzle, rain drops, *ice particles*, snow, etc.) have different fall velocities, resulting in different contribution to the Doppler spectrum. Therefore, the Doppler spectrum, see Fig. 21.5, can be used to derive further microphysical information. The different modes of the Doppler spectrum can be associated with the combined effects from different particle types. This includes snow with a mean fall speed of about −1.5 m/s and cloud droplets or small ice particle likely representing the surrounding upward air velocity of 1 m/s at 3 298 m height. Melting snow and small raindrops in the *melting layer* may explain the Doppler spectrum at 3 058 m; and rain that at 2 248 m height. Further information can be retrieved from higher order moments such as the skewness (third moment) and the kurtosis (forth moment, Kollias et al. (2011)).

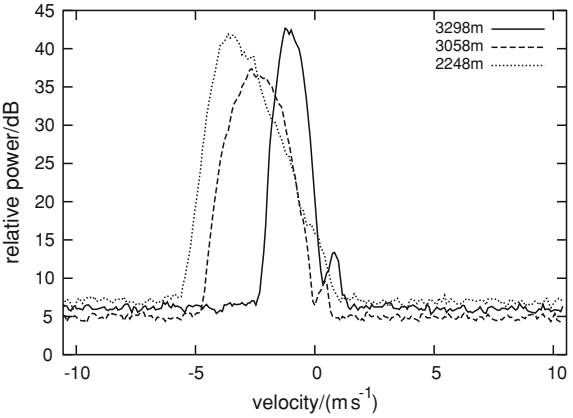

Fig. 21.5 Doppler spectra of power versus vertical velocity derived from a vertical pointing cloud radar, measurements at three altitudes above ground. 3 298 m: above the melting layer; 3 058 m: just beneath the melting layer; 2 248 m: inside rain

21.5 Summary and Outlook

Weather radar is the only system which allows three-dimensional insight into precipitation systems with high spatial and temporal resolution. Most knowledge about the dynamic and microphysical mechanism of deep convection has been gained through the application of polarimetric Doppler radar systems.

The synergy of different remote sensing measurement systems like radar, microwave *radiometer* and *lidar* will considerably help to increase insights into the microphysics of clouds and precipitation. Warning and nowcasting products gain from the combination of radar observations with *satellite* and lightning observations. Currently, the *airports* at Munich and Frankfurt/Main are equipped with a combined X-band radar and lidar system to provide all-weather monitoring of the wind field.

Weather radar research has a long tradition at IPA. The polarimetric Doppler radar POLDIRAD provided new insights in thunderstorm research and weather radar applications. Despite its age of more than 25 years, the radar is still a unique and powerful research instrument due to continuous upgrading of the system. The cloud radar at Schneefernerhaus will be part of the European observation network CLOUDNET (www.cloud-net.org). This network aims to provide measurements and retrieval techniques for cloud properties and to compare measurements with mesoscale numerical models. The radar and the other instruments qualify the station as a ground truth site for satellite missions like Cloudsat, Calipso, or the upcoming EarthCARE and GPM missions. The German research aircraft HALO will fly the HALO microwave package (HAMP) system, a combination of the MIRA-36 cloud radar and five microwave radiometers covering different frequency bands between 22 and 183 GHz. Together with a water vapor lidar, this instrumentation is a powerful tool for developing and validating existing and upcoming satellite missions.

Several trends are visible for future radar technology internationally. First, meteorology needs much faster updates of the observations than are possible with the conventional radar systems with parabolic dishes today. Updates of the volume scan are possible every 5 min presently; however, for tornado warnings, updates every 10–30 s would be desirable. Active phased-array antennas can considerably speed up the scanning and will allow for techniques like adaptive beam-forming or selective scanning of regions of interest. Second, there is further need to provide information at low altitude levels. Present operational weather radar networks include radars with 100–200 km separation. Their radar beam is often above low-level features at far ranges. Additional short-range X-band radar systems may fill the gaps. Third, it is now time to replace the inflexible and antiquated transmitter technology by up-to-date solid-state technology. Following the current development in weather radar technology, IPA, together with the DLR Microwaves and Radar Institute and industry, will consider a new radar system like a mobile X-band radar with a phased-array antenna.

References

Friedrich, K., Hagen, M.: Evaluation of wind vectors measured by a bistatic Doppler radar network. J. Atmos. Oceanic Technol. **21**, 1840–1854 (2004)

Hagen, M.: On the appearance of a cold front with a narrow rainband in the vicinity of the Alps. Meteor. Atmos. Phys. **48**, 231–248 (1992)

Hagen, M., van Baelen, J., Richard, E.: Influence of the wind profile on the initiation of convection in mountainous terrain. Quart. J. Roy. Meteor. Soc. **137**, 224–235 (2011)

Höller, H., Bringi, V.N., Hubbert, J., Hagen, M., Meischner, P.F.: Life cycle and precipitation formation in a hybrid-type hailstorm revealed by polarimetric and Doppler radar-measurements. J. Atmos. Sci. **51**, 2500–2522 (1994)

Kollias, P., Rémillard, J., Luke, E., Szyrmer, W.: Cloud radar Doppler spectra in drizzling stratiform clouds: 1. Forward modeling and remote sensing applications. J. Geophys. Res. **116**, D13201 (2011). doi:10.1029/2010JD015237

Meischner, P. (ed.): Weather Radar: principles and Advanced Applications, p. 337. Springer, Berlin (2004)

Meischner, P.F., Bringi, V.N., Heimann, D., Höller, H.: A squall line in Southern Germany: kinematics and precipitation formation as deduced by advanced polarimetric and Doppler radar measurements. Mon. Weather. Rev. **119**, 678–701 (1991)

Müller, H.G., Singler, A.: Radarbeobachtungen an konvektiven Niederschlägen im Voralpengebiet. Meteorol. Rdsch. **25**, 45–51 (1972)

Sauvageot, H.: Radar Meteorology, p. 384. Artech House, Boston (1992)

Schmidt, K., M. Hagen, H. Höller, E. Richard, Volkert, H.: Detailed flow, hydrometeor and lightning characteristics of an isolated thunderstorm during COPS. Submitted to Atmos. Chem. Phys. Discuss. 12, 9717–9769 (2012). doi:10.5194/acpd-12-9717-2012

Schroth, A.C., Chandra, M.S., Meischner, P.F.: A C band coherent polarimetric radar for propagation and cloud physics research. J. Atmos. Oceanic Technol. **5**, 803–822 (1988)

Weather Nowcasting and Short Term Forecasting

22

Arnold Tafferner and Caroline Forster

Abstract

This article is about present weather and its immediate development, on the challenge of how to observe it, and how to forecast it in the short term. It touches on the problems meteorologists have in delivering reliable estimates of, e.g., which path a thunderstorm will take during its track, whether it will bring hail or just rain, or when there will be freezing conditions at an airport with subsequent problems for air traffic on ground, arrival and departure. Some illustrative examples are given, showing how the problems are tackled and how integrated forecasting systems, in particular, can be successful in meeting the challenge.

22.1 Introduction

How will the weather be in the next few hours? This is a question often asked by people planning outdoor activities, by organizers of an outdoor event or sport activity, by farmers, by airport operators and flight regulators, by emergency and warning services and by many others for whom the weather in the very near future impacts their activities and decision making. Accurate and reliable *weather*

A. Tafferner (✉) · C. Forster
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Arnold.Tafferner@dlr.de

C. Forster
e-mail: Caroline.Forster@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_22,

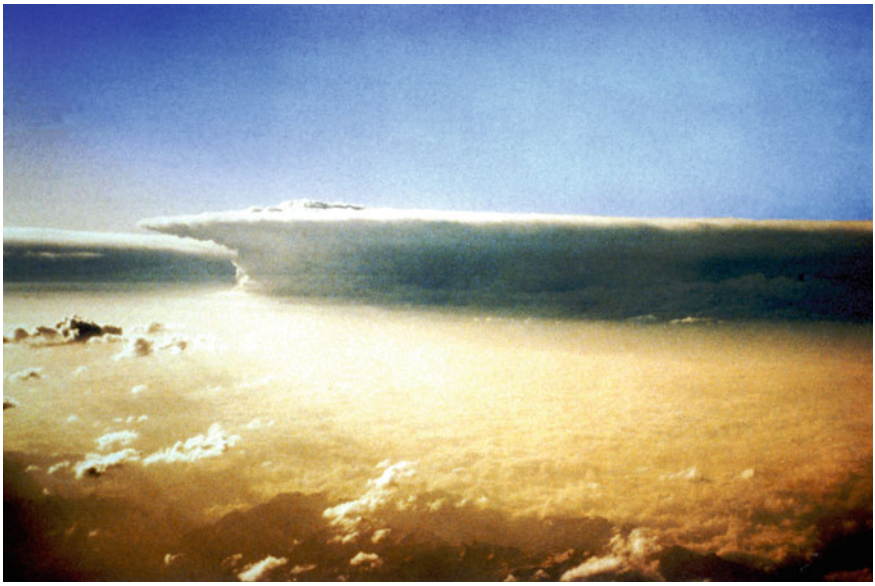

Fig. 22.1 Photograph of thunderstorm cloud over the northern Alpine forelands at 17 UTC 12 July 1984 taken from an airplane at 10 600 m near Innsbruck by the pilot, F. D. Oltmanns. Height of the cloud approximately 12 km, width of the anvil cloud 100 km, diameter of heavy precipitation core about 40 km, as estimated by the pilot from his onboard radar. An overshooting top (to the *left* of figure *center*) is apparent over the anvil cloud

forecast for this time range, referred to in meteorology as *nowcasting*, is especially needed in areas where safety, infrastructure protection, flood prevention, air pollution, energy provision and air and ground transport are concerned. Quite naturally, severe weather events like *thunderstorms* with associated weather phenomena like hail, lightning strokes, strong wind and tornadoes, frontal passages with wind gusts, and heavy precipitation leading to widespread flooding or localized flash floods, all of which can be quite destructive, are of prime interest to the public. Unfortunately, the occurrence, path, intensity, growth and decay of these weather phenomena are quite difficult to predict. Although many of these events have been observed during the past 25 years since modern technology like *polarimetric radar* (POLDIRAD) and satellite instruments (*Meteosat*-1, launched in 1977) have become available, no event is quite the same as another, which makes it impossible to estimate the evolution of an actual event from one in the past.

The development of the Munich hailstorm on 12 July 1984 (Fig. 22.1) which caused about 2 billion DM in economic loss could be traced back to its first development near Lake Constance. After an event, meteorologists can explain what happened when all available meteorological data have been collected and analyzed (Heimann and Kurz 1985). However, will an upshooting *cumulus* cloud near the same location on a hot summer day develop into another destructive storm

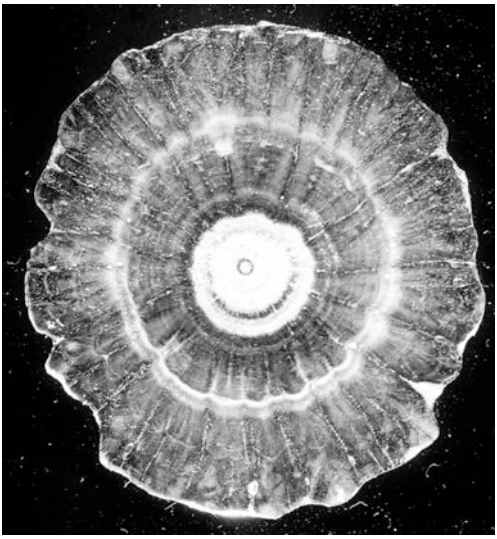

Fig. 22.2 Crosscut through a hailstone collected at IPA from a hailstorm on 12 July 1984, diameter 5 cm (photographed by A. Waldvogel, ETH Zürich)

of that kind, or will it simply evolve to an average thunderstorm? This is the challenge meteorologists have to deal with. The difficulties are a result of various circumstances. First of all, extreme weather events arise from a complicated interplay of various factors where local influences such as available moisture, solar insulation and orographic lift sometimes interact with the large scale environment, which has to be in a favorable condition to support the growth of these disturbances. Furthermore, extreme weather phenomena appear in all kind of shapes and sizes, in the case of thunderstorms from small-scale single cells with diameters of about 2 km up to multicell storm complexes 100 km in size. Their lifetimes vary from about 20 min for single cells up to several hours in the case of moving thunderstorm lines. In addition, these systems do not grow steadily in a linear sense from small to large: their development results from a sudden *instability* when a certain threshold has been exceeded. After initiation, cloud and precipitation processes set in which influence the dynamical storm structure, which again interacts with the environmental flow. In Fig. 22.1 one can notice a so-called overshooting top over the wide flat thunderstorm *anvil*, indicating a vigorous upward motion which penetrates into the *stratosphere*. Such vigorous developments can produce hail. In this case hailstones up to 9 cm in diameter have been observed, leading to fatalities, destroying house roofs and cars, and inflicting widespread damage to agriculture. Figure 22.2 shows a crosscut through a hailstone 5 cm in diameter collected at the DLR Institute of Atmospheric Physics (IPA). Obviously, the environmental flow and local conditions were favorable to producing an exceptional thunderstorm in this case (details in Heimann and Kurz (1985) and Höller and Reinhardt (1986)).

Besides thunderstorms, other weather phenomena can also be very difficult to forecast short term. Even in the case of quasi-steady flow, e.g., a large-scale *cyclone* residing over the Alps, it is difficult to predict which valleys will be hit by the moist airstreams embedded within the rotating system. A lateral shift of only 20 km of such a moist airstream will decide whether this or another valley will be flooded. Besides these highly dynamical systems, comparatively tranquil conditions can also be difficult to predict, e.g., the development and decay of fog, which is of prime interest at airports as it influences arrival and departure procedures. For winter weather conditions, it is again small changes in the parameters temperature and humidity which can turn a harmless cloud into a hazardous cloud with freezing precipitation, affecting road transport and aircraft in flight and requiring de-icing procedures at airports.

Considering the complex task nowcasting has to fulfill, the following definition provided by Conway (1998) is appropriate: "nowcasting is forecasting with local detail, by any method, over a period from the present to a few hours ahead; this includes a detailed description of the present weather." The latter is made possible by fully utilizing all observations available, in particular *remote sensing* data from radars, satellites, wind profilers, lightning observations, etc. (e.g., Browning (1982). Today, nowcasting is expanded to include the blending of extrapolation techniques, statistical techniques, heuristic techniques and *numerical weather prediction* (NWP). Here, heuristic is defined as forecast rules based on such things as experiment, numerical simulation, theory, expert systems, fuzzy logic and forecaster rules of thumb COST (2004).

From these definitions it is clear that there is no strict separation between nowcasting and short term forecasting. Grossly speaking, nowcasting is understood as extrapolation of observed weather features into the future, be it their location and/or intensity, whereas short term forecasting is reserved for forecasting with *numerical models*. Also, nowcasting encompasses the time range from now to the very near future, e.g. 1 h in the case of thunderstorms, whereas short term forecasting may range up to 12 h, depending on the weather phenomenon under consideration. It is also obvious that a successful nowcast of a weather event will best rely on a combination of data and algorithms in order to be able to cope with the complexity of the phenomenon (Tafferner et al. 2008). Depending on the required forecast length, a NWP model may contribute to the short term forecast.

Figure 22.3 illustrates the situation. Mere extrapolation of the present weather state cannot cope with quickly changing weather systems, especially at small scale, therefore nowcast accuracy drops rapidly already within the first hour. Recall that thunderstorms, e.g., have a typical lifetime of only about 30 min. Explicit storm models, i.e., NWP models with high grid resolution of about 1 km or less, together with data *assimilation* on the *mesoscale* have successfully demonstrated that they are able to simulate the evolution of storms. These models can augment the accuracy especially beyond 1 h forecast time and fill the gap between extrapolation and large scale models. For the immediate time up to about 60 min, extrapolation techniques are more useful since output from explicit storm models is delayed due to the time needed for data processing, assimilation and forecasting.

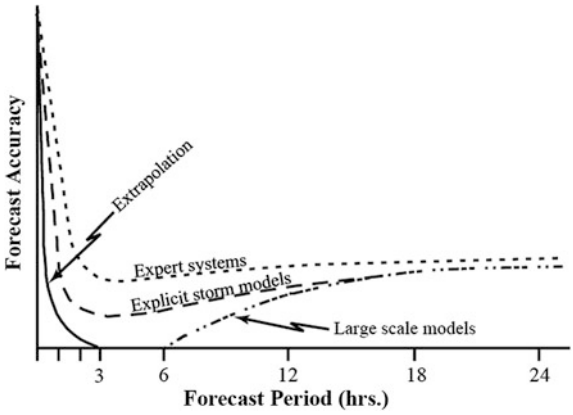

Fig. 22.3 Qualitative assessment of forecast accuracy as a function of forecast period for convective precipitation on a spatial scale of a few kilometers (Wilson et al. 1998)

Expert systems are expected to do best, as they combine various observation data, nowcasting algorithms and NWP model output in a sophisticated manner. In the following sections this will be explained with some real world examples.

22.2 Nowcasting Techniques

Four basically different nowcasting methods can be distinguished:

- Decision making based on a conceptual model.
- Extrapolation under steady state assumption or size trending.
- Numerical prediction.
- Predicting initiation/dissipation.

22.2.1 Conceptual Model

In meteorology, a common approach to understanding weather events is to analyze past observation data, to carry out *field campaigns* with special observation instruments where a weather phenomenon is most likely to happen, and to simulate events in a numerical model. With the help of basic thermodynamic principles it is then possible to derive conceptual models of the life cycle of these events (Browning 1982). During an event forecasters can then use such a conceptual model to estimate the short term evolution from actual observation data.

22.2.2 Extrapolation

Extrapolation is useful when a weather phenomenon has already developed; then the change in time allows estimation of the future path and development. It can be performed by anybody, e.g., by inspection of a sequence of radar images. Put

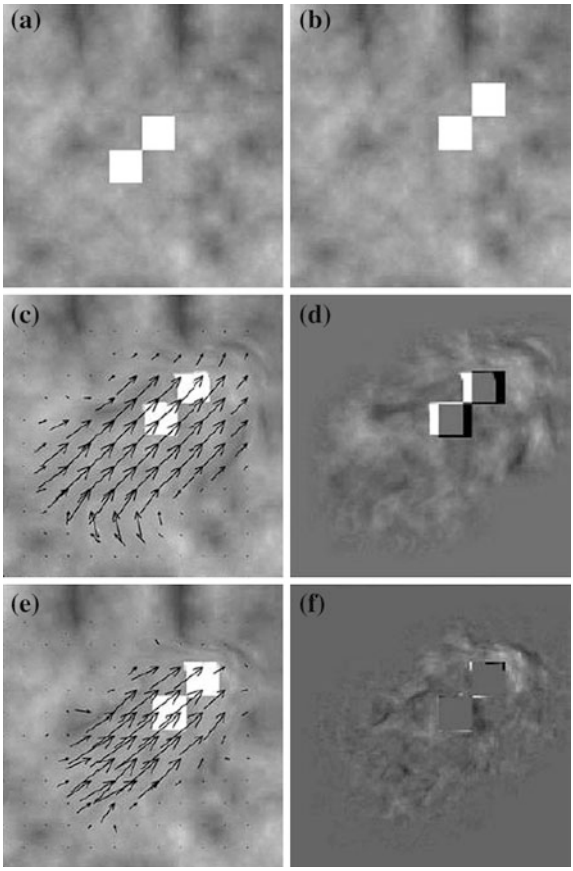

Fig. 22.4 Demonstration of the image matching procedure for use in monitoring and nowcasting. Images **a** and **b** differ with respect to the position of two white squares which represent image features. The vectors define the displacement required for best correlation between successive images. After calculating the vectors for coarse displacements (**c**), an approximate match (**d**) is obtained. The fine-scale step (**e**) results in an almost perfect match of the squares (**f**). (Zinner et al. 2008)

together as an animation, one can visually extrapolate the movement of individual image features for the very near future. This is actually done by forecasters when they give advice to customers, e.g., hydrological services. However, this process can be automated and algorithms have been developed which correlate individual image features between consecutive images and calculate future positions. At IPA a so-called pyramidal image matcher has been developed (Mannstein et al. 1999) which transforms one image into another. In a stepwise procedure which first matches large image features and then smaller detail later on, displacement vectors for every image pixel are obtained. When applied to one of the images, these vectors distort it in such a way as to achieve the best possible match with the other image. Figure 22.4 illustrates this process. The images a and b differ with respect to the position of two white squares which represent image features. The task of the matching procedure is to displace the squares in the left image to bring them into the position seen in the right image. Note that the matching procedure does not recognize the features as squares; it only handles values at every image pixel. In a first step, both images are smoothed to a coarse image resolution to evaluate

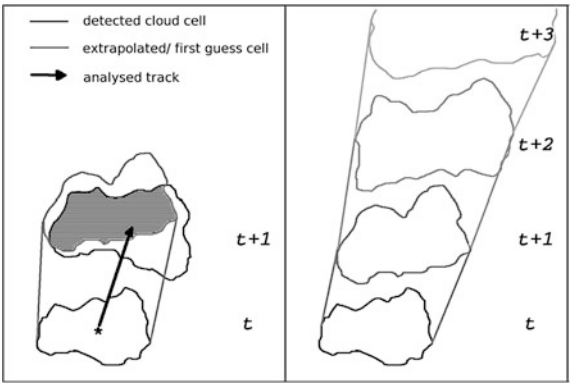

Fig. 22.5 Depiction of cell tracking (*left*) and cell nowcasting (*right*). A detected cloud cell at time t is extrapolated to a first guess position at time t + 1 (arbitrary unit). It overlaps with the detected cell at time t + 1 resulting in a track vector, as indicated by the black arrow. A detected cell at time t (black contour, *right*) is extrapolated for three consecutive time steps using displacement vectors from the image matching procedure (see Fig. 22.4), thereby size and form of the nowcasted cells can change. (adapted from Zinner et al. (2008))

large scale differences. Displacement vectors for every pixel are calculated which provide the maximum correlation among the pixels of both images for the smoothed image. Figure c shows the smoothed image a overlaid with the displacement vectors. Note that only 1 out of 100 vectors is displayed to aid in visibility. After displacement according to these vectors an approximate match (d) is obtained. The following matching step for the fine-scale differences (without smoothing) results in revised displacement vectors (e) which yield an almost perfect match (f). For details see Zinner et al. (2008).

The displacement vectors can be used for various purposes. One application is the tracking of image features in the past, another is the extrapolation of image features to a future time, see Fig. 22.5. On the left, a detected cell at time t, which in the real world could be a cloud cell or a precipitation cell, is extrapolated by use of the displacement vectors to yield a first-guess position and shape (grey contours) at time t + 1. At time t + 1, the observed cell overlaps with the first-guess cell. Therefore, a vector can be drawn connecting the two detected cell centers. In the case of more than one overlapping cell at time t + 1, the one with the maximum overlap is chosen for the tracking. On the right, nowcasted cells are depicted for three consecutive time steps. Note that the cell not only changes its position, but also its size and shape. This is due to the fact that displacement vectors act on every pixel within a cell (not drawn). The image matching procedure has been applied for tracking and nowcasting thunderstorm cells in satellite and radar images. Figure 22.6 shows an example of thunderstorm nowcasting using the radar tracking and *monitoring* algorithm Rad-TRAM (Kober and Tafferner 2009) for the situation around Munich airport on 21 May 2009. Using consecutive radar *reflectivity* composite images, cells of heavy precipitation are detected by applying a

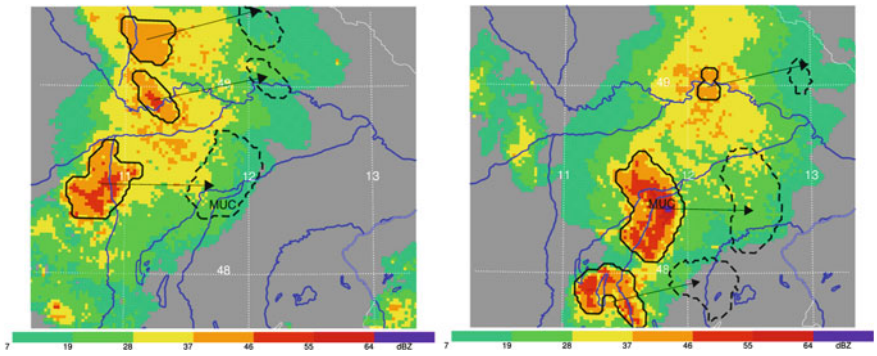

Fig. 22.6 Nowcasting example with Rad-TRAM radar reflectivity data on a map of a 300 × 300 km^2 region near Munich airport (MUC) with major rivers marked in *blue*. Detected thunderstorm cells are identified by black contours overlaid on radar data (*color shading*) at 16:15 UTC 21 May 2009 (*left*) and 17:15 UTC (*right*). Nowcasted cell positions 1 h after the observation are shown by dashed contours. *Arrows* point from actual to nowcasted cell positions

threshold of 37 dBZ, which has been found to fit best to thunderstorm cells. At 16 UTC (Fig. 22.6 *left*) three cells are marked as black contours overlaid on the radar image. Also displayed are the nowcasted cells after 60 min to the right of these cells and past tracks for the centers of two cells. Note that the nowcast of one of the cells indicates that the airport will be hit. One hour later (Fig. 22.6 *right*) this was indeed the case as the detected cell at 'MUC' shows. Note that the position and size of the nowcasted cell fit quite well to the observed cell.

For satellite imagery the thunderstorm tracking and monitoring system Cb-TRAM has been developed (Zinner et al. 2008). Cb-TRAM is a fully automated tracking and nowcasting algorithm. Intense convective cells are detected, tracked and discriminated with respect to onset, rapid development, and mature phase. In addition, short range forecasts are provided. The detection is based on a combination of Meteosat *SEVIRI* (Spinning Enhanced Visible and Infra-Red Imager) data from the broadband high resolution visible (HRV), the infrared 6.2 μm (water vapor), and the infrared 10.8 and 12.0 μm channels. The detected thunderstorm objects contain thunderstorm location (polygons), nowcast contours and some additional parameters, e.g., cell center, cloud top temperature and *trend*.

Figure 22.7 shows an application for the aircraft incident of 1 October 2009 when an Airbus A319 on its way from Forli to Palermo flew into a thunderstorm cell which caused hail damage to the radome (Fig. 22.8). The satellite image over part of the western Mediterranean is overlaid with detected and nowcast thunderstorm cells. Marked also are lightning observations from the LINET network and the path of the aircraft. Cb-TRAM nowcasts the propagation of one thunderstorm cell to the east crossing the flight path north of Palermo. If this information had been provided to the pilot in real time before his aircraft approached Palermo it could have made him aware of the hazardous situation well before the hailstorm hit his aircraft.

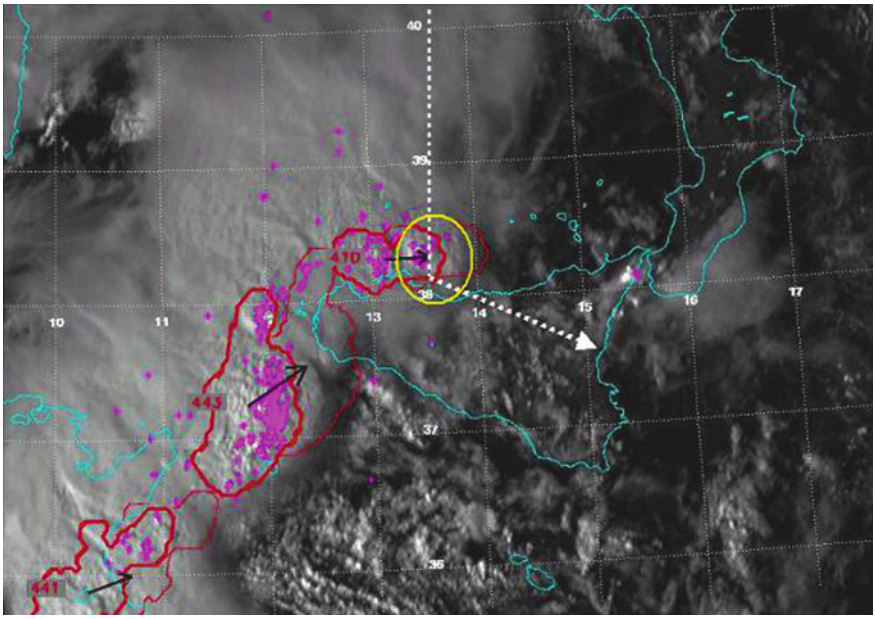

Fig. 22.7 Nowcasting example with Cb-TRAM. High resolution visible Meteosat satellite image overlaid with Cb-TRAM contours. Detected cells *bold red*, nowcasted *thin red*. *Dotted white lines* indicate the estimated flight path of an airbus A319 on its way from Forli (northern Italy) to Palermo and further on to Catania on 1 October 2009. The region where the aircraft encountered hail in the thunderstorm is marked by the *yellow circle* (Tafferner et al. 2010).

Fig. 22.8 Hail damage in the radome of the A319 after landing (photo: ATRDRIVER)

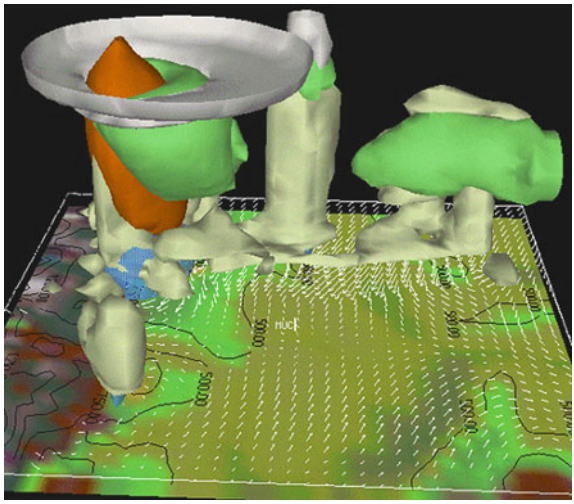

Fig. 22.9 Numerical simulation of a thunderstorm line propagating through the Alpine forelands on 12 August 2004. Perspective view from east to west of three-dimensional hydrometeor concentration simulation results at 19 UTC. Alps to the *left*, MUC in the *center*, approximate domain size 250 × 250 km. Different colors of the convective cells distinguish different hydrometeors: rain (*blue*), cloud water (*white*), graupel (*orange*), snow (*green*), cloud ice (*silver*). The vectors indicate the wind field at 800 m above sea level (Tafferner et al. 2008)

Both applications, Rad-TRAM as well as Cb-TRAM, were operated in real time at IPA during two summer campaigns at Munich airport in the years 2010 and 2011. Output in the form of weather objects has been provided in real time to aviation customers including Deutsche Flugsicherung (DFS), Lufthansa Hub Control Center, and Munich Operations.

22.2.3 Numerical Prediction

Numerical prediction at high resolution has the advantage that all physical processes relevant to a certain weather phenomenon are taken into account. These include turbulent mixing within the *boundary layer*, cloud and precipitation processes, *radiation* and *advection*. Figure 22.9 shows a result of a simulation of a thunderstorm passage through the Alpine forelands with the MM5 model. In the perspective 3D-view of *hydrometeors* the convective cells are clearly recognizable, especially the fully developed thunderstorm at the Alpine baseline, which was indeed observed at about that time and location. Note the anvil shaped top which is visible in the cloud ice distribution. In the wind field at 800 m above sea level (about 200 m above ground), the approaching gust *front* ahead of the convective cells is apparent. Despite many successful simulations, mesoscale weather features 10–100 km in size are often not predicted, or predicted at wrong locations or times for the reasons mentioned in the introduction. Another challenge is to

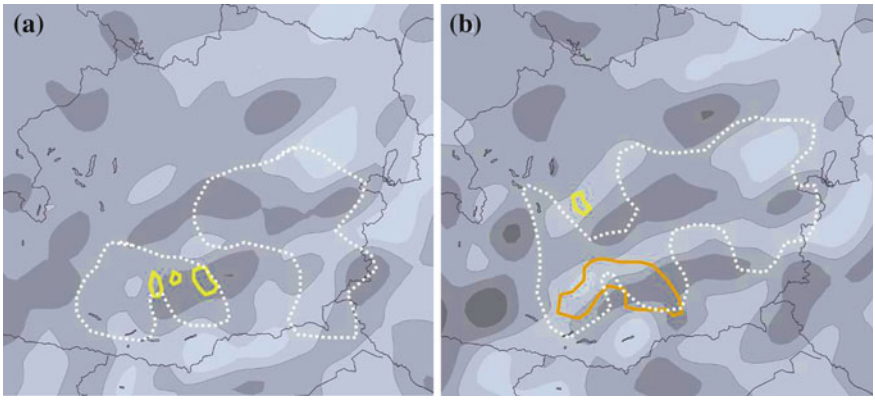

Fig. 22.10 Convective initiation over eastern Austria on 25 May 2009 at 11 UTC (**a**) and 12 UTC (**b**). *Dark shaded* areas indicate upward air motion (*darker* = *stronger*) as forecast by the COSMO-EU model at 500 hPa level. *White dashed* contours enclose regions of moisture flux convergence at the ground as analyzed by the VERA system. *Yellow* contours mark initial convective cells as diagnosed by Cb-TRAM; the *orange* contour marks the location of a rapidly developing thunderstorm cell (**b**)

capture the life cycle of the phenomenon, i.e., initiation, development with correct intensity, and decay. However, numerical forecasts on this scale can be useful. Instead of demanding successful direct simulation of every detail of a phenomenon, derived indicators such as areas of convective precipitation above a certain threshold or output quantities like thunderstorm indices can give guidance to where and when a weather feature is most likely to happen. Furthermore, instead of relying on only one forecast, a multitude of forecasts can be run, a so-called ensemble, with perturbed initial and boundary conditions or different model physics, with the aim of achieving a probability distribution of the events under consideration. Similarly, a so-called time-lagged *ensemble* can be built from model runs starting from different times. At IPA such an approach is followed with COSMO-MUC, where a subdomain of the COSMO-DE model (Baldauf et al. 2011) is centered over Munich airport. Besides the smaller domain, COSMO-MUC is started hourly instead of three-hourly, which has the advantage that local data like radar and Aircraft Meteorological Data Relay (AMDAR) reports are assimilated more frequently, with the benefit of better adjustment to local changes (Dengler et al. 2011).

22.2.4 Predicting Initiation/Dissipation

Predicting the initiation of a weather event can be a quite difficult task. As stated above, certain weather phenomena, in particular those leading to extreme events, do not grow slowly and steadily from a small perturbation, but often rapidly after the release of an atmospheric instability. For instance, a small cumulus cloud can

grow into a mature thunderstorm cell within 10 min after it has reached the level of free *convection*. Therefore, forecasters attempting to predict the initiation of a process do not rely on one data source alone, but inspect various information sources, e.g., satellite imagery, radar imagery, local measurements, NWP output, and output from special analysis systems in particular. An example is provided in Fig. 22.10. Within a region in eastern Austria several data sources indicate the possible development of a convective cell (Fig. 22.10a). The surface analysis of the Vienna Enhanced Resolution Analysis tool (VERA), Steinacker et al. (1997), indicates convergence of moist air on the ground (inside the white dashed contours) exactly at the location where the Cb-TRAM analysis also shows initial cloud cells (solid contours). It can be expected that these convective cells will grow further as there is lift of the air column from above, as forecasted by the COSMO model (dark shaded areas). An unstable air mass is prevailing over the area, as indicated by the vertical profile of equivalent *potential temperature* (not shown). This implies that all the necessary ingredients for the occurrence of *deep convection*, i.e., low level moisture, instability and forcing, have come together in one place. One hour later a rapidly developing thunderstorm cell appeared in that location (solid contour in Fig. 22.10b) which later developed into a mature thunderstorm. At IPA, an algorithm is currently under development for combining the data described above in order to automatically predict convection initiation (Stich et al. 2011).

Sometimes the decay of a certain weather feature is also of interest. For airports there is demand for exact forecasts of when a weather phenomenon will disappear, e.g., a thunderstorm or low level fog. Here too, a combination of data, e.g., from radar and lightning observations and a life cycle model, can indicate when a weather feature starts to decay (Bretl 2010).

22.3 Integrated Systems for Nowcasting and Short Term Forecasting

The nowcasting techniques described above have been designed individually for particular purposes independent of one another. Combining these within one integrated system will provide a greatly enhanced benefit in monitoring and nowcasting capability, i.e., an integrated system can process and contrast the assertions of the individual tools regarding, for example, the exact location of a particular weather system as well as its intensity and movement. The result is more reliable prediction of the future state of a weather system than if only one data source or nowcasting tool were used (Tafferner et al. 2008). Certainly, in order to be useful an integrated system for nowcasting and short term forecasting must be constructed in a way that processes large data amounts within a very short time. The prime aim of such a system is to reduce the complexity of weather to a description of the event that supports users in decision making. This requirement is quite obvious in extreme weather situations, where individual weather parameters

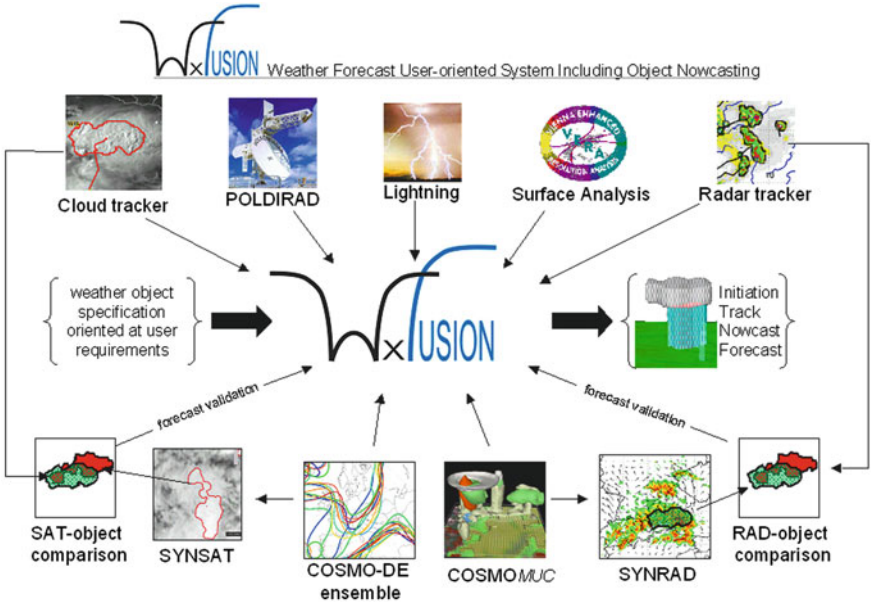

Fig. 22.11 Integrated system WxFUSION. Observation data and output from nowcasting tools (*top row* elements) is combined with output from numerical models (*bottom row*) in order to track, nowcast and forecast a specific weather object (adapted from Tafferner et al. (2008) and Forster and Tafferner (2009))

change quickly and not much time is available to analyze a complex and large data set visually or by hand. Therefore, there is a need for automation in order to fulfill this requirement. Several integrated systems automatically combining different data sources and providing *weather hazard* information especially tailored to the needs of decision makers at airports have been developed and successfully tested in the USA. Examples are the Integrated Terminal Weather System (ITWS) from the Massachusetts Institute of Technology (Evans and Ducot 1994) and the Auto-Nowcast System developed at NCAR (Mueller et al. 2003).

At IPA the integrated system WxFUSION is under development. The acronym is read as "weather fusion" and stands for "Weather Forecast User-oriented System Including Object Nowcasting". The aim of the system is to integrate all available data to describe as well as possible the current and future state of a weather phenomenon. In particular, nowcasting data will be combined with short term forecasting data from numerical model output in order to cope with the poor quality of extrapolation for longer time ranges.

Figure 22.11 illustrates the system, with various data sources and system components represented by symbols. The elements in the top row are the previously mentioned cloud tracker Cb-TRAM, the radar tracker Rad-TRAM, and the VERA analysis. In addition, polarimetric radar (POLDIRAD) and lightning data from the LINET network operated by nowcast GmbH are available. Together these

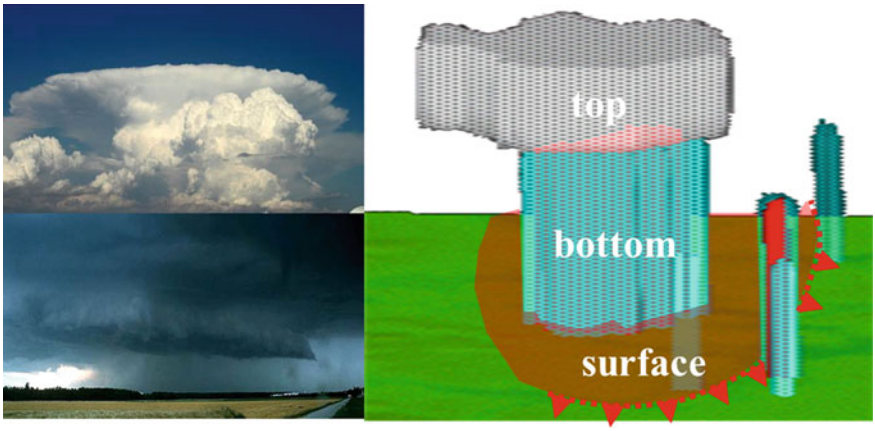

Fig. 22.12 *Left* actual thunderstorm as seen by an observer (*upper* photograph by M. Köhler). *Right* thunderstorm weather objects constructed from nowcasting tools (Tafferner et al. 2012)

elements compose the observation and nowcasting part of WxFUSION. The bottom row of elements comprises numerical model forecasts and forecast validation components (object comparison). Forecast data are available from the COSMO-DE ensemble and from the high-resolution time-lagged ensemble COSMO-MUC. From these model outputs, synthetic satellite (SYNSAT) and radar (SYNRAD) images are generated which are used as input to Cb-TRAM and Rad-TRAM. Synthetic objects are thus obtained which can be compared to their observed counterparts (object comparison). The best match determines which members from the ensemble will be used for fusion with nowcast data.

A weather object specification oriented to user requirements has to be provided as further input to WxFUSION. As already mentioned above, users of weather forecasts need this information in a form useful for decision making. Users do not want, and often do not have, the background knowledge to interpret a multitude of complex meteorological data. Instead, they need this information in a reduced form that is easy to understand and tailored to their needs. Figure 22.12 makes this reduction of complexity visible. On the left side two photographs of two different thunderstorms are composited in order to make typical weather features visible in one image. In the upper part there is a fully developed cloud anvil typical of mature thunderstorms together with a new convective cloud developing in front of it. In the lower part, the typical cloud wall in front indicates the forced ingestion of moist air into the storm; in the rear heavy precipitation with corresponding downburst can be seen. Correspondingly, using Cb-TRAM, which detects the upper part, and using Rad-TRAM, which detects the bottom part, a composite thunderstorm object can be constructed as shown on the right. Also indicated is a surface object describing the *squall line* generated from the thunderstorm outflow. Therefore, the information about the thunderstorm given to users has a top and a bottom part, each of which consists of a polygon with bottom and top, and some

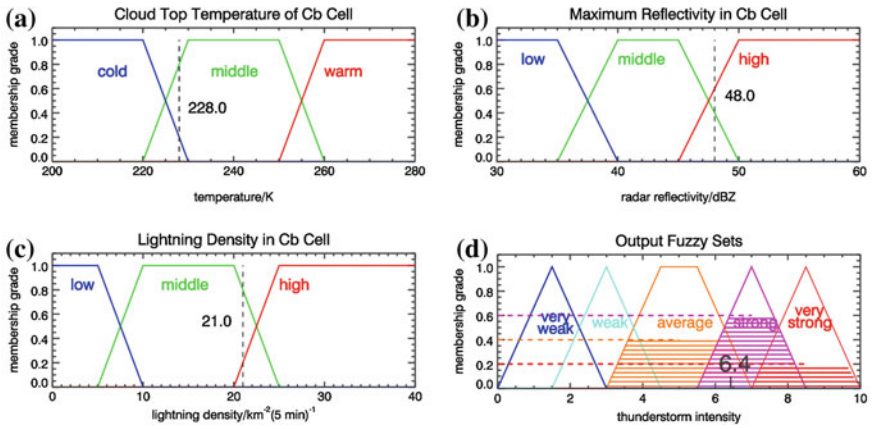

Fig. 22.13 Fuzzy logic procedure with input fuzzy sets (**a**–**c**) and output fuzzy sets (**d**). The lines indicate thunderstorm properties for an example with low, middle and high intensity. *Vertical dashed* lines in (**a**–**c**) denote observed values for cloud top temperature, radar reflectivity and lightning flash-frequency density, respectively. Output fuzzy sets for five different thunderstorm intensity classes (on an intensity scale from 0 to 10) from very weak to very strong (**d**). The shaded areas indicate the contribution of the input parameters to the different classes which after evaluation achieves a thunderstorm intensity of 6.4 in this example. (Tafferner et al. 2012)

descriptive parameters such as intensity, trend, rain rate, etc. which are deduced from other observation data.

The heart of WxFUSION is the fusion module (center of Fig. 22.11). Its task is to retrieve from all available nowcast and forecast tools those data which are needed to calculate the required parameters of the weather object on output. For example, a thunderstorm weather object on output (symbolized on the right in Fig. 22.11) should contain information on initiation, actual state as detected, nowcast up to 1 h and forecast for later hours based on model forecasts.

The *data fusion* itself is undertaken employing fuzzy logic. The fuzzy logic procedure mimics the data analysis of an experienced forecaster. By evaluating and combining the outputs of nowcasting procedures and numerical prediction together with his knowledge about thunderstorm development (conceptual model) he composes a picture in his mind of what is going to happen. In contrast to decision trees which build on true/false branches when information contents are evaluated, fuzzy logic is a decision finding technique that allows a gradual transition between true and false, i.e., it deals with parameter ranges instead of fixed thresholds. Every contribution can be weighted with regard to its importance for the weather phenomenon by using so-called fuzzy functions. A final decision is then determined through an adequate combination of all weighted contributions. As an example, the task to estimate the intensity of a thunderstorm can be considered. Cloud top temperature, radar reflectivity and lightning density are chosen as contributing elements. Figure 22.13 shows an example of the fuzzy logic

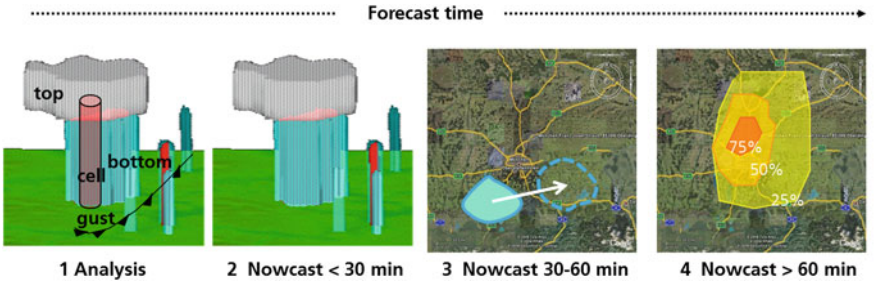

Fig. 22.14 Nowcasting: decrease of descriptive detail over forecast time (Tafferner et al. 2012)

evaluation. The influence (weight) of all data is taken into account by appropriately chosen fuzzy functions for low, middle or high membership grade. An observed value, e.g., a cloud top temperature of 228 K as shown by the dashed line in the top diagram (Fig. 22.13a), with a membership grade of 0.8 belongs to "cold" and with a grade of 0.2 to "middle." The other parameters are evaluated similarly. After collecting and evaluating all possible data combinations, a thunderstorm intensity value of 6.4 on a scale from 0 to 10 is found from the output fuzzy sets, which are composed of five different thunderstorm intensity classes from very light to very strong (Fig. 22.13b). In conclusion, the thunderstorm falls within the class "strong."

The calculated intensity is included as a descriptive parameter in the weather object together with output from the already mentioned nowcasting tools. Further detail such as trend and lightning density will be included in the fuzzy logic procedure by including recent research results (Bretl 2010; Meyer 2010).

For estimating future development, only output from the nowcasting tools and numerical model output is available. It therefore makes sense to relax the detailed description of a weather phenomenon based on observation data to a more probabilistic estimate of how the event will evolve. This is illustrated schematically in Fig. 22.14. Whereas a three-dimensional description of a thunderstorm is possible at analysis time together with small scale additional features, like the position of a gust front, after 30 min only the location of the thunderstorm is nowcast. Beyond 1 h time, this is relaxed to a probability of occurrence. Considering the typical lifetime of thunderstorms, the uncertainties in numerical forecasting and the experience gained, it would not make sense to forecast thunderstorms deterministically beyond 1 h. Work is under way to combine output of nowcasting tools with numerical forecasts (Köhler 2011) in order to increase forecast accuracy, as indicated in Fig. 22.3 (expert system). Forecast evaluation plays an important role in this effort (Forster and Tafferner 2009).

WxFUSION can also be extended to handle winter weather conditions. Of particular importance is the occurrence of in-flight and ground icing conditions at airports. Here, data from surface observations of precipitation type and intensity, surface conditions (dry, liquid, frozen), hydrometeor distributions within clouds observed by POLDIRAD, and aircraft observations of temperature and humidity

can be fused using fuzzy logic as well. With respect to in-flight icing, a related system named Advanced Diagnosis and Warning System for Aircraft Icing Environments (ADWICE) has been developed (Tafferner et al. 2003; Leifeld 2004) and is running operationally at the German weather service Deutscher Wetterdienst (DWD). Work is underway to combine ADWICE algorithms with the additional data sources available within WxFUSION to build a corresponding winter weather module.

Of course, there are also elsewhere and worldwide efforts underway to nowcast and short term forecast not only thunderstorms but also other weather phenomena, such as wind storms (Föhn), *clear-air turbulence*, flash floods, environmental pollution, dispersion of *aerosols*, and winter weather conditions, each of which deserves its own nowcasting approach. Prominent centers of study are the NCAR Research Application Laboratory in Boulder, Colorado and the Massachusetts Institute of Technology (both USA), the MetOffice in Exeter (UK), Météo France in Toulouse, and DWD in Offenbach, to name just a few. Data fusion techniques similar to the one presented above are being developed at these centers.

22.4 Concluding Remarks

The examples presented in this article have been restricted to thunderstorms and corresponding research carried out at IPA. It is apparent that both nowcasting techniques and expert systems will continue to be developed and improved. Great impetus can be expected from the availability of ensemble models in the meso-scale, especially for providing more accurate estimates on the short term of probabilities of occurrence of the mentioned weather phenomena. In the foreseeable future it is most likely that deterministic and probabilistic nowcasts in the 0–6 h time period can be provided to users with sufficient accuracy in space and time in order to mitigate losses of life and property and to enhance benefits.

Acknowledgments The work presented here was funded by the DLR project Wetter & Fliegen (2008–2012) and benefited from the EU projects RISK-AWARE, FLYSAFE. Thanks are due to nowcast GmbH for providing lightning observations, and to the Department of Meteorology at Vienna University for the VERA system.

References

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., Reinhardt, T.: Operational convective-scale numerical Weather prediction with the COSMO model: description and sensitivities. Mon. Weather Rev. **139**, 3887–3905 (2011)

Bretl, S.: Untersuchung des lebenszyklus von gewittern in mitteleuropa mit hilfe von fernerkundungs- und modelldaten. Diplomarbeit, Ludwig-Maximilians-Universität München, München (2010)

Browning, K. A.: Nowcasting, p. 256. Academic Press, London (1982)

Conway, B. J.: An overview of nowcasting techniques. In: SAF Training Workshop—Nowcasting and Very Short Range Forecasting, EUMETSAT (1998)

COST 722, 2004: WWRP/NWP Expert Meeting, FMI, Helsinki, 9–11 June 2004
Dengler, K., Anger, J., Keil, C.: Validation of time-lagged ensemble forecasts relevant for predicting aircraft wake vortices. Meteorol. Z. **20**, 625–634 (2011)
Evans, J., Ducot, E.: The integrated terminal weather system (ITWS). MIT Lincoln Laboratory J. **7**(2), 449 (1994)
Forster, C., Tafferner, A.: An integrated user-oriented weather forecast system for air traffic using real-time observations and model data. In: Proceedings of the European Air and Space Conference (CEAS), Manchester, UK, 26–29 October 2009
Heimann, D., Kurz, M.: The Munich Hailstorm of July 12, 1984–a discussion of the synoptic situation. Beitr. Phys. Atmos. **58**, 528–544 (1985)
Höller, H., Reinhardt, M.E.: The Munich Hailstorm of July 12, 1984–convective development and preliminary hailstone analysis. Beitr. Phys. Atmos. **59**(1), 1–12 (1986)
Kober, K., Tafferner, A.: Tracking and nowcasting of convective cells using remote sensing data from radar and satellite. Meteorol. Z. **1**(18), 075–084 (2009). doi:10.1127/0941-2948/2009/359
Köhler, M.: Untersuchung der Auslösung von Gewittern während der "Wetter und Fliegen" Sommerkampagne 2010. Masterarbeit, DLR/IPA (2011)
Leifeld, C.: Weiterentwicklung des Nowcastingsystems ADWICE zur Erkennung vereisungsgefährdeter Lufträume, Berichte des Deutschen Wetterdienstes, Offenbach am Main, vol. 224, pp. 118 (2004)
Mannstein, H., Meyer, R., Wendling, P.: Operational detection of contrails from NOAA-AVHRR-Data. Int. J. Rem. Sensing **20**, 1641–1660 (1999)
Meyer, V.: Thunderstorm tracking and monitoring on the basis of three dimensional lightning data and conventional and polarimetric radar data. Dissertation, Ludwig-Maximilians-Universität München, Mai (2010)
Mueller, C., Saxen, T., Roberts, R., Wilson, J., Betancourt, T., Dettling, S., Oien, N., Yee, J.: NCAR auto-nowcast system. Weather Forecast **18**(4), 545–561 (2003)
Steinacker, R., Pöttschacher, W., Dorninger, M.: Enhanced resolution analysis of the atmosphere over the Alps using the fingerprint technique. Annalen der Meteorologie **35**, 235–237 (1997)
Stich, D., Forster, C., Zinner, T., Tafferner, A.: Convection initiation—nowcasting by data fusion and its verification.In: European Conferences on Severe Storms (ECSS 2011), Palma de Mallorca, Balearic Islands, Spain, 3–7 October 2011
Tafferner, A., Hauf, T., Leifeld, C., Hafner, T., Leykauf, H., Voigt, U.: ADWICE—advanced diagnosis and warning system for aircraft icing environments. Weather Forecast **18**(2), 184–203 (2003)
Tafferner, A., Forster, C., Hagen, M., Keil, C., Zinner, T., Volkert, H.: Development and propagation of severe thunderstorms in the Upper Danube catchment area: towards an integrated nowcasting and forecasting system using real-time data and high-resolution simulations. Meteorol. Atmos. Phys. **101**, 211–227 (2008). doi:10.1007/s00703-008-0322-7
Tafferner, A., Forster, C., Hagen, M., Hauf, T., Lunnon, B., Mirza, A., Guillou, Y., Zinner, T.: Improved thunderstorm weather information for pilots through ground and satellite based observing systems. In: 14th conference on Aviation, Range, and Aerospace Meteorology, 90th AMS Annual Meeting, Atlanta, 17–21 January 2010
Tafferner, A., Forster, C., Gerz, T.: Concatenating weather monitoring and forecast: the WxFUSION concept. In: Gerz, T., Schwarz, C. (eds.) The DLR Project Wetter & Fliegen. Final Report DLR-FB 2012-02, pp. 25–30, 2012
Wilson, J.W., Crook, N.A., Mueller, C.K., Sun, J., Dixon, M.: Nowcasting thunderstorms: a status report. Bull. Am. Meteorol. Soc. **79**(10), 2079–2093 (1998)
Zinner, T., Mannstein, H., Tafferner, A.: Cb-TRAM: tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data. Meteorol. Atmos. Phys. **101**, (2008). doi:10.1007/s00703-008-0290-y

Lagrangian Modeling of Transport and Dispersion of Air Constituents

23

Robert Baumann and Hans Schlager

Abstract

Numerical models simulating the transport and dispersion of gaseous or particulate pollutants support the understanding of long-distance transport processes and are a valuable tool in various scientific and civil applications. Here the basic principles of Lagrangian trajectory and dispersion models and the major contributing physical processes are outlined: advection, dispersion, deposition, wash-out and settling and how they are implemented in the case of the HYSPLIT model, which is described here as a typical representative of this model class. Three recent applications of the model follow.

23.1 Introduction

Lagrangian trajectory and *dispersion models* such as Hybrid Single Particle Lagrangian Integrated Trajectory Model (*HYSPLIT*) (Draxler and Hess 1997; 1998), *FLEXPART* (Stohl et al. 2005) and *LAGRANTO* (Wernli and Davies 1997) gained importance in the past years as forecast and planning tools for scientific aircraft flight missions as well as for the analysis of airborne or ground-based *trace gas* or *aerosol* measurements. The term Lagrangian means that the physics of the

R. Baumann (✉) · H. Schlager
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Robert.Baumann@dlr.de

H. Schlager
e-mail: Hans.Schlager@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_23,

dispersion (or any other process modeled) is carried out in a reference frame moving along with the trajectory, i.e., as it would be seen by an observer moving in a balloon. Hence *advection* and dispersion can be calculated independently. In contrast to the Lagrangian approach, an *Eulerian model* considers all physical processes as exchanges of properties (e.g., mass, energy, momentum) between neighboring cells of a three-dimensional grid. In the Eulerian approach the processes are formulated in an Earth-fixed coordinate system as viewed by a resting observer. However, Eulerian models tend to blur sharp gradients of mass concentrations too fast unless a very fine resolution grid is used, implying, however, an increased demand for computational resources.

At the Institute of Atmospheric Physics (IPA), Lagrangian models have been used, for example, to predict the transport of *volcanic ashes* and gases from the *Eyjafjallajökull* (in 2010), Etna (in 2011), and Stromboli (in 2011) emissions in order to sample the pollutant plume with the *research aircraft Falcon* downwind even when it was not visible. In Lagrangian experiments the models have been used to plan the optimal time and location of a tracer release [e.g., in the experiments RECONCILE 2010 and Stratospheric Ozone: Halogen Impacts in a Varying Atmosphere (SHIVA) 2011]. In this kind of experiment a trace substance with extremely low natural background concentration, e.g., perfluromethylcyclopentane, C_6F_{12} (PMCP) is released either from ground or from an aircraft in order to mark an air mass, which is then sampled multiple times within the next few hours or days. As a post-mission analysis tool, dispersion models can be used to compare measured concentrations with expected concentration when the *source* strength is known, or to estimate the source strength when only the position of the source is known. The models can also be operated in a backward-in-time mode, meaning traveling back from the location of measurement to the location of the source. In this case the backward-in-time trajectories are usually folded with known *anthropogenic* or natural *emission* inventories of the pollutant species in order to determine the most probable emission source area. Besides their scientific use, these models also play a role as a tool in civil emergency management in cases of nuclear accidents, chemical accidents, or *wildfires*; for air quality management; or to re-route *air traffic* during major volcanic *eruptions*.

23.2 Transport Model Types

A trajectory model simulates the transport of gases or aerosol particles by following the local *wind* like an observer in a balloon would do starting at the location of the source (advection). The trajectory is constructed piecewise from short sections representing the air transport within an appropriately small time step by simply multiplying the time step with the wind velocity at the position and time of each section. The wind field itself is not part of the trajectory model but has to be provided as input, represented as a three-dimensional vector (u, v, w) on a four-dimensional coordinate grid in space (x, y, z) and time t. Usually the wind data

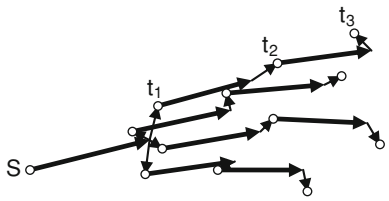

Fig. 23.1 Illustration of the principle of dispersion simulation. Shown is the path of four individual particles starting at a common point S during three time steps: the *thicker*, *longer arrows* represent the advection with the (local) mean wind. The *thinner*, *shorter arrows* depict the random dispersion steps

comes from any operational meteorological forecast system, e.g., *European Centre for Medium-Range Weather Forecasts* (ECMWF) or Global Forecast System (GFS). Such a trajectory can be considered "deterministic" with respect to the underlying wind field. Trajectories are reversible and can be calculated both forward-in-time from the source downwind as well as backward-in-time from the receptor upwind.

A Lagrangian dispersion model is based on a trajectory model and additionally simulates the effect of *dilution* by (mostly *turbulence* driven) mixing with the ambient air. Two basic "flavors" of Lagrangian dispersion models are in widespread use, *particle model* or *puff model*.

In a particle model the total amount of a trace gas or aerosol release is split into a large number of parcels, usually named "particles" (not to be confused with much smaller aerosol particles). At each time step a number of particles are released at the source location, each representing a certain mass of the pollutant. Each particle is followed individually downwind in small time steps, and the dispersion in a turbulent *atmosphere* is simulated by a nondeterministic, stochastic component of motion which displaces particles from each other even when they start at the same point in space and time, as depicted in Fig. 23.1. The concentration of a pollutant species at any point in space and time is then calculated by subdividing the simulated domain into a 3D grid, summing the masses of particles occupying a grid cell at that time and dividing this mass sum by the volume of the cell. It is worth mentioning that the size of the concentration grid cells is selectable independently of the size of the meteorological grid cells. Particularly the concentration grid can be made much finer than the meteorological grid in order to reproduce sharp gradients near small sources, because the process of dispersion is modeled without any spatial *discretization*. Depending on the size of the source, the desired spatial resolution, and the duration of the emission, several ten thousand to several million particles are necessary for "smooth" representation of the spatial distribution.

While in a particle model each particle is treated as punctual mass, in a puff model a number of three-dimensional Gaussian-shaped puffs are released from the source. These puffs grow with time to represent the dilution of the pollutant and are transported with the mean wind. However, the meteorological parameters

determining the puff growth rate can change locally and therefore puffs must be split into smaller ones from time to time whenever they exceed the size of the meteorological data grid cells, particularly in the vertical direction. In the course of a *simulation*, growing puffs may overlap, too. The concentration is evaluated at discrete grid points by summing the analytic concentration profile of each puff at the grid points. In principle, the number of puffs needed to represent a smooth concentration field will be much smaller than the number of punctual particles. However, the higher complexity of the splitting algorithm and the concentration calculation significantly reduces this benefit in terms of computational time. Some models, e.g., HYSPLIT, also offer a hybrid operation mode option using either a particle approach for the vertical dimension and the puff mode for the horizontal, or a switch from one mode to another after a certain simulation time.

23.3 Model Physics

The following description is based on the model system *HYSPLIT 4* (Draxler and Hess 1997) using the particle model mode unless mentioned otherwise. HYSPLIT, which is freely available (http://www.arl.noaa.gov/HYSPLIT_info.php), was developed at the Air Resources Laboratory of the National Oceanographic and Atmospheric Administration. It is, for example, being used operationally at the Volcanic Ash Advisory Centers (VAAC) in Washington, DC and Darwin. It has been used at IPA since 2010 for several projects, mainly as a tool to plan flight missions and tracer releases during *field campaigns* (RECONCILE, CONCERT, SHIVA) and for the special Falcon missions during the Eyjafjallajökull volcanic eruption 2010 and the Fukushima nuclear accident 2011. Its ease of use and installation, flexible capabilities, and economical use of computer resources made it attractive as a base for campaign forecast systems. Although usable stand-alone and "out of the box," the model's visualization and post-processing capabilities have been significantly expanded to fit particular needs. Many other transport and dispersion models are however similar in their basic approach, although they may differ in details of implementation, *parameterization*, and applicability to different length and time scales.

23.3.1 Advection

The advection term describes the motion of air masses due to the mean wind, i.e., the wind speed and direction averaged over the volume of the grid cells and the time interval of the meteorological input data. The time integrated advection term builds the deterministic particle trajectory. The position $P = (x, y, z)$ of each particle is updated during each integration time step Δt in two iterations using

$$\begin{aligned} P'(t+\Delta t) &= P(t) + V(P(t),t)\ \Delta t \\ P(t+\Delta t) &= P(t) + 0.5[V(P(t),t) + V(P'(t+\Delta t), t+\Delta t)]\ \Delta t. \end{aligned} \tag{23.1}$$

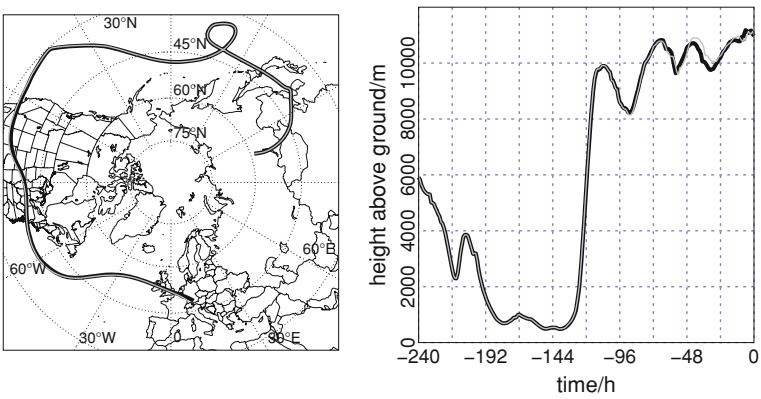

Fig. 23.2 Reversibility of forward/backward trajectories. The *thick black line* shows a trajectory starting in 11 km altitude over Germany and traveling 10 days backward-in-time to eastern Asia. The *thin gray line* drawn on *top* is the forward trajectory, starting at the end point and end time of the backward-trajectory. The destination point (again over Germany) differs from the initial starting point by about 170 km horizontally and less than 50 m vertically

$P'(t + \Delta t)$ is an intermediate first-guess of the new position and $P(t + \Delta t)$ the final new position. The velocity vector $V = (u, v, w)$ of the mean wind field is linearly interpolated in both space and time of the meteorological input data grid. In *mesoscale* to global scale simulations the time step Δt is in the range between a few minutes and half an hour. The model can adjust Δt automatically so that the advection distance per time step is less than a certain fraction (≤ 0.75) of the smaller of meteorological grid spacing and concentration grid spacing.

Figure 23.2 depicts, as a measure of the integration errors, the reversibility of a ten day trajectory. The trajectory was calculated first backward-in-time from Germany to eastern Asia and then forward-in-time from this endpoint. The situation selected for this example can be considered challenging for trajectory calculations, because the path crosses a *warm conveyor belt* with a strong vertical updraft and a *vortex*, which means that small deviations from the path will quickly grow because of wind shear and the divergent nature of the flow field (the trajectory is one of the trajectories shown later in Fig. 23.5). With a constant time step $\Delta t = 5$ min, forward and backward integration matched within ≈ 170 km horizontally on this $2 \times 27\,000$ km long travel path. In less critical situations much larger time steps would suffice for similar accuracy. The final position error in this example corresponds to a mean velocity error of 0.1 m/s. Thus, the numerical errors of this integration scheme can be considered negligible compared to the *uncertainties* of the wind field used as input. The overall accuracy of trajectory calculations typically depends more on the quality of the input data than on the choice of a specific trajectory model or integration scheme (Stohl et al. 2001). The spatial and temporal resolution of the input data are important factors, as is the meteorological situation. Forecast data is less accurate than re-analysis

data. Close to the surface of complex terrain or within the *atmospheric boundary layer* trajectories are less accurate than in the *upper troposphere*. A review of error sources can be found, e.g., in Stohl (1998).

23.3.2 Vertical Motion

Preferably the mean vertical motion of the particles is calculated using the vertical velocity component w provided by the meteorological input data set. If this input field is not available, the model computes the mean vertical velocity w by vertical integration of the horizontal velocity divergence

$$w(z) = \int_0^z \frac{\partial w}{\partial z} dz = -\int_0^z \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} dz. \tag{23.2}$$

This follows from the wind field's being divergence-free in order to conserve mass (although this approximation is valid only for incompressible flow regimes). Alternatively, for specific applications, the trajectory could also be forced to follow a path of constant pressure, *potential temperature* or density.

23.3.3 Dispersion

Dispersion is the process of mixing with ambient air leading to successive dilution of pollutants. In the *free atmosphere*, molecular diffusion, driven purely by the concentration gradient, plays almost no role compared to turbulent diffusion. The latter is the most complex physical process of the transport model. Since there is no analytical solution of fluid dynamics, the process has to be approximated by empirical or semi-empirical relationships (based on so-called similarity theory) between the measurable state of the atmosphere and the velocity variance of turbulent, sub-grid scale motions (*sub-grid scale* motions means deviations from interpolated values of the wind field between grid points). Vertical mixing is parameterized as functions of height with respect to boundary layer height (also known as *mixed layer* depth), z_i, and various *atmospheric stability* parameters derived from vertical *profiles* of the *temperature* and the wind, or taken from the meteorological dataset. Horizontal mixing can be either linked with vertical mixing or derived from the velocity deformation of the wind field.

In a particle model, dispersion is implemented as a stochastic process. After computing the advection and the mean vertical motion according to Eqs. (23.1) and (23.2), a random displacement is added to the mean (deterministic) position of particle k,

$$x_i(k, t + \Delta t) = x_{i,mean}(k, t + \Delta t) + u_i'(k, t + \Delta t)\, \Delta t, \tag{23.3}$$

where $i = (1, 2, 3)$ denote the x-, y-, and z-components. The turbulent velocity components $u' \equiv u'_1$, $v' \equiv u'_2$, $w' \equiv u'_3$ depend on the local standard deviation of turbulent velocity, $\sigma_i\,(x, y, z, t)$ (calculated for each grid cell of the meteorological data grid), an auto-correlation coefficient R, and a sequence of Gaussian-distributed random numbers λ with zero mean and variance 1. These sequences (as functions of time) are unique for each particle k and component i.

$$u'_i(k, t + \Delta t) = u'_i(k, t)\, R_i(\Delta t) + \sigma_i\, \lambda_i(k, t) \left(1 - R_i(\Delta t)^2\right)^{1/2} \quad \text{for } i = (1, 2). \tag{23.4}$$

The vertical component ($i = 3$) is computed similarly, but contains an additional velocity variance gradient term to avoid accumulation of particles in low-turbulence regions, see Draxler and Hess (1997). The autocorrelation R depends on the time step Δt and the Lagrangian *time scale* T_{L}, which is a measure of the *persistence* of fluid motion:

$$R_i(\Delta t) = e^{-\Delta t / T_{Li}}. \tag{23.5}$$

HYSPLIT uses $T_{\mathrm{Lw}} = 200$ s and $T_{\mathrm{Lu}} = T_{\mathrm{Lv}} = 10\,800$ s, while for example FLEXPART uses a sophisticated parameterization of the timescales. These values result in random-walk behavior in the vertical for time steps larger than $\approx$ 5 min ($R_w << 1$), i.e., $w(k,\ t + \Delta t) \cong \sigma_w\ \lambda_w(k,\ t)$. However, the autocorrelation of horizontal fluctuations is close to one ($R_u \cong 1\text{–}\Delta t/T_{\mathrm{Lu}}$) for all usual time step values, leading to random fluctuations of the horizontal velocity on time scales on the order of hours.

The velocity variances σ_{i}^2 are taken from one of the following sources, listed in order of preference. If available, the meteorological model's TKE field (turbulent kinetic energy $E = 0.5\ (u^2 + v^2 + w^2)$) is used directly to calculate the three components assuming constant anisotropy in the stable or neutral boundary layer and height-dependent anisotropy in the unstable boundary layer. However, in global meteorological models (e.g., GFS) TKE is usually not available. For the atmospheric boundary layer, an approach of Kantha and Clayson (2000) is used in such cases by default to compute the variances as a function of the friction velocity u^*, the convective velocity scale w^* and normalized height z/z_{i}. As a result, the variances are proportional to $(1\text{–}z/z_{\mathrm{i}})^{3/2}$ in the stable or neutral boundary layer. In the unstable boundary layer, $\sigma_u^2 = \sigma_v^2 =$ constant and $\sigma_w^2 \propto (z/z_{\mathrm{i}})^{2/3}(1\text{–}z/z_{\mathrm{i}})^{2/3}$. The boundary layer height z_{i} is taken from the meteorological data set. The velocity scales u^* and w^* are computed from the surface sensible heat flux and the surface momentum flux, both taken from the meteorological data set, when available. If the fluxes are not included in the data set, they have to be parameterized too, using vertical profiles of temperature and wind. An alternative, older approach calculates the vertical diffusivity coefficient, K_v from various stability parameters and the horizontal diffusivity coefficient, K_h, from the velocity deformation

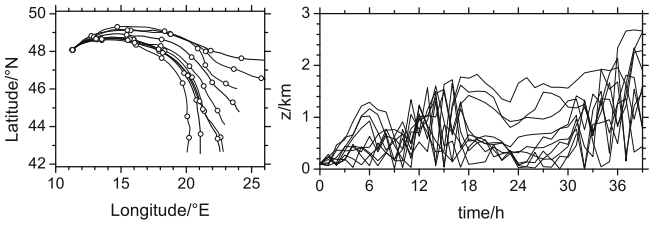

Fig. 23.3 Examples of individual particle trajectories including the dispersion (Eqs. 23.3–23.5). All particles start at the same location 100 m above ground at midnight local time on a summer day. *Circles* in the *left graph* indicate positions every 6 h. In the *right graph* the diurnal cycle of the vertical variances can be observed with lower variances at night (t = 0–6 h and 18–30 h in the graph) and higher variances during the day (6–18 h and 30 h onwards)

$\propto [(\partial v/\partial x + \partial u/\partial y)^2 + (\partial u/\partial x - \partial v/\partial y)^2\ ]^{1/2}$, then calculates $\sigma_i = (K_i/T_{Li})^{0.5}$. The details of these parameterizations are beyond the scope of this article and can be found in Draxler and Hess (1997, 1998). Figure 23.3 shows an example of particle trajectories including the dispersion simulation, which also demonstrates the *diurnal cycle* of turbulent variances.

23.3.4 Gravitational Settling

Nongaseous pollutants are subject to *gravitational settling* (also known as *sedimentation*), which has two effects: first it causes a descent of the "particles" (here in both the physical and the modeling sense), which increases with the size and mass density of the physical particles. This descent can exceed typical mean mesoscale vertical velocities (Eq. 23.2) for particle sizes larger than some tens of micrometers, see Fig. 23.4. Second, it contributes to the process of surface *deposition,* see next section. The settling velocity is calculated following Van der Hoven (1968)

$$V_s = g d_{ve}^2 C_c (\rho_p - \rho_a)(18\mu\alpha)^{-1}, \tag{23.6}$$

using gravity g, volume-equivalent particle diameter d_{ve}, mass density of particles and air, ρ_p and ρ_a, and the dynamic viscosity of air, μ ($\mu \approx 1.8 \times 10^{-5}$ kg m^{-1}s^{-1} for standard conditions). The Cunningham slip correction, C_c, corrects for very small $d \ll 10$ μm and is a function of (Λ/d), where Λ is the molecular mean free path ($\Lambda \approx 66$ nm at standard conditions), and α is a factor between ≈ 1 and ≈ 2 depending on the shape of the particles [see, e.g., Hinds (1999)]. Gravitational settling becomes very important for example when modeling volcanic ash transport.

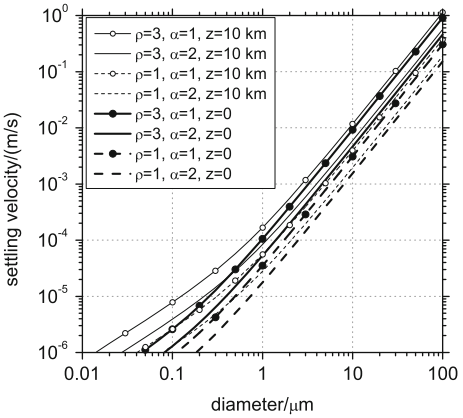

Fig. 23.4 Gravitational settling velocity as function of the volume-equivalent particle diameter (= the diameter of a sphere with the same volume), particle mass density ρ (in g cm^{-3}), and altitude z. $\alpha = 1$ corresponds to spherical particles (with largest settling velocity); most other geometries are in the range between $\alpha = 1$ and $\alpha = 2$

23.3.5 Removal Processes

During atmospheric transport, pollutants are, in the model, continuously removed from the system by three main processes: *dry deposition*, *wet deposition* and radioactive decay (chemical conversions will not be discussed here). All removal processes are implemented as decay of the model particles' (or model puffs') individual mass m_k at each time step, described by a sum of reciprocal time constants β_j,

$$m_k(t + \Delta t) = m_k(t) \exp\left[-\Delta t\left(\beta_{dry} + \beta_{inc} + \beta_{bel} + \beta_{rad}\right)\right]. \qquad (23.7)$$

Dry deposition is non-zero only if a model particle's (or puff's) center is within the lowest vertical model layer, i.e., $z < z_1$. Then, $\beta_{\mathrm{dry}} = V_{\mathrm{d}}/z_1$, where V_{d} is called the deposition velocity. Note that z_1 may be arbitrary, as long as the mean concentration of particles can be considered constant between ground and z_1: a smaller z_1 means fewer particles to be affected during that time step, but the inversely increased decay rate β results in the same total mass loss. For gases, V_{d} can either be set explicitly as input parameter or computed using the so-called resistance method, which depends on meteorological conditions as well as on surface properties [for details see Draxler and Hess (1997)]. For particulate pollutants, V_{d} can also be defined to be equal to the gravitational settling velocity V_{s}.

Wet deposition calculation needs information on the *precipitation* rate, P, taken from the meteorological data, as well as on the heights of *cloud* top and bottom. In HYSPLIT *cloud bottom* is parameterized as the vertical level where the *relative humidity* RH first exceeds 80 %, and cloud top as the level where RH drops below 60 %. Then the *wet deposition* rate of particles within clouds, β_{inc}, is computed using a wet removal velocity $V_{\mathrm{inc}} = SP$, where S is the scavenging ratio. S is the (constant) ratio of the pollutant's concentration in water to the concentration in air. Removal of particles below clouds is modeled directly using a constant β_{bel}. Wet deposition of water-soluble gases is applied to all layers below the cloud top and

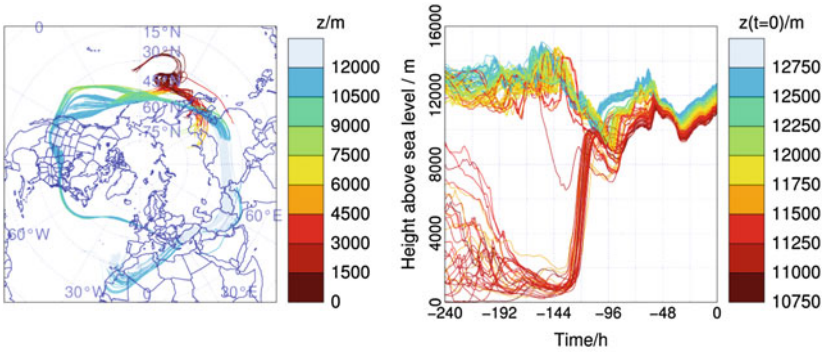

Fig. 23.5 10-d backward trajectories from a site of enhanced SO_2 observations in the upper troposphere over southern Germany on 3 November 2010. Some of the trajectories reveal the transport of probably polluted air from the boundary layer over eastern Asia to the upper troposphere by a warm conveyor belt over the western Pacific about 120 h before the observation. The *line color* indicates the local height of the trajectory (*left picture*) respectively the height at t = 0 (*right picture*)

depends on Henry's Law constant H of the gas, which describes the ratio of its equilibrium concentration in water to that in air, and on the precipitation rate P.

Finally, the radioactive decay rate is given by the half-life $T_{1/2}$ using $\beta_{\text{rad}} = \ln(2)/T_{1/2}$.

23.4 Examples of Application

23.4.1 Origin and Transport Pathways of SO_2 Detected in the Lower Stratosphere

In the frame of the HALO-TECHNO campaign in November 2010, which included the first scientific *HALO* flights, a SO_2 pollution layer was detected in the *lower stratosphere* (LS) above southern Germany during the flight on 3 November 2010. Besides SO_2, the *mixing ratios* of HNO_3 and *water vapor* were also enhanced in this pollution layer. In addition, the measured *LIDAR backscatter* signal was increased (Schlager et al. 2012). HYSPLIT simulations were used to investigate the origin and transport pathway of the LS pollution layer. Figure 23.5 depicts 10-d HYSPLIT backward-trajectory simulations for the air parcels observed in the LS pollution layer. About six days prior to the HALO measurements these backward trajectories passed through the *planetary boundary layer* of Eastern China, Japan and Korea, regions including strong anthropogenic SO_2 and NOx sources. Thus, here the air masses of interest most probably took up significant amounts of SO_2 and NOx pollutants. About 120 h prior to the HALO observation these air masses experienced an uplift to about 10 km altitude caused by a warm conveyor belt (WCB) within a time span of only about 12 h. After

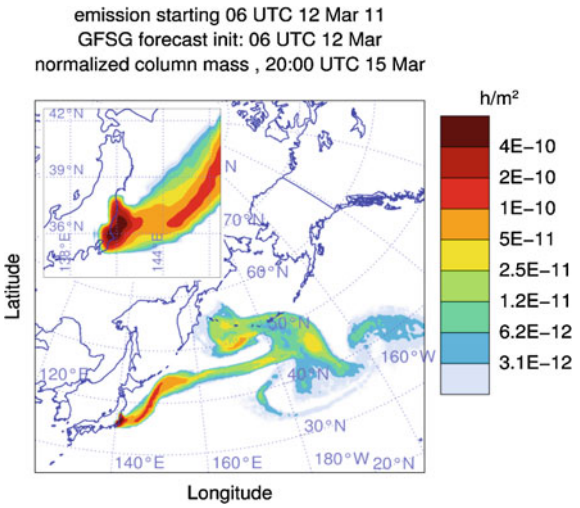

Fig. 23.6 First qualitative dispersion simulation of Fukushima plume as forecasted on 12 March 2011. The picture shows the vertical column mass in free air (e.g., in Bq/m^2) divided by the release rate (e.g., in Bq/h) valid at 20 UTC 15 March (*note* radioactive decay and deposition have been neglected in this early simulation)

uplift, the air masses were advected to Central Europe by the *jet stream* while slowly further ascending. The high-altitude transport mainly took part in the LS. Here, according to *chemistry* and aerosol box simulations, SO_2 and NOx experienced OH-induced partial conversion to gaseous *sulfuric acid* and near complete conversion to HNO_3, respectively. Thereby, *sulfate* aerosol droplets were also formed in the LS pollution layer by sulfuric acid nucleation and *condensation*. These sulfate aerosols could grow to sizes sufficient to explain the increased LIDAR *backscatter ratio*. This was the first observation of a direct injection of SO_2 from East Asia into the lower stratosphere by a WCB.

23.4.2 Fukushima Radioactive Plume over Europe

The accident at the Fukushima Daiichi nuclear power plant complex in March 2011 led to a release of numerous radionuclides, including some which were transported and detected almost around the whole northern *hemisphere*, such as I-131, Cs-137 and the radioactive gas Xe-133. The radioactive releases were a consequence of the main cooling system's failure after a disastrous tsunami flooded the power plant, caused by the Tohoku earthquake at 05:46 UTC 11 March. Already on 12 March, a few hours after the first *hydrogen* explosion in reactor block 1 at 06:36 UTC made the public aware of the hazardousness of the situation in Japan, IPA made first-guess HYSPLIT simulations of the transport and dispersion of the radioactive plume. The simulation was based on the assumption

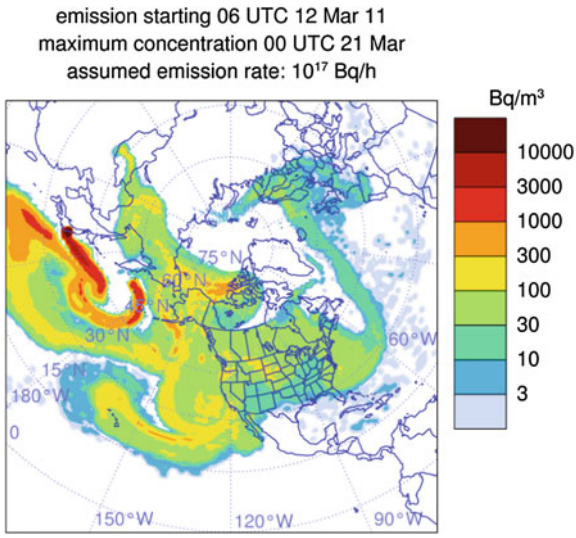

Fig. 23.7 First estimate of highest expected concentrations (at any height) within the atmosphere at 0 UTC 21 March, the time of highest concentrations over Europe in this simulation. The peak values over Europe occurred in the upper troposphere. *Note* in order to deliver a worst-case scenario, the simulation neglected radioactive decay

of a continuous release beginning at 06 UTC 12 March and showed that the plume was essentially drifting to the east over the Pacific within the first few days, however hitting the Tokyo area on 15 March, see Fig. 23.6.

This engagement was also in response to concerns that the plume could reach the upper airspace over Europe in a relatively short time and represent a threat to air traffic. Because of its gaseous state and its high concentration in spent nuclear fuel elements, Xe-133 is one of the first radionuclides to be released into the atmosphere in major quantities after damage to a reactor block. As a noble gas, it is neither washed out of the atmosphere nor deposited, and is therefore an ideal candidate for *long-range transport* from the Fukushima accident site. Further, its *lifetime* (half-life: 5.25 days) is comparable to the transport time from Fukushima to Europe, allowing later measurements to be unambiguously attributed to the source in Fukushima.

In order to timely quantify the radioactive dose rate expected in the European airspace before arrival, a quick estimate of the source strength was needed. Unfortunately, information on the emission rate was rare and disputed in the first days after the accident, and early official statements claimed that only a few percent of the fuel rods might be damaged. This uncertainty was also attributed to the fact that most of time the plume was blown seawards, far from land based *monitoring* stations. Therefore, first estimates were based on unofficial information about the typical nuclide inventory of spent nuclear fuel. The release rate was assumed to be in the range of 0.1–1 %/h of the total core inventory of the three reactor blocks which were active at the time of the emergency shutdown, which corresponds to about 10^{16}–10^{17}/h Bq Xe-133. To make the worst-case *scenario* more conservative, the radioactive decay was first neglected and continuous emission was assumed, which effectively overestimates the effect of accumulation

Table 23.1 Day of first detection at selected CTBTO observation stations and arrival times according to HYSPLIT simulation

Station number and location		Day of first detection	Arrival time in model
60	Petropavlovsk (Russia)	14 March	03 UTC 15 March
70	Sacramento (California)	16 March	21 UTC 16 March
14	Vancouver (Canada)	18 March	06 UTC 18 March
71	Sand point (Alaska)	18 March	06 UTC 18 March
74	Ashland (Kansas)	18 March	09 UTC 18 March
75	Charlottesville (Virginia)	19 March	06 UTC 19 March
17	St. John's (Canada)	19 March	06 UTC 19 March
79	Oahu (Hawaii)	19 March	15 UTC 19 March
72	Melbourne (Florida)	20 March	21 UTC 20 March
34	Reykjavik (Iceland)	20 March	06 UTC 23 March
63	Stockholm (Sweden)	22 March	09 UTC 22 March
61	Dubna (Russia)	22 March	12 UTC 22 March
53	Ilha de São Miguel	23 March	18 UTC 23 March
33	Schauinsland (Germany)	24 March	12 UTC 24 March

The CTBTO detection days were taken from http://www.bfs.de/ de/ion/imis/spurenmessungen.html and http://www.bgr.bund.de/DE/Themen/Erdbeben-Gefaehrdungsanalysen/Seismologie/Kernwaffenteststopp/Verifikation/Atmosphaer-Transport/Besondere%20Ereignisse/atm_fukushima_inhalt.html as of Jan 13th, 2012

at a distant target. Recently, inverse FLEXPART model simulations (Stohl et al. 2011) based on a global network of ground-based monitoring stations estimated a total release of (1.34–2.0) $\times$ 10^{19} Bq Xe-133 within the 120-h-period 11–15 March 2011 in the course of a series of pressure ventings, explosions, and fires in four reactor blocks. Thus, considering the neglected radioactive decay, the initial upper limit turned out to be quite realistic with respect to the average emission rate.

Figure 23.7 shows the maximum concentration in the vertical column at the time of highest concentrations over northern Europe, calculated for an emission rate of 10^{17} Bq/h (without decay), beginning at 06 UTC 12 March. Over Scandinavia the concentration peaks well below 100 Bq/m^3, and over the rest of Europe well below 30 Bq/m^3, and the highest values occurred in the upper troposphere. A value of 100 Bq/m^3 of Xe-133 corresponds to a dose rate of 0.47 nSv/h. In comparison, the average background level of cosmic *radiation* is 32 nSv/h at mean sea level (doubling for every 1500 m of height), and the total average dose rate by natural sources is about 240 nSv/h in Germany. Therefore, no risk for passengers and aircraft over Europe was predicted.

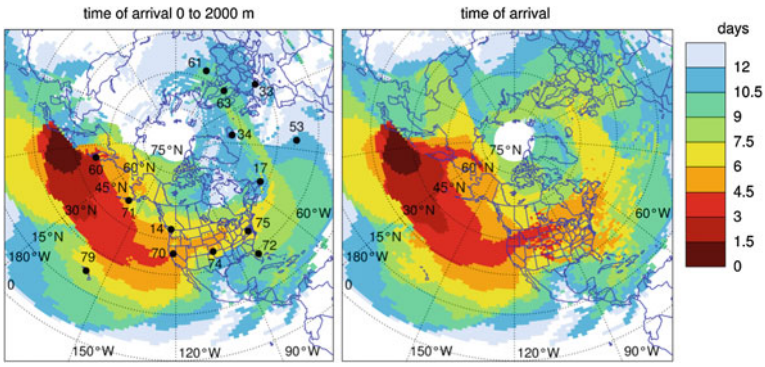

Fig. 23.8 Time of arrival of the radioactive plume from the Fukushima nuclear power plant accident. Simulation of a continuous gas emission (no deposition, no settling), beginning at 06 UTC 12 March (the time of the first major hydrogen explosion), using HYSPLIT with GFS 1° data. The *color* indicates the time (relative to 06 UTC 12 March), when a statistically significant concentration is exceeded the first time. *Left* arrival in ground vicinity up to 2 km altitude (boundary layer). Numbers denote CTBTO station numbers (Table 23.1). *Right* arrival at any height

According to the HYSPLIT simulation the radioactive plume was expected to arrive over Europe first in the upper troposphere over Great Britain and Scandinavia on 19 March, seven days after the first explosion at the Fukushima nuclear power plant. This shorter transport time was possible because part of the plume was lifted to the upper troposphere by a WCB caused by a Pacific *cyclone* west of Japan after it traveled about two days in the *marine boundary layer*. In the upper troposphere the plume was quickly transported with the jet stream over North America and the North Atlantic to Europe. Figure 23.8 depicts the arrival time with respect to the beginning of the emission. It can be seen that the arrival in North America and Europe in the mid and upper troposphere is earlier by one and a half days or more than within the lower troposphere below 2 km altitude. Table 23.1 compares the modeled arrival times with the dates of first detection at a subset of the stations belonging to the worldwide monitoring network of the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO). The simulation was run with 72 000 particles/h emitted beginning at 06 UTC 12 March. Arrival time was defined as the first 3-h interval when $\geq$4 simulation particles are within a cylinder of 50 km radius and 2 km height centered at the station. Despite the unaccounted time structure of the emissions, the simulation is in very good agreement with the measurements, predicting the correct day of arrival in 12 of 14 cases.

To evaluate the HYSPLIT simulations, two Falcon measurements flights were performed in the airspace over Germany on 23 March and 14 April 2011. During these flights whole air samples were taken to analyze Xe-133 activity. In addition, on-line measurements were taken of SO_2, HNO_3, total and nonvolatile condensation nuclei with a *chemical ionization* ion trap *mass spectrometer* and a set of

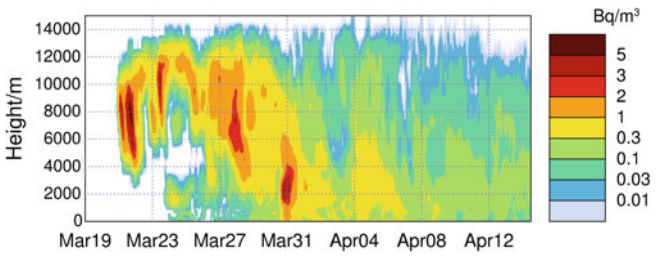

Fig. 23.9 Vertical concentration profile of Xe-133 as a function of time over the location 48°N, 11°E. The simulation shown here assumes a time-limited emission starting at 12 UTC 14 March and lasting for 50 h, and considers radioactive decay with a half-life of 5.25 days. The concentration is scaled for a total release of 10^{19} Bq (10 EBq) on 14–16 March

condensation particle counters, respectively. The on-line measurements of tracer for polluted boundary layer air were used during the flights to identify air masses in the upper troposphere which may originate from the boundary layer in Japan and to take samples preferentially in these layers.

The analysis of Xe-133 activity in the air samples was conducted at the Max Planck Institute for Nuclear Physics (MPI-K) in Heidelberg using a miniaturized proportional counter with an extremely low instrumental background count rate and a sophisticated Xe-133 extraction method, both developed at MPI-K. This was necessary because, due to the pressing time, evacuated stainless steel canisters could only be filled with a sampling volume of 1 and 2 liters in the Falcon cabin, which needed no certification.

The measured Xe-133 activity concentrations in the upper troposphere over southern Germany during the flight on 23 March at altitudes of 8.1, 9.2, and 11.8 km are 0.492, 0.988, and 0.180 Bq/m^3 at standard pressure (STP; 0.191, 0.349, and 0.044 Bq/m^3 at ambient pressure). On 14 April Xe-133 activity concentrations were detected of 0.049 Bq/m^3 STP (0.021 Bq/m^3 at ambient pressure) in the sample taken at altitudes between 6 and 8 km. The values of 23 March are lower by a factor of 85–680 with respect to the upper limit estimated in Fig. 23.7 for the same region, but two days earlier. Figure 23.9 depicts a time series of the vertical activity concentration profile at ambient pressure over a point in southern Germany near the location of the different samplings. First, it can be seen that the plume reaches the mid and upper troposphere from 4 to 12 km altitude about three days earlier than it reaches the ground. Second, the figure reveals strong variations in the upper tropospheric concentration during 21–24 March from temporarily below 0.01 to over 3 Bq/m^3. However, comparing a few single locations and times in a fairly turbulent flow over such a long time scale and distance with a few point measurements at the same locations and times is not expected to correlate perfectly. The analysis, which is not yet finished at the time of writing, must therefore be refined in future considering the temporal structure of the emission rate and averaging the simulated concentration over *statistically significant* areas and time intervals.

23.4.3 HYSPLIT Simulations for Planning and Analysis of Lagrangian Tracer Experiments

There are several motivations for Lagrangian tracer experiments in atmospheric research, including studies of pollutant transport and transformation, understanding atmospheric dispersion mechanisms on different scales, and evaluation of transport and dispersion models. The approach in airborne Lagrangian experiments is to measure the dispersion and change of *composition* of a single air mass as it moves, using successive aircraft flights. Lagrangian tracer experiments can be performed using conventional ambient tracer, i.e., atmospheric trace gases that exhibit spatially variable distributions in a region of interest and are long-living compared to the transport times of interest. However, the ability to quantify microphysical and chemical processes in an air mass during transport could be greatly enhanced with the use of an artificial tracer for nonambiguous air mass identification and information on mixing processes. The most suitable artificial tracers are *perfluorocarbons* (PFCs) because of their extremely low ambient atmospheric concentrations (10^{-15} mixing ratio). PFCs are hydrocarbons consisting of four, five and six atom carbon rings, in which all hydrogen atoms are substituted by fluorine atoms. They are nontoxic, nonreactive, nonscavenged, and nondepositing, and contribute neither to *stratospheric ozone depletion* nor to the production of tropospheric *air pollution*. PFCs are potent *greenhouse* gases; however, due to the very small amount used in a tracer experiment (e.g., 10 kg) the *global warming* impact of such an experiment is negligible.

Recently, a PFC tracer release and detection system was developed at IPA and a first successful proof of concept tracer experiment in the upper troposphere was performed during the POLARCAT project. High-resolution forecasts of the dispersion of the released tracer over the experiment time period are essential for the planning and successful performance of a Lagrangian experiment. During the SHIVA campaign in 2011 in Malaysia HYSPLIT simulations were used to plan a Lagrangian experiment where the PFC tracer was released from the research vessel Sonne and the corresponding air mass was sampled two times with the Falcon aircraft downstream of the Sonne after 5 and 24 h, respectively. The objective of this experiment was to study the dispersion of of halocarbons produced from seaweed along the coast of Borneo in the marine boundary layer and the entrainment of these emissions in the free troposphere.

Figure 23.10 shows the HYSPLIT forecast of PFC (C_6F_{12}) tracer concentration after the release of 10 kg from the Sonne on 21 November 2011 over a dispersion period of 24 h. Depicted is the maximum concentration in the PFC tracer plume. Also included is the Falcon flight path in the plume during the first and second sampling flights. The HYPLIT forecast was regularly updated every 6 h based on the latest available GFS data to plan the flight track of the sampling flights.

Figure 23.11 shows a photo of the Falcon passing the Sonne during the experiment. Tracer release on the Sonne was performed from a liquid PFC reservoir using an atomizer spray nozzle device. Plume sampling during the

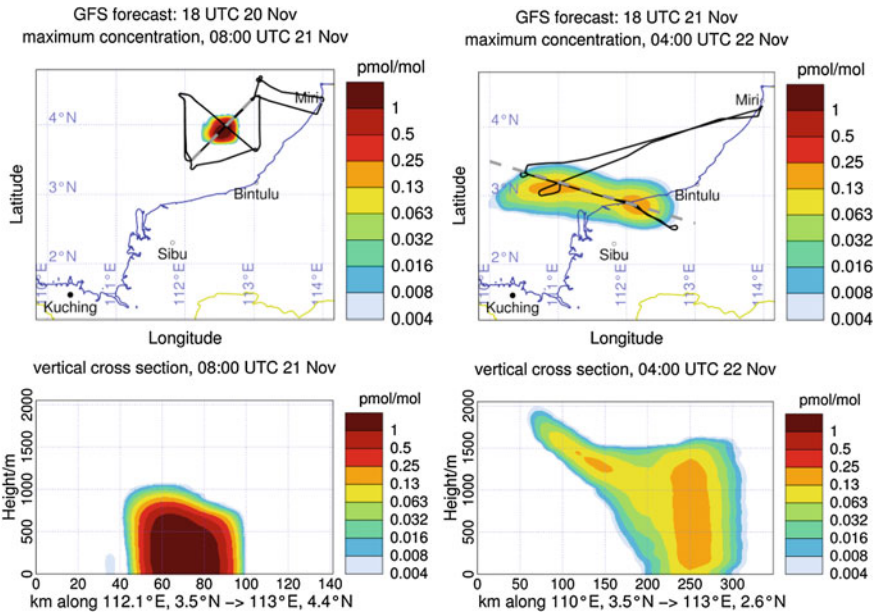

Fig. 23.10 Forecast of PFC tracer dispersion concentration at the northern coast of Borneo, 5 h (*left*) and 25 h (*right*) after the release of the tracer. The flight track of the two Falcon sampling flights are also shown (*black lines*). The *thick dashed lines* indicate the location of vertical cross sections

Fig. 23.11 The Falcon aircraft cruising in front of the research vessel "Sonne"

Falcon flights was conducted using an adsorption tube sampler built at IPA. Here, sample air (about 2 normal liters) is passed through adsorption tubes sequentially for 30–60 s and PFC tracer molecules are trapped and concentrated in the tubes

onto a molecular sieve adsorbent. The volume of air passed through the tubes is recorded. Analysis of the PFC amount in the tubes is performed after the campaign in the IPA laboratory by thermal desorption technique and *gas chromatography*.

23.5 Conclusions

Lagrangian trajectory and dispersion models have been proven to be a valuable tool for planning scientific aircraft flight missions and interpreting aircraft *in situ measurements*. They have been applied in cases ranging from regional scale tracer experiments to global scale emission scenarios. They can be used to predict pollutant concentrations in space and time, estimate unknown emission rates, determine probable emission locations, and calculate pollutant ages. Current models consider most of the essential physical processes, such as advection, turbulent dispersion, sedimentation, deposition, outwash and radioactive decay. Their accuracy, however, is not only determined by the model itself, but also by the quality of the meteorological input data. Especially in local to regional scale applications over complex terrain or in convective situations, high-resolution data of local weather models should be used.

References

Draxler, R.R., Hess, G.D.: Description of the HYSPLIT_4 modeling system, NOAA Technical Memorandum ERL ARL-224, 24 pp. Silver Spring (1997)

Draxler, R.R., Hess, G.D.: An overview of the HYSPLIT_4 modeling system of trajectories, dispersion, and deposition. Aust. Meteorol. Mag. **47**, 295–308 (1998)

Hinds, W.C.: Aerosol Technology—Properties, Behavior, and Measurement of Airborne Particles, 504 pp. Wiley-Interscience, New York (1999)

Kantha, L.H., Clayson, C.A.: Small Scale Processes in Geophysical Fluid Flows. International Geophysics Series, 883 pp. Academic Press, San Diego (2000)

Schlager, H., Arnold, F., Aufmhoff, H., Baumann, R., Pirjola, L., Roiger, A., Sailer, T., Wirth, M., Schumann, U.: First detection of East-Asian anthropogenic SO_2 lifted to the lower stratosphere by a warm conveyor belt: HALO-aircraft measurements above Europe, prepared for Geophys. Res. Lett (2012)

Stohl, A.: Computation, accuracy and applications of trajectories—a review and bibliography. Atmos. Environ. **32**(6), 947–966 (1998). doi:10.1016/S1352-2310(97)00457-3

Stohl, A., Haimberger, L., Scheele, M.P., Wernli, H.: An intercomparison of results from three trajectory models. Meteorol. Appl. **8**, 127–135 (2001). doi:10.1017/S1350482701002018

Stohl, A., Forster, C., Frank, A., Seibert, P., Wotawa, G.: Technical note: the Lagrangian particle dispersion model FLEXPART version 6.2. Atmos. Chem. Phys. **5**, 2461–2474 (2005). doi:10.5194/acp-5-2461-2005

Stohl, A., Seibert, P., Wotawa, G., Arnold, D., Burkhart, J.F., Eckhardt, S., Tapia, C., Vargas, A., Yasunari, T.J.: Xenon-133 and Caesium-137 releases into the atmosphere from the Fukushima Dai-ichi nuclear power plant: determination of the source term, atmospheric dispersion, and deposition. Atmos. Chem. Phys. Discuss. **11**, 28319–28394 (2011). doi:10.5194/acpd-11-28319-2011

Van der Hoven, I.: Deposition of particles and gases. In: Slade, D. (ed.) Meteorology and Atomic Energy, 445 pp. TID-24190, NTIS, Springfield (1968)

Wernli, H., Davies, H.C.: A Lagrangian-based analysis of extratropical cyclones. I: the method and some applications. Quart. J. Roy. Meteor. Soc. **123**, 467–489 (1997). doi:10.1002/qj.49712353811

Radiative Transfer: Methods and Applications

24

Bernhard Mayer, Claudia Emde, Robert Buras and Arve Kylling

Abstract

Solar and terrestrial radiation is the driver of atmospheric dynamics and chemistry and can be exploited by remote sensing algorithms to determine atmospheric composition. For this purpose, accurate radiative transfer models are needed. Here, a modern radiative transfer tool developed over many years at the Institute of Atmospheric Physics is explained. As an application, the remote sensing of cloud microphysics using the angular distribution of reflected solar radiance in the rainbow and backscatter glory is shown, with special emphasis on the polarization of radiation.

B. Mayer (✉)
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: bernhard.mayer@lmu.de

B. Mayer · C. Emde · R. Buras
Ludwig-Maximilians-Universität München (LMU), Meteorological Institute Munich (MIM), Theresienstraße 37, 80333 München, Germany
e-mail: claudia.emde@lmu.de

R. Buras
e-mail: robert.buras@lmu.de

A. Kylling
NILU-Norwegian Institute for Air Research, PO Box 100, 2027 Kjeller, Norway
e-mail: Arve.Kylling@nilu.no

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_24,

24.1 Radiative Transfer Modeling

Solar radiation is the one and only relevant external energy source driving Earth's *weather* and *climate*. Solar and *thermal radiation* are the main ways energy is exchanged between the Earth's surface and atmosphere and, together with sensible and *latent heat* flux, they determine the energy budget of the surface and the atmosphere. Radiation also strongly affects *atmospheric chemistry* and thus air quality, since some of the most important reactions are *photolysis* reactions where molecules like NO_2 or O_3 are split after the *absorption* of a *photon*. Each atmospheric constituent, whether gas, *aerosol*, *water droplet* or *ice particle*, interacts with radiation, putting its individual signature on the radiation field either by *scattering* or absorbing radiation. Reflected solar or emitted thermal radiation is exploited by various *remote sensing* algorithms to determine cloud and aerosol properties as well as *trace gas* concentrations. To correctly consider these processes and retrieve information from measured *radiances*, accurate *radiative transfer models* are needed.

Radiative transfer has a long history at Ludwig Maximilians University (LMU) Munich and DLR Oberpfaffenhofen. In the 1960s Fritz Möller (LMU) pioneered the development of radiative-convective models which allowed quantitative calculation of the effect of changing atmospheric CO_2 concentrations on the Earth's *surface temperature*. In the following decades more and more sophisticated radiative transfer models were developed to meet increasing demands, which included the calculation of UV radiation at the Earth surface; resource and site assessments of solar energy; development of *climate model parameterizations*; *satellite*, aircraft, and ground-based remote sensing of water and ice clouds; and satellite sensor *simulation*. The focus for much of this development has been the *libRadtran* package (Mayer and Kylling 2005), which is the institute's current "work horse" and which is freely available to the community at http://www.libradtran.org. The institute's most important radiative transfer equation solver—because of its flexibility and accuracy—is the *Monte Carlo code for the physically correct tracing of photons in cloudy atmospheres*, MYSTIC (Mayer 2009; Mayer et al. 2010; Emde and Mayer 2007; Emde et al. 2010, 2011; Buras and Mayer 2011). libRadtran and MYSTIC have been used for various purposes, as illustrated in more than 200 peer-reviewed publications to date. In this chapter it will be shown that even after so many years of development, there are still challenges and open issues. For that purpose a remote sensing application is selected which goes far beyond "traditional" passive remote sensing and makes use of distinct features in the *scattering phase function* of clouds (the *rainbow* and the backscatter *glory*, see Fig. 24.1) and an effect mostly neglected in radiative transfer applications but highly useful for remote sensing: *polarization*. For nonraining clouds with smaller droplets the rainbow is usually called a *cloudbow*.

Fig. 24.1 *Left* first and second order rainbow, photograph by Bernhard Mayer on Reunion Island. *Right* backscatter glory, photograph by Hermann Mannstein from a flight above a stratocumulus deck

24.2 Cloud Remote Sensing Using Special Observation Geometries

The most common application of passive cloud remote sensing is evaluating radiance observations from a satellite or aircraft looking downward onto a cloud (Fig. 24.2, left). Although clouds are clearly inhomogeneous in all directions, it turns out that in many cases they may be considered homogeneous within the field-of-view of the satellite (5 km, in the case of *Meteosat*) and net horizontal photon transport may be neglected. These "plane-parallel" and "independent-pixel" approximations are the backbone of all operational *retrieval* algorithms and allow determination of bulk properties such as vertically integrated *optical thickness*, cloud top altitude, and effective droplet or *ice crystal* size. While these are very valuable data, more detailed information is required to address, e.g., the indirect aerosol effect on clouds or to constrain current cloud models. In order to gain deeper insight and to study vertical profiles rather than bulk properties, more recent exploratory approaches attempt to retrieve the optical properties by imaging clouds from the side (Fig. 24.2, right), either from aircraft, from the ground, or

from a slant-viewing satellite instrument (Zinner et al. 2008). In these applications, the plane-parallel and independent-pixel assumptions break down completely, and radically new approaches using three-dimensional radiative transfer are required. In addition, specific viewing geometries may be used to gain deeper insights into cloud *microphysics*: e.g., details of the backscatter glory or the rainbow carry information about the width of the droplet *size distribution*, which is not at all available from any other viewing direction. It will be shown in the following what can be learned from such observations, in particular if the polarization state of the radiation is utilized additionally.

The general solution of *Maxwell's equation* for *electromagnetic radiation* is a transversal wave where the electric field vector is perpendicular to the propagation direction. The field vector is conveniently split into two components:

$$\begin{aligned} E_l &= a_l e^{i(2\pi ft - kz - \epsilon_l)} \\ E_r &= a_r e^{i(2\pi ft - kz - \epsilon_r)} \end{aligned} \tag{24.1}$$

where E_l and E_r are the parallel and perpendicular components of the electric field vector, a_l and a_r are the amplitudes, ϵ_l and ϵ_r the phases, t the time, f the frequency of the radiation, $k = 2\pi/\lambda$ the wave number, and z is the location. The ellipse described by the tip of the electromagnetic field vector at any fixed location is called a polarization ellipse and it contains all information about the state of polarization. Special cases of the ellipse are a straight line, corresponding to linear polarization, and a circle, corresponding to circular polarization. Sunlight is unpolarized, i.e., the individual polarization ellipses are uncorrelated and the electric field vector exhibits no preferred direction.

The energy transported by the radiation—the intensity—is proportional to the time-averaged square of the electric field strength. Four components are required to quantify the polarization state in terms of intensity. These components form the *Stokes Vector* (see, e.g., Hansen and Travis (1974)):

$$\begin{aligned} I &= <E_l E_l^* + E_r E_r^*> \\ Q &= <E_l E_l^* - E_r E_r^*> \end{aligned} \tag{24.2}$$

$$\begin{aligned} U &= <E_l E_r^* + E_r E_l^*> \\ V &= i<E_l E_r^* - E_r E_l^*> \end{aligned} \tag{24.3}$$

where I is the scalar (unpolarized) radiance, Q and U describe the state of linear polarization, V gives the circular polarization, and the asterisk denotes the complex conjugate. The Stokes vector can be measured using a polarizing filter and a retarder.

Interaction between matter and radiation is quantified in terms of the *extinction coefficient*, the *single scattering albedo*, and the *scattering matrix*. Extinction is the reduction of the intensity of a parallel beam, either by absorption or scattering. For a single molecule or droplet the *extinction cross-section* σ_{ext} can be specified,

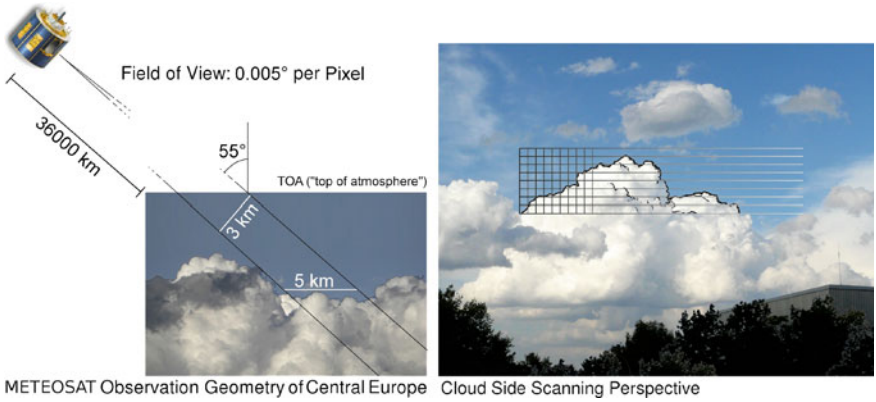

Fig. 24.2 Remote sensing perspectives. *Left* satellite or aircraft; *right* ground

which has the dimension of an area. For droplets or particles, the extinction efficiency is a common quantity which is the ratio of the extinction cross-section and the geometrical cross-section of the droplet:

$$Q_{\text{ext}} = \frac{\sigma_{\text{ext}}}{\pi r^2} \tag{24.4}$$

where r is the radius of the droplet or molecule. Since the atmosphere is a medium with many absorbing or scattering molecules, droplets, or particles, the volume extinction coefficient is usually employed:

$$k_{\text{ext}} = n\sigma_{\text{ext}} \tag{24.5}$$

where n is the volume concentration $[\text{cm}^{-3}]$ of the constituent under consideration. To distinguish between scattering and absorption, the single scattering albedo which gives the fraction of the scattering is defined as:

$$\omega_0 = \frac{k_{\text{sca}}}{k_{\text{ext}}}. \tag{24.6}$$

The angular distribution of the scattered radiation is described with the 4×4 phase matrix $\mathbf{P}$, which in general depends on incoming $\mathbf{n}^{\text{inc}}$ and scattered $\mathbf{n}^{\text{sca}}$ directions:

$$\mathbf{I}^{\text{sca}} = \mathbf{P}(\mathbf{n}^{\text{sca}}, \mathbf{n}^{\text{inc}})\mathbf{I}^{\text{inc}}. \tag{24.7}$$

Here, $\mathbf{I}^{\text{inc}}$ is the Stokes vector of the incident radiation, which is converted to $\mathbf{I}^{\text{sca}}$ during the scattering process. For applications where polarization is not considered, the phase matrix is reduced to a scalar phase function $p(\mathbf{n}^{\text{sca}}, \mathbf{n}^{\text{inc}})$ which describes

the probability that radiation coming from direction $\mathbf{n}^{\mathrm{inc}}$ is scattered into direction $\mathbf{n}^{\mathrm{sca}}$. If polarization is taken into account, the change of polarization state needs to be considered as well, e.g., that initially unpolarized radiation (with $Q = U = V = 0$) is polarized after the scattering.

For randomly oriented or spherical particles, the scattering matrix is much simpler than in the general case since various matrix elements are 0 or at least identical. In particular, for single scattering of unpolarized incident light at randomly oriented particles the intensity and the polarization of the scattered light can be directly calculated:

$$\begin{aligned} I^{\mathrm{sca}} &= P_{11}(\Theta) I^{\mathrm{inc}} \\ Q^{\mathrm{sca}} &= P_{12}(\Theta) I^{\mathrm{inc}} \end{aligned} \tag{24.8}$$

since U^{sca} and V^{sca} are both zero. The latter become nonzero when multiple scattering is included. In the following the focus is on the scattering phase function $P_{11}(\Theta)$ and the polarized scattering phase function $P_{12}(\Theta)$ since features like the cloudbow or the backscatter glory are single scattering phenomena. Here Θ is the angle between incident and scattered directions, commonly denoted as the scattering angle.

Figure 24.3 shows phase matrices computed for various gamma size distributions of cloud droplets. The upper panels show phase matrices calculated for a wavelength of 555 nm for three different effective radii. The small panels zoom into the cloudbow and the backscatter glory regions. It can be noted that the maximum of the cloudbow depends on the *effective radius*: it shifts towards a smaller scattering angle for a larger effective radius and the absolute value of the maximum increases. These features are present in both P_{11} and P_{12}. The glory region shows an opposite behavior: for larger effective radii the value of the maximum decreases and it is shifted towards larger scattering angles. P_{12} is negative in the cloudbow region, which implies that the perpendicular polarized part of the scattered light is larger than the parallel polarized part, whereas P_{12} is positive in the glory region.

The lower panels of Fig. 24.3 show phase matrices for an effective radius of 10 μm and three effective widths of the size distribution; $v_{\mathrm{eff}} = 0.01$ corresponds to a very narrow size distribution and $v_{\mathrm{eff}} = 0.25$ to a rather broad size distribution. The figure shows that the phase functions are very sensitive to the width of the size distribution in the cloudbow and glory regions. For narrow size distributions the polarized phase function shows oscillations in the region next to the cloudbow at scattering angles about 150° which are damped for wider size distributions. For the widest size distribution in this example the oscillations vanish completely. The maximum in the glory region is most well pronounced for the narrow size distribution and it broadens for wider size distributions. Also, the side maxima vanish with increasing width.

Since the scattering phase matrix is very sensitive to size distribution parameters in the cloudbow and glory regions, observations of scattered light in these geometries can provide new insights into cloud microphysics. The signal under cloudy conditions is usually dominated by multiple scattering, therefore highly

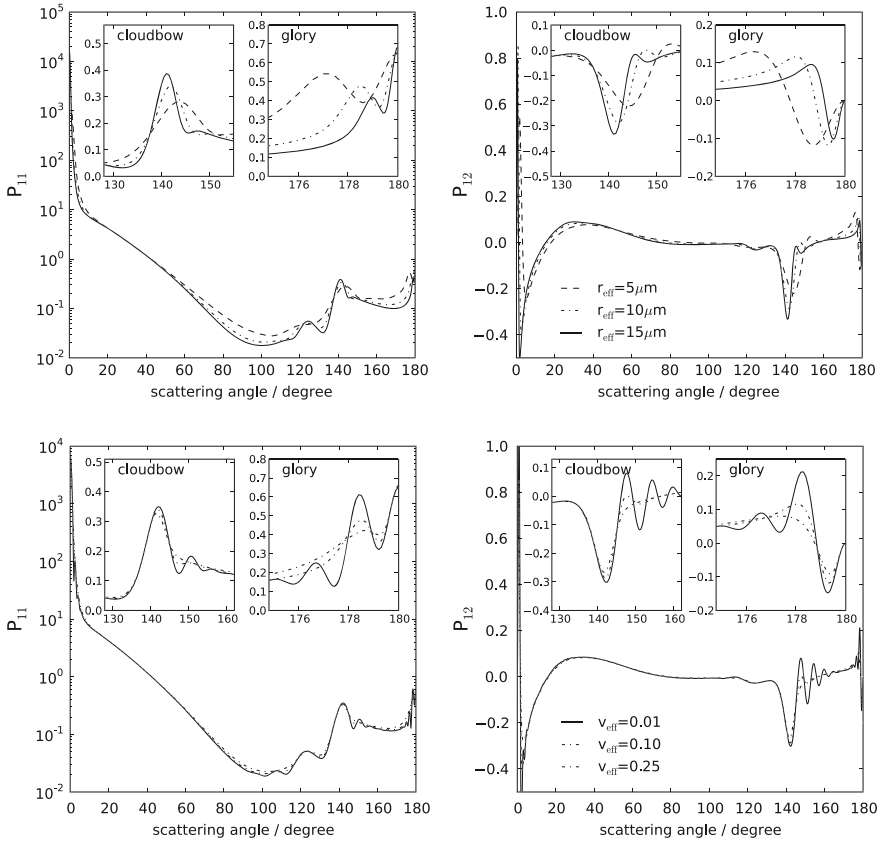

Fig. 24.3 Phase function P_{11} (*left*) and polarized phase function P_{12} (*right*) at 555 nm for liquid water. The *top panels* show calculations for three effective radii for a constant size distribution width ($v_{\text{eff}} = 0.1$), and the *lower panels* show calculations for three effective variances where the effective radius was kept constant ($r_{\text{eff}} = 10\,\mu\text{m}$). In the small panels the cloudbow and glory regions are zoomed in. The figures show that scattering in these regions is highly sensitive to the size distribution of cloud droplets

sophisticated radiative transfer codes are required to develop retrieval algorithms. In particular they need to handle polarization and they need to resolve the angular distribution of the scattered fields very accurately.

24.3 Monte Carlo Radiative Transfer

Various methods have been developed to solve the radiative transfer equation, that is, to calculate radiances and *irradiances* for a given atmospheric distribution of scatterers and absorbers (molecules, aerosol particles, water and ice cloud droplets and particles) over a partially reflecting Earth surface. Most of these assume that

the optical properties of the atmosphere vary only with the vertical coordinate—the so-called one-dimensional radiative transfer codes. The *Monte Carlo* method is the most versatile solution which is not restricted to one-dimensional geometries and which allows explicit consideration of arbitrary effects and processes, including three-dimensional clouds, inhomogeneous surface properties (topography and angular *reflection*), and polarization. For this reason, at IPA and LMU the Monte Carlo method has been adopted as the method of choice for detailed and accurate radiative transfer calculations and MYSTIC has been developed, which is currently one of the most versatile codes available.

Monte Carlo is not only the most flexible method, but also the most straightforward: individual photons are traced from their origin (either solar, or *thermal emission* by the atmosphere or surface) until they eventually leave the atmosphere or are absorbed by the surface. Interaction with the atmosphere and surface is simulated by random scattering and absorption events where the probabilities for these processes as well as the new direction after scattering are defined by the optical properties of the atmosphere: absorption and scattering coefficients, as well as the phase matrix which defines new direction and polarization state after a scattering event. As an example, how the location of the next scattering is determined and how absorption is considered, are illustrated below:

Photons travel along straight lines until they are scattered or absorbed. First, only scattering is considered. The probability for survival p_{sur} of a photon along its straight path is determined by the *Lambert-Beer law*:

$$p_{\text{sur}}(\tau) = \exp(-\tau) = \exp\left(-\int_0^s k_{\text{sca}}(s')ds'\right) \tag{24.9}$$

where the optical thickness τ is the integral of the volume scattering coefficient k_{sca} along the photon's path. The probability $p_{\text{sca}}d\tau$ that a photon is scattered between τ and $\tau + d\tau$ is:

$$p_{\text{sca}}d\tau = p_{\text{sur}}(\tau) - p_{\text{sur}}(\tau + d\tau) = -\frac{dp_{\text{sur}}}{d\tau}d\tau \tag{24.10}$$

Inserting (24.9) into (24.10) yields

$$p_{\text{sca}} = \exp(-\tau) \tag{24.11}$$

That is, the probability that the photon is scattered between τ and $\tau + d\tau$ is simply the probability that the photon survived from 0 to τ multiplied with the *optical depth* of the layer, $d\tau$. To determine the distance to the next scattering event, a random optical thickness must be drawn so that, when this step is repeated many times, the histogram of the optical thickness equals the probability density p_{sca}. This can be done applying Von Neumann's Golden Rule of Sampling: first, the cumulative probability density is calculated:

$$P_{\mathrm{sca}}(\tau) = \int_0^{\tau} p_{\mathrm{sca}}(\tau')d\tau' = 1 - \exp(-\tau) \tag{24.12}$$

The reason for using the cumulative *probability density function* is that it is, e.g., equally likely that τ falls into the intervals $P(\tau)\epsilon[0, 0.1]$ or $P(\tau)\epsilon[0.3, 0.4]$, namely 10 %. One can therefore simply assign a real random number between 0 and 1 to P_{sca} and convert this to a random optical thickness by inverting (24.12):

$$\tau = -\ln(1 - P_{\mathrm{sca}}). \tag{24.13}$$

Figure 24.4 illustrates how that works: the left plot shows 100 000 random numbers ρ equally distributed between 0 and 1, produced by a random number generator. When converted to optical thickness (middle plot) according to (24.13), the distribution is no longer even, but most photons are scattered at optical thicknesses between 0 and 2 while only a few photons make it to significantly larger optical thicknesses. The histogram of the thus-sampled optical thicknesses (right plot) shows that after 100 000 photons the distribution closely follows the Lambert-Beer law (dashed line). Since photons are traced in physical space rather than in optical thickness space, the optical thickness has to be translated to a length using the known distribution of the scattering coefficient.

Using the Golden Rule of Sampling, we can sample from arbitrary (normalized) probability densities $p(x)$, by calculating the cumulative probability density $P(x) = \int_0^x p(x')dx'$ and inverting it:

$$x = P^{-1}(\rho) \tag{24.14}$$

where ρ is a random number between 0 and 1. This procedure is the fundamental process in any Monte Carlo model: for our application there are a variety of probability density functions including the thermal emission of the photon in the atmosphere, the path length of the photon before it is scattered, the angular distribution of scattered light, or the directional reflectance of the surface. All of them can be handled by (24.14), but only very few of them can be solved analytically, such as the step width in (24.13). For those which cannot, we need to tabulate x as a function of ρ according to (24.14).

Absorption may be considered either by killing the photon or, alternatively, by reducing the photon weight which is set to 1 initially. In MYSTIC the latter was chosen because this turns out to be more efficient in many applications. The photon weight considers that a fraction of the photon, again described by the Lambert-Beer law, is absorbed along the path between two scattering events:

$$w = \exp(-\tau_{\mathrm{abs}}) = \exp\left(-\int_0^s k_{\mathrm{abs}}(s')ds'\right) \tag{24.15}$$

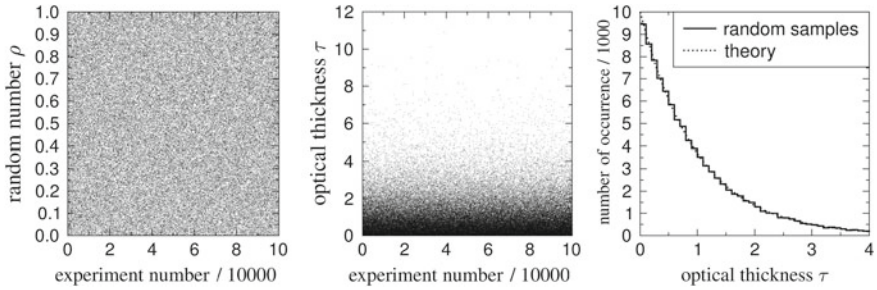

Fig. 24.4 Illustration of random sampling from the Lambert-Beer law: *left* 100 000 random numbers between 0 and 1; *middle* random numbers converted to optical thickness using the inverted cumulative probability density; *right* histogram of the thus-sampled optical thicknesses

where the absorption optical thickness is the absorption coefficient k_{abs} integrated along the path.

To implement polarization, one possibility is to assign a random polarization state to each traced photon entering the atmosphere. Another more efficient possibility is to interpret each traced photon as a photon package and to assign a weight vector corresponding to the Stokes parameters to each package. The second method was implemented in MYSTIC. The scattered Stokes weight vectors are calculated according to (24.7). This equation is valid in the scattering coordinate system, which is defined by the plane of incident and scattered directions and a vector orthogonal to this plane. Hence each scattering process has to be calculated in a different coordinate system and the appropriate coordinate transformations need to be included in the model. Reference is made to Emde et al. (2011) for more details about polarized radiative transfer in MYSTIC.

In order to obtain irradiances, radiances, *heating rates* or other radiation quantities, many photons are traced and the respective photon weights are summed at the desired locations. Since the individual contribution of each photon is determined by a statistical process, the relative uncertainty (relative standard deviation) of the result is proportional to the inverse square root of the number of photons traced. Accurate results may therefore be computationally expensive, depending on the application. To keep computational times reasonable a number of variance reduction techniques have been implemented: basically, photons are "forced" to reach the desired detector; the bias introduced is corrected by adjusting the photon weight. Variance reduction methods can be mathematically shown to yield results identical to the brute force tracing of large numbers of photons, but with significantly reduced statistical noise and thus computing time for the same level of accuracy. A typical example of a MYSTIC simulation is shown in Fig. 24.5.

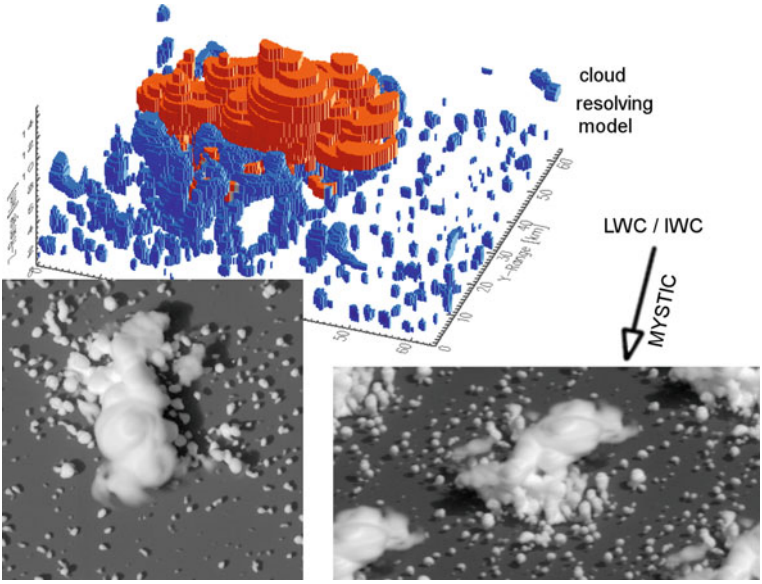

Fig. 24.5 MYSTIC simulations of high spatial resolution satellite radiances for downward (*bottom left*) and slant viewing (*bottom right*) geometries. The liquid and ice water distribution provided by a cloud-resolving model (*top*) were used as input to MYSTIC

24.4 Applications

In the following, two examples are given for remote sensing applications using the MYSTIC model: backscatter glory (unpolarized) and rainbow (polarized).

24.4.1 The Backscatter Glory

The backscatter glory occurs within a few degrees around the exact backscatter direction, see Fig. 24.1. Figure 24.6 illustrates observations by the *Compact Airborne Spectrographic Imager* (CASI) of FU Berlin, mounted on an airplane flying above a marine stratocumulus deck. Since the instrument scans across track, we obtain a cross-section through the glory across the flight track, rather than the expected circular shape. During the flights the viewing direction was exactly aligned to the sun, which is confirmed by the fact that even the aircraft shadow in the center of the glory can be identified.

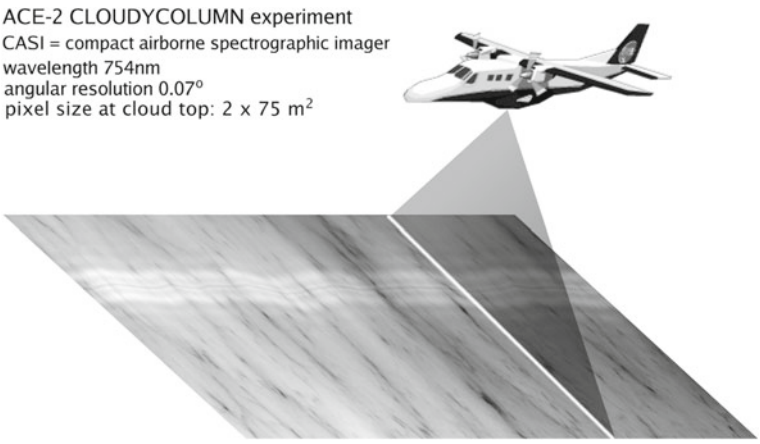

Fig. 24.6 Scanning geometry for the CASI instrument together with along track observations

Figure 24.7 (left) summarizes the information content of a glory scan. The absolute *reflectivity* allows determination of the optical thickness of the cloud, which is also exploited by "traditional" remote sensing methods. The effective radius is easily derived from the angular diameter of the first circle—this is certainly one of the most simple remote sensing methods available. It may even be applied to simple photographs of the glory if the focal length of the camera is known (the focal length and other relevant information is usually stored in the EXIF data of digital camera images and can be used to convert pixel numbers to angles). In general, the larger the effective droplet radius, the smaller the angular diameter of the glory (see also Fig. 24.3). Information about the width of the size distribution is mainly contained in the side maxima, that is, the second and higher order rings: wide size distribution cause these extra rings to be smeared out until they eventually vanish. To quantitatively exploit these features, a set of reflectivities was precalculated for a number of optical thicknesses, effective radii and widths of the size distribution. By a straightforward least square fit of the calculated reflectivity to the observed one, the set of parameters which minimized the difference between both was determined, see Fig. 24.7 (right). Figure 24.8 shows the thus-determined parameters along the flight track. For details of the method please refer to Mayer et al. (2004).

24.4.2 Cloudbow

In order to investigate the potential for retrieving information about cloud microphysics from polarized measurements we simulated such measurements assuming various cloud optical thicknesses and different size distribution parameters. A homogeneous cloud layer at 2–3 km altitude was assumed, as well as a

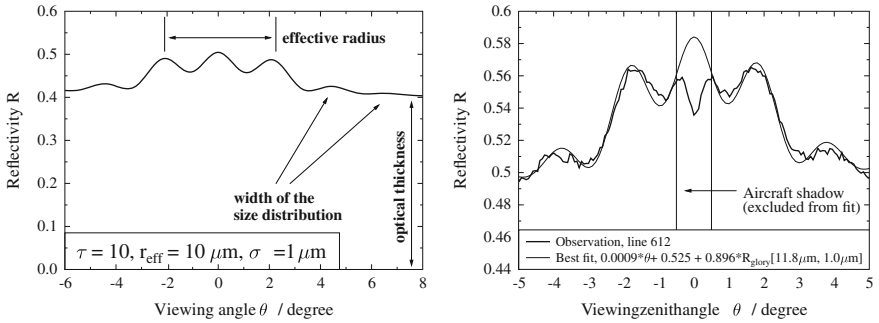

Fig. 24.7 Information content of the backscatter glory (*left*) and fit of a calculated radiance scan to the observation (*right*)

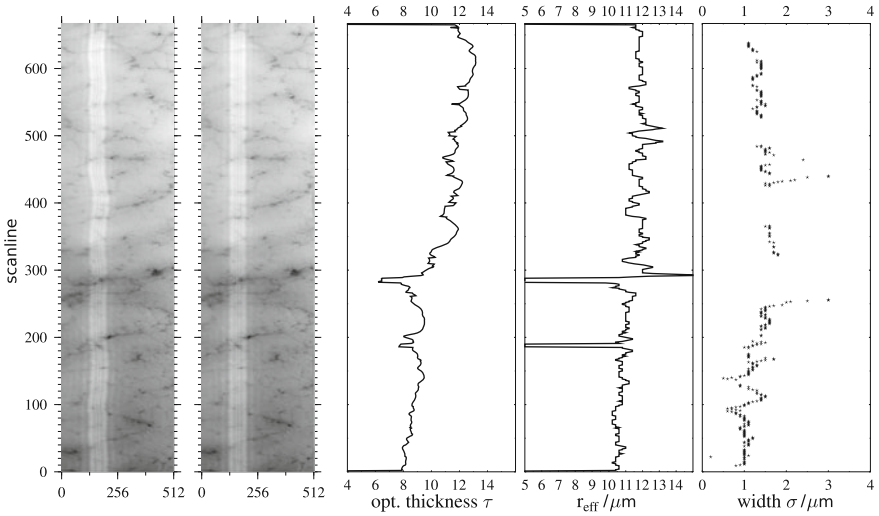

Fig. 24.8 Results retrieved from the backscatter glory. From *left* to *right*: original observation; stratified observation; optical thickness; effective radius; width of the size distribution (defined here as the standard deviation from the average radius)

solar zenith angle of 40°. The sensor is located above the atmosphere and scans from $-10°$ to $+20°$ with respect to nadir in the solar principle plane, corresponding to scattering angles from $130-160°$.

The results are shown in Fig. 24.9. The left panels show simulations with different cloud optical thicknesses keeping the drop size distribution constant. As expected, the scalar radiance increases monotonically with optical thickness.

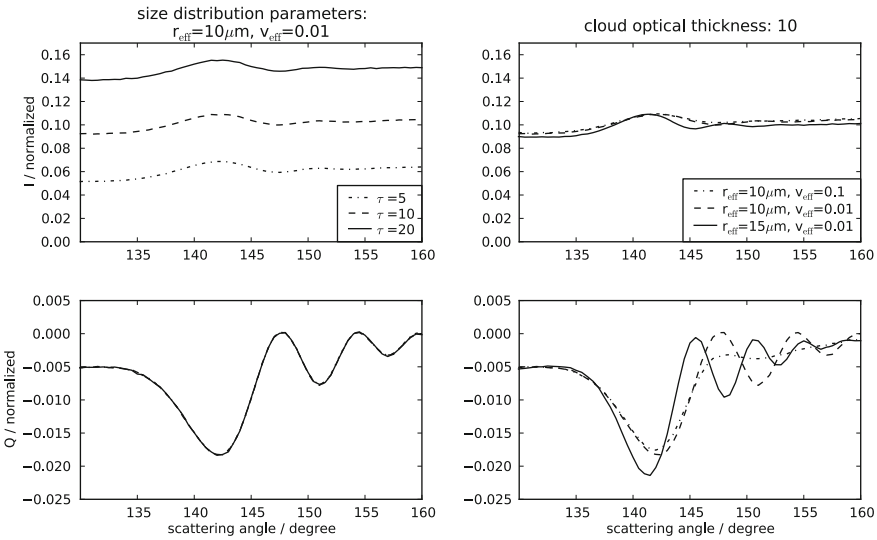

Fig. 24.9 Simulations of polarized radiances in the principal plane for the cloudbow region. The *top panels* show the scalar radiance I and the *bottom panels* show the Stokes vector component Q. The *left panels* are for constant size distribution parameters $(r_{\text{eff}} = 10\mu\text{m}, v_{\text{eff}} = 0.01)$; the different lines correspond to different cloud optical thicknesses. The *right panels* are for a constant cloud optical thickness and varying size distribution parameters

This is the reason why various cloud retrieval algorithms use the reflectance in the visible wavelength region to obtain information about cloud optical thickness. The polarization difference Q remains constant for optical thicknesses larger than approximately 5. The reason is that the polarization state of photons changes with each scattering and after a certain number of scattering events the polarization state is random. Therefore, all photons that are multiply scattered do not contribute to Q, which implies that Q does not include information about cloud optical thickness.

The right panels show simulations with different size distribution parameters keeping the cloud optical thickness constant. Here the scalar radiance I remains almost unaffected, whereas Q shows large variations. Since Q is hardly influenced by higher orders of scattering, it is very closely connected to the phase matrix element P_{12}, which is highly sensitive to size distribution parameters. At LMU and IPA a new cloud retrieval algorithm is currently being developed which makes use of the high sensitivity of polarization measurements to cloud size distribution parameters. This algorithm will yield not only the cloud effective radius but also the width of the size distribution.

References

Buras, R., Mayer, B.: Efficient unbiased variance reduction techniques for Monte Carlo simulations of radiative transfer in cloudy atmospheres: the solution. J. Quant. Spectrosc. Radiat. Transf. **112**, 434–447 (2011). doi:10.1016/j.jqsrt.2010.10.005

Emde, C., Mayer, B.: Simulation of solar radiation during a total solar eclipse: a challenge for radiative transfer. Atmos. Chem. Phys. **7**, 2259–2270 (2007). doi:10.5194/acp-7-2259-2007

Emde, C., Buras, R., Mayer, B. et al.: The impact of aerosols on polarized sky radiance: model development, validation, and applications. Atmos. Chem. Phys. **10**, 383–396 (2010). doi:10.5194/acp-10-383-2010

Emde, C., Buras, R., Mayer, B.: ALIS: an efficient method to compute high spectral resolution polarized solar radiances using the Monte Carlo approach. J. Quant. Spectrosc. Radiat. Transf. **112**, 1622–1631 (2011). doi:10.1016/j.jqsrt.2011.03.018

Hansen, J., Travis, L.: Light scattering in planetary atmospheres. Space Sci. Rev. **16**, 527–610 (1974). doi:10.1007/BF00168069

Mayer, B.: Radiative transfer in the cloudy atmosphere. Eur. Phys. J. Conf. **1**, 75–99 (2009). doi:10.1140/epjconf/e2009-00912-1

Mayer, B., Kylling, A.: Technical note: the libRadtran software package for radiative transfer calculations: description and examples of use. Atmos. Chem. Phys. **5**, 1855–1877 (2005). doi:10.5194/acp-5-1855-2005

Mayer, B., Schröder, M., Preusker, R. et al.: Remote sensing of water cloud droplet size distributions using the backscatter glory: a case study. Atmos. Chem. Phys. **4**, 1255–1263 (2004). doi:10.5194/acp-4-1255-2004

Mayer, B., Hoch, S., Whiteman, C.: Validating the MYSTIC three-dimensional radiative transfer model with observations from the complex topography of Arizona's Meteor Crater. Atmos. Chem. Phys. **10**, 8685–8696 (2010). doi:10.5194/acp-10-8685-2010

Zinner, T., Marshak, A., Lang, S. et al.: Remote sensing of cloud sides of deep convection: towards a three-dimensional retrieval of cloud particle size profiles. Atmos. Chem. Phys. **8**, 4741–4757 (2008). doi:10.5194/acp-8-4741-2008

Ice Cloud Properties From Space

25

Luca Bugliaro, Hermann Mannstein and Stephan Kox

Abstract

The Spinning Enhanced Visible and InfraRed Imager SEVIRI radiometer aboard the geostationary Meteosat Second Generation MSG satellite enables the quantitative identification and characterization of clouds from space with high temporal resolution. In this chapter we illustrate basic features of this satellite instrument and present some recent advances related to the detection of ice clouds and the determination of their optical and microphysical properties at IPA.

25.1 Introduction

25.1.1 Clouds

In everyday life, *clouds* strongly influence human behavior as they affect recreational as well as commercial activities in two ways: they modulate *solar radiation* at the Earth surface and they are responsible for *precipitation*. Of course, clouds play an eminent role in solar energy production and are a fundamental aspect of the hydrological cycle. However, the most striking feature about clouds is their

L. Bugliaro (✉) · H. Mannstein · S. Kox
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Luca.Bugliaro@dlr.de

H. Mannstein
e-mail: Hermann.Mannstein@dlr.de

S. Kox
e-mail: Stephan.Kox@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_25,

impact on *Earth's energy budget* and so on *climate*. Low water clouds effectively reflect solar radiation because they are usually optically thick. At the same time, they have a very small effect on Earth's *thermal emission* since they are located in warm layers of the atmosphere close to the ground. Thin *cirrus* clouds in contrast are weak solar radiation reflectors but trap *thermal radiation* emitted by the warmer Earth's surface and re-emit radiation according to their own (low) temperature and *Planck's law*. Therefore, more radiation is captured in the atmosphere than in the cirrus-free case in such a way that their net *instantaneous radiative forcing* is positive and they can enhance the *greenhouse effect*. The topic of cloud impact on climate has large relevance, but is at the same time very complicated, as testified by the fact that *cloud feedbacks* represent the largest uncertainty in *climate models*. For this reason, and because aircraft condensation trails (*contrails*) are anthropogenic ice clouds that can evolve to extended cirrus, IPA soon started to investigate such high clouds.

25.1.2 Satellite Remote Sensing of Clouds

Soon after the launch of Sputnik 1 by the Soviet Union, which marked the official beginning of the space age on 4 October 1957, NASA launched the world's first weather *satellite* TIROS 1 on 1 April 1960. The first black-and-white "camera" for weather observations on a geostationary satellite flew on NASA's ATS 1 and produced spectacular full-disc images of Earth with a spatial resolution of about 3 km. One of the first applications of this camera was tracking of clouds to determine cloud motion vectors. The first operational weather satellite was NOAA's geosynchronous GOES-1 carrying the Visible Infrared Spin Scan Radiometer VISSR. Thanks to its two channels, one visible and one infrared, it provided day and night imagery of cloud conditions over the full disc and enabled the determination of cloud heights and temperatures. The first European geostationary satellite Meteosat-1, launched in 1977, demonstrated the usefulness for cloud remote sensing of image acquisition in additional intervals of the electromagnetic spectrum. Its three channel instrument, the Meteosat Visible and Infrared Imager MVIRI, also comprised a *water vapor* channel in the infrared that could be used for the extraction of cirrus clouds and cloud top height determination. Present geostationary satellites come with up to 12 spectral channels, typical spatial resolutions around 3 km, and temporal resolution of 15–30 min. Selected channels can reach even higher spatial resolutions and some reduced Earth disc subsets can be scanned faster. Thus, they represent the ideal tool for daytime and nighttime investigation of cloud processes and their role in climate.

Although almost exclusively geostationary platforms have been mentioned in this context, polar orbiting satellites are also important. They usually acquire data with better spatial resolution, are able to cover the pole regions, and traditionally comprise more spectral channels than do geostationary satellites. Although they do not allow investigation of cloud cycles, some polar orbiting instruments have

produced long time series of geophysical properties relevant for *climate change* detection. So, the NOAA Very High Resolution Radiometer VHRR and its successor, the Advanced VHRR (AVHRR) have led, for instance, to continuous, 30-year-long cloud climatologies with a spatial resolution at nadir of 1.1 km. The Moderate Resolution Imaging Spectrometer MODIS, aboard NASA's polar orbiting sun-synchronous Earth Observing System (EOS) spacecrafts Terra and Aqua, provides data just from the past ten years but with high radiometric sensitivity (12 bits) in 36 spectral bands ranging from 0.4 to 14.4 μm with spatial resolutions of 250 m, 500 m or 1 km.

In the following, the physical and optical properties of clouds will be reviewed and MSG/SEVIRI will be introduced in some detail, before the IPA ice cloud *retrievals* are presented. Finally, a comparison to the airborne *high spectral resolution lidar* of IPA is used to estimate the accuracy of these space borne algorithms.

25.2 Cloud Properties

25.2.1 Physical Properties of Clouds

Clouds consist of an agglomeration of suspended liquid water droplets or *ice crystals* and are formed when air parcels become saturated. Formation mechanisms include cooling of the air, for instance by *adiabatic*, convective, or orographic lifting, and adding water vapor to the air, e.g., by evaporation of the oceans. High spatial variability down to the centimeter and millimeter scale characterizes the microscopic cloud. The macroscopic size of clouds ranges from typically 1 km for *cumulus* to approximately 100 km for *cumulonimbus* clouds in the *tropics* up to several thousand km for frontal cloud structures. The microphysical description of clouds encompasses the indication of the three-dimensional distribution of *cloud particle* size and, for ice crystals, particle shape. Liquid water droplets usually extend from approximately 5 to 100 μm, while ice crystals can be larger and reach a maximum dimension of 5 000 μm. Larger droplets and particles are usually classified as precipitation and neglected in the description of clouds. From cloud *particle size distribution* $n(r)$ one can compute the condensed water content *CWC* as well as the particle number $N = \int_0^\infty n(r)\,dr$ and the mean volumetric radius r_{vol}, which represents the radius of an equivalent sphere with the same amount of water:

$$CWC = \rho_w V_{tot} = \rho_w \int_0^\infty n(r)\, V_{part}(r)\, dr =: \rho_w \frac{4\pi}{3} r_{vol}^3 .$$

V_{part} is the volume of one particle with "radius" r and ρ_w the density of water, liquid or ice, depending on the cloud phase.

25.2.2 Optical Properties of Clouds

This discussion is restricted to the interaction of clouds with solar radiation in the spectral range 0.4–2.2 μm where most solar satellite channels are located. The interaction consists in the *scattering* and/or *absorption* of radiation by water droplets and ice crystals. One first notices that the mean free *photon* path in a thick cloud is of the order of 10 m. Thus, microphysical variability at smaller scales is of no concern for remote sensing of clouds with passive sensors. Furthermore, multiple scattering enhances horizontal photon transport and this induces a "radiative smoothing" of cloud structures, i.e., sharp variations of cloud bulk properties are not mirrored in the cloud radiation distribution that appears smoother. The resulting "smoothing scale" amounts to some hundreds of meters and represents a physical limitation on the spatial resolution of remote sensing instruments.

In the solar spectral range a description of cloud-radiation interaction can be reduced to two main quantities:

(a) The extinction *optical thickness* $\tau = \int_{z_{bot}}^{z_{top}} \beta_{ext}(z)\, dz$ is the integral of the *extinction coefficient* $\beta_{ext} = \int_0^\infty n(r)\sigma_{ext}\, dr$ from cloud bottom height z_{bot} to cloud top height z_{top}, and σ_{ext} is the *extinction cross section* of one particle; since β_{ext} has units of km^{-1}, τ is dimensionless;

(b) The cloud particle *effective radius* $r_{eff} = 3V_m/(4A_{pm})$, where V_m is the mean particle volume and A_{pm} the mean projected cross-sectional area of the particle. It is typically expressed in μm.

Scattering properties are almost completely described with these quantities τ and r_{eff}, and details (shape and width) of the cloud particle size distribution $n(r)$ do not play a significant role. This is mainly because cloud particles are much larger than the *wavelength* of light, at least of visible light. In this case, the extinction cross-section σ_{ext} of a particle is twice its projected cross-sectional area A_p and

$$\beta_{ext} = \int_0^\infty n(r)\sigma_{ext}(r)\, dr \cong \int_0^\infty n(r)\, 2A_p(r)\, dr = 2\, A_{pm}\, N = \frac{3\, V_m\, N}{2\, r_{eff}} = \frac{3\, V_{tot}}{2\, r_{eff}} = \frac{3\, CWC}{2\, \rho_w\, r_{eff}}\ .$$

Using the effective radius r_{eff} we can determine the extinction coefficient β_{ext} and in turn this determines the optical thickness τ. To completely describe the scattering processes, two further quantities are needed: the scattering coefficient β_{sca} (or equivalently the *single scattering albedo* $\omega_0 = \beta_{sca}/\beta_{ext}$ if the extinction coefficient is known) and the *scattering phase function*. The single scattering albedo ω_0 ranges between 0 (pure absorption, no scattering) and 1 (pure scattering, no absorption) and indicates the relative importance of the scattering and absorption processes. The scattering phase function $p(\Omega', \Omega)$ describes the probability that a photon coming from direction (i.e., from the solid angle) Ω' is

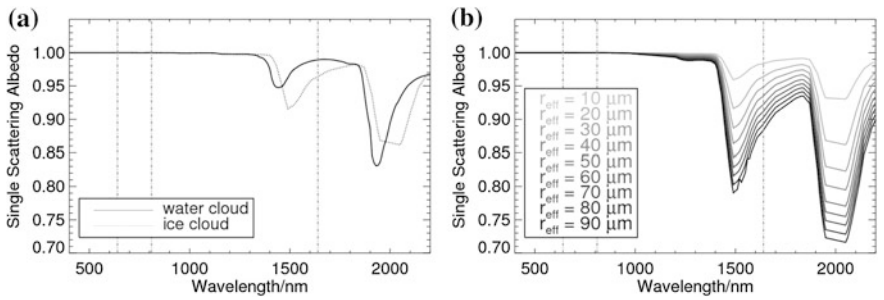

Fig. 25.1 **a** Single scattering albedo of water droplets and ice particles of effective radius 20 µm as a function of wavelength. **b** Single scattering albedo of ice particles of effective radius ranging from 10 to 90 µm as a function of wavelength. *Vertical dashed lines* correspond to the central wavelength of the first three solar SEVIRI channels VIS006, VIS008 and IR_016

scattered into the solid angle Ω. It shows very little dependence on shape and width of the cloud particle size distribution. Exceptions are represented by particular geometries like the *cloudbow* and the glory.

25.2.3 Spectral Properties of Clouds

The basic radiative processes of scattering and absorption, and thus the corresponding radiative quantities introduced in the last section, show a spectral dependency that can be used to characterize clouds from space. Figure 25.1a depicts the single scattering albedo of a distribution of liquid water droplets with effective radius 20 µm and the single scattering albedo of a distribution of ice crystals according to Baum et al. (2005a, b) with the same effective particle size. Both cloud types essentially scatter radiation in the visible spectral range and even beyond. Starting at 1.0 µm, the single scattering albedos decrease and the two curves for liquid water and ice behave differently, but both show two further minima: one close to 1.4–1.5 µm and one close to 1.9–2.0 µm.

The dependency of single scattering albedo on effective radius and wavelength is shown in Fig. 25.1b for ice clouds only. The larger the effective radius the smaller the single scattering albedo and consequently the higher the amount of radiation absorbed in two spectral intervals around 1.5 and 2.0 µm. In the visible spectral range all the curves coincide. Finally, comparing panel (a) with panel (b) in Fig. 25.1 one has to remember that an effective radius of 20 µm is rather large for liquid water clouds, with typical effective radii between 10 and 15 µm. For ice clouds instead, 20 µm represent a moderate size and a much stronger absorption can be observed in ice clouds with a growing effective radius (minimum $\omega_0 \cong 0.7$ for wavelengths around 2.0 µm).

Table 25.1 Met-8/SEVIRI channels: effective channel central wavelengths λ_c and spectral interval limits in μm, sampling distance in km, and gas absorption properties (data from Schmetz et al. (2002))

Channel	$\lambda_c/\mu m$	Spectral interval/ μm	Sampling distance/ km	Gas absorption	Absorbing gases
VIS006	0.64	0.56–0.71	3	Low	O_3, H_2O
VIS008	0.81	0.74–0.88	3	Low	H_2O
IR_016	1.64	1.50–1.78	3	Low	H_2O
IR_039	3.90	3.48–4.36	3	Low	H_2O, CO_2, CH_4, N_2
WV_062	6.25	5.35–7.15	3	High	H_2O
WV_073	7.35	6.85–7.85	3	High	H_2O
IR_087	8.70	8.30–9.10	3	Low	H_2O
IR_097	9.66	9.38–9.94	3	Medium	O_3
IR_108	10.80	9.80–11.80	3	Low	H_2O
IR_120	12.00	11.00–13.00	3	Low	H_2O
IR_134	13.40	12.40–14.40	3	Medium	CO_2
HRV	0.75	0.40–1.10	1	Low	O_3, H_2O

25.3 Meteosat Second Generation

The geostationary *Meteosat* Second Generation (MSG) program started on 28 August 2002 with the launch of MSG-1, renamed Meteosat-8 (Met-8) when it became operational, continued with MSG-2 (Met-9) on 21 December 2005, and will last at least until 2018 with two more satellites, MSG-3 and MSG-4. Met-8 and Met-9 have been operational since 29 January 2004 and 18 July 2006, respectively. One of the satellites is always in charge of the operational service at 0° E that provides full disc imagery every 15 min. The other serves as backup platform and/or for rapid scan service during which Europe and Northern Africa are observed with a repetition time of 5 min. The MSG satellites carry three instruments, the small communication payload Geostationary Search and Rescue GEOSAR, the Geostationary Earth Radiation Budget GERB and the main sensor, the Spinning Enhanced Visible and InfraRed Imager *SEVIRI*.

SEVIRI's 12 spectral channels, four in the solar and eight in the thermal range, extend from 0.63 to 13.4 μm and are listed in Table 25.1. All but one channel have a sampling distance of 3 km at the subsatellite point while the High Resolution Visible channel HRV has a sampling distance of 1 km, but lower spatial coverage. Thermal channels are calibrated on board while solar bands have no onboard *calibration* facility and must rely on cross-calibration with calibrated satellite sensors or on vicarious calibration procedures using radiative transfer computations for stable targets like deserts and deep convective clouds or well-known

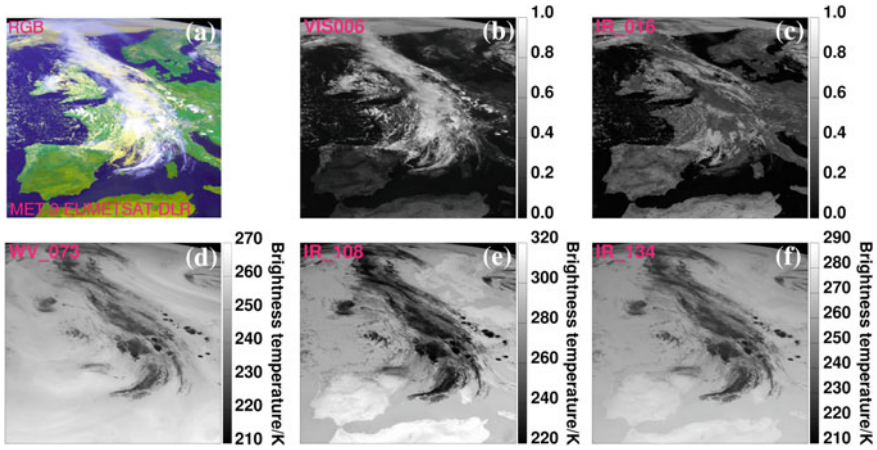

Fig. 25.2 Met-9/SEVIRI slot at 12:00 UTC 3 July 2008, Central Europe: **a** false color composite with channels VIS006, VIS008 and IR_108 inverted; **b–c** reflectivity in VIS006 and IR_016; **d–f** brightness temperatures in WV_073, IR_108 and IR_134

surfaces like the sea. Vicarious calibration for SEVIRI relies strongly on desert targets and is thought to be less accurate than thermal onboard calibration. Due to the Earth curvature and the slant view path of the sensor the spatial resolution degrades when moving away from the subsatellite point (located at the Equator) to higher or lower latitudes/longitudes. Over Europe, it amounts to approximately $3.5 \times 5\,\mathrm{km}^2$. From the measured *radiance* L it is customary to build reflectivities $R = \pi L/(E_0 \cos(\theta_0))$ (E_0 the *solar constant* at top of atmosphere, θ_0 the solar zenith angle) for solar channels and equivalent *brightness temperatures* T by inverting Planck's radiation law under the assumption of a *black body*. Reflectivities are a measure of the fraction of reflected solar radiation, and the equivalent brightness temperature is the temperature that the object observed would have if it were a black body and if the atmosphere had no influence on radiation. The characteristics of the different SEVIRI channels are illustrated in Fig. 25.2 showing a *front* passing over Central Europe: panel (a) shows a false color composite using VIS006, VIS008 and IR_108 inverted, panels (b) and (c) depict solar reflectivities in VIS006 and IR_016 respectively, while panels (d–f) represent equivalent brightness temperatures for the water vapor channel WV_073, the window channel IR_108 and the CO_2 channel IR_134. Please note that all clouds are bright in VIS006 (Fig. 25.2b), while high clouds are dark in IR_016 (Fig. 25.2c). In the visible range clouds mainly reflect according to their optical thickness, while ice crystals absorb radiation in IR_016 (Fig. 25.1). In WV_073 (Fig. 25.2d) only high clouds can be observed since they "emerge" from the underlying water vapor layers. Absorption of radiation by H_2O in IR_108 is very weak so one can observe relatively precise cloud top temperatures (the difference usually does not exceed 5 K) when clouds are thick (Fig. 25.2e). In fact, clouds can be most easily distinguished when their temperature clearly differs from that of

the underlying surface. The identification of low (warm) clouds or thin high clouds is more difficult. In IR_134 (Fig. 25.2f) absorption of radiation by CO_2 along the entire atmospheric column leads to a blurred picture with less pronounced land contours and colder temperatures than in IR_108. The main advantage of this channel is that it contains information about many layers of the *troposphere* due to the homogeneous vertical mixing of *carbon dioxide*.

Because of its spatial resolution and the above considerations, a single SEVIRI pixel is expected to often contain highly variable clouds, both in the horizontal and the vertical direction. In general, little information can be derived about the vertical cloud structure. Nevertheless, enough information is contained in the thermal channels due to the varying vertical *temperature profile* in the troposphere to apply various techniques to derive cloud top height. In contrast, solar channels usually show almost no or very small dependence on the vertical distribution of atmospheric gases and also on the vertical location of clouds. This is beneficial for cloud remote sensing since it reduces the amount of quantities affecting solar observations and enables the derivation of properties like cloud optical thickness (see below).

25.4 Ice Cloud Retrieval

As already hinted above, the detection of thin ice clouds is a challenging task since they are hardly visible both in the solar range where they reflect only very limited amounts of radiation and in the thermal range where they only slightly affect the observed brightness temperature due to their transparency. Thus, dedicated techniques have been developed in the course of the years. Particularly important is the so-called *split window* technique that considers brightness temperature differences between two thermal window channels caused by different ice extinction properties (see below). The determination of thin ice cloud optical properties using solar channels is ambitious as well since solar reflectivity is low and depends strongly in that case on surface albedo.

In the next two subsections the basic principles of IPA algorithms for cirrus cloud detection and the derivation of cloud optical thickness and effective radius are presented. Illustration of their outcome is based on the same time and area as in Fig. 25.2.

25.4.1 Ice Cloud Detection: MeCiDA

The Meteosat Cirrus Detection Algorithm MeCiDA (Krebs et al. 2007) and its improved version MeCiDA2 (Ewald et al. 2012) use exclusively SEVIRI thermal channels to identify cirrus clouds in SEVIRI imagery. MeCiDA is composed of six more or less independent tests and every one of them has been designed to detect both thin and thick cirrus clouds. MeCiDA detects a cirrus cloud when at least one of its tests provides a positive result. Every test uses either modified split window tests or morphological tests or a combination to identify cirrus-contaminated

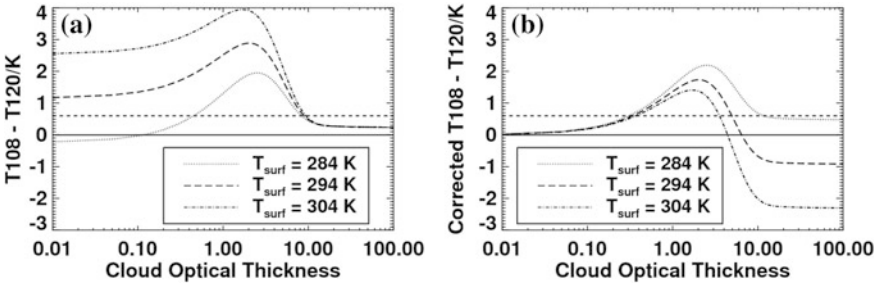

Fig. 25.3 **a** Brightness temperature differences between channels IR_108 and IR_120; **b** Corrected brightness temperature differences between channel IR_108 and IR_120. The threshold of 0.6 K above which this BTD test detects a cirrus clouds is plotted as a short *dashed line* in both panels. The three *curves* correspond to three different *surface temperatures* T_{surf}

pixels. Furthermore, simple temperature tests are employed to detect thick and cold cirrus clouds. Figure 25.3 illustrates the basic concept of the "classical" (see for instance the cloud detection scheme APOLLO (Kriebel et al. 2003)) and the modified split window test using the IR_108 and the IR_120 SEVIRI channels. First, consider Fig. 25.3a and the curve corresponding to the surface temperature $T_{surf} = 294$ K (long dashes). The tiny brightness temperature difference BTD between these two channels observed for optical thickness 0.01 is mainly due to the different absorption properties of water vapor in the atmosphere at these two wavelengths since cloud transmittance is very high. On the opposite side, for large optical thickness all the radiation emitted by the surface is absorbed within the cloud, and the cloud top emits radiation according to its own temperature and Planck's radiation law. Since *emissivity* is 1 in both channels, the very tiny difference observed is caused by the very limited amount of water vapor present in the *upper troposphere* and above. In between, one sees that BTDs first increase with optical thickness because of the increasing weight of the different single scattering properties. At the same time the cloud becomes more and more opaque to radiation and approaches the behavior of a black body for large τ. However, this behavior is strongly affected by surface temperature, as shown by the two additional lines plotted in Fig. 25.3a and it is not easy to select a threshold value for discriminating cirrus-free from cirrus-contaminated pixels. For this reason, corrected BTDs are introduced where the BTD of cirrus-free conditions is subtracted from the observed BTDs. In MeCiDA, this correction is extracted from the data itself by assuming that the warmest pixel in a given surrounding area corresponds to cirrus-free conditions. This approach minimizes the effect of surface, atmosphere, and viewing geometry on test performance. As demonstrated in Fig. 25.3b, all three curves now start close to zero for small optical thicknesses and a single threshold value of 0.6 K (short dashed line in Fig. 25.3a, b) can be adopted to identify thin to moderate ice clouds. This value has been determined by consideration of a comprehensive data set of radiative transfer calculations.

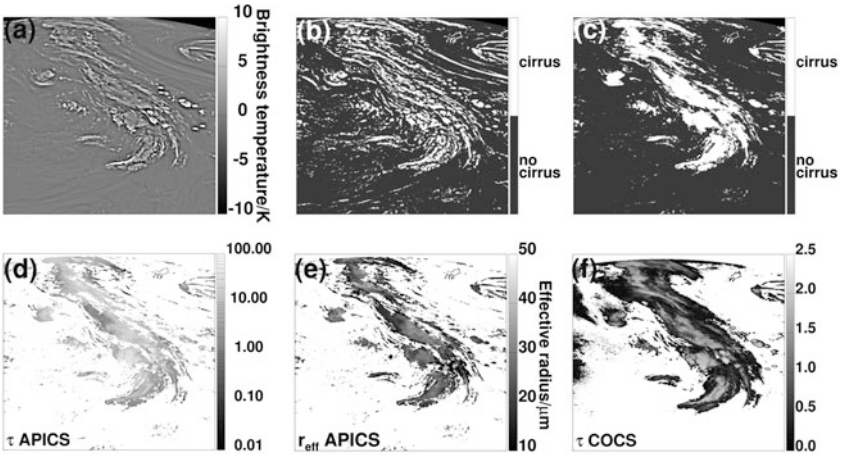

Fig. 25.4 Met-9/SEVIRI slot at 12:00 UTC 3 July 2008, Central Europe: **a** difference between the boxcar filter of T073 and T073; **b** result of the high-pass filter: cirrus is detected when the boxcar filter of T073 is larger than T073 by 0.5 K; **c** cirrus cloud mask obtained with MeCiDA; **d** cirrus optical thickness from APICS; **e** cirrus effective particle radius from APICS; **f** cirrus optical thickness from COCS

The morphological tests exploit the fact that in water vapor channels only high clouds are visible because the absorption of radiation by water vapor is so intense that low clouds and the surface do not affect the brightness temperature (see above). These high clouds are horizontally inhomogenous and contrast the smooth structures produced by water vapor in these spectral bands. High clouds can be extracted by applying appropriate high-pass filters. Figure 25.4a shows the first step of the implementation of one high-pass filter for WV_073: A moving average (also called boxcar filter) over an area of 19 × 19 pixels is applied to the original brightness temperatures (Fig. 25.2d) and the original image is then subtracted from it. Thresholding this difference at 0.5 K yields the cirrus mask in Fig. 25.4b. For morphological tests, thresholds are primarily selected by visual inspection. The combination of all single tests yields the desired cirrus mask in Fig. 25.4c.

25.4.2 Ice Cloud Optical and Microphysical Properties: APICS

An evaluation of two solar reflectivities for the quantitative determination of cloud optical thickness τ and effective particle radius r_{eff} is possible since solar satellite measurements only depend on these two quantities and only slightly on the shape and width of the particle size distribution. In turn, this means that the shape and width of the particle size distribution are not accessible. Standard assumptions made to derive τ and r_{eff} are the plane parallel and independent column assumptions. The first, the plane parallel assumption (PPA), states that clouds

inside one pixel are homogeneous and one-dimensional (1D). The second one, the independent column assumption (ICA), declares the net horizontal radiation exchange between pixels to be zero. While the error induced by the PPA increases with pixel size, as more and more natural cloud variability is neglected, the ICA error decreases with pixel size since shadows and cloud side reflections are washed out. Thus, there is optimal resolution in the range 1–5 km where the sum of both errors is minimal. This fits very well the physical restriction on the spatial resolution of passive satellite remote sensing instruments dictated by the smoothing of radiation inside clouds that also amounts to approximately 1 km.

The determination of optical thickness and effective radius is implemented by exploiting two SEVIRI channels, one in the visible spectral range where dependency on τ is largest and one in a spectral region where sensitivity to r_{eff} is given. To this end, the VIS006 and the IR_016 channels were selected (see vertical lines in Fig. 25.1). The former is independent of *vegetation* greenness, the latter is sensitive to both cloud scattering and absorption, i.e., to both τ and r_{eff}. A table of reflectivity values in both channels as a function of solar and satellite viewing angle, relative azimuth angle, surface albedo, optical thickness and effective radius has been precalculated with the 1D radiative transfer model *libRadtran* (Mayer and Kylling 2005). A model atmosphere has been selected with given vertical gas profiles and cloud bottom and top heights. The first three parameters (solar zenith angle, satellite zenith angle, relative azimuth angle) can be determined by astronomical and geometrical considerations; if the fourth one (albedo) can be extracted from an ancillary data set, then only two free parameters are left (τ and r_{eff}) which can be determined by comparing measured values in the above mentioned channels with tabled reflectivities. When good agreement is achieved, cloud properties have been found that enable accurate reproduction of the VIS006 and IR_016 measurements and thus represent a possible state of the atmosphere consistent with the two measurements used. This is the so-called Nakajima-King technique (Nakajima and King 1990) that has been implemented into the IPA Algorithm for the Physical Investigation of Clouds with SEVIRI, APICS (Bugliaro et al. 2011). Figure 25.4d and e show the optical thickness and effective radius derived from SEVIRI with APICS for the pixels previously identified as ice clouds by MeCiDA (Fig. 25.4c). The selected ice models are again those of Baum et al. (2005a, b) that currently represent the state-of-the-art.

25.4.3 Synergistic Retrieval: COCS

When the focus is exclusively on ice clouds the instrument of choice for their investigation is a lidar. Launched April 2006, the Cloud-Aerosol Lidar with Orthogonal Polarization CALIOP aboard NASA-CNES's Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations mission CALIPSO has gathered vertical profiles of clouds and aerosols on a global scale. This near-nadir viewing two-wavelength polarization-sensitive lidar is particularly sensitive to high thin cirrus clouds for which *cloud top height* and cloud optical thickness are derived by

Winker et al. (2009). In order to merge CALIOP's high sensitivity to cirrus properties with SEVIRI's high repeat cycle and extended view of Earth, an artificial neural network has been trained using CALIOP's cirrus properties and collocated SEVIRI thermal measurements. This data set contains eight million observations with cirrus clouds alone, cirrus clouds on top of water clouds, and cirrus-free conditions. As a result, the Cirrus Optical properties derived from CALIOP and SEVIRI (COCS) algorithm enables night and day derivation of cirrus cloud properties over the entire Earth disc using only SEVIRI brightness temperatures and brightness temperature differences. Of course, information about cloud top height is contained in passive thermal measurements due to the vertical temperature profile of the troposphere. The feasibility of the derivation of optical thickness for thin cirrus clouds from passive thermal observations emerges directly from Fig. 25.3. The use of all SEVIRI thermal channels corroborates the retrieval. Please note that a distinction in two steps like the above for MeCiDA and APICS is not necessary since cirrus detection and property determination take place at the same time. In addition, thermal measurements are thought to be more sensitive to ice crystals due to the absorption higher than in the solar spectral range, and at the same time the angular dependency of the physical processes involved (emission and scattering) is easier to account for. One further highlight of this approach is that only optical properties of the upmost cirrus layer can be derived, even if low clouds are present. The main drawbacks are the limited optical thickness range obtainable and the complicate dependency on surface temperature and atmospheric vertical temperature and water vapor profiles. The cirrus optical thickness resulting from COCS is shown in Fig. 25.4f. Please note the different range than in Fig. 25.4d and the larger extent of thin cirrus clouds (dark gray in the plot).

25.5 Validation

Validation of cloud properties is a difficult task. Even if cloud cover can be validated for instance against *synoptic* observations, this comparison is always hampered by different viewing geometries and perspectives. For cloud optical thickness and cloud effective radius the situation is even harder since detailed validation is impossible even with *in-situ* observations due to the nature of the investigated quantities. However, the technical progress of recent years has brought about the design and installation on board the DLR research aircraft *Falcon* of a High Spectral Resolution Lidar HSRL. This kind of instrument is very sensitive to cirrus clouds and is able to provide precise measurements of thin cirrus cloud optical thickness without any assumptions about particle size distribution and particle shape. This HSRL has been operated during the DLR project "Partikel und Zirrus" (PAZI) in order to investigate *contrail cirrus* properties, and also in support of the Lindenberg Upper Air Methods Intercomparisons LUAMI field campaign organized by the Meteorologisches Observatorium Lindenberg/Richard Aßmann Observatorium (MOL-RAO). The Falcon path on 18 October 2008 is

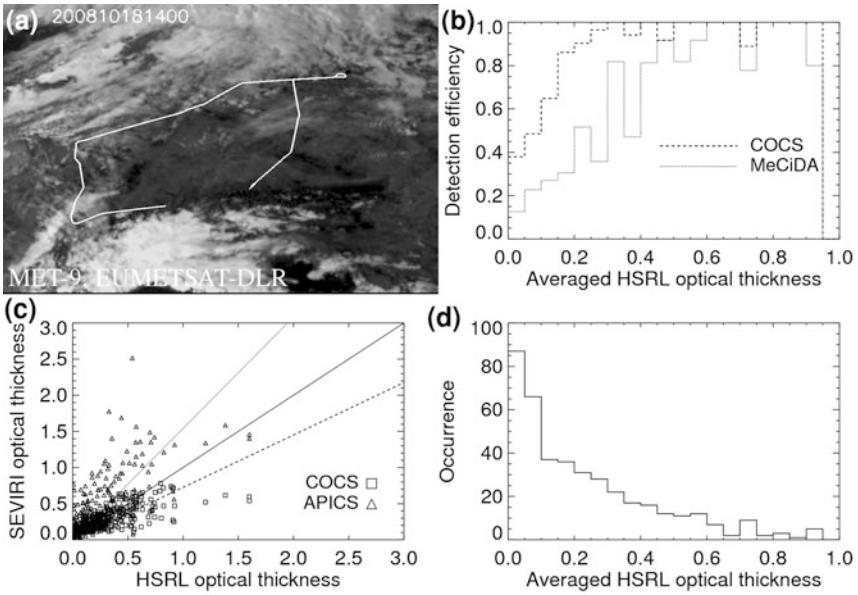

Fig. 25.5 **a** Met-8/SEVIRI slot at 15:15 UTC on 18 October 2008, Central Europe: VIS006 with Falcon path. **b** MeCiDA/COCS detection efficiency using the HSRL data collected during PAZI. **c** Scatter plot of HSRL optical thickness and APICS/COCS optical thickness. **d** HSRL data collected during PAZI that could be used in (**b**)

shown in the grayscale image of Fig. 25.5a as a thick white line. The aircraft first flew along the north side of the Alps; over Switzerland it deviated towards the North Sea and navigated over the Rhine Valley. Over northern Germany it then flew back towards Oberpfaffenhofen, passing over Lindenberg Observatorium. During the flight that started approximately at 12:35 UTC and ended at 16:15 UTC, various cirrus clouds above 6 km height could be measured. Their optical thickness was usually very low ($\tau < 1$) with some larger peaks up to approximately 3.0. Fortunately, during this time the rapid scan service of SEVIRI was active so that the 5 min temporal resolution could be exploited. This means that the largest acquisition time difference between lidar and imager amounted to at most 2.5 min. This greatly increased the probability that the two sensors were observing the same cloud. Nevertheless, the footprint of the lidar on the cirrus amounted to ca. 2 m when the Falcon was flying 200 m above the clouds, and its high repetition rate of 100 Hz together with an approximate flight speed of 720 km/h yielded a sampling distance of 2 m. To reduce noise the data were averaged over approximately 1 s such that one measurement was representative for ca. 200 m of the cloud. In order to make HSRL and SEVIRI data comparable, collocation of the two data sets was first performed, and finally mean HSRL optical thicknesses for single SEVIRI pixel were computed.

In the following, the focus is first on cirrus detection and then on the accuracy of cirrus optical thickness as provided by MeCiDA, APICS and COCS. More details can be found in Bugliaro et al. (2012), Kox et al. (2012).

25.5.1 Detection Efficiency

HSRL measurements are very well suited for detecting thin cirrus clouds since lidar sensitivity is very high, in particular much higher than that of the SEVIRI imager. However, the airborne lidar observations only cover a narrow strip inside every SEVIRI pixel. With this HSRL data base the detection efficiency of MeCiDA was determined as the ratio between the number of cirrus pixels correctly detected by MeCiDA to the number of available cirrus-covered SEVIRI pixels as classified by HSRL. This can be computed in various HSRL optical thickness intervals when the lidar measurement is not saturated and is plotted in Fig. 25.5b (dotted line). Figure 25.5d gives the distribution of HSRL optical thicknesses at the SEVIRI scale used in panel (b). One can observe that the sensitivity of MeCiDA to thin cirrus increases from 0.13 to 0.82 when optical thickness goes from 0.012 (the lower HSRL detection bound) to 0.4. For thicker cirrus the detection efficiency is very close to 0.82 or higher and reaches 1.0 for the first time at optical thickness 0.6. The fact that the population of the optical thickness intervals used here decreases rapidly explains the effect of a detection efficiency different from 1.0 for $\tau > 0.6$. Cirrus clouds of this optical thickness or larger occurred fewer than 10 times per interval along the Falcon flight path. Finally, the HSRL data set only contains 23 cirrus-free SEVIRI pixels, and all of them are identified as cirrus-free by MeCiDA as well, which yields a false alarm rate of 0.00.

For COCS, detection efficiency, determined the same way as for MeCiDA, is plotted in Fig. 25.5b as a dashed line. COCS's sensitivity to thin cirrus is much higher (almost 0.4 for $\tau \leq 0.05$) and reaches 1.0 already for optical thickness 0.3. For thicker cirrus, COCS's detection efficiency is between approximately 0.9 and 1.0. One pixel out of 23 cirrus-free SEVIRI pixels according to HSRL has been classified as cirrus covered by COCS.

25.5.2 Accuracy of Ice Cloud Optical Thickness

HSRL and SEVIRI algorithms are inherently different due to the different kind of measurements. In particular, with HSRL it is possible to isolate single cloud layers and derive their optical properties. For PAZI, all cirrus (i.e., ice) clouds above 6 km were evaluated with the HSRL and lower clouds could be neglected. For APICS, the situation is different since a cloud model is assumed that comprises only one cloud layer inside every SEVIRI pixel and this cloud is either a water or an ice cloud. Thus, situations like thin cirrus on top of water clouds cannot be treated satisfactorily and have to be identified and expunged from the comparison

data set. Furthermore, the low solar conditions encountered during the PAZI flight represent another challenge for the APICS algorithm since 1D radiative transfer models are usually inaccurate for solar zenith angles larger than ca. 75°. To reduce the impact of this effect, a spherical radiative transfer solver was used that should be accurate enough to evaluate pixels with solar zenith angles up to 85°. To further make the retrieval as realistic as possible, an *ozone* correction was introduced to account for ozone absorption in VIS006 (see Table 25.1) and a surface albedo for the SEVIRI pixels visited by the Falcon was derived from SEVIRI observations by considering the 20 October 2008 situation that presented cloud-free conditions.

The quantitative comparison of cirrus optical thickness from HSRL and APICS is shown in Fig. 25.5c (triangles). The correlation coefficient is 0.64 and APICS overestimates cirrus optical thickness by a factor 1.54, corresponding to the slope of the dotted regression line. Such behavior could be explained by the choice of the selected ice particle *habit* that represents a mixture of different ice crystal shapes. Although this *parameterization* is of course valuable for global investigations of ice cloud properties where many ice forms can be found at the same time, for the evaluation of a selected field campaign like this one this choice might not be the most appropriate one. In fact, variability of derived cirrus optical thickness due to different assumptions on ice particle habit is at least as large as the deviation observed between APICS and HSRL. One should be reminded, however, that ice crystal shape is not accessible with SEVIRI. Relating to the overestimation of τ, it is also found in comparisons of the ice optical thickness of MODIS against CALIOP, unfortunately not a HSRL. It is believed to stem from a too high forward scattering of the assumed ice particle phase functions.

Figure 25.5c also shows the cirrus optical thickness as derived by COCS (squares). As already mentioned above, a restriction to pure single cloud layer pixels is not necessary. Up to optical thickness 0.4–0.5, COCS reproduces very well the lidar based values. For larger optical thicknesses, COCS tend to underestimate HSRL and results in an overall underestimation by a factor 0.73 (slope of the dashed regression line). These two optical thickness regimes could be related to the twofold retrievals that enter the CALIOP products used for the COCS development. The "constrained" retrieval is applied to small optical thickness values and needs no assumption about the *lidar ratio*; the second one, the "unconstrained" retrieval, is applied to larger values and needs to assume some lidar ratio value.

25.6 Summary and Outlook

The long term tradition of satellite remote sensing of clouds at IPA has been applied to detailed quantitative evaluation of the most advanced geostationary space borne instrument to date, MSG/SEVIRI. It has been shown that it is possible to derive ice cloud cover and ice cloud optical thickness with satisfactory accuracy. Moreover, algorithms to derive further quantities like reflected shortwave radiation and

outgoing longwave radiation have been developed at IPA as well, and are available for a deeper characterization of clouds and their *climate impact*. Extensive time series are currently being processed in order to investigate the impact of *aviation* on cloudiness as well as for intercomparison with climate models.

In the near future, new satellites, both geostationary and polar orbiting, will carry new instruments and instrument combinations that will enable new and/or more detailed investigations of clouds. In particular, the Advanced Baseline Imager ABI aboard GOES-R, the Flexible Combined Imager FCI on Meteosat Third Generation MTG-I for geostationary platforms, the Visible/Infrared Imager VII together with the Multi-viewing Multi-channel Multi-polarisation Imager 3MI aboard EPS Second Generation, the Visible/Infrared Imager/Radiometer Suite VIIRS aboard NASA's NPP, and ESA's EarthCARE mission with a lidar, a *radar*, an imager and a broadband radiometer on the same platform all come to mind.

References

Baum, B., Heymsfield, A., Yang, P., et al.: Bulk scattering models for the remote sensing of ice clouds. Part 1: Microphysical data and models. J. Appl. Meteorol. **44**, 1885–1895 (2005a). doi:10.1175/JAM2308.1

Baum, B., Yang, P., Heymsfield, A., et al: Bulk scattering models for the remote sensing of ice clouds. Part 2: Narrowband models. J. Appl. Meteorol. **44**, 1896–1911 (2005b). doi:10.1175/JAM2309.1

Bugliaro, L., Zinner, T., Keil, C. et al.: Validation of cloud property retrievals with simulated satellite radiances: a case study for SEVIRI. Atmos. Chem. Phys. **11**(12), 5603–5624 (2011). doi:10.5194/acp-11-5603-2011

Bugliaro, L., Ostler, A. Wirth, M., et al.: Validation of cirrus optical thickness derived from SEVIRI with an airborne high spectral resolution lidar. Atmos. Meas. Tech. (2012, in preparation)

Ewald, F., Bugliaro, L., Mannstein, H., et al.: Improved cirrus detection for SEVIRI and validation with MODIS. Atmos. Meas. Tech. (2012) (submitted)

Kox, S., Mannstein, M., Bugliaro, L., et al.: A novel neural network approach for the retrieval of cirrus clouds from SEVIRI exploiting CALIOP. Atmos. Meas. Tech. (2012, in preparation)

Krebs, W., Mannstein, H., Bugliaro, L. et al.: Technical note: a new day- and night-time meteosat second generation cirrus detection algorithm MeCiDA. Atmos. Chem. Phys. **7**(24), 6145–6159 (2007). doi:10.5194/acp-7-6145-2007

Kriebel, K.-T., Gesell, G., Kästner, M. et al.: The cloud analysis tool APOLLO: improvements and validation. Int. J. Remote Sens. **24**(12), 2389–2408 (2003). doi:10.1080/01431160210163065

Mayer, B., Kylling, A.: Technical note: the libRadtran software package for radiative transfer calculations: description and examples of use. Atmos. Chem. Phys. **5**, 1855–1877 (2005). doi:10.5194/acp-5-1855-2005

Nakajima, T., King, M. Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory. J. Atmos. Sci. **47**, 1878–1893 (1990). doi:10.1175/1520-0469(1990)047<1878:DOTOTA>2.0.CO;2

Schmetz, J., Pili, P. Tjemkes, S., et al.: An introduction to Meteosat Second Generation (MSG). Bull. Am. Meteorol. Soc. **83**(7), 977–992 (2002). doi:10.1175/1520-0477(2002)083<0977:AITMSG>2.3.CO;2

Winker, D.M., Vaughan, M.A., Omar, A. et al.: Overview of the CALIPSO mission and CALIOP data processing algorithms. J. Atmos. Ocean. Technol. **26**, 2310–2323 (2009). doi:10.1175/2009JTECHA1281.1

Contrail Detection in Satellite Images

26

Hermann Mannstein, Margarita Vázquez-Navarro, Kaspar Graf, David P. Duda and Ulrich Schumann

Abstract

Methods for detecting linear contrail pixels in satellite infrared images are described. An objective contrail detection algorithm has been developed and extensively applied to data from various polar and geostationary satellite sensors. The method uses the contrast in brightness temperatures near 11 and 12 μm wavelengths and detects linear contrails using image processing techniques. The paper discusses the development of the algorithms, detection efficiency, false alarm rate, some of the results, and their validation. The contrail detection algorithm detects only a fraction of all contrail cirrus. Progress is expected from combining spatiotemporal satellite data in correlation with traffic and meteorological data.

H. Mannstein (✉) · M. Vázquez-Navarro · K. Graf
D. P. Duda · U. Schumann
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Hermann.Mannstein@dlr.de

M. Vázquez-Navarro
e-mail: Margarita.Vazquez@dlr.de

K. Graf
e-mail: Kaspar.Graf@dlr.de

D. P. Duda
Science Systems and Applications, Inc.,
Hampton, VA 23681, USA
e-mail: david.p.duda@nasa.gov

U. Schumann
e-mail: Ulrich.Schumann@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_26,

26.1 Introduction

Direct evidence for the impact of *air traffic* on *cloud cover* is given by the condensation trails, those linear cloud structures visible in all regions with air traffic. For the human eye and brain it is easy to identify these non-natural clouds both from ground and also in *satellite images*, at least in the early stages of their life-cycle, and the *aircraft* at the tip of a fresh *contrail* proves that these clouds are caused by aviation (Fig. 26.1). The identifying feature of a contrail is its linear shape caused by the high speed of the aircraft compared to the gradients of wind velocity. Most contrails are found in the *upper troposphere* above $\sim$8 km altitude at low *temperatures*. They consist of tiny *ice particles* and interact with the *radiation* field, similar to natural *cirrus clouds*. In contrast to lower level clouds the optical thin cirrus clouds have a net warming effect on the Earth-atmosphere system. The *absorption* of *thermal radiation* from lower atmospheric levels and the ground over the whole day is higher than the enhanced reflection of sunlight during daytime. In order to estimate the impact of aviation on climate via the pathway of contrail formation, a reasonable knowledge of their occurrence, coverage and properties is needed, which can either be inferred from measurements or derived from modeling. Satellite data provide the possibility to measure the extent, the optical properties and the impact on the radiation field of those contrails which can be discriminated from natural cirrus clouds by their shape: the linear contrails. The impact of the large remaining part of *aviation induced cloudiness* (AIC) has to be inferred by other means.

26.2 Historical Development

Satellite images and data give a more or less undisturbed view of contrails. Already in 1973, the first year of operation of the ERTS 1 (later renamed Landsat 1) satellite, Joseph et al. (1975) analyzed several single contrails over the Mediterranean south of Cyprus from two images with a nominal pixel resolution of 90 m and derived the *optical depth*, *particle number density* and *spreading rate*. In 1977 Kaminski and Krüger analyzed contrail occurrence over Germany by visual inspection of printouts of Very High Resolution Radiometer (VHRR) data of the National Oceanic and Atmospheric Administration NOAA-3, -4 and -5 weather satellites and estimated regionally up to 3 % coverage by anthropogenic clouds due to contrails over western Germany. They related the observed contrails to tropopause height and found that most contrails occurred when the *tropopause* was relatively high for this region (11–13 km). In their report they also referred to interviews with astronauts and cosmonauts, who stated consistently that the massive occurrence of contrails over western Europe was unique on the whole globe.

Bakan et al. (1994) analyzed contrail cloudiness over Europe and the eastern part of the North Atlantic Ocean for two periods (September 1979–December 1981 and September 1989–August 1992) by visual inspection of photographic prints of

Fig. 26.1 Young, linear contrails at different altitudes, seen from *below*. The inlet enlarges the four-engine jet aircraft forming the youngest contrail shortly behind the engines (*Photo* H. Mannstein)

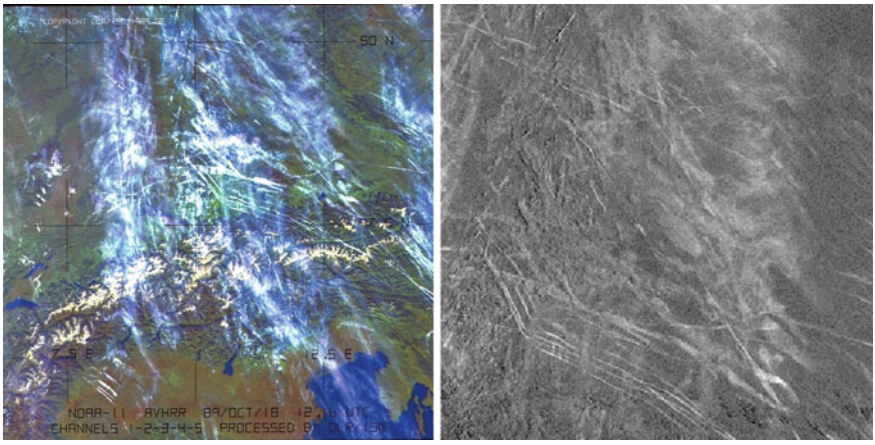

Fig. 26.2 *Left* False color NOAA-11 AVHRR image of a contrail cluster over southern Germany from 18 October 1989. *Right* Thin cirrus clouds and contrails are enhanced by the thermal *split window* channel difference between 11 and 12 μm (original 1 km resolution)

NOAA/AVHRR thermal infrared images received at the Dundee Satellite Receiving Station. Along the transatlantic flight corridor around 50°N they found maximum values of almost 2 %, and over western Europe 0.5 % contrail cover on average. In this study contrails defined by length and width as well as 'contrail areas' are

considered for the estimate of contrail cloudiness. A strong yearly cycle appears with a maximum (<2 %) in spring and summer over the Atlantic and a smaller maximum (<1 %) in winter over southwestern Europe. This study was for many years the basis for model calculations of the climate impact of linear contrails.

With the development of interactive digital image processing systems, contrast stretching techniques and channel combinations like the *brightness temperature* difference between the 11 and 12 μm channel (Fig. 26.2, right) became available. In particular this channel difference enhances the possibility to detect contrails and all other optically thin cirrus clouds composed of small ice particles. In the APOLLO cloud retrieval scheme (Saunders and and Kriebel 1988) and following satellite retrieval methods for clouds from weather satellites it is used as a test for thin cirrus.

26.3 The DLR *Contrail Detection* Algorithm

The visual analysis of satellite images is highly subjective and very time consuming. A pixel based distinction between contrail and surrounding area, which is a requirement for the measurement of impact on the radiation field, is nearly impossible. Therefore, at the DLR Institute of Atmospheric Physics an operational and objective contrail detection algorithm was developed (CDA, Mannstein et al. 1999) and first presented at several conferences in 1997. The DLR-CDA is the result of a development lasting several years with testing of different methods. A first version was described in Schumann and Wendling (1990). It was developed in order to identify without human interference pixels belonging to linear contrails in data from the Advanced Very High Resolution Radiometer (AVHRR) carried by operational NOAA polar orbiting weather satellites. In addition to the thermal infrared window channel at 11 μm, which was used in prior visual analyses, the CDA also used the brightness temperature difference, which was available with the AVHRR/2 instruments first flown on NOAA-7, launched 24 August 1981. The AVHRR instrument has a sampling distance of ∼1 km both in along track and—at nadir—in cross track direction, but this full resolution data was only available within the reach of a receiving station. Therefore, contrail studies based on AVHRR data were restricted to areas covered by those satellite receiving stations which kept a data archive.

The algorithm detects contrails based on spectral and morphological information. Figure 26.3 shows model calculations for different instruments of the response of the brightness temperature difference to ice clouds with increasing optical depth in a cloud-free summer *standard atmosphere*. The algorithm uses the experience gained from visual interpretation aided by interactive image processing techniques. To avoid differences between daytime and nighttime observations, only information from thermal infrared channels is used. All data is converted into equivalent blackbody temperatures according to the instrument characteristics. Single pixel errors are eliminated, missing or corrupted lines as well as the image margins are marked and not considered in the results.

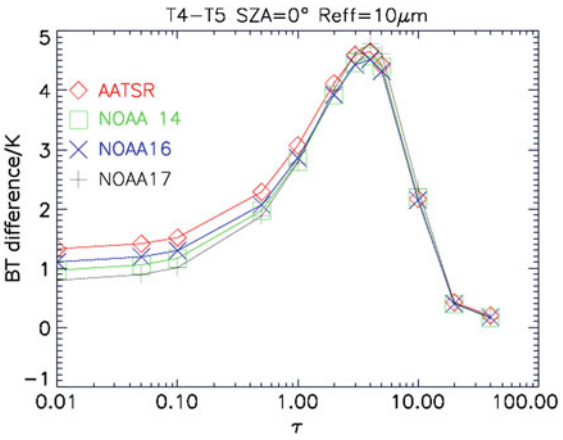

Fig. 26.3 Model calculation of the *brightness temperature* difference between the 11 and 12 μm channel for several sensors as a function of cirrus (and contrail) optical depth embedded in a cloud-free standard atmosphere. The effective radius of the ice particles is assumed to be 10 μm

The main CDA processing steps involve:

- a gradient filter applied to the temperature image in order to reject linear edges of clouds and surface features,
- an adaptive high pass filter, which mimics the ability of the human eye to detect even faint lines on a homogeneous background,
- a linear combination of the filtered temperature and the temperature difference images,
- a directional filter applied for several directions to the filtered data in order to identify elongated objects,
- a series of tests for each detected object in order to discriminate between contrails and natural features.

In order to estimate the area with *potential contrail cover*, the areas between adjacent contrails were filled (Fig. 26.4, right). With this method it was found that in central Europe in 15 % of space and time the atmospheric conditions at the main *flight levels* are suited to the formation of *persistent contrails*.

Meanwhile, the CDA has been adapted and applied to data from other satellite instruments like Moderate Resolution Imaging Spectrometer (MODIS), Along-track Scanning Radiometer (ATSR-2) on ESA-ERS2, and Advanced Along-Track Scanning Radiometer (AATSR) on ENVISAT, and to the Spinning Enhanced Visible and Infrared Imager (*SEVIRI*) instrument on the *geostationary* weather satellites *Meteosat*-8 and -9. The modifications did not change the general structure of the algorithm. An adaptation to new sensors was possible with changes of parameters and the proper selection of input data. The basic structure of the CDA has to be tuned by the variation of $\sim$20 parameters like filter sizes, number of different directions, thresholds for the object identification, and also for the tests. The full description of the algorithm and its tuning for the NOAA-14 AVHRR is given in Mannstein et al. (1999).

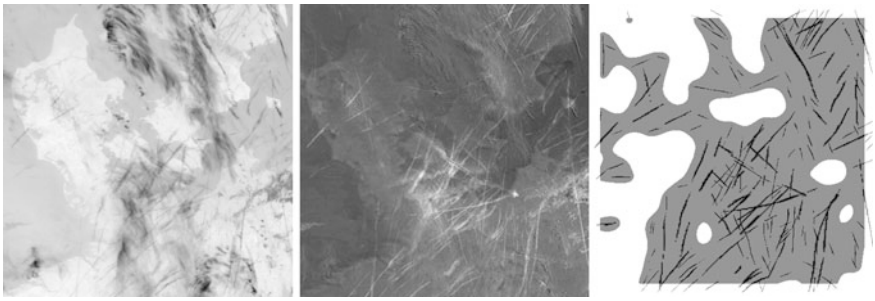

Fig. 26.4 Input data and result of the CDA: brightness temperature at 11 μm (*left*), temperature difference between 11 and 12 μm (*middle*) and detected contrails (*right*). (from Mannstein et al. 1999, reprinted by permission of the publisher, Taylor and Francis Ltd., http://www.tandf.co.uk/journals)

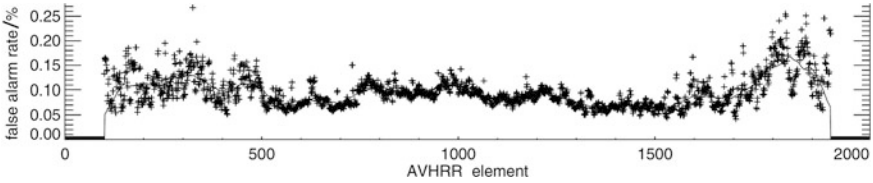

Fig. 26.5 *False alarm rate* as a function of the position within the scan line derived from 42 scenes (~200 000 lines) of NOAA-14 AVHRR data over the Southern Pacific (New Zealand), April 1997

26.4 Validation of the Contrail Detection Algorithm CDA

The aim of tuning the CDA was always to reach a very low *false alarm rate* (FAR) combined with sufficient *detection efficiency* (DE). The mean FAR can be assessed without human interpretation, because large regions on Earth have no air traffic and thus no linear contrails should be detected in these areas. For NOAA-14 AVHRR, a FAR of 0.1 % (Fig. 26.5) was derived. In a later study based on aviation data for the year 2004, it was realized that FAR and DE differ substantially for the various AVHRR instruments (see Table 26.1). The main reason seems to be the varying misalignment between the two *split window* channels on the different instruments and differences in the *spectral response*. As 90 % of the false alarms are elongated cirrus clouds, FAR increases with cirrus cover.

A much harder challenge is determining detection efficiency. The true coverage by linear contrails is unknown for several reasons: there exists no conclusive definition for '*linear contrails*,' no accepted lower limit for the effective optical depth, no lower limit for the length and no upper limit for the width of a contrail. Therefore, the DE had to be derived from comparing to visual interpretations of satellite images or to ground-based observations, which cover only small regions.

Table 26.1 False alarm rate (in %) of the CDA applied to AVHRR data from different satellites

Satellite, NOAA-	12	15	16	17
All	0.140	0.205	0.099	0.210
Day	0.144	0.188	0.114	0.218
Night	0.131	0.259	0.090	0.201
Best estimate	0.130	0.180	0.085	0.195

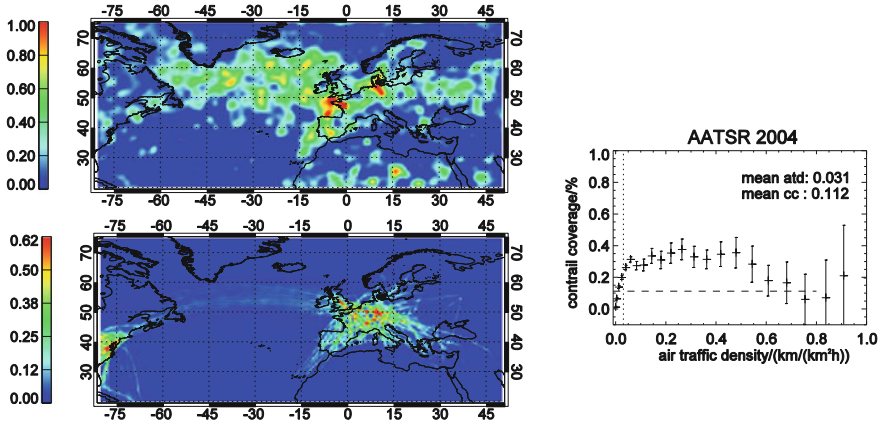

Fig. 26.6 Contrail coverage (in %) observed in AATSR data 2004 (*left top*), air traffic density (in km^{-1} h^{-1}) one hour before the satellite observation (*left bottom*), and contrail coverage as a function of air traffic density (*right*). The mean value of the false alarm rate is 0.10 % (*dashed line right*)

Figure 26.7 compares the result of two independent observers deriving the contrail cover of 50 scenes by assigning length and width to the detected contrails. The mean values differ by a factor of 2, but for single scenes much higher differences show up. There are similar differences between individuals if the analysis is done by marking the contrails and using different information sources. An extreme example is given in Fig. 26.8. Here we compare the visual analysis of an image obtained from channel combination of the MEdium spectral Resolution Imaging Spectrometer (MERIS), which provides a spatial resolution of 260 × 300 m^2 in the visible and near infrared range to the Advanced Along Track Scanning Radiometer (AATSR) temperature difference between the 12 and 11 μm channel, and to the results of the CDA. It is obvious that the temperature difference at 1 km resolution shows more and wider contrails. The DE of the CDA is higher over homogeneous background like ocean. In very heterogeneous areas like the Alps contrails cannot be detected by this method (Fig. 26.11). A comparison of contrail cover to air traffic density measured in flight path per area and time shows that the DE saturates at an air traffic density of 0.2 km/(km^2h) (Fig. 26.6, right). At higher air traffic density, the overlap of single contrails restrains the detection.

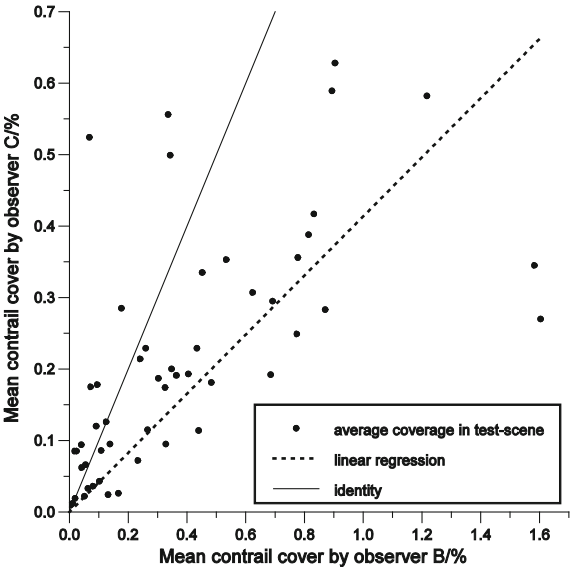

Fig. 26.7 Contrail coverage in 50 AVHRR scenes. Two independent observers assigned length and width to the detected contrails

Another source of information can be obtained from ground-based sky cameras. In Mannstein et al. (2010) we find for the CDA a DE of 8 % of all contrails visible from ground and 28 % of contrails with a width between 1 and 5 km that are also visible in the AVHRR satellite data. Based on this analysis we conclude that the DE for contrails (not for contrail cover) in Bakan et al. (1994) should have been about 20 %. The CDA has also been adapted to use SEVIRI data from the geostationary Meteosat satellites. For this data we find a DE of 9 %, but due to the much lower spatial resolution (4×6 km^2) only 15 % of the contrails visible from ground could also be identified in the temperature difference images. A relatively high DE (58 % of the detectable contrails) was only possible by accepting a high false alarm rate in the order of 1 %.

26.5 Results

With the strong increase of local data handling and computing capacity in the 1990s it became possible to evaluate the large amounts of satellite data which are necessary for statistically significant results. The first studies used the NOAA AVHRR data received locally at DLR's remote sensing data center at Oberpfaffenhofen (Figs. 26.9, 26.10 and 26.11) (Meyer et al. 2002). The spatial distribution of contrails is closely related to airways (Fig. 26.11), which indicates that mainly young contrails with an age less then 1 or 2 h are detected by the CDA. This is also supported by the mean width (1.8 km) of the detected linear contrails. At the same time the half-life of cirrus clouds originating from convective events was found to be $\sim$2.5 h, which led to the conclusion that only a small part

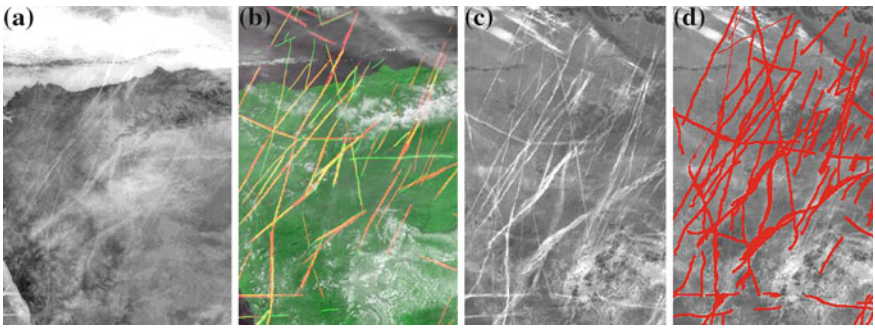

Fig. 26.8 **a**–**d** (from *left* to *right*) A *contrail outbreak* over Spain observed by MERIS (**a** and **b**) and AATSR (**c** and **d**) both on ENVISAT. **a** Normalized difference between near infrared and visible MERIS data in a pixel resolution of 260 × 300 m^2; **b** MERIS false color image with an overlay of visual/manual analysis of contrails (*green*), the result of the CDA applied to AATSR (*red*) and pixels marked by both (*yellow*); **c** AATSR split window temperature difference 12 to 11 μm; **d** Visual/manual analysis of the AATSR temperature difference image (*red*)

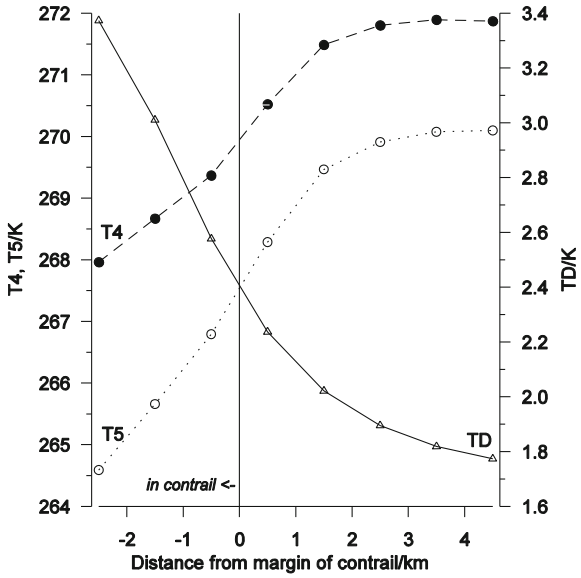

Fig. 26.9 Mean temperature (*left axis*) and temperature difference (*right axis*) of the 11 μm (T4) and 12 μm (T5) channel of NOAA-14 AVHRR as a function of distance to the edge of all detected contrails

(∼10 %) of the total direct impact of aviation on cirrus cover and its radiative forcing is indicated by linear contrails as detected by the CDA.

The observation that in a tropical area (Thailand), despite a reasonable amount of air traffic, no contrails were visible during daytime triggered a study based on AVHRR data obtained from the Asian Institute of Technology in Bangkok and the University of Tokyo. These regional studies provided data for the successful validation of global model results (Meyer et al. 2007). In the ESA-funded project

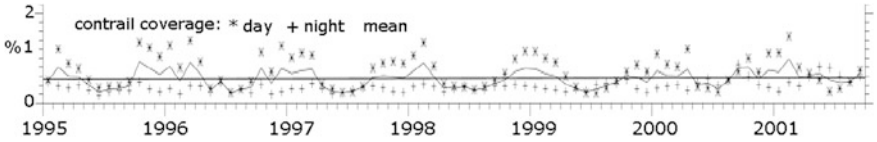

Fig. 26.10 Time-series of monthly mean contrail coverage derived from NOAA-14. Data refers to Central and Western Europe, 6°W–26°E, 42–54°N

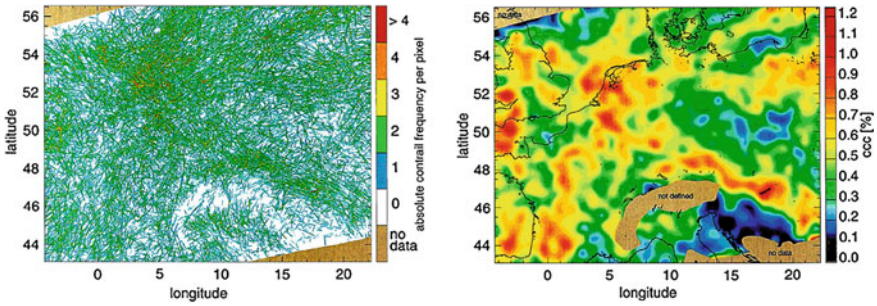

Fig. 26.11 Annual AVHRR-derived contrails (*left*) and heterogeneity-corrected contrail coverage (ccc) at noon derived from 357 NOAA-14 noon overpasses in 1996

'contrails' it became possible to analyze AVHRR data for the years 1985, 1990, 1995, 2000, and 2004 from different receiving stations together with ATSR2 and AATSR data for an area reaching from the North American east coast to Europe (80°W–50°E, 20–50°N, see Fig. 26.6). In spite of thorough analysis of FAR and DE for the different sensors, no conclusive trend of contrail cover for this area, which includes one-third of the global air traffic, could be given. Figure 26.12 summarizes the results of this study.

Further increases in computing capacity in the early twenty first century along with the availability of global coverage from MODIS now allow for near-global analysis of contrails from satellite imagery. The group lead by Pat Minnis at NASA Langley used the DLR CDA for different contrail studies over the USA (Minnis et al. 2005; Palikonda et al. 2005), and recently modified the CDA for MODIS data to produce a one-year climatology of contrail cover over the northern hemisphere. In addition to changes in the CDA parameters necessary to adapt the CDA to the MODIS sensors, the modified CDA uses additional thermal infrared channels (6.8, 8.6, and 13.3 μm channels) on MODIS to develop a series of tests that help screen out lower cloud streets and surface features. Fourier-transform-based filtering and remapping the raw imagery to a common map were added to account for multiple detector striping and other imaging issues with MODIS, and to extend the range of the CDA to larger satellite viewing angles (Fig. 26.13). To verify the contrail cover determined from the modified CDA, the study uses a

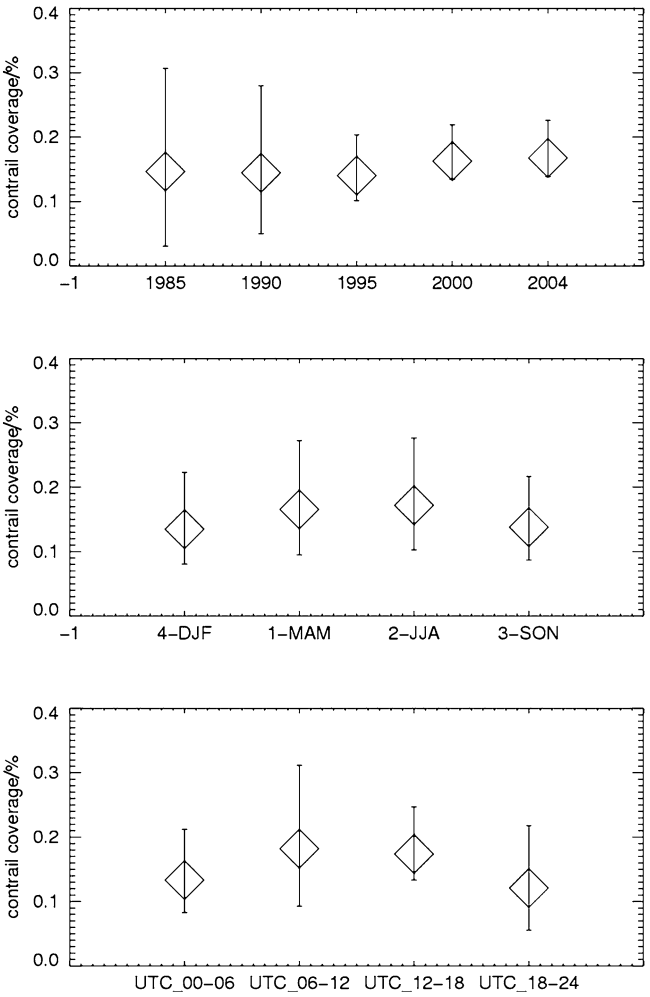

Fig. 26.12 Mean contrail coverage for different years (*top*), seasons (*middle*), and times (*bottom*) derived from statistically-corrected AVHRR and ATSR data for the area between 80°W–50°E, 20–50°N. The whiskers indicate a high and low estimate according to uncertainties in false alarm rate, detection efficiency and data amounts of the different sensors

comprehensive commercial aviation flight track database of northern hemisphere flights supplied by the U.S. Federal Aviation Administration (FAA), along with additional wind information from *numerical weather prediction* to account for contrail advection and spreading.

With some modification, the CDA has also been applied to data from SEVIRI on the geostationary Meteosat satellite. SEVIRI provides the infrared channels necessary for contrail and cirrus detection at a temporal resolution of 5 min and a spatial resolution of 3×3 km^2 at nadir. Due to the coarse spatial resolution many

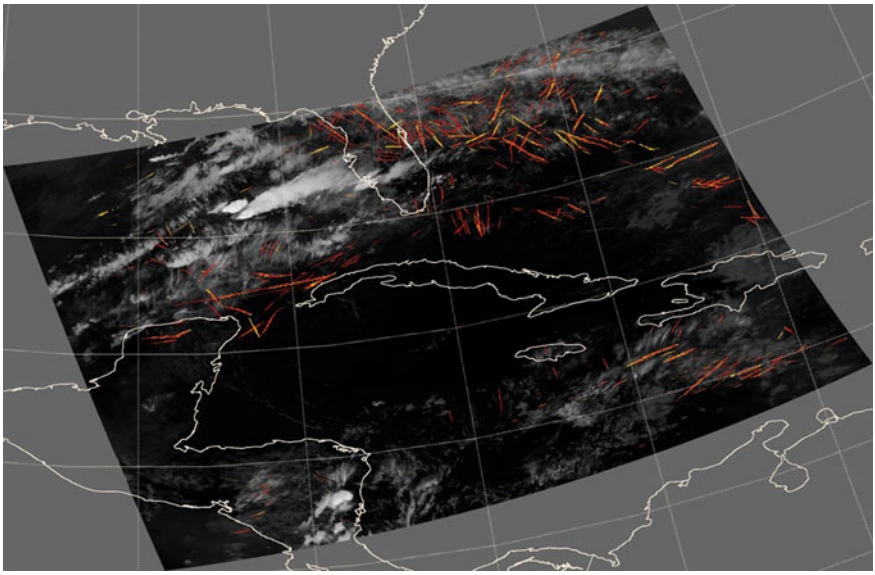

Fig. 26.13 Contrail detection in MODIS data. The *yellow* and *red* pixels indicate contrail pixels found with a CDA method with different sensitivity

natural cirrus clouds with a linear appearance are misclassified as contrails, resulting in a FAR of ~ 1 %. Nevertheless, this data shows a strong daily cycle over the Atlantic which follows the air traffic with a delay of ~ 2 h (Fig. 26.14). The high rate of misdetections can be avoided by using satellite data with higher resolution to identify the contrails.

The Automatic Contrail *Tracking Algorithm* (ACTA) combines the good spatial resolution of polar orbiting satellites with the good temporal resolution of geostationary sensors. It starts by using the results of the CDA on images obtained from a polar orbiting sensor such as MODIS and tracks the identified contrails in 5 min steps on the Meteosat-SEVIRI rapid scan images. Because the structures are already identified as contrails, the constraints for their identification can be relaxed substantially. As a result, a good part of the *lifetime* of the contrails can be observed. Figure 26.15 shows an example of a contrail tracked for several hours at the Bay of Biscay. Without the temporal information, the identification of this cloud as *anthropogenic* would have been impossible in the late stages, because by then it had completely lost its trademark linear shape. The frequency distribution of the lifetime of all contrails tracked over Europe during almost a year is shown in the right part of Fig. 26.15. It can be seen that variability is very high. Most contrails can only be identified for less than 1 h. The distribution of longer living contrails seems to follow an exponential decay with a half-life of 2–3 h. Due to the good spectral coverage of SEVIRI it is possible to estimate the radiative forcing of the detected contrails by comparison to the surroundings. The tracked contrails show a mean net *radiative forcing* of 5.9 ± 0.06 W/m^2, at daytime a cooling of -20.5 ± 0.4 W/m^2 over water

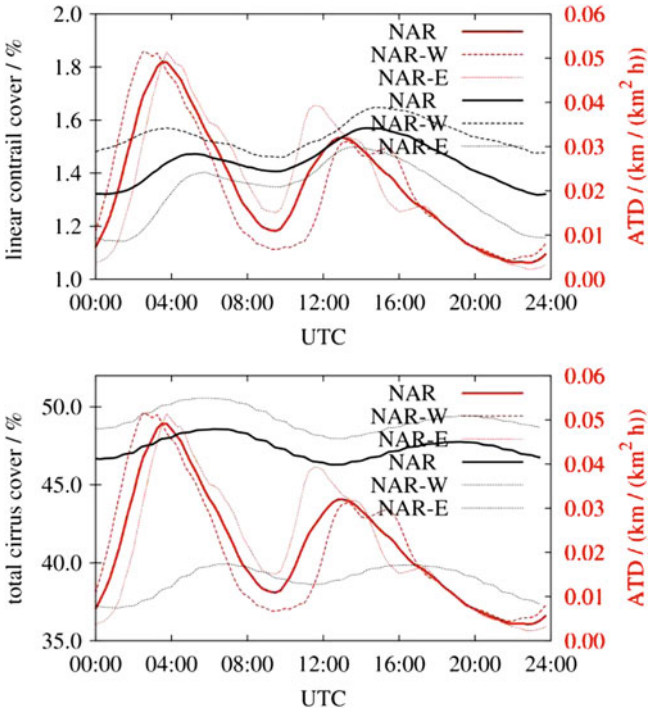

Fig. 26.14 Mean *diurnal cycle* of linear contrail coverage (*top*) and cirrus coverage (*bottom*) over the North Atlantic region. The contrail coverage detected by the CDA follows the air traffic density (*red*) with a delay of ∼2 h, the cirrus coverage with a delay of 3–4 h. Partition of the data in an eastern (NAR-E, *thin*) and a western (NAR-W, *dotted*) half part supports the hypothesis that air traffic is responsible for the timing of this daily variation

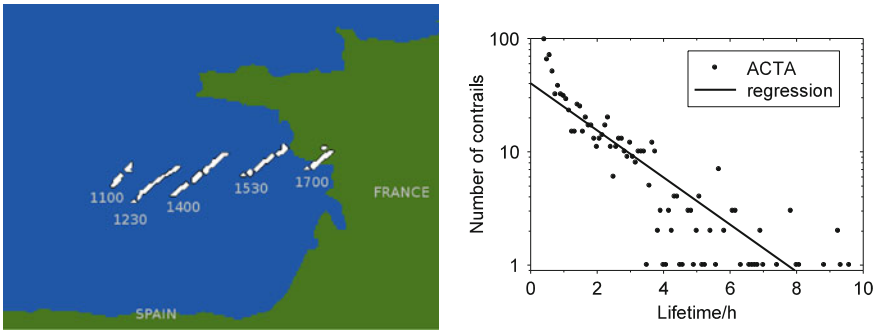

Fig. 26.15 *Left* A single contrail tracked over the Bay of Biscay for more than 6 h (at given UTC times) within Meteosat SEVIRI rapid scan data. *Right* number of contrails out of all contrails tracked with the ACTA algorithm over Europe during a year and regression line versus *lifetime*

and -13.2 ± 0.4 W/m^2 over land, while at night the warming amounts to 23.2 ± 0.2 W/m^2 over water and to 22.3 ± 0.4 W/m^2 over land.

26.6 Conclusions

Contrail detection from satellite data determines a lower limit for the changes in cirrus cover. The term 'detectable under optimal conditions at the spatial resolution of the sensor' describes the working definition for 'linear contrails' used in the contrail studies reported here. These linear contrails represent only the tip of the iceberg. About half of the fresh contrails visible from ground cannot be detected in weather satellite data of ~ 1 km spatial resolution. Aged contrails, which lose their linear shape and become more and more similar to natural cirrus clouds can be tracked in data from geostationary satellites with a high temporal resolution, but this still does not cover the whole life cycle of contrail cirrus. Here, *fingerprint* methods which combine spatiotemporal satellite data with traffic and meteorology data must be employed (Graf et al. 2009).

References

Bakan, S., Betancor, M., Gayler, V., Grassl, H.: Contrail frequency over Europe from NOAA satellite images. Ann. Geophysicae **12**, 962–968 (1994)

Graf, K., Mannstein, H., Mayer, B., Schumann, U.: Some evidence of aviation fingerprint in diurnal cycle of cirrus over the North Atlantic. In: Proceedings of the 2nd International Conference on Transport, Atmosphere and Climate (TAC-2), pp. 180–185. DLR-FB 2010-10, ISSN 1434-8454, Aachen, 22–25 June 2009

Joseph, J.H., Levin, Z., Mekler, Y., Ohring, G., Otterman, J.: Study of contrails observed from ERTS I satellite imagery. J. Geophys. Res. **80**, 366–372 (1975)

Mannstein, H., Meyer, R., Wendling, P.: Operational detection of contrails from NOAA-AVHRR data. Int. J. Remote Sens. **20**, 1641–1660 (1999)

Mannstein, H., Brömser, A., Bugliaro, L.: Ground-based observations for the validation of contrails and cirrus detection in satellite imagery. Atmos. Meas. Tech. **3**, 655–669 (2010). doi: 10.5194/amt-3-655-2010

Meyer, R., Mannstein, H., Meerkötter, R., Schumann, U., Wendling, P.: Regional radiative forcing by line-shaped contrails derived from satellite data. J. Geophys. Res. **107**, ACL 17-11–ACL 17-15 (2002). doi: 10.1029/2001jd000426

Meyer, R., Buell, R., Leiter, C., Mannstein, H., Pechtl, S., Oki, T., Wendling, P.: Contrail observations over Southern and Eastern Asia in NOAA/AVHRR data and comparisons to contrail simulations in a GCM. Int. J. Rem. Sens. **28**, 2049–2069 (2007). doi: 10.1080/01431160600641707

Minnis, P., Palikonda, R., Walter, B.J., Ayers, J.K., Mannstein, H.: Contrail properties over the eastern North Pacific from AVHRR data. Meteorol. Z. **14**, 515–523 (2005). doi: 10.1127/0941-2948/2005/0056

Palikonda, R., Minnis, P., Duda, D.P., Mannstein, H.: Contrail coverage derived from 2001 AVHRR data over the continental United States of America and surrounding areas. Meteorol. Z. **14**, 525–536 (2005). doi:10.1127/0941-2948/2005/0051

Saunders, R.W., Kriebel, K.T.: An improved method for detecting clear sky and cloudy radiances from AVHRR data. Int. J. Rem. Sens. **9**, 123–150 (1988)

Schumann, U., Wendling, P.: Determination of contrails from satellite data and observational results. In: Schumann, U. (ed.) Air Traffic and the Environment—Background, Tendencies and Potential Global Atmospheric Effects, pp. 138–153. Lecture Notes in Engineering. Springer, Berlin (1990)

Lidar for Aerosol Remote Sensing

27

Matthias Wiegner

Abstract

Measurements of aerosols are urgently required for understanding and modeling their role in the climate system and for investigating interactions between aerosols, clouds and radiation. Lidar (light detection and ranging) is an active remote sensing method for aerosol analysis which provides range resolved information. In this paper the different aerosol properties—geometrical, optical and microphysical—that can be derived from lidars are briefly described. In particular, the role of the so-called lidar ratio is discussed.

27.1 Introduction

Atmospheric aerosols not only influence the radiation budget, air chemistry and the hydrological cycle, but can also have an adverse influence on human health and the environment. Due to the large variety of aerosol sources and sinks and complex modification processes during their lifetime it is obvious that the optical and microphysical properties of the aerosols and their vertical layering may change in time and space. As a consequence, monitoring their distribution and characterizing their properties is a challenging task requiring advanced observation techniques. In particular, it is necessary to have a measuring system that can provided range resolved information.

M. Wiegner (✉)
Ludwig-Maximilians-Universität München (LMU), Meteorological Institute Munich (MIM), Theresienstraße 37, 80333 München, Germany
e-mail: m.wiegner@lmu.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_27,

Lidar (Light detection and ranging) is certainly the best suited experimental setup for achieving this goal. It has been used in atmospheric physics for more than 40 years, and aerosol *remote sensing* is one of the main applications. The main reasons are that the *optical depth* of aerosols is normally low enough to prohibit complete attenuation of the emitted radiation, and that lidar can provide data during daytime and nighttime.

For radiative studies in general three categories of parameters are used to characterize atmospheric aerosols: spatial distribution, and the optical and microphysical properties of the particles. The first category concerns the most basic set of information, i.e., the vertical aerosol distribution and the extent of elevated layers as a function of time. The second category includes extinction and *scattering coefficients* and information on the angular distribution of scattering (*phase function*, asymmetry parameter). For deeper understanding of the interaction of particles with radiation, microphysical properties such as particle size, shape, mass concentration and *refractive index* are useful. Several (but not all) of these parameters can be derived from lidar measurements. The achievable accuracy and resolution depend on the physical concept of the measurement: a distinction is made between "*backscatter lidars*," "*Raman lidars*" and "*high spectral resolution lidars*" (HSRL). *Polarimetric* measurements can be considered as an add-on. "*Differential absorption lidars*" are not considered, although in principle they can be used for aerosol remote sensing in the same way as backscatter lidars.

The main part of the paper will focus on different aerosol products that can be derived by lidar, starting with a brief introduction to the basics of lidar methodology and closing with a short overview of recent lidar applications.

27.2 Lidar Concept

The physical concept underlying a backscatter lidar is the emission of short *laser* pulses (at wavelength λ_0)—thus, it is called "active" remote sensing—and the time resolved detection of photons (of the same wavelength) backscattered from air molecules and particles (aerosols and clouds). Using the velocity of light, the time domain is transformed into the spatial domain. The spatial resolution is described as "range-bin" Δz. As the received signal depends on the amount and properties of the aerosols along the line of sight, lidar can in principle be used to characterize aerosol distribution.

The number of received photons from a range-bin at range z depends on the characteristics of the lidar system, e.g., the number of photons emitted per pulse and the length of the sampling interval, described by the lidar constant C_L. Furthermore, it depends on atmospheric conditions, i.e., on the efficiency of the backscattering of particles and air molecules at range z, and on the probability that emitted photons reach z and again the lidar after being backscattered, without being absorbed or scattered in arbitrary directions in between. Accordingly, one frequently used formulation of the basic equation of lidar remote sensing for a

monostatic (i.e., emitter and receiver are at the "same" place) backscatter lidar is given in Eq. 27.1. Note that the nomenclature in the literature is not standardized, and "the signal" may be expressed, e.g., in terms of number of photons, photoelectrons or power.

$$P(z,\lambda_0) = C_L \frac{\beta_m(z,\lambda_0) + \beta_p(z,\lambda_0)}{z^2} \exp\left\{ -2 \int_0^z [\alpha_m(z',\lambda_0) + \alpha_p(z',\lambda_0)]\, dz' \right\} \tag{27.1}$$

Here, $P(z,\lambda_0)$ is the number of backscattered photons from a range-bin at range z at wavelength λ_0, and C_L the above mentioned instrument constant. Backscatter and *extinction coefficients* are denoted as β (typically given in km^{-1} sr^{-1}) and α (typically in km^{-1}), respectively; subscripts p and m are added to distinguish between the contributions of particles and molecules. The exponential term is the two-way transmission of the atmosphere between the lidar and range z describing the above mentioned probabilities. Eq. 27.1 implicitly means that all photons that contribute to the signal, do not experience any further scattering between the lidar and z. Therefore, this equation is also referred to as the lidar equation in single scattering approximation. Numerical simulations have shown that multiple scattering can indeed be neglected in case of aerosol observations from ground, however, in cases of *satellite observations* multiple scattering can contribute to the signal.

Most aerosol lidars apply Nd:YAG lasers with a fundamental wavelength of 1 064 nm. Using frequency doubling and tripling crystals, simultaneous emission at 532 and 355 nm can be obtained. The specifications of the receiver are often quite different as they are optimized for specific scientific applications and measurement conditions; note that many lidars are research oriented. The receiver field of view is typically 1 mrad (FWHM) and slightly larger than the beam divergence of the laser. The pulse repetition frequency is in the order of 10–100 Hz, the spatial resolution in the order of a few meters to a few tens of meters.

As already mentioned, measurements of tropospheric aerosols are among the main scientific objectives of lidar. For this purpose, measurements should cover the range from the surface to the tropopause, or to the stratosphere if volcanic aerosols are to be observed. The lower limit should be as close as possible to the surface to minimize the loss of information in altitudes where the aerosol concentration is maximum. The near end of the range is primarily determined by the overlapping of the receiver field of view and the divergence of the emitted beam. Typically, full overlap is realized after several hundred meters up to more than one kilometer. As a consequence, significant parts of the *boundary layer* aerosol are "invisible" for a lidar. To overcome this drawback there are several options: measurements under different zenith angles can be combined, an optical design with an adjustable field of view can be implemented, or two or more receiving telescopes with different fields of view can be used.

The far end of the measurement range is important even for boundary layer observations because the upper aerosol-free troposphere is used for calibrating the lidar signals (the so called "Rayleigh calibration" or "Rayleigh fit"). Rayleigh calibration is required because C_L is not known with sufficient accuracy.

27.3 Retrieval of Geometrical Properties

Monitoring the spatial and temporal distribution of aerosols is comparably easy. This becomes obvious after introduction of the transmission T (and the optical depth τ of the atmosphere)

$$T(z,\lambda_0) = e^{-\tau(z,\lambda_0)} = \exp\left\{-\int_0^z \alpha(z',\lambda_0)\, dz'\right\} \quad \text{with} \quad \alpha = \alpha_p + \alpha_m$$

The exponential term in Eq. (27.1) can then be expressed as the two-way transmission $T^2(z)$ and the range corrected signal $P(z)z^2$

$$P(z)z^2 \propto \beta^* = \beta(z)\; T^2(z) \tag{27.2}$$

is accordingly proportional to the attenuated backscatter β^*.

For the interpretation of Eq. 27.2 it is useful to discuss the transmission: $T^2(z)$ only slowly decreases with range z starting from $T^2(0) = 1$. Keeping in mind that optical depth typically decreases with wavelength, the decrease of $T^2(z)$ is the smaller the larger the wavelength is. For example, for $\lambda_0 = 1\,064$ nm and under typical atmospheric conditions, the two-way transmission remains larger than approximately 0.9, and β^* is a good approximation for β. According to the Angström approach

$$\alpha_p \propto \lambda^{-\kappa} \quad \text{and} \quad \alpha_m \propto \lambda^{-4.08} \tag{27.3}$$

with typically $\kappa < 2$ it becomes obvious that the influence of air molecules compared to aerosols is drastically reduced at long wavelengths. As a consequence, $\beta = \beta_m + \beta_p$ is dominated by β_p so that any spatial discontinuity in the aerosol distribution is clearly visible in β^*. Note that β_m is a smooth function with height.

As a result, the stratification of aerosols, e.g., the top of the *planetary boundary layer* or the residual layer, and the height and vertical extent of elevated layers, can easily be monitored from range corrected signals. Most state-of-the-art backscatter lidars include a detection channel at 1 064 nm so that a temporal resolution of a few minutes can be achieved; shorter wavelengths are less suitable but sufficient for most applications. An example is given in Fig. 27.1. Shown is the range corrected signal (arbitrary units, color coded) at 1 064 nm as derived at Maisach (25 km northwest of Munich) from the three-wavelength Raman *depolarization*

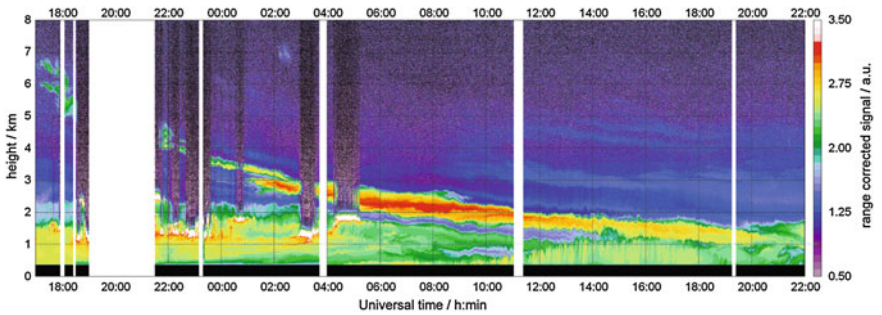

Fig. 27.1 Range corrected signal (log. scale, a.u.) at 1 064 nm versus height above mean sea level as derived from MULIS:, 17:00 UTC 16 April 2010 until, 22:00 UTC 17 April 2010, at Maisach (25 km northwest of Munich) showing the advent of the ash layer (see text) from the Eyjafjallajökull eruption. MULIS is the three-wavelength Raman depolarization lidar of MIM. The lowermost 350 m are influenced by the incomplete overlap (shown in *black*)

lidar MULIS of the Meteorological Institute at Munich (MIM) two days after the *eruption* of the Icelandic *volcano Eyjafjallajökull* (Wiegner et al. 2011). The aerosol stratification is immediately clear without any complex numerical *inversion* scheme: a pronounced layer appeared at about 7 km at 17:00 UTC 16 April 2010, and descended to about 4 km at midnight. The layer further descended but was still separated from the residual layer until the afternoon of 17 April. After 17:00 UTC mixing with the boundary layer took place. Figure 27.1 also shows faint aerosol structures throughout the free troposphere from 09:00 UTC 17 April. It should be emphasized that it is intuitively clear that the pronounced layer was the *ash* plume of the volcanic eruption, but from the range corrected signal alone there is no strict evidence.

The fact that backscattering at $\lambda_0 \approx 1\,\mu m$ is very well suited for monitoring aerosol distribution and transport led to an increasing awareness of lidar technology for the surveillance of atmospheric aerosols. It gave rise to the implementation of ceilometer networks; the German Weather Service recently set up a network of almost sixty *ceilometers* in Germany. Ceilometers are single-wavelength backscatter lidars with quite low laser pulse energies but higher pulse repetition rate. They are eye-safe and can thus be operated unattended and continuously. They are going to be used to monitor the temporal development of the boundary layer and the dispersion of aerosol plumes.

27.4 Retrieval of Optical Properties

For radiative studies the provision of range corrected signals (previous section) is not sufficient; quantitative evaluation of optical properties is conclusive. From Eq. (27.1) it is clear that the aerosol related *optical property* is either α_p or β_p. Equation (27.1) however includes four unknown meteorological variables—α_m,

β_m, α_p, and β_p. Thus, it is obvious that the quantity of interest cannot be retrieved without further information.

From Rayleigh theory it is known, that the optical properties of air molecules, $\alpha_m(z,\lambda)$ and $\beta_m(z,\lambda)$, can be calculated with high accuracy from air density profiles. Air density is determined according to the ideal gas law using profiles of pressure and temperature from *radiosonde* ascents. Therefore, they are considered as known in the lidar equation. Nevertheless, the lidar equation is underdetermined as two unknowns remain: α_p and β_p. This problem is formally solved by introducing a relationship between α_p and β_p, the so called *lidar ratio* S_p.

$$S_p(\lambda) = \frac{\alpha_p(\lambda)}{\beta_p(\lambda)} \tag{27.4}$$

Thus, the main problem of the inversion of the lidar equation is the need to find the correct lidar ratio; note that S_p depends on the aerosol type (see below).

27.4.1 Backscatter Lidar

A backscatter lidar is the least complicated and therefore by far the most frequently used configuration. It is based on elastic scattering under 180 degrees. The signals are described by the lidar Eq. (27.1) that can be solved analytically by the so called "*Klett method*" (Klett 1981; Fernald 1984). For α_p the solution reads (omitting λ_0 for simplicity):

$$\alpha_p(z) = \frac{Z(z)}{N(z)} - \frac{S_p(z)}{S_m} \cdot \alpha_m(z) \tag{27.5}$$

with

$$Z(z) = S_p(z) \cdot z^2 \cdot P(z) \cdot \exp\left\{ 2 \int_z^{z_{\mathrm{ref}}} \left[\frac{S_p(z')}{S_m} - 1 \right] \alpha_m(z') dz' \right\} \tag{27.6}$$

and

$$N(z) = \frac{S_p(z_{\mathrm{ref}})\ z_{\mathrm{ref}}^2\ P(z_{\mathrm{ref}})}{[S_p(z_{\mathrm{ref}})/S_m]\ \alpha_m(z_{\mathrm{ref}}) + \alpha_p(z_{\mathrm{ref}})} + 2 \int_z^{z_{\mathrm{ref}}} Z(z') dz' \tag{27.7}$$

It should be emphasized that analogously to Eq. 27.5 a solution for $\beta_p(z)$ can be found.

The solution for α_p requires a boundary value $\alpha_p(z_{\mathrm{ref}})$ at a reference height z_{ref}, as can be seen in Eq. 27.7 and the lidar ratio S_p (Eqs. 27.6 and 27.7). The reference

height is typically set to an altitude in the upper free troposphere where no aerosols are present. The reference height is found by comparing the measured signal and a modeled signal according to the Rayleigh atmosphere.

The lidar ratio of the air molecules S_m is known from Rayleigh theory and in good approximation wavelength-independent

$$S_m = \frac{8\pi}{3}\left(1 + \frac{\rho^n(\lambda)}{2}\right) \quad \text{with} \quad \rho^n = \begin{cases} 0.03016 & \text{at} \quad \lambda = 355\,\text{nm} \\ 0.02837 & \text{at} \quad \lambda = 532\,\text{nm} \\ 0.02719 & \text{at} \quad \lambda = 1064\,\text{nm} \end{cases}$$

To elucidate the complexity of S_p of the particles it is convenient to express it in terms of well-known parameters, the phase function p (normalized to 1) and the *single scattering albedo* ω_0 (ratio of scattering and extinction coefficient). Then

$$S_p(\lambda) = \frac{1}{\omega_0(\lambda) p(180^\circ, \lambda)} \tag{27.8}$$

As ω_0 and p depend on the refractive index m of the particles and their *size distribution* $n(r)$, it follows from Eq. 27.8 that S_p also depends on m and the particle size. Furthermore, the *microphysics* can change with height and *relative humidity*, thus, the same is true for S_p. However, most of this information is not available under normal conditions. As a consequence, the exact values of S_p along the lidar's line of sight are never known and one has to rely on plausible estimates. These can be based on model calculations or independent measurements.

In case of model calculations, the refractive index m and the radius r of the particle are required as input. For spherical particles the *Mie theory* is applied, but in cases of *nonspherical* particles more complex scattering models must be used and adequate parameters for the description of the shape must be defined. Depending on the shape and the size of the particle different approaches are available, e.g., the T-matrix method, discrete dipole approximation, the finite difference time domain method, or the geometrical optics approach. As already mentioned the microphysical properties and the meteorological conditions are unknown under normal conditions. Moreover, most of these models are computationally very expensive. As a consequence, the calculation of lidar ratios (and other aerosol properties) focusses rather on basic research than on routine application.

The alternative approach to get S_p-estimates relies on independent measurements (see next section). It has been found that the assignment of a type (e.g., dust, marine, biomass burning) to the aerosol (often referred to as "aerosol typing") is a feasible way to constrain the possible range of S_p. If the meteorological situation of the site and back trajectories are considered, aerosol typing is considered as sufficiently reliable. For wavelengths of 355 and 532 nm the lidar ratio typically is between 30 and 80 sr: for marine aerosols, lidar ratios are lower than 30 sr, for desert dust S_p is around 60 sr, and for biomass burning it is up to 80 sr. However, even if classification is possible, an uncertainty range of the order of ± 10 sr remains.

A consequence of the *uncertainty* of the lidar ratio estimate is uncertainty about the retrieved aerosol optical properties (α_p or β_p) in the order of up to 30 %. Therefore, α_p-profiles should always be treated with care and the assumption used for the S_p-estimate should be critically examined if backscatter lidars are deployed.

27.4.2 Raman Lidar

A *Raman lidar* (Ansmann et al. 1992) takes advantage of inelastic ("Raman") scattering of photons by molecules of known density, normally N_2. These photons are wavelength shifted and can be recorded separately from the elastically backscattered photons at the emitted wavelength. The wavelengths corresponding to emitted radiation at 532 and 355 nm are 607 and 387 nm, respectively. The height dependent Raman *backscatter coefficients* can be calculated from the known air density.

To exploit Raman lidar data, two equations are considered. The well-known equation for elastic backscattering at λ_0, Eq. 27.1, and a second equation for the corresponding Raman wavelength λ_R:

$$P(z,\lambda_R) = C_R \frac{\beta_R(z,\lambda_0)}{z^2} \exp\left\{ -\int_0^z \left[\alpha_p(z',\lambda_0) + \alpha_m(z',\lambda_0) + \alpha_p(z',\lambda_R) + \alpha_m(z',\lambda_R)\right] dz' \right\} \tag{27.9}$$

with β_R as the Raman backscatter coefficient. Equation 27.9 describes the signal at the Raman-shifted wavelength λ_R, which is governed by the transmission at λ_0 along the path from the lidar to the location of the scattering process and by the transmission at λ_R on the way back to the receiver. Equations 27.1 and 27.9 can be solved for the two unknowns $\alpha_p(\lambda_0)$ and $\beta_p(\lambda_0)$.

If more than one wavelength can be exploited, the wavelength-dependence of α_p can be used as a measure for the particle size. For this purpose the *Angström* exponent κ (see Eq. 27.3) is determined. As a rule of thumb, $0<\kappa<0.5$ indicates large particles (dust, humidified particles), whereas $1<\kappa<2$ is typical for small particles.

As the Raman lidar technique allows determination of α_p and β_p independently, it is possible to retrieve the lidar ratio. This is a fundamental difference to a backscatter lidar. Thus, Raman lidars are used to establish S_p-climatologies of different aerosol types, e.g., in the framework of the European aerosol research lidar network (EARLINET, see Bösenberg et al. 2003) and such dedicated *field campaigns* as the Saharan mineral dust experiment (SAMUM). These data sets can serve as estimate of the lidar ratio required for the Klett method.

As a consequence of the very low Raman backscatter coefficients, Raman lidars require powerful lasers, long averaging times of up to one or two hours, and very effective suppression of the solar background radiation. Therefore, most systems are restricted to nighttime operation only, and the temporal and spatial resolution of the retrieved aerosol profiles is limited. Technical solutions to provide daytime operation are underway.

27.4.3 Polarization Lidar

The previous sections have shown that the classification of aerosols ("aerosol typing") is a promising approach to find estimates of lidar ratios. Polarimetric measurements might also support aerosol typing. As it is known from electrodynamics that spherical particles do not change their original state of polarization, whereas nonspherical particles do, it is useful to implement two polarimetric channels to the lidar: one is oriented in the same plane as the emitted linear polarized radiation, the second perpendicular to it. Thus, polarimetric information can be used to distinguish between these two classes of particles.

For the sake of simplicity we do not want to consider cross-talk between the two channels. Then, the lidar Eq. (27.1) can be written separately for the co-polar channel

$$P_{\parallel}(z) = C_{\parallel}\beta_{\parallel}(z)\frac{1}{z^2}\exp\left\{-2\int_0^z \alpha(z')dz'\right\} \tag{27.10}$$

and the cross-polar channel

$$P_{\perp}(z) = C_{\perp}\beta_{\perp}(z)\frac{1}{z^2}\exp\left\{-2\int_0^z \alpha(z')dz'\right\} \tag{27.11}$$

In literature, sometimes $\beta_{\parallel}$ and $\beta_{\perp}$ are replaced by $(1 - d/2)\ \beta(z)$ and $(d/2)\beta(z)$, respectively. The depolarization parameter d can be calculated from scattering theory.

A convenient measure of the polarization state is the so-called volume linear depolarization ratio δ_v.

$$\delta_v = \frac{\beta_{\perp}}{\beta_{\parallel}} = \frac{d/2}{1-(d/2)} = \frac{C_{\parallel}}{C_{\perp}}\frac{P_{\perp}}{P_{\parallel}} = C_{\delta}\frac{P_{\perp}}{P_{\parallel}} \tag{27.12}$$

It can directly be derived as the ratio of the two lidar measurements provided that the calibration constant C_{δ} has been determined (Freudenthaler et al. 2009). No further information, in particular no lidar ratio, is required. However, δ_v is not an aerosol property as it depends on the relative contributions of aerosols and air molecules. Nevertheless, it can provide a first indication of the aerosol type, in

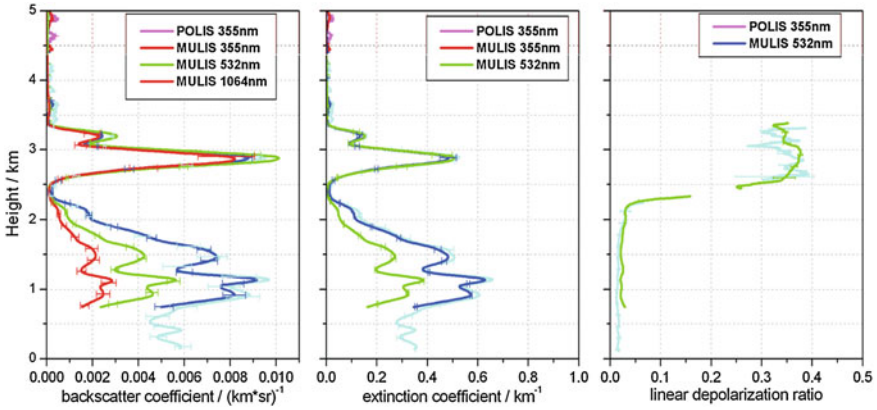

Fig. 27.2 Profiles of particle optical properties: backscatter coefficient, extinction coefficient and linear depolarization ratio as a function of height above ground; derived from MULIS and POLIS measurements at Maisach on 17 April 2010 between 01:30 and 02:40 UTC. The error bars indicate systematic errors (modified after Groß et al. 2011)

particular in cases where aerosols are the dominating component (e.g., in Saharan dust layers). A more accurate parameter describing aerosol properties is the particle linear depolarization ratio δ_p. It can be determined according to

$$\delta_p = (\delta_v + 1)\left(\frac{\beta_m(\delta_m - \delta_v)}{\beta_p(1 + \delta_m)} + 1\right)^{-1} - 1 \qquad (27.13)$$

The derivation of δ_p is more demanding: it requires knowledge of δ_v and β_p at the corresponding wavelengths, the latter normally retrieved from the Klett method with an estimated lidar ratio. δ_m is the depolarization ratio of air molecules and—again—known from Rayleigh theory; it is between 0.0036 and 0.0146 depending on the filter bandwidth of the detection channel. Typical values of δ_p for desert dust are between $0.25 < \delta_p < 0.35$ for wavelengths at 355 and 532 nm. Measurements over Munich during the *Eyjafjallajökull* eruption showed that for volcanic ash layers the values were even larger at $0.35 < \delta_p < 0.40$. Any other aerosol type shows significantly lower depolarization ratios. Though the data base of high quality depolarization measurements is still limited it is generally accepted that δ_p has a large potential for aerosol typing, and thus can help in better estimating lidar ratios for the Klett method. Furthermore, it was found that under certain conditions it is possible to estimate the mixing state of aerosols from δ_p-retrievals.

27.4.4 Optical Characterization of Aerosols

As an example of the potential of different lidar configurations, the characterization of the Eyjafjallajökull plume over Maisach (cf. Fig. 27.1) is again discussed.

With advanced lidar systems, i.e., systems including elastic backscatter detection, polarimetric measurements, and Raman channels at different wavelengths, a large set of optical properties of aerosols as shown in Fig. 27.2 can be retrieved. Here, results from the two MIM lidars, MULIS and POLIS (portable two-channel lidar), are combined, primarily to get depolarization ratios at two wavelengths.

Shown are profiles of β_p, α_p, and δ_p (from left to right) at three and two wavelengths, respectively. The β_p- and α_p-profiles are derived from the Klett method applying lidar ratios from the Raman method. It can be seen that β_p and α_p of the elevated layer are virtually wavelength-independent, whereas a significant decrease with wavelength occurs below 2 km. This is a clear indication that the particles of the lower layer are much smaller than in the upper layer, where the *Angström* exponent κ is close to zero. The profile of δ_p also reveals a significant difference between the elevated layer and lower altitudes. The large depolarization ratio suggests nonspherical particles. Both findings with respect to the size and shape of the particles clearly suggest that the elevated layer was indeed the ash plume of the volcano.

In case only a three-wavelength backscatter lidar is available, the characterization of the aerosols is limited to the β_p-profiles (left panel). As a consequence the vertical extent of the layer can be assessed, and from the wavelength dependence of β_p a rough estimate of the size of the particles. Note, that the Angström exponent is *per definitionem* related to α_p. However, as long as the lidar ratio does not change much with the wavelength, the wavelength-dependence of β_p can be used as an approximation. In case only a single-wavelength system is available (simple backscatter lidar, *ceilometer*) only the vertical extent of the layer can be found.

27.5 Retrieval of Microphysical Parameters

In the previous sections it was briefly demonstrated that optical properties can be used to estimate microphysical properties: the wavelength-dependence of α_p, expressed in terms of κ, is an indication of particle size, δ_p an indication of particle shape. Recently, measurements of optical properties, i.e., α_p and β_p at several wavelengths, were also used to retrieve the microphysical properties of the particles in a more rigorous way. Of special interest are the volume distribution $v(r)$ and the refractive index m, the latter primarily to estimate the single scattering albedo of atmospheric aerosols. A further quantity, the mass concentration M, came into focus during the *Eyjafjallajökull* eruption as this parameter is of highest relevance for *air traffic* safety.

The link between optical and microphysical properties is given by scattering theory according to

$$\alpha_p(\lambda) = \int_0^\infty Q_\alpha(r, \lambda, m) \frac{3}{4r} v(r) dr \tag{27.14}$$

$$\beta_p(\lambda) = \int_0^\infty Q_\beta(r, \lambda, m) \frac{3}{4r} v(r) dr \tag{27.15}$$

Q_α and Q_β are the extinction and backscatter efficiencies at wavelength λ for particles with a complex refractive index m and size r. In the case of spherical particles these functions are calculated from *Mie theory*, in the case of *nonspherical* particles more complex scattering models must be used (cf. Sect. 27.4.1). In the case of M, the density of the particles is required as an additional parameter. The inversion is an ill-posed problem, and thus the numerical implementation is critical. Further complications arise from the limited number of measured optical parameters (i.e., the left side of the equations) and the magnitude of the measurement errors.

As already mentioned, state-of-the-art multiwavelength Raman lidar measurements provide α_p at two wavelengths (355, 532 nm) and β_p at three wavelengths (355, 532, 1 064 nm). Measurements of these five optical properties are the minimum requirement for the inversion of Eqs. (27.14) and (27.15) for the volume distribution and the refractive index (Müller et al. 2001). Several case studies on the basis of Mie theory have shown that the inversion in principle is feasible, nevertheless, the accuracy of the retrieved parameters is often insufficient. Thus, the establishment of robust inversion schemes is a current research topic. Among others, more efficient numerical solutions are tested, the additional consideration of depolarization ratios is investigated, implying the application of scattering models for nonspherical particles and the combination of lidar and *sun- and sky-radiometers* data. The latter provide multiwavelength direct and diffuse radiances; however, in contrast to the lidar data, this information is "vertically integrated." This combined approach was applied to the retrieval of M in the case of the *Eyjafjallajökull* plume when Gasteiger et al. (2011) retrieved conversion factors from α_p to M. They demonstrated that the conversion factors critically depend on the size distribution of the particles and their shape. As a result the mass concentration can be retrieved with an uncertainty range of approximately ± 50 %.

27.6 Examples of Applications

As already mentioned, long term observations of aerosol distributions ("climatologies") as well as process studies to understand interactions between aerosols, clouds and radiation are required to improve knowledge of the *climate system*. Both issues can be supported by lidar. A few examples of current applications are briefly discussed below.

27.6.1 Closure Studies

Closure experiments are important, e.g., for understanding the interaction between radiation and aerosol particles, and they are the basis for testing and validating new

remote sensing schemes. To obtain a complete set of information, the vertical distribution of aerosol optical properties must be measured. Key contributions come from airborne *in-situ measurements* and lidars—deployed from ground and aircraft. This approach was realized in a series of aerosol characterization experiments, e.g., ACE-1, which was devoted to the pristine marine atmosphere, ACE-2, where anthropogenic aerosols from Europe and transported desert dust were investigated, ACE-Asia which focused on anthropogenic aerosols over the Asian Pacific region, and most recently SAMUM, which concentrated on mineral dust aerosols.

In the framework of SAMUM, three ground-based *Raman lidar* systems (of MIM and IfT, Leipzig) and the airborne high spectral resolution lidar of DLR were operated. The scientific goal of the measurements included several of the topics mentioned in the previous sections: assessment of the lidar ratio of Saharan dust and the changes due to aging, investigation of the spectral dependence of the depolarization ratios, *validation* of scattering models for nonspherical particles, and development of strategies to characterize aerosol mixtures by means of lidar measurements.

27.6.2 Networks

To get long term observations for the assessment of the variability and trends, lidar networks can provide substantial contributions. As advanced lidar systems are a quite expensive technology, the number of stations and the temporal sampling will always be limited and the representativeness might suffer. Thus, recent activities to establish lidar networks focus rather on research (improvement of data evaluation, development of synergistic methodologies, standardization of hardware, automatization) than on monitoring.

The first ground-based lidar network following this concept was established in Germany in 1997 with five stations, including MIM. This small network was working on a regular basis, had agreed on common data products and formats, and had for the first time undergone strict quality assurance on hardware and software level. In 2000 the network was extended to the "European Aerosol Research Lidar Network" (EARLINET) and now includes 27 stations.

During the last few years other networks were established outside Europe, e.g., the "Asian Dust Network," "REALM" (Eastern United States), the "CIS-LiNet" (former Soviet Union) and "ALINE" (South America). The objectives, the degree of cooperation and maturity, and strategies for data exchange are not yet harmonized; however, the *World Meteorological Organization* initiated first steps to combine the different continental networks under the umbrella of "GALION" (GAW Aerosol Lidar Observation Network). However, in view of monitoring, networks of "low cost" standardized systems that can be operated unattended and continuously must be implemented. Concepts for the combination of a limited number of high-end systems and a denser network of ceilometer-type instruments are currently under investigation.

The lidar networks mentioned above have been active for about one decade, and primarily concentrate on tropospheric aerosols. In the case of stratospheric aerosols, long-term observations are already available. Due to the long residence time and the larger spatial homogeneity of stratospheric aerosols, the requirements with respect to the number of stations and the temporal sampling are relaxed. The most consistent time series is available for Garmisch-Partenkirchen (Germany) and covers 30 years (Jäger 2005). It documents a close correlation between stratospheric aerosol load and volcanic eruptions.

27.6.3 Lidar in Space

Though ground-based networks have proved to be of great benefit for characterizing aerosols on larger scales, spaceborne observations are indispensable to ultimately provide global aerosol data sets. The first spaceborne lidar experiment was the Space Shuttle mission LITE, (Lidar in Space Technology Experiment) conducted by NASA in September 1994. LITE had a three-wavelengths *backscatter lidar* on board with the inherent difficulties in deriving quantitative aerosol parameters (see Sect. 27.4.1), but it successfully demonstrated the feasibility of active aerosol monitoring from space.

The first long term mission is *CALIPSO* (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations Experiment) launched in April 2006 and still in operation. Again, a backscatter lidar is used, including polarimetric measurements at 532 nm. To take advantage of complementary measurements CALIPSO is flown as a part of a multisatellite constellation (A-train). In parallel to the spaceborne measurements, extensive ground truth campaigns are conducted to verify the aerosol retrievals. In particular, the accuracy of the assumed lidar ratios is critically reviewed.

The first European long-term spaceborne mission including an aerosol lidar will be *EarthCARE* (Earth Clouds, Aerosols and Radiation Explorer). It is planned to implement a high spectral resolution lidar. As a consequence, no estimates of lidar ratios will be required. The tentative launch year is 2015.

27.7 Summary

The development of lidars for atmospheric research in the 1970s constituted a new era of aerosol remote sensing by providing range resolved information. At that time, however, aerosol distributions could only be determined qualitatively. Only the advent of powerful lasers and improved detection systems in the 1990s allowed the quantitative derivation of aerosol optical properties. After the turn of the century it was demonstrated that under favorable conditions even microphysical properties can be inferred from advanced lidar systems.

Applications of lidars in atmospheric science are numerous: From the beginning lidar measurements were mostly applied in dedicated field campaigns and aircraft missions (e.g., Mörl et al. 1981). Such studies were primarily set up to improve

and validate aerosol retrievals (e.g., from space), to understand the *radiative forcing* of aerosols and interactions with clouds. To increase the representativeness of aerosol profiling and to investigate spatial and temporal long term trends, lidar networks were established in the last decade. The most prominent example is EARLINET, a European network consisting of more than twenty sites with state-of-the-art lidar systems. To what extent secondary networks of *ceilometers* as currently being implemented by national weather services can improve knowledge of the four-dimensional distribution of aerosols is presently under discussion. There is a strong demand for such data for *numerical weather prediction* and in contingency cases, e.g., volcanic eruptions. In this context the provision of aerosol profiles in near-real-time is of special importance.

In parallel to the growing number of ground-based lidar systems, spaceborne missions were defined to provide global coverage. *CALIPSO*, with a dual-wavelength backscatter lidar, was launched in 2006 providing new perspectives in aerosol remote sensing, in particular when different satellite sensors are ultimately exploited synergistically.

In the next few years, efforts must be intensified to make use of lidar data for numerical weather prediction and climate modeling, and to exploit synergies between lidar, radar and passive remote sensing. With respect to aerosols, investigations of their indirect effects, in particular, will move into the focus of attention.

References

Ansmann, A., Wandinger, U., Riebesell, M., Weitkamp, C., Michaelis, W.: Independent measurement of extinction and backscatter profiles in cirrus clouds by using a combined raman elastic-backscatter lidar. Appl. Opt. **31**(33), 7113–7131 (1992)

Bösenberg, J., V. Matthias, A., Amodeo, V., Amoiridis, A., Ansmann, J., Baldasano, M., Balin, I., Balis, D., Böckmann, C., Boselli, A., Carlsson, G., Chaikovsky, A., Chourdakis, G., Comeron, A., De Tomasi, F., Eixmann, R., Freudenthaler, V., Giehl, H., Grigorov, I., Hagard, A., Iarlori, M., Kirsche, A., Kolarov, G., Komguem, L., Kreipl, S., Kumpf, W., Larcheveque, G., Linné, H., Matthey, R., Mattis, I., Mekler, A., Mironova, I., Mitev, V., Mona, L., Müller, D., Music, S., Nickovic, S., Pandolfi, M., Papayannis, A., Pappalardo, G., Pelon, J., Perez, C., Perrone, R.M., Persson, R., Resendes, D.P., Rizi, V., Rocadenbosch, F., Rodrigues, J.A., Sauvage, L., Schneidenbach, L., Schumacher, R., Shcherbakov, V., Simeonov, V., Sobolewski, P., Spinelli, N., Stachlewska, I., Stoyanov, D., Trickl, T., Tsaknakis, G., Vaughan, G., Wandinger, U., Wang, X., Wiegner, M., Zavrtanik, M., Zerefos, C.: EARLINET: A European Aerosol Research Lidar Network to Establish an Aerosol Climatology. MPI-Report 348, Max-Planck-Institut für Meteorologie, Hamburg, Germany, p. 192, ISSN 0937-1060 (2003)

Fernald, F.G.: Analysis of atmospheric lidar observations: some comments. Appl. Opt. **23**(5), 652–653 (1984)

Freudenthaler, V., Esselborn, M., Wiegner, M., Heese, B., Tesche, M., Ansmann, A., Müller, D., Althausen, D., Wirth, M., Fix, A., Ehret, G., Knippertz, P., Toledano, C., Gasteiger, J., Garhammer, M., Seefeldner, M.: Depolarization-ratio profiling at several wavelengths in pure Saharan dust during SAMUM 2006. Tellus B **61**, 165–179 (2009). doi:10.1111/j.1600-0889.2008.00396.x.

Gasteiger, J., Groß, S., Freudenthaler, V., Wiegner, M.: Volcanic ash from Iceland over Munich: mass concentration retrieved from ground-based remote sensing measurements. Atmos. Chem. Phys. **11**(5), 2209–2223 (2011). doi:10.5194/acp-11-2209-2011

Groß, S., Freudenthaler, V., Wiegner, M., Gasteiger, J., Geiß, A., Schnell, F.: Dual-wavelength linear depolarization ratio of volcanic aerosols: lidar measurements of the Eyjafjallajökull plume over Maisach, Germany. Atmos. Environ. **48**, 85–96 (2011). doi:10.1016/j.atmosenv.2011.06.017

Jäger, H.: Long-term record of lidar observations of the stratospheric aerosol layer at Garmisch-Partenkirchen. J. Geophys. Res. **110**, D08106 (2005). doi:10.1029/2004JD005506

Klett, J.D.: Stable analytical inversion solution for processing lidar returns. Appl. Opt. **20**, 211–220 (1981)

Mörl, P., Reinhardt, M.E., Renger, W., Schellhase, R.: The use of the airborne lidar ALEX-F for aerosol tracing in the lower troposphere. Contr. Atmos. Phys. **45**, 403–410 (1981)

Müller, D., Wandinger, U., Althausen, D., Fiebig, M.: Comprehensive particle characterization from three-wavelength Raman-lidar observations: case study. Appl. Opt. **40**(27), 4863–4869 (2001)

Wiegner, M., Gasteiger, J., Groß, S., Schnell, F., Freudenthaler, V., Forkel, R.: Characterization of the Eyjafjallajökull ash-plume: potential of lidar remote sensing. Phys. Chem. Earth (2011). doi:10.1016/j.pce.2011.01.006

Measuring Water Vapor with Differential Absorption Lidar

28

Martin Wirth

Abstract

Despite the need for global measurements of water vapor profiles with low bias and high vertical resolution there is currently no operational remote sensing system that would deliver such data. A possible solution to this problem is offered by the differential absorption lidar (DIAL) approach. The basic principle of operation will be described and some background on atmospheric light absorption by water vapor will be given. DLR's airborne water vapor DIAL system WALES represents the currently most advanced system worldwide using a multiwavelength technique to cover the troposphere and lower stratosphere simultaneously. A few examples of measurements made with this system will illustrate the power of this active remote sensing method.

28.1 Introduction

Water vapor (H_2O) is an essential species of the earth's atmosphere. Besides its ability to condense in the form of *clouds*, it is the most important contributor to the *greenhouse effect* and therefore a major regulative component of the earth's *climate*. Evaporation and transpiration (via plants) at the earth's surface are the fundamental processes producing large amounts of water vapor which drive the weather and the global *hydrological cycle*. Due to this strong source at the surface, water vapor is basically concentrated within the *atmospheric boundary layer* (ABL) at 0–2 km.

M. Wirth (✉)
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Martin.Wirth@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_28,

Above, the mixing ratio strongly decreases with height, dropping from values of 0.1–3 % within the ABL to values below four parts per million in the *stratosphere*. This strong decrease of moisture with height over about four orders of magnitude and the large gradients around the top of the ABL and the *tropopause* region require a tremendous technical effort to measure water vapor accurately and with adequate horizontal and vertical resolution.

Up to now, most regular observations of water vapor as used in the data *assimilation* systems of *numerical weather prediction* centers are based on *radiosonde* measurements. However, radiosonde launches are sparse in the southern hemisphere and there are practically no radiosondes over the oceans. In addition, current operational radiosondes do not deliver good water vapor data at higher altitudes or in dry regions (e.g., subsidence zones or intrusion regions). For these regions it is necessary to use *satellite* data for the assimilation. At present only passive space borne *remote sensing instruments* such as infrared sounders and microwave imagers are used for this purpose. While microwave imagers provide the total water vapor column (or precipitable water) infrared sounders can be used to determine the vertical water vapor distribution at a coarse vertical resolution of 2–3 km. However, both sensors have severe problems under cloudy conditions and need first guess profiles as well as tremendous bias correction efforts to operate at reasonable accuracy. In the near future satellite navigation systems in combination with low orbit satellites for occultation measurements promise to provide valuable measurements of refractivity, including moisture data from the low/mid *troposphere* at good vertical resolution. The disadvantage of these systems lies in their limited horizontal resolution ($\sim$300 km) as well as the limited vertical range, typically starting at 5 km altitude.

Besides numerical weather prediction there are also topics of basic atmospheric research which would strongly benefit from high resolution remote sensing of water vapor, like turbulent transport of latent energy in the boundary layer, stratosphere/troposphere exchange of water, or the role of super-saturated regions in *cirrus* cloud formation.

A promising alternative to the current passive remote sensing instruments is provided by a special light detection and ranging (*lidar*) technique, the so called *DIfferential Absorption Lidar* (DIAL). The basic idea of this method, as explained in the next paragraph, emerged during the early 1960 s shortly after the invention of the *laser*, but Dobbins and LaGrone (1969) and Schotland (1974) were the first to describe the method in detail in a scientific journal. The first experimental results from ground based measurements of H_2O were published about a decade later (Browell et al. 1979; Werner and Herrmann 1981; Cahen et al. 1982; Zuev et al. 1983). These first systems proved the feasibility, but all suffered from various problems, mostly related to the laser source. Either the lasers were not tunable to optimal absorption lines, lacked the necessary output power for fast and highly resolved measurements, or their radiation had broad spectral components leading to a severe measurement bias. It took another decade to develop high power, tunable, pulsed lasers with suitable spectral properties. Since the DIAL method is especially well suited to the retrieval of atmospheric water vapor when employed in a downward looking mode, i.e., looking from dry into moister regions of the

atmosphere, the next natural step was the development of airborne systems. Up to now, three such airborne H_2O-DIAL systems have been developed at DLR (Ehret et al. 1993; Poberaj et al. 2002; Wirth et al. 2009), one in the USA (NASA, Browell et al. 1998) and one in France (CNRS, Bruneau et al. 2001), each using a different laser concept. However, only the DLR systems are capable of measuring H_2O in the *upper troposphere* and *lower stratosphere*. Based on experience with airborne instruments, several ESA studies have shown the feasibility of a space borne instrument on a medium sized platform (ESA 2004).

28.2 Basics of the Differential Absorption Lidar Method

As the light pulse emitted from a lidar system propagates through the atmosphere its intensity is reduced (extinct) by *scattering* and absorption by the molecular constituents of the air and by *aerosol* and *cloud particles*. Part of the light is scattered backward and propagates back to the lidar system, again undergoing the same extinction process. In this way the detected signal of a lidar depends on the local ability of the atmosphere to scatter light in a backward direction and on integral light extinction along the complete light path between the system and the probed volume. Extinction by aerosol, clouds and molecular scattering only shows a relatively weak dependence on wavelength, the strongest being the λ^{-4} dependence of the molecular scattering. In sharp contrast, the absorption of light by molecules in the visible to infrared spectral region manifests itself in the form of very narrow *spectral absorption lines*.

By probing the atmosphere with a lidar at two properly selected wavelengths this fact can be utilized to measure the atmospheric absorption which is solely due to one *trace gas* component of the air. The two wavelengths must be close enough so that the scattering and extinction by aerosols, clouds and N_2/O_2 molecules are virtually the same. For example, the scattering and absorption of the atmosphere at 935 nm changes only by 0.2 % over a wavelength interval of 0.5 nm if special trace gas absorptions are neglected. One wavelength is set to a special absorption line of the trace gas under consideration (usually called the online wavelength). The other one is set to a nearby wavelength with much less, but not necessarily zero, absorption. Care has to be taken that all other trace gases show the same absorption for the two wavelengths. Optimally, these additional absorptions would be zero, but for the DIAL principle to work only equality is required. Under these requirements the mean trace gas concentration ρ in molecules per cubic meter within a certain height interval is given by the DIAL equation:

$$\rho(\bar{r}) = \frac{1}{2 \cdot (\sigma_{on} - \sigma_{off}) \cdot (\Delta r)} \ln\left(\frac{S_{on}(r_1) \cdot S_{off}(r_2)}{S_{on}(r_2) \cdot S_{off}(r_1)}\right) \tag{28.1}$$

Here $\bar{r}$ denotes the center point of a range interval from r_1 to r_2 of length $\Delta r = r_2 - r_1$. $S_{on/off}$ are the lidar signal strengths at the online and offline wavelength. The ability of the probed molecule to absorb light is described by the

absorption cross sections $\sigma_{on/off}$ at the two wavelengths, which give the equivalent geometrical area (in cm^2) which would blank out the same amount of incident light. Figure 28.1 illustrates the measurement principle graphically. No calibration or system constants enter into this equation, only the measured raw signals and the absorption cross sections which are obtained from laboratory measurements. Therefore DIAL is called a self-calibrating technique.

For a DIAL to work in a satisfactory manner the magnitude of the absorption cross section σ must be within a certain range. If the absorption is too low, as for example in the upper few km of the signals shown in Fig. 28.1, it is difficult to measure it with high accuracy. If, on the other hand, the absorption is very high, the lidar signal is extinct to below the detection limit within short range. Under the simplified assumption that the measurement error is solely caused by the *photon* statistics of the received signal it can be shown that the sensitivity of a DIAL instrument increases until an *optical thickness* of 1.1 is reached (Remsberg and Gordley 1978). For higher values the instrument error begins to rise again with increasing optical thickness. Therefore, the absorption of the probed trace gas has to be chosen in such a way that the relation:

$$\sigma_{on} \approx \frac{1.1}{\int_O^R \rho(r)dr} \tag{28.2}$$

holds for a typical trace gas profile $\rho(r)$ and an envisaged maximum range R.

For atmospheric water vapor with its large concentration range a problem arises here. If σ is chosen such that the measurement range extends from a typical flight altitude of 12 km down to the ground, then one is forced to select a relatively weak absorption. This guarantees good measurement conditions for the high moisture values of the boundary layer, but results in large errors for the much lower concentrations in the mid and upper troposphere. The solution to this problem is to use more than one online wavelength. For H_2O it has turned out that at least three wavelengths have to be used to cover the whole mid-latitude troposphere. If all atmospheric conditions from the *tropics* to the *Arctic* are to be covered in one setup, one has to use at least five different wavelengths, where only three of them are used at the same time. As will be seen in the next paragraph for water vapor it is possible to find five online wavelengths with suitable absorption coefficients and a low absorption reference wavelength within an interval of 0.5 nm in the 935 nm wavelength region.

The measurements on the different wavelengths are then combined into a single H_2O profile by weighted addition of the single retrievals. The optimal relative weights are functions of the standard deviations of the single profiles and can be calculated from a statistical analysis of the raw data without further assumptions.

The next paragraph will present the structure of the absorption cross section of water vapor in more detail. Especially it will become clear that one silent assumption that was made in the current section, namely that σ is a constant for fixed wavelength, is not true in general.

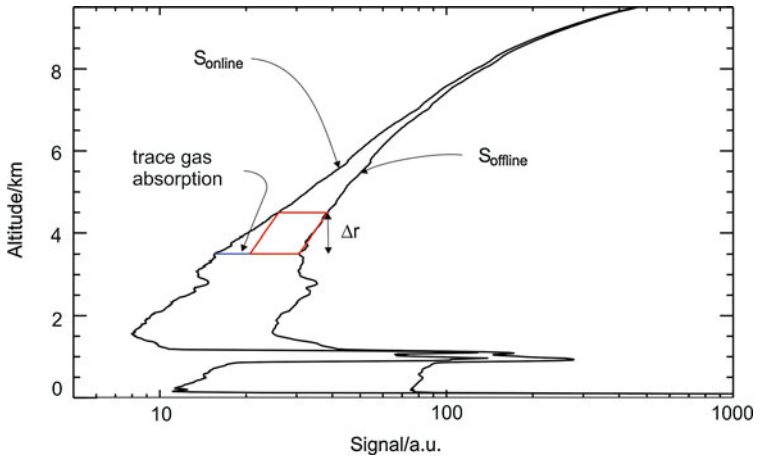

Fig. 28.1 Principle of DIAL: The same atmospheric volume is probed by two laser pulses (S_{online} and $S_{offline}$). For a proper wavelength selection the only difference between the two signals is due to absorption by one trace gas component of the air. By comparing the two signals over a height interval Δr the trace gas absorption within this interval is determined without any additional system or calibration parameter

28.3 Spectral Characteristics of Water Vapor Absorption

Water vapor is an asymmetric-top molecule. Thus the electronic, vibrational, and rotational spectrum consists of a large number of nonregularly spaced lines. The electronic absorption bands are in the far ultraviolet spectral range. The vibrational–rotational spectrum contains three main bands in the infrared at 2.73, 3.1, and 6.2 μm wavelength. In addition, overtones, combinations and hot bands exist from the infrared to the visible spectral region. Due to the high concentration of H_2O in the troposphere even weak lines from combination bands at 720, 830 and 935 nm wavelength in the near infrared cause considerable attenuation of light. The fine structure of the vibrational–rotational spectrum of H_2O is extremely complex. Each of the bands consists of thousands of lines. Figure 28.2 gives an overview of the line strengths in the visible to near infrared region.

The small scale spectrum of a single line strongly depends on atmospheric *temperature* and *pressure*. The temperature determines the *thermodynamic* population of the ground state of the transition, which alters the total strength of the absorption line. The second temperature effect comes from the high speed random motion of the molecules that leads to a *Doppler shift* of the absorption line as compared to a molecule at rest. Since the thermal speed distribution has a Gaussian shape (Maxwell distribution) the mean spectral line of an ensemble of H_2O molecules also resembles a Gaussian function. Collisions with other air molecules lead to a further broadening effect of the spectral line form by limiting the duration of the emission. This results in a *Lorentzian line-shape* which is, at the same time, accompanied by a shift of the center frequency caused by the influence of nearby

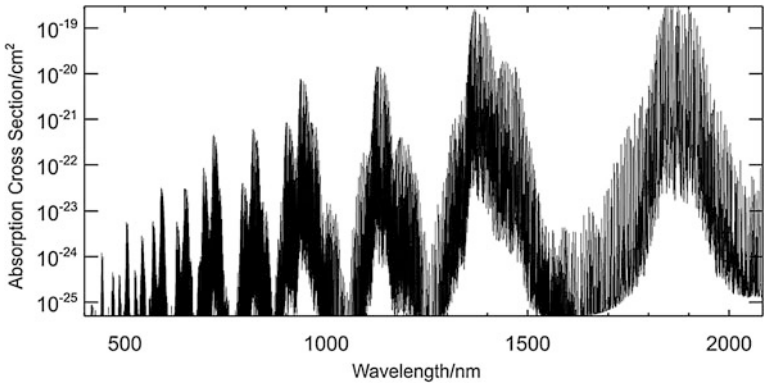

Fig. 28.2 Absorption spectrum of water vapor in the visible and near infrared spectral region calculated for atmospheric conditions at 10 km altitude. It consists of several bands which in turn are made up from thousands of single lines. The total range shown here contains about 39 000 lines. Line parameter data was taken from the HITRAN 2008 database (Rothman et al. 2009)

molecules onto the energies of the transition states. Since this broadening and shifting effect is to the first order proportional to the ambient pressure it is called the *pressure broadening*. The Doppler and pressure broadening effects combine into a so-called *Voigt profile* by convolution of the two functional dependences. At high altitudes (16 km and above) the Doppler effect dominates and the line shape is nearly Gaussian. Near the ground the pressure broadening effect dominates. Figure 28.3 illustrates this behavior.

For a proper selection of absorption lines suitable for DIAL measurements several things have to be considered: First, the line strength has to be adequate for the expected water vapor concentration (cf. the discussion in paragraph 2), second there has to be a reference wavelength very close by with much lower absorption, and third the above-discussed temperature and pressure dependences have to be low enough as not to introduce errors from uncertainties in the auxiliary atmospheric parameters. Furthermore, there are technical constraints like the availability of low noise detectors and high power laser sources. All these factors taken together lead almost naturally to the selection of the absorption band around 935 nm for closer consideration. The bands at smaller wavelengths have no lines strong enough to allow measurements in the upper troposphere/lower stratosphere. For the bands above 1.1 μm there are no good (i.e., nearly quantum noise limited) detectors available and the signal from the molecular backscattering drops considerably (λ^{-4}-dependence). If one searches for closely grouped multiplets of temperature and pressure insensitive lines which cover about three orders of magnitude in strength, only a few wavelength regions are left in the 935 nm band. One of these regions is shown in Fig. 28.4 where lines suitable for H_2O retrieval covering the whole troposphere can be found within a wavelength interval of about 0.5 nm.

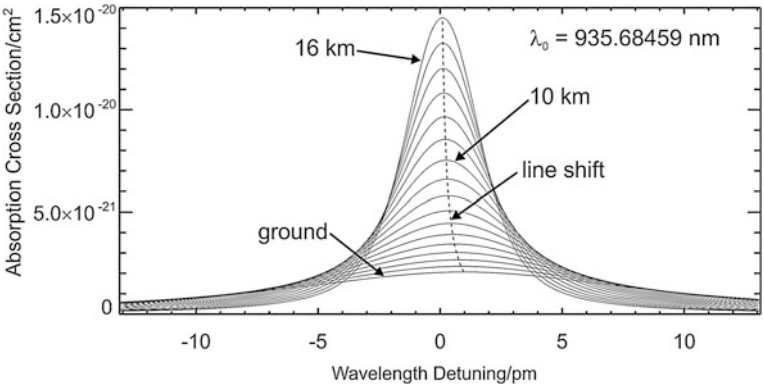

Fig. 28.3 Pressure and temperature dependence of the shape of a single H_2O absorption line. At high altitudes (e.g., 16 km) the line-shape is dominated by Doppler-broadening due to the thermal movement of the molecules. The line-shape is nearly Gaussian. Near the ground the pressure broadening effect dominates, giving rise to a Lorentzian frequency dependence. Here, in addition to the broadening a shift of the center wavelength is induced by the molecular collisions

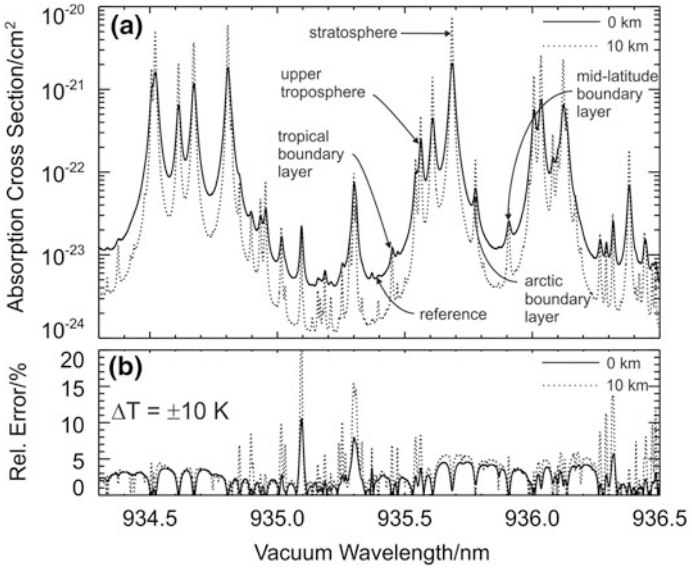

Fig. 28.4 **a** Absorption lines of water vapor in the 935 nm regions as used for the WALES airborne demonstrator. Absorption cross section data was calculated from the HITRAN 2008 database (Rothman et al. 2009) for sea level conditions (*solid line*) and at 10 km altitude (*dashed line*) using the US-standard atmosphere. Possible wavelengths of operation for different atmospheric regions are indicated by arrows. **b** Relative error in water vapor retrieval if the absorption cross sections are calculated with a temperature profile which is wrong by 10 K. By proper selection of the absorption line used the error can be kept below 1 % even in this extreme case

Table 28.1 Main system parameters of the airborne water vapor DIAL WALES

Transmitter parameter	Value	Receiver parameter	Value
Repetition rate	200 Hz	Telescope diameter	480 mm
Pulse energy @ 1 064 nm	120 mJ	Field of view	1.5 mrad
Pulse energy @ 532 nm	75 mJ	Filter bandwidth	1 nm
Pulse energy @ 935 nm	45 mJ	Detector type 1 064/935 nm	Avalanche photo diode
Pulse length	8 ns	Detector type 532 nm	Photo multiplier
Beam divergence	1 mrad	Pulse response (FWHM)	100 ns
Wavelength stability @ 935 nm	<0.1 pm	Sampling rate	10 MHz (15 m)

The very small widths of the absorption lines put stringent requirements on the bandwidth and stability of the laser source. For stratospheric retrievals, where the broadening effects are lowest, the wavelength stability has to be better than 0.2 pm. This holds for measurements at the line center. If one would use a wavelength in the slope of a line, which is perfectly possible from the general measurement principle, the stability requirement is tightened by about one order of magnitude. Also, the pressure and temperature dependence goes up at the line slope for the majority of lines. Thus one usually stays with positions at the line center.

28.4 Examples of Water Vapor DIAL Measurements

This section presents examples of water vapor measurements which were acquired by DLR's WALES instrument. This system was developed during the last decade as an airborne demonstrator for a possible space-borne mission. It only uses technology like diode pumped solid state lasers and nonlinear conversion techniques which are suitable for in-orbit deployment. The pulse energy, the pulse repetition rate and the overall electrical-to-optical efficiency are the same as envisaged for a future space instrument. It operates in the 935 nm region and is capable of emitting at four wavelengths at the same time for water vapor measurements. To generate the pulses at 935 nm an *Optical Parametric Oscillator* (OPO) is used. In a nonlinear optical process using potassium titanyl phosphate (KTP) crystals the photons of a frequency-doubled Nd:YAG laser are split into two lower energy photons, one of them at the selected measurement wavelength. The main advantage of this technique is that the splitting ratio and thus the output wavelength can be tuned easily over a wide range by tilting the nonlinear optical crystal. Additionally, there are aerosol detection channels operating at 1 064 and 532 nm wavelength where the *depolarization* and aerosol extinction (only at 532 nm) of the atmosphere are measured. Table 28.1 summarizes the technical parameters of this system (see Wirth et al. 2009 for a more comprehensive description).

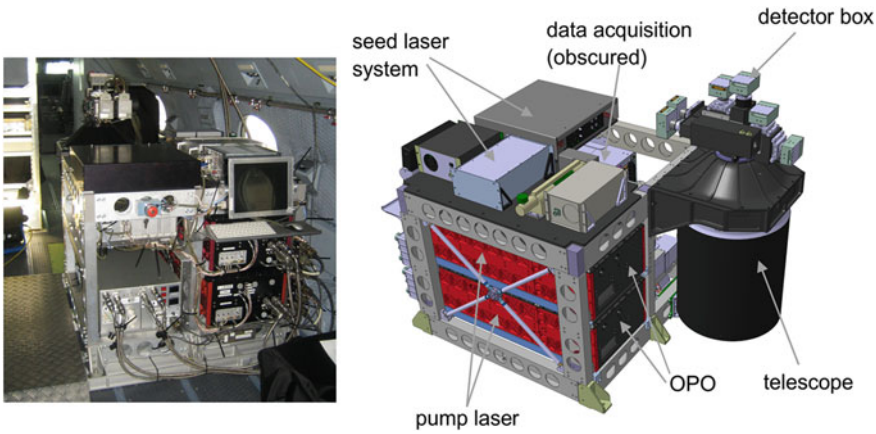

Fig. 28.5 The combined water vapor and aerosol lidar WALES: main building blocks (*right*) and integration into the HALO research aircraft (*left*)

The WALES system has been operated for more than 300 flight hours on DLR's *Falcon* F20 research aircraft. During the last year the system was adapted to the new German research aircraft HALO and demonstrated its full capabilities within the new environment during a technical test mission in late 2010. Figure 28.5 shows the complete instrument rack, which has a dimension of about one cubic meter.

The water vapor measurements of this system have a typical resolution of 200 m in the vertical and 6 km in the horizontal. Due to the much higher resolution of the raw data signals (15 m vertical and 40 m horizontal) it is also possible to achieve a much higher resolution for special applications like boundary layer flux measurements, but at the cost of a limited range and/or a higher statistical noise level.

The main systematic measurement error sources of a DIAL system are uncertainties of the absorption cross sections from laboratory data, uncertainties of the pressure and temperature profiles used in the retrieval, and errors in the wavelength calibration of the system. For WALES the sum of these errors is estimated to be lower than 5 %. The other error type dominant for a lidar system is the statistical error introduced by the finite number of received photons. This noise source, which follows a Poisson statistic, depends strongly on the range, the horizontal and vertical resolution of the calculated water vapor, and on atmospheric *backscatter* and extinction. Therefore, this error cannot be given as a general number. For each individual profile it can be calculated from the measured signals themselves and is typically tabulated together with the water vapor data.

These theoretical error estimates were checked by measurements against other high precision hygrometers. Figure 28.6 shows two examples. The first is a comparison with the cryogenic frost point hygrometer CFH (Vömel et al. 2007) launched on a balloon in close coordination with an overpass of the DLR Falcon aircraft. Since the balloon needed about half an hour to reach 10 km altitude and the overpass of the aircraft took about 5 min, a backtracking scheme was applied

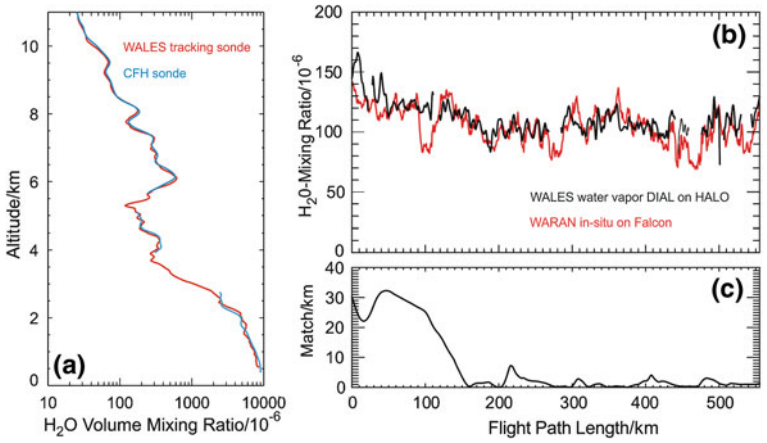

Fig. 28.6 Comparisons of the WALES instrument with the cryogenic frost point hygrometer CFH (Vömel et al. 2007) mounted on a balloon (**a**) and a diode laser spectrometer on a second aircraft (**b**). The best match distance between the two aircraft as a function of the flight path length is shown in (**c**)

to find a best match between the air masses probed by the balloon probe and the DIAL system. The measurements agree within the error bound given for each instrument. The second example (Fig. 28.6b) shows a comparison with an in situ diode laser spectrometer called WARAN on a coordinated flight of the HALO aircraft together with DLR's Falcon F20. The WALES system was mounted on HALO, which flew at an altitude of 14.5 km over an extensive field of cirrus clouds. At the same time the Falcon tried to follow the flight track of HALO, flying within the cirrus cloud at about 11 km altitude. Figure 28.6c shows the minimum distance between the air masses probed by the two instruments as a function of the flight path length. It can be seen that even small structures in the water vapor field as seen by the in situ instrument are tracked by the remote sensing DIAL instrument, at least when the probed air volumes are close enough.

The example shown in Fig. 28.7 comes from a flight over central Europe where the WALES system was mounted on the DLR Falcon aircraft. The most apparent feature in the data is a very dry layer at about 3 km altitude with only 600 m vertical extent, which is stretched out all over the measurement region. The water vapor mixing ratios below 50×10^{-6} within this layer are typical for the uppermost troposphere or lowermost stratosphere. The exact minimum values could not be determined by the DIAL measurements because the weakest of all three absorption lines had to be used to evaluate the data. The, in principle, much more suitable stronger absorption lines could not be used because the laser beam was not able to penetrate the 100 times moister layers above the dry region. This dry layer touched the surface within the Alps. More to the north the dry layer was dissolving and stretching out vertically. To clearly detect this dry layer is an extreme benchmark for all remote sensing water vapor profiling instruments and thus an excellent test case showing the benefits of the DIAL method.

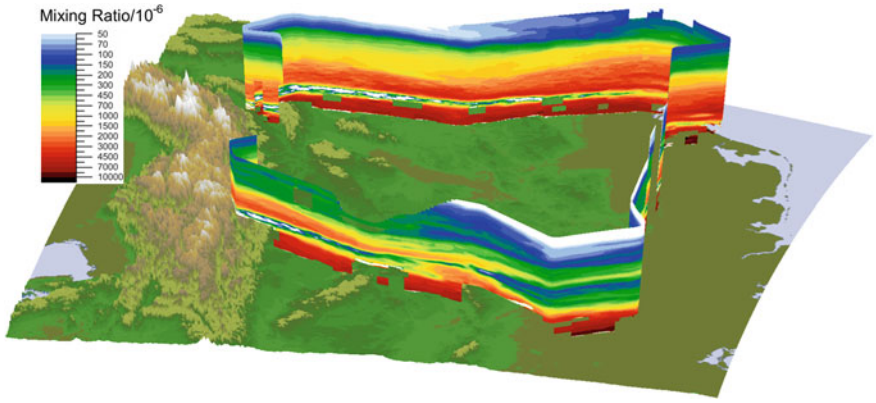

Fig. 28.7 Two dimensional cut through the water vapor distribution over central Europe on 17 October 2008. The data displayed was measured by the WALES water vapor DIAL onboard the DLR Falcon F20 aircraft. The Alps are shown to the *left* and the North Sea to the *right*

28.5 Summary

After a long period of development that was mainly hampered by the lack of tunable high power, narrow band (single mode) pulsed laser sources, nowadays the measurement of atmospheric trace gases, especially water vapor, by differential absorption lidar has become the most accurate range-resolved remote-sensing technique. The DIAL method is able to deliver two dimensional cuts through the atmospheric moisture field with a horizontal and vertical resolution comparable to in-situ instruments. The radiometric-calibration-free measurement principle guarantees low bias and long-term comparable measurements. Even today the technology is ready for the development of a medium sized satellite instrument which would deliver global profiles of water vapor with the low bias and high vertical resolution needed for numerical weather prediction. With the ongoing advances in laser technology (Löhring et al. 2010) even a small satellite mission seems to be possible within the foreseeable future.

Acknowledgments The author would like to thank Holger Vömel of the German Weather Service (DWD) for kindly providing the cryogenic frost point hygrometer (CFH) data.

References

Browell, E.V., Wilkerson, T.D., McIlrath, T.J.: Water vapor differential absorption lidar development and evaluation. Appl. Opt. **18**, 3474–3483 (1979). doi:10.1364/AO.18.003474

Browell, E.V., Ismail, S., Grant, W.B.: Differential absorption lidar (DIAL) measurements from air and space. Appl. Phys. B **67**, 399–410 (1998). doi:10.1007/s003400050523

Bruneau, D., Quaglia, P., Flamant, C., Meissonnier, M., Jacques Pelon, J.: Airborne lidar LEANDRE II for water-vapor profiling in the troposphere. I. System description. Appl. Opt. **40**, 3450–3461 (2001). doi:10.1364/AO.40.003450

Cahen, C., Megie, G., Flamant, P.: Lidar monitoring of the water vapor cycle in the troposphere. J. Appl. Meteorol. Clim. **21**, 1506–1515 (1982). doi:10.1175/1520-0450(1982)021<1506:LMOTWV>2.0.CO;2

Dobbins, D.L., LaGrone, A.H.: Number density determination in the atmosphere of O2, H2O, and CO2 gas constituents by use of a high intensity laser beam. Radio Sci. **4**, 407–411 (1969). doi:10.1029/RS004i005p00407

Ehret, G., Kiemle, C., Renger, W., Simmet, G.: Airborne remote sensing of tropospheric water vapor with a near-infrared differential absorption lidar system. Appl. Opt. **32**, 4534–4551 (1993). doi:10.1364/AO.32.004534

ESA: Report for mission selection: WALES-water vapour lidar experiment in space, SP-1279(3) (2004). ISBN 92-9092-962-6

Löhring, J., Meissner, A., Hoffmann, D., Fix, A., Ehret, G., Alpers, M.: Diode-pumped single-frequency-Nd:YGG-MOPA for water-vapor DIAL measurements: design, setup and performance. Appl. Phys. B **102**, 917–935 (2010). doi:10.1007/s00340-010-4314-1

Poberaj, G., Fix, A., Assion, A., Wirth, M., Kiemle, C., Ehret, G.: Airborne all-solid-state DIAL for water vapour measurements in the tropopause region: system description and assessment of accuracy. Special issue "trends in laser sources, spectroscopic techniques and their applications to trace gas detection". Appl. Phys. B **75**, 165–172 (2002). doi:10.1007/s00340-002-0965-x

Remsberg, E.E., Gordley, L.L.: Analysis of differential absorption lidar from the space shuttle. Appl. Opt. **17**, 624–630 (1978). doi:10.1364/AO.17.000624

Rothman, L.S., et al.: The HITRAN 2008 molecular spectroscopic database. J. Quant. Spectr. Rad. Transf. **110**, 533–572 (2009). doi:10.1016/j.jqsrt.2009.02.013

Schotland, R.M.: Errors in the lidar measurements of atmospheric gases by differential absorption. J. Appl. Meteorol. **13**, 71–77 (1974). doi:10.1175/1520-0450(1974)013<0071:EITLMO>2.0.CO;2

Vömel, H., David, D.E., Smith, K.K.: Accuracy of tropospheric and stratospheric water vapor measurements by the cryogenic frost point hygrometer: Instrumental details and observations. J. Geophys. Res. **112**, D08305 (2007). doi:10.1029/2006JD007224

Werner, C., Herrmann, H.: Lidar Measurements of the Vertical Absolute Humidity Distribution in the Bounday Layer. J. Appl. Meteorol. **20**, 476–481 (1981). doi:10.1175/1520-0450(1981)020<0476:LMOTVA>2.0.CO;2

Wirth, M., Fix, A., Mahnke, P., Schwarzer, H., Schrandt, F., Ehret, G.: The airborne multi-wavelength water vapor differential absorption lidar WALES: system design and performance. Appl. Phys. B **96**, 201–213 (2009). doi:10.1007/s00340-009-3365-7

Zuev, V.V., Zuev, V.E., Makushkin, Y.S., Marichev, V.N., Mitsel, A.A.: Laser sounding of atmospheric humidity: experiment. Appl. Opt. **22**, 3742–3746 (1983). doi:10.1364/AO.22.003742

Aerosol Classification by Advanced Backscatter Lidar Techniques

29

Silke Groß, Martin Wirth and Michael Esselborn

Abstract

The high spectral resolution lidar (HSRL) method based on an iodine absorption filter and a frequency doubled pulsed Nd:YAG laser is presented. This method has the capability to directly measure the extinction and backscatter coefficients of aerosols and clouds. Measurements of an airborne HSRL system from four different field experiments are used to build up an aerosol classification. Two examples show the potential of this aerosol classification to distinguish between different aerosol types.

29.1 Introduction

For a better understanding of the effects of *aerosols* on the *climate system* it is important to obtain highly accurate information on the aerosol optical properties (e.g., *extinction coefficient*, *single scattering albedo* and *phase function*) as well as

S. Groß (✉) · M. Wirth
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Silke.Gross@dlr.de

M. Wirth
e-mail: Martin.Wirth@dlr.de

M. Esselborn
European Southern Observatory, Technology Division, Karl-Schwarzschild-Str 2,
85748 Garching, Germany
e-mail: messelbo@eso.org

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_29,

on their temporal and spatial distribution. Passive satellite instruments provide global coverage of aerosol distribution with high temporal resolution. However, to retrieve aerosol optical properties from passive instruments certain assumptions have to be made, causing errors in the retrieved data products. Furthermore, the passive remote satellite measurements provide only low or even no vertical resolution. Active *remote sensing* with *lidar* offers measurements with high vertical and temporal resolution and is therefore highly appropriate for studying atmospheric aerosols. To establish a global map of the vertical distribution of aerosols, the *backscatter lidar* CALIOP was launched on board the CALIPSO satellite in 2006 (Winker et al. 2007). But the main problem with this kind of lidar is that it only measures the atmospheric backscatter, a quantity where several *scattering* and *absorption* parameters are entangled in a nontrivial way. Thus, with such a system it is not possible to directly retrieve the climate-relevant extinction coefficient and thus the *aerosol optical depth* (AOD). For the derivation of the extinction coefficient further assumptions concerning the aerosol type have to be made. However, this procedure can be affected by misinterpretation and thus lead to large errors in the retrieved AOD. Therefore, a method to distinguish between different types of aerosols is highly valuable. Advanced lidar systems like Raman or *high spectral resolution lidar* are able to directly retrieve the extinction coefficient without further assumptions, and provide information that can be used to evaluate measurements from simpler backscatter lidar systems. Additionally, they provide information on optical parameters that can be used for an aerosol classification.

29.2 Basics of the High Spectral Resolution Lidar Method

As the light pulse emitted from a lidar propagates through the *atmosphere* its intensity is reduced by scattering and absorption by the molecular constituents of the air and by aerosol and *cloud particles*. Part of the light is also scattered in a backward direction and propagates back to the lidar system, again undergoing the same extinction process. In this way the signal of a lidar depends on the local ability of the atmosphere to scatter light and on the integral light extinction along the complete light path between the system and the probed volume. This can be seen in a more formal way from the so called lidar equation, which gives the power P received from an atmospheric volume at distance r:

$$P(r) = \left(\frac{E_0 c}{2} \cdot \frac{A}{r^2}\right) \cdot \left(\beta_m(r) \cdot T_a^2(r) + \beta_p(r) \cdot T_a^2(r)\right) \tag{29.1}$$

Here the first term in brackets contains system-specific parameters: the *laser* pulse energy E_0 and the area of the receiving telescope A (c is the speed of light). The atmospheric parameters are described by the *backscatter coefficients* for light scattering from air molecules β_m and particles β_p as well as the total atmospheric *transmission* T_a from the system to the probed volume. For a monochromatic light

source and in the absence of multiple scattering T_a can be written in the form (Beer's law):

$$T_a(r) = \exp\left(-\int_0^r \left(\alpha_m(r') + \alpha_p(r')\right) \cdot dr' \right) \tag{29.2}$$

where $\alpha_m(r)$ and $\alpha_p(r)$ are the molecular or particle extinction coefficients.

The molecular coefficients β_m and α_m are proportional to the air density and can be calculated from the molecular scattering cross-sections obtained from laboratory measurements and the pressure and temperature profile from *numerical weather prediction* (NWP) models or independent collocated measurements, e.g., from *radiosondes*. Given typical temperature and pressure errors of NWP the resulting uncertainties in β_m and α_m are below 1 %.

In sharp contrast to the molecular case, the coefficients describing light scattering by particles β_p and α_p are extremely sensitive to the size, shape and *refractive index* of the particles. Especially there is no simple general relation between the two parameters, as for molecular (*Rayleigh*) *scattering*. Thus in the lidar equation we are left with two unknown atmospheric parameters related to only one measured quantity. For the signal inversion of a simple backscatter lidar, the height dependent ratio α_p/β_p (the so called *lidar ratio*) is therefore usually intelligently guessed. It is constant for a specific type of aerosol or cloud and does not depend on the particle concentration. The type information has then, of course, to come from other sources like *in situ measurements* and a database of already analyzed aerosol types. This is a main error source and the major drawback of a conventional backscatter lidar.

If on the other hand one could devise a way to measure β_p and α_p independently, the above reasoning can be inverted and the aerosol or cloud type inferred (or greatly restricted) from the lidar measurement. Two solutions are commonly in use today, the one is the *Raman lidar* technique and the other the so-called high spectral resolution lidar (HSRL) method. The Raman lidar utilizes the at least two orders of magnitude smaller *Raman scattering* from nitrogen or oxygen molecules, which results in large systems, long integration times of up to 1 or 2 h, and a restricted daylight capability. It is therefore not reasonable for aircraft applications and is mainly used in ground based systems. It will not be further discussed in this article.

The HSRL method uses the fact that air molecules have a much lower mass and therefore a much higher thermal velocity than aerosol or cloud particles. This high random velocity leads to a significant spectral broadening of the laser light scattered by the molecules due to the Doppler effect. If plotted against the wavelength shift, the spectrum of the light scattered back to the lidar receiver looks like the dashed curve shown in Fig. 29.1. It consists of a broad, nearly Gaussian-shaped part coming from scattering by molecules and a narrow central peak from particle scattering. Now the HSRL trick is to use an optical notch filter to cut out the central part of the spectrum, allowing only photons which were scattered by molecules to reach the detector.

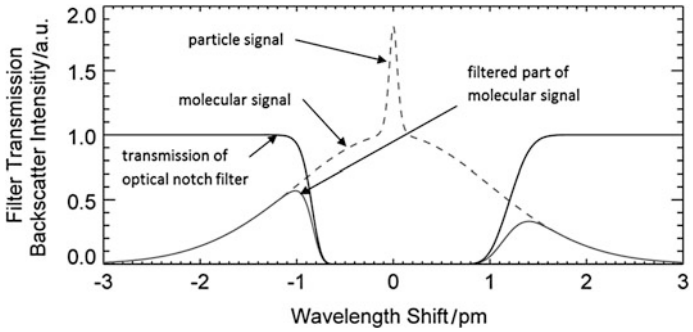

Fig. 29.1 Spectral signatures of the light backscattered from the atmosphere. The radiation scattered by air molecules is strongly Doppler-broadened due to their low mass and thus high thermal velocity. The light scattered by aerosol and cloud particles shows a spectrally much narrower distribution due to their much higher mass. By filtering the backscattered signal with an optical notch filter a signal is obtained which is proportional to the molecular scattering of the atmosphere alone. The example shown here is for light of 532 nm wavelength and a J_2 vapor filter (the vacuum wavelength of the iodine line used is 532.26002 nm)

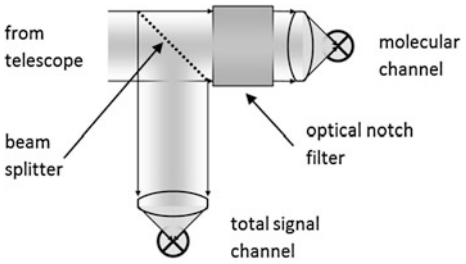

Fig. 29.2 Receiver setup of a high spectral resolution lidar. The light collected by the telescope is split into two branches. One part goes directly to an optical detector; the other part is fed through an optical notch filter, usually an atomic or molecular vapor filter

The lidar equation for the filtered signal is then reduced to the first term of Eq. 29.1:

$$P_{HSRL}(r) = C \cdot \left(\frac{E_0 c}{2} \cdot \frac{A}{r^2}\right) \cdot \beta_m(r) \cdot T_a^2(r) \tag{29.3}$$

where the calibration constant C accounts for the fact that also part of the molecular scattered light is rejected by the notch filter (cf. Fig. 29.1). The big advantage of this reduced equation is that from the particle scattering parameters it only contains α_p (wrapped into T_a). Since all other parameters are known or can be calculated with good accuracy, the atmospheric transmission between the lidar and an arbitrary point along the laser beam can be directly measured. From this data product it is straightforward to calculate α_p by simple numerical differentiation with a relative systematic error of usually less than 5 % at a vertical resolution of 200–500 m. To put it more graphically, the HSRL signal only shows the shadow of clouds and aerosol cast on the smooth and well-known background of molecular

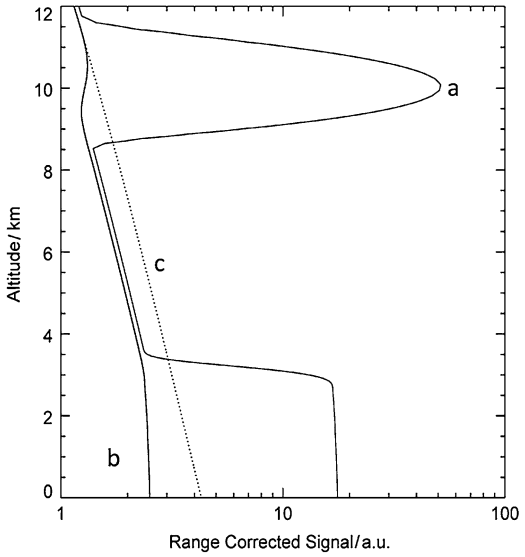

Fig. 29.3 Schematic of range-corrected signals of a downward looking HSRL as a function of altitude. **a** Total backscatter signal including molecular scattering, scattering from a cirrus cloud at 10 km altitude, and from aerosol within the planetary boundary layer below 3.5 km. **b** Backscatter signal from which the part scattered by cloud and aerosol particles has been removed by an optical notch filter. **c** Calculated backscatter signal for a pure molecular atmosphere without any cloud or aerosol particles

scattering. With a required integration time of only a few seconds this method is well suited to aircraft applications.

Since we not only want to measure the extinction coefficient α_p but also the backscatter coefficient β_p the signal in a HSRL receiver is split into two parts, one for the total signal and one with the optical notch filter as shown in Fig. 29.2. As the transmission from the HSRL channel is known, it is now easy to solve the lidar equation for β_p. Figure 29.3 shows how the signals for such a receiver would look.

The idea described above dates back to the end of the 1960s (Fiocco and Dewolf 1968), but the first systems used optical *interference* devices (Fabry–Perot etalons) for the notch filter, which has many disadvantages like only moderate aerosol suppression, low transmission, and extreme sensitivity to environmental pressure and temperature changes and to the mechanical alignment relative to the incoming beam. Hence such systems were never used in real atmospheric science applications but remained at the proof of concept stage. This situation changed with the discovery that an iodine (J_2) molecular vapor cell provides a suitable notch filter exactly at the wavelength of a frequency-doubled Nd:YAG laser, a laser type in widespread use in atmospheric backscatter lidar systems (Piironen and Eloranta 1994). Atomic or molecular vapor filters have none of the above mentioned drawback of etalons and are therefore optimally suited for rough environmental conditions as encountered in aircraft applications. Their only downside is that they are not tunable and a suitable high-power single-mode laser system has to be found at their absorption wavelength.

The worldwide first airborne iodine-filter based HSRL was flown by the DLR Institute of Atmospheric Physics during the LACE campaign in 1998 (Wandinger 2002). Since then several improvements have been made to this system (Esselborn

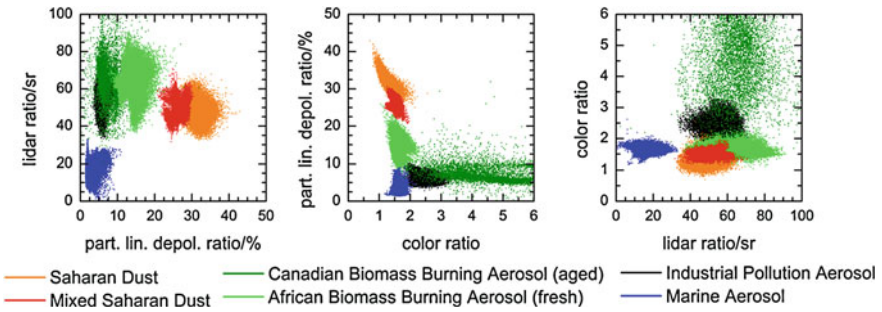

Fig. 29.4 Scatter plots of all pairwise combinations of three intensive lidar quantities for different aerosol types. Data are shown from four airborne campaigns over central Europe, the Saharan desert and the Cape Verde islands in the Atlantic Ocean

et al. 2008), finally resulting in the combined aerosol and *water vapor differential absorption lidar* WALES (Wirth et al. 2009). The examples shown later on were mainly obtained with this system, which provides aerosol channels at 1 064 and 532 nm wavelength, the latter equipped with an additional HSRL channel. At both wavelengths the *depolarization* of the atmospheric scattering is also measured, allowing for a closer characterization of the probed aerosol and cloud particles (Esselborn et al. 2009).

29.3 Aerosol Classification Using Lidar Measurements

The aerosol classification by HSRL is based on so-called intensive aerosol properties: the lidar ratio, the particle linear depolarization ratio and the *color ratio* of backscattering. While the extinction and backscatter coefficient depend on the amount of aerosols (extensive properties), the intensive aerosol properties are ratios of extensive lidar quantities and therefore independent of the aerosol concentration. As the values of the intensive lidar quantities are quite different for different aerosol types (see Fig. 29.4), a combination is very appropriate for distinguishing between different aerosol types. For this to work a database has to be built where the optical parameters of different aerosol types are catalogued. To unambiguously identify the aerosol type, a combination of airborne in situ measurements, calculated backward trajectories, and satellite images is used. An optimal strategy is to perform measurements of lidar together with a (nearly) simultaneous airborne in situ analysis of microphysical properties and chemical composition. The sampling strategy generally consisted of two parts: remote sounding at high altitude to observe and identify interesting aerosol layers by HSRL and then in situ sounding within the layers at several height levels (Weinzierl 2011). Furthermore, temporally resolved backward trajectories are compared with periods of possible aerosol emission (e.g., *mineral dust* release or

fire burning times) derived from satellite information (Groß et al. 2011). Figure 29.4 shows examples of measurements by a HSRL from four different field campaigns.

In nature, aerosol layers do not always consist of pure aerosol types. Mixtures of two or more components are frequently found. For a two-component mixture the values of the lidar ratio and the particle linear depolarization ratio can easily be calculated if both values for the pure aerosol types are known. Therefore, the contribution of the single aerosol type to the total extinction coefficient of particles is used (Groß et al. 2011). The total extinction and backscatter coefficient of particles of a two-component mixture are

$$\alpha_{p,mix} = \alpha_1 + \alpha_2 = (1 - x) \cdot \alpha_{p,mix} + x \cdot \alpha_{p,mix} \tag{29.4}$$

and

$$\beta_{p,mix} = \beta_1 + \beta_2 = \frac{(1 - x) \cdot \alpha_{p,mix}}{S_{p,1}} + \frac{x \cdot \alpha_{p,mix}}{S_{p,2}} \tag{29.5}$$

with α_1 and α_2, and β_1 and β_2 being the extinction coefficient and backscatter coefficient, respectively, of component 1 and component 2, and $x = \alpha_2/\alpha_{p,mix}$ being the fraction of component 2 to the total extinction coefficient of particles. The lidar ratio $S_{p,mix} = \alpha_{p,mix}/\beta_{p,mix}$ of a two-component mixture is then given by

$$S_{p,mix} = \frac{1}{\frac{1-x}{S_{p,1}} + \frac{x}{S_{p,2}}} \tag{29.6}$$

where $S_{p,1}$ and $S_{p,2}$ are the lidar ratios of the two pure aerosol components in the mixture. The particle linear depolarization ratio of a two-component mixture can be calculated from

$$\delta_{p,mix} = \frac{\frac{1-x}{S_{p,1}} \cdot \frac{\delta_{p,1}}{1+\delta_{p,1}} + \frac{x}{S_{p,2}} \cdot \frac{\delta_{p,2}}{1+\delta_{p,2}}}{\frac{1-x}{S_{p,1}} \cdot \frac{1}{1+\delta_{p,1}} + \frac{x}{S_{p,2}} \cdot \frac{1}{1+\delta_{p,2}}} \tag{29.7}$$

with $\delta_{p,1}$ and $\delta_{p,2}$ being the particle linear depolarization ratio of the two aerosol components considered. In Fig. 29.5 mixing lines of *biomass burning* aerosols and mineral dust (thick lines), and of mineral dust and marine aerosols (thin lines) are shown, which do not follow a linear characteristic.

29.4 Examples of Airborne HSRL Measurements

Taking into account the above described aerosol classification, different aerosol types can be identified by HSRL measurements as is shown in the following two examples (Fig. 29.6). The measurements were performed in the framework of the

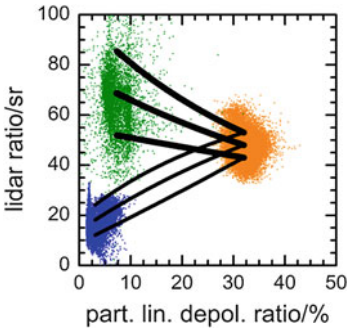

Fig. 29.5 Characteristic lidar quantities of biomass burning aerosol (*left top*), Saharan dust (*right middle*) and marine aerosol (*left bottom*), as well as mixing lines of pairs of two aerosol types. The *mixing lines* are shown for the mean values of the aerosol properties (*central curve*), as well as for cases where uncertainties about the mean values are considered

Saharan mineral dust experiment (SAMUM), Heintzenberg (2009). SAMUM is a closure experiment and aimed at the characterization of physical, chemical and optical properties of Saharan dust, a major component of the atmospheric aerosol load.

The first example (Fig. 29.6—left panels) is for a pure dust case on 19 May 2006. The measurement was conducted during the first campaign of the SAMUM project, which took place in Ouarzazate, Morocco, in May and June 2006 to investigate the properties of pure fresh Saharan dust close to the source region. From the HSRL *backscatter ratio* $((\beta_a + \beta_m)/\beta_m)$ at 532 nm an aerosol layer from the ground up to 5.5 km above mean sea level is obvious. However, the backscatter ratio does not provide further information about the observed aerosol type. Considering the information of the above discussed intensive parameters, the aerosol layer can clearly be identified as a fresh Saharan dust layer. The Saharan desert could be identified as the source region of these air masses from backward trajectory calculations.

The second field experiment of the SAMUM project took place in Praia, Cape Verde, in January and February 2008. The aim of this campaign was to study Saharan dust at the beginning of long-way transport across the Atlantic Ocean, to investigate possible transformation of the Saharan dust properties during transport, and to observe the mixing of Saharan dust with other aerosol types. Figure 29.6 (right panels) shows a case from 4 February 2008. Looking at the HSRL backscatter ratio at 532 nm, three different aerosol layers can be identified. But from the backscatter ratio alone one is not able to assign a certain aerosol type or mixture to each of the layers. Using the intensive lidar properties the components of the different layers can be identified, clearly showing a marine aerosol layer in the lower 500 m topped by a mixed layer of Saharan dust and biomass burning aerosol and a lofted layer of African biomass burning aerosol. This conclusion is supported by backward trajectories showing a strong influence of marine air

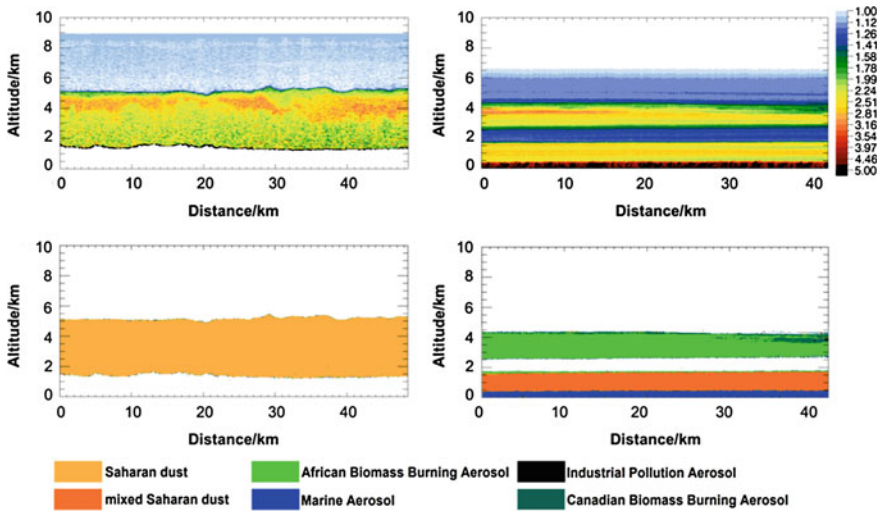

Fig. 29.6 HSRL backscatter ratio measurement at 532 nm on 19 May 2006 over Morocco (*upper left panel*) and on 4 February 2008 over the Atlantic Ocean near the Cape Verde islands (*upper right panel*), and aerosol-type mask of both measurements (*lower panels*)

masses in the lower layer and the location of the middle and upper layer in the Saharan desert and central African regions, respectively.

29.5 Conclusion

For many radiative transfer models and remote sensing applications an aerosol type has to be defined. This aerosol type is described by a variety of chemical, optical, and microphysical properties. These aerosol properties are independent of the aerosol concentration and partly show large differences for different types. The choice of the right aerosol type is therefore a critical process, as it can lead to large errors in the retrieved data products. A method based on HSRL measurements of a set of intensive aerosol properties, in particular the lidar ratio, the particle linear depolarization ratio and the color ratio of backscatter, was presented to distinguish between different aerosol types. This is in particular useful for discriminating hazardous aerosol types such as *volcanic ash*, which pose a risk for flight safety, from less harmful aerosols. Furthermore, the differentiation is valuable for satellite remote sensing. Future satellite missions of the European Space Agency (ESA) will deploy HSRL systems. For example, the ATLID (Atmospheric Lidar) instrument will be part of the EarthCARE (Earth Clouds, Aerosol and Radiation Explorer) mission. Using the presented method, the HSRL measurements on EarthCARE will not only provide the climate relevant properties, the extinction coefficient and aerosol optical depth, but also a global distribution of the vertical and spatial structure of aerosols and clouds. Together with other active and passive

remote sensing instruments on board, this may lead to improved satellite-retrieved data products and thereby to improved understanding of the *climate impact* of aerosols.

References

Esselborn, M., Wirth, M., Fix, A., Tesche, M., Ehret, G.: Airborne high spectral resolution lidar for measuring aerosol extinction and backscatter coefficients. Appl. Opt. **47**, 346–358 (2008). doi:10.1364/AO.47.000346

Esselborn, M., Wirth, M., Fix, A., Weinzierl, B., Rasp, K., Tesche, M., Petzold, A.: Spatial distribution and optical properties of saharan dust observed by airborne high spectral resolution lidar during samum 2006. Tellus B **61**, 131–143 (2009). doi:10.1111/j.1600-0889.2008.00394.x

Fiocco, G., Dewolf, J.B.: Frequency spectrum of laser echoes from atmospheric constituents and determination of the aerosol content of air. J. Atmos. Sci. **25**(3), 488–496 (1968). doi:10.1175/1520-0469(1968)025<0488:FSOLEF>2.0.CO;2

Groß, S., Tesche, M., Freudenthaler, V., Toledano, C., Wiegner, M., Ansmann, A., Althausen, D., Seefeldner, M.: Characterization of Saharan dust, marine aerosols and mixtures of biomass-burning aerosols and dust by means of multi-wavelength depolarization and Raman lidar measurements during SAMUM 2. Tellus B **63**, 706–724 (2011). doi:10.1111/j.1600-0889.2011.00556.x

Heintzenberg, J.: The SAMUM-1 experiment over Southern Morocco: overview and introduction. Tellus B **61**, 2–11 (2009). doi:10.1111/j.1600-0889.2008.00403.x

Piironen, P., Eloranta, E.W.: Demonstration of a high-spectral-resolution lidar based on an iodine absorption filter. Opt. Lett. **19**, 234–236 (1994). doi:10.1364/OL.19.000234

Wandinger, U., et al.: Optical and microphysical characterization of biomass-burning and industrial-pollution aerosols from multiwavelength lidar and aircraft measurements. J. Geophys. Res. **107**(D21), 8125 (2002), doi: 10.1029/2000JD000202

Weinzierl, B., et al.: Microphysical and optical properties of dust and tropical biomass burning aerosol layers in the Capverde region—an overview of the airborne in situ and lidar measurements during SAMUM-2. Tellus B **63**(4), 589–618 (2011)

Winker, D.M., Hunt, W.H., McGill, M.J.: Initial performance assessment of CALIOP. Geophys. Res. Lett. **34**, L19803 (2007). doi:10.1029/2007GL030135

Wirth, M., Fix, A., Mahnke, P., Schwarzer, H., Schrandt, F., Ehret, G.: The airborne multi-wavelength water vapor differential absorption lidar WALES: system design and performance. Appl. Phys. B **96**, 201–213 (2009). doi:10.1007/s00340-009-3365-7

Wind Lidar for Atmospheric Research

30

Oliver Reitebuch

Abstract

Wind lidars use the optical Doppler effect to measure the atmospheric wind with high spatial and temporal resolution. This offers a wide range of applications from aircraft wake vortex detection and characterization to measurement of turbulent quantities like fluxes to resolving small and mesoscale atmospheric flows, and even global wind profiling from space in the future. Principles of wind lidars including two complementary techniques and applications for atmospheric research are discussed.

30.1 Introduction

Wind observations with *lidar* (light detection and ranging) systems were used by scientists from the DLR Institute of Atmospheric Physics (IPA) to study atmospheric flow on a large variety of scales: from atmospheric turbulence to aircraft wake vortices or flows around wind turbines to mesoscale atmospheric flows, e.g., within mountain valleys or land-sea breeze circulation, to large scale flows like jet streams or typhoon outflows. Different types of wind lidars allow measurements in different altitude regions of the atmosphere from the atmospheric boundary layer with high aerosol content to the upper troposphere, stratosphere and even mesosphere with molecules as backscattering targets. Optical *remote sensing* with lidar

O. Reitebuch (✉)
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Oliver.Reitebuch@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_30,

is a very powerful method for measuring the atmospheric wind vector. It is possible to measure the horizontal wind vector as well as the vertical wind component with different viewing geometries. Wind lidars are operated from ground, balloon, ship, or aircraft and will be deployed on satellites in the future.

The movement of the air is described by the wind vector $\mathbf{V} = (U, V, W)$, which is composed of three components, in contrast to most other atmospheric variables such as temperature, humidity or pressure, which are scalar parameters. Usually the wind vector is described by a vertical component W (up and down along the axis of Earth gravity, positive for an updraft wind) and two horizontal components U, V. The horizontal wind vector ("velocity") can be either described by the magnitude of the vector V_h ["wind speed"; $V_h = \sqrt{(U^2 + V^2)}$] and by the wind direction θ (angle between the horizontal wind vector and north; $\theta = \arctan(U/V)$; $\theta = 0°$ for northerly wind, wind from north towards south, $\theta = 90°$ for easterly wind, wind from east to west), or it is decomposed into the two components along the west-east axis U (zonal wind component, positive if wind is blowing from west towards east, which is a westerly wind) and along the south-north axis V (meridional wind component, positive if wind is blowing from south to north, which is a southerly wind).

30.2 Principle of Wind Lidar Methods

Most wind lidars rely on the use of the optical Doppler effect to measure wind speed, thus they are usually named *Doppler wind lidars* (DWL). Nevertheless, it should be mentioned that correlation techniques from two successive images (similar to the derivation of cloud-drift winds) or from two different range measurements (e.g., used by the police for speed measurements) are also used for specific applications to determine wind speed.

DWL instruments determine the component of the wind vector which is projected onto the *laser* beam propagation direction—the so called *line-of-sight* (LOS) with its unit pointing vector **D**. The vector **D** can be described by the angle α, between the LOS direction and the zenith (for ground DWL) or nadir (for airborne and space-borne DWL) and the azimuth angle β between the LOS direction and the north direction in the horizontal plane. This LOS wind speed component v_{LOS}, or v for short, (in units of m/s) is obtained by the projection of the wind vector **V** on the LOS direction **D**:

$$v_{LOS} = \vec{D} \cdot \vec{V} = \sin(\alpha) \cdot \sin(\beta) \cdot U + \sin(\alpha) \cdot \cos(\beta) \cdot V + \cos(\alpha) \cdot W$$
$$v_{LOS} = \sin(\alpha) \cdot V_{h,||} + \cos(\alpha) \cdot W$$

When only the horizontal wind speed $V_{h,||}$ in the plane of the LOS direction is considered, then v_{LOS} is determined by the off-nadir or off-zenith angle α (as indicated in Fig. 30.2). Thus for a measurement of the vertical wind W, the laser beam has to be directed vertically ($\alpha = 0°$), whereas for the horizontal wind

vector, a minimum of two preferably orthogonal LOS directions tilted towards the horizontal plane with an angle of α are used. If the LOS direction and the wind vector are perpendicular, then no LOS wind speed is observed. Depending on the pointing direction **D** towards zenith from ground or nadir from space, the sign of the above equations has to be adapted to be consistent with the meteorological definition of the sign for U, V, W and the sign definition for v_{LOS} in the *Doppler shift* equations below.

A DWL determines the wind speed of the air by using the collective motion of a large number of aerosols, cloud particles and molecules within the air volume as tracers. Different methods are applied for using aerosol and cloud particles as tracers with a *coherent detection DWL* and molecules as tracers with a *direct-detection DWL* (which is also called "incoherent" detection DWL). A direct-detection lidar relies on the measurement of the signal intensity or number of photons, whereas a coherent lidar measures in addition the phase and frequency of the backscattered signal. Overviews of DWL principles, techniques and applications can be found in Vaughan et al. (1996), Werner (2004), and Henderson et al. (2005).

30.2.1 The Doppler Effect

Most wind lidars rely on the Doppler effect, which was first described by the Austrian physicist Christian Doppler in 1842. The Doppler effect can be observed for *acoustic waves*, which is used in remote sensing with a Doppler sodar (sonic detection and ranging), for electromagnetic waves, which is used for Doppler *radar* (radio detection and ranging), and for lidar. It describes the change in *frequency* (or *wavelength*) that is present when the source of the wave and the observer are in relative motion to each other. The correspondence of the frequency f_0 and the wavelength λ_0 of an electromagnetic wave is obtained using the *speed of light* c ($c = 2.9979 \times 10^8$ m/s):

$$f_0 = \frac{c}{\lambda_0}$$

The change in frequency Δf or wavelength $\Delta\lambda$ is referred to as the Doppler shift and can be positive or negative, depending on the direction of the relative movement. For the case that a laser is emitting a light wave with frequency f_0, an observer moving with velocity v would detect the frequency f_1 with:

$$f_1 = f_0 \cdot \left(1 + \frac{v}{c}\right)$$

The velocity $v = v_{LOS}$ describes the movement of source and observer relative to each other. The Doppler frequency shift Δf or wavelength shift $\Delta\lambda$ is obtained from:

Table 30.1 Doppler shift for a line-of-sight speed of 1 m/s for typical wavelengths of a DWL

λ_0/μm	f_0/THz	$\Delta\lambda$/fm	Δf/MHz
0.355	845	2.37	5.63
2.02	148	13.5	0.987
10.6	28.3	70.7	0.189

$$\Delta f = f_1 - f_0 = f_0 \cdot \frac{\mathrm{v}}{c}$$
$$\Delta\lambda = \lambda_1 - \lambda_0 = -\lambda_0 \cdot \frac{\mathrm{v}}{c}$$

The Doppler wavelength shift $\Delta\lambda$ is positive (called the "red shift", because the wavelength is shifted towards larger wavelengths λ_1, which are towards red spectral colors) if the velocity v is negative and the distance between source and observer increases. If the velocity v is positive, then the observer is moving towards the source, and the wavelength is decreasing ("blue shift") and the frequency increasing.

For a Doppler lidar (but also for Doppler radar and sodar) the Doppler effect occurs two times. First, when the light wave from the emitting laser source with a frequency of f_0 is perceived by the moving, scattering particle (aerosol, cloud, molecule) and thus is exposed to a light wave of frequency f_1; the second Doppler effect occurs when the moving particle re-emits the scattered wave with frequency f_1, which is perceived by the lidar receiver as a wave with frequency f_2:

$$f_2 = f_1 \cdot \left(1 + \frac{\mathrm{v}}{c}\right) = f_0 \cdot \left(1 + \frac{\mathrm{v}}{c}\right)^2 = f_0 \cdot \left(1 + 2 \cdot \frac{\mathrm{v}}{c} + \frac{\mathrm{v}^2}{c^2}\right) \approx f_0 \cdot \left(1 + 2 \cdot \frac{\mathrm{v}}{c}\right)$$

The term v^2/c^2 can be neglected, as it is a factor of 10^{-7}–10^{-8} smaller than the term v/c for usual atmospheric wind speeds v compared to the speed of light c. The Doppler frequency and wavelength shift can be finally written as:

$$\Delta f = f_2 - f_0 = 2 \cdot f_0 \cdot \frac{\mathrm{v}}{c}$$
$$\Delta\lambda = \lambda_2 - \lambda_0 = -2 \cdot \lambda_0 \cdot \frac{\mathrm{v}}{c}$$

The relative Doppler shift expressed as $\Delta f/f_0$ or $\Delta\lambda/\lambda_0$ is in the order of 10^{-7}–10^{-8} for electromagnetic waves and typical atmospheric wind speeds. It is very demanding to detect such small changes in the frequency of a light wave, which sets high challenges for the laser sources and optical receivers used for a DWL. For illustration, the Doppler shifts for typical wavelengths of a DWL are provided in Table 30.1. A 1 m/s LOS wind speed would change the wavelength of a DWL from 2.02 to 2.0200000135 μm, by only 13.5 fm (1 fm = 10^{-15} m).

30.2.2 Spectral Line Shape from Molecular and Particle Backscatter

The collective, mean movement of all molecules, aerosol, and cloud particles within a certain atmospheric volume is referred to as wind. This is superimposed by the random, thermal motion of the molecules and particles ("Brownian motion") and causes so-called *Doppler broadening*. This motion can be described by the *Maxwell–Boltzmann velocity distribution*, which is dependent on the atmospheric temperature T and the mass m of a single molecule (or particle). The mean velocity of the air molecules (mixture of gases with mean mass $m = 4.79 \times 10^{-26}$ kg) is quite high with a value of 459 m/s for a temperature of 15 °C. If only one component in the LOS direction of the velocity distribution is considered, then it can be described by a Gaussian distribution with a standard deviation in velocity space σ_v or wavelength space σ_λ after applying the Doppler shift equation:

$$\sigma_v = \sqrt{\frac{k \cdot T}{m}} \qquad \sigma_\lambda = \frac{2 \cdot \lambda_0}{c} \sqrt{\frac{k \cdot T}{m}}$$

with Boltzmann constant k ($k = 1.38 \times 10^{-23}$ J/K). The Gaussian approximation is only valid for dilute gases. For typical atmospheric densities an additional *Brillouin scattering* effect on moving pressure fluctuations has to be considered, which is described by Rayleigh-Brillouin scattering (Witschas et al. 2010). The Gaussian velocity distribution is quite broad for air molecules with $\sigma_v = 288$ m/s and a *full-width at half-maximum* (FWHM) of 678 m/s for 15 °C. (FWHM $= 2\sqrt{(2 \ln(2))} = 2.355$ for a Gaussian distribution with unit standard deviation.) For aerosol and cloud particles the width of the velocity distribution is several orders of magnitude smaller due to the larger mass of aerosols and cloud particles. The assumptions of typical particle radii of 0.1–1 μm, which are of sizes relevant for Mie scattering at laser wavelengths, and of spherical water particles (density of 10^3 kg/m^3) lead to typical masses of 4.2×10^{-18}–4.2×10^{-15} kg. Thus Doppler broadening is a factor of 10^4–10^6 smaller for aerosol particles compared to molecules.

The mean of the Gaussian velocity distribution is given by the mean, collective movement of all molecules, which is referred to as LOS wind speed. For the case that a laser emits a wavelength λ_0, the backscattered light from molecules is spectrally broadly distributed, which is illustrated in Fig. 30.1, where the distribution of Doppler shifts is shown for a typical DWL wavelength of 355 nm. A broad bandwidth spectrum is obtained from molecules, whereas a narrow bandwidth spectrum is obtained from aerosol and cloud particles. In the case of a mean wind speed of the air volume, both the aerosol and the molecular spectrum are shifted in parallel along the wavelength axis.

The measurement of the wind speed is directly related to the measurement of the mean (or first moment) of the velocity or Doppler wavelength shift distribution. Either the broad bandwidth molecular backscatter return or the narrow

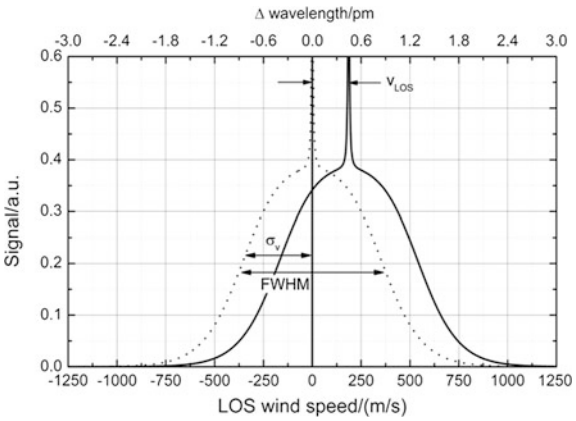

Fig. 30.1 Spectral distribution of LOS wind speed in m/s of broad bandwidth molecular and small bandwidth aerosol signal in arbitrary units (a.u.); corresponding Doppler wavelength shifts in pm for an emitted wavelength of $\lambda_0 = 355$ nm (1 pm equivalent to 422 m/s), standard deviation σ_v and FWHM are indicated; one spectrum (*dotted lines*) is shown for a mean LOS wind speed of 0 m/s, while the other spectrum (*bold line*) is shown for 180 m/s LOS wind speed (equivalent to a Doppler shift of 0.44 pm)

bandwidth aerosol backscatter return is used for wind determination. Although both techniques differ fundamentally in their technical realization (see next sections), the fundamental, theoretical limit of the minimum random error is given by the spectral distribution. This theoretical, statistical limit of the random error ε_v for the wind speed estimation only depends on the spectral width σ_v and the number of backscattered photons N_{ph} in the case of pure Poisson distributed photon noise:

$$\varepsilon_v = \frac{\sigma_v}{\sqrt{N_{photons}}} = \frac{\sigma_v}{\sqrt{N_{pulses} \cdot N_{photons/pulse}}}$$

The spectral width of the aerosol backscatter is dominated by the spectral width of the emitted laser pulse, due to the negligible Doppler broadening of the aerosol. Typically, this laser width is a factor of 100 smaller than the width of the broadband Rayleigh spectrum. Thus a factor of 10^4 more photons are needed to obtain the same random error if the Rayleigh spectrum is used for wind measurement. This can be achieved by averaging the backscattered signal from a number of laser pulses N_{pulse} and by increasing the detected number of photons per pulse $N_{photons/pulse}$, e.g., by higher laser pulse energies.

30.3 Wind Lidar Methods

IPA is one of few institutes worldwide where both DWL methods—namely coherent and direct-detection—are developed and operated. Some key parameters of airborne DWL at IPA are shown in Table 30.2 and are further discussed in the

Table 30.2 Key parameters of airborne Doppler wind lidar (DWL) at DLR

	Wind infrared Doppler lidar WIND	2-μm DWL	ALADIN airborne demonstrator A2D
Developer	DLR, CNRS, CNES, Meteo-France	DLR, Lockheed Martin Coherent Technologies	DLR, EADS-Astrium
Maiden flight	1999	2001	2005
Major campaigns	MAP 1999 (Alps) ESCOMPTE 2001 (France) AMMA 2006 (West Africa)	VERTIKATOR 2002 (Alps) A-TReC 2003 (Iceland) T-PARC 2008 (Japan) several aircraft wake vortex campaigns, e.g. A380 Eyjafjalla volcanic ash 2010 (Europe, Iceland)	Aeolus pre-launch validation 2007, 2008, 2009 (central Europe, Iceland)
Lidar principle	Coherent, heterodyne detection	Coherent, heterodyne detection	Direct-detection with double-edge and fringe imaging technique
Atmospheric backscatter	Aerosol and cloud particles	Aerosol and cloud particles	Molecules, aerosol and cloud particles
Laser type and wavelength	CO_2 gas laser 10.6 μm	Solid-state Tm:LuAG 2.02254 μm	Solid-state Nd:YAG 354.9 nm
Laser energy, repetition rate, pulse length	100–200 mJ 10 Hz 2–3 μs	1–2 mJ 500 Hz 400 ns (FWHM)	50–60 mJ 50 Hz 25 ns (FWHM)
Telescope Ø (cm)	20	10.8	20
Scanner	Conical scan with 30°	Conical scan with 20°, arbitrary scan patterns with double wedge	No scanner, only fixed LOS under 20° slant angle
Vertical resolution (m)	250	100	250
Power-aperture product mean ($mW \cdot m^2$)	47	7	86
Precision of wind observation	0.5–1.5 m/s horizontal wind speed	0.3–1.0 m/s horizontal wind speed	1–2 m/s LOS wind speed

next section. All were deployed on the DLR Falcon aircraft and used for numerous atmospheric and validation campaigns. Table 30.3 provides a comparative overview of coherent and direct-detection wind lidar. Coherent DWL usually operates in the eye-safe, infrared region of the spectrum with wavelengths of 1.5–1.6, 2 and 10 μm, whereas direct-detection DWL uses ultraviolet (355 nm), visible (532 nm) or near-infrared (1 064 nm) wavelengths. Coherent DWL are used for applications

Table 30.3 Comparative overview of different wind lidar techniques

	Coherent detection	Direct-detection
Wavelengths and lasers	CO_2 gas lasers 10.6 μm Tm:LuAG or Tm,Ho:YAG at 2 μm Er-doped fiber at 1.5–1.6 μm	Nd:YAG at 1 064 nm, frequency-doubled at 532 nm, and -tripled at 355 nm
Atmospheric backscatter	Aerosol and cloud particles narrow bandwidth line shape	Molecules with Rayleigh-Brillouin lineshape (broad bandwidth), aerosol and cloud return
Receiver	Mainly heterodyne detection some earlier systems used homodyne detection	(a) filter techniques, e.g., double edge or iodine filter (b) fringe imaging, e.g., Fizeau (c) two-beam interferometer, e.g., Mach–Zehnder
Dominating noise sources	Photon noise of local oscillator, speckle noise, atmospheric turbulence	Solar background light and signal photon noise
Calibration	No calibration needed, as relation between Doppler frequency shift and detected signal is obtained by Fourier transformation.	Calibration needed to obtain relationship between Doppler frequency shift and detected signal
Need for additional parameters	No additional parameters needed for wind retrieval	Atmospheric temperature and pressure needed for correction of retrieved wind from molecular backscatter
Technology readiness	Systems commercially available	Research systems, satellite lidar under development
Main applications	Need for high precision and high temporal resolution, aircraft wake vortex, turbulence, wind turbines, boundary layer winds, vertical wind speed, and turbulent fluxes	Wind sensing for higher altitudes of upper troposphere, stratosphere and mesosphere

where sufficient aerosol backscatter is present, e.g., for atmospheric boundary layer winds, whereas direct-detection DWL is the system of choice for higher altitudes up to the mesosphere. Coherent DWL are commercially available today, can be very compact and reliable, and high temporal resolution and high precision can be achieved, which even allows the determination of vertical winds, turbulent quantities, or aircraft wake vortices. Direct-detection DWL using molecular backscatter are the systems of choice for wind measurements in the upper troposphere, stratosphere and even up to the mesosphere.

30.3.1 Coherent Doppler Wind Lidar

The coherent DWL is the most common system concept today and is available as a reliable and easy to operate (“turn-key”) instrument from commercial companies. A coherent DWL is purely using the narrow bandwidth aerosol and cloud particle

backscatter return. Its principle relies on the optical mixing of two optical beams—one returned from the atmosphere and one from an instrumental internal source—and detection of the optical beat signal. Coherence is a characteristic of an electromagnetic wave which permits *interference* and is related to the *wave phase*. Waves radiating with a fixed relationship between their phases are coherent, which is the case for laser emissions, whether cw or pulsed. In contrast radiation from a light bulb is not coherent, because of the spontaneous, thermal generation of light waves with random relation between their phases. In principle, every type of lidar uses coherent emission with a laser source. A specific type of DWL is called coherent, because it uses in addition a coherent laser in the optical receiver and detection part of the instrument. The coherent DWL makes use of an additional cw laser source (called "local oscillator") and the coherence of the backscatter signal from the atmosphere to obtain a beat frequency signal from the interfering waves during optical mixing.

The Doppler frequency is determined by mixing two optical beams on an optical element. Here, the atmospheric signal $E_s(t)$ with its frequency $f_S = f_2 = f_0 + \Delta f$ and phase φ_S is optically mixed with the cw laser beam of the local oscillator $E_{LO}(t)$ with frequency f_{LO}. Both electrical field vectors of the light beams $E_s(t)$ and $E_{LO}(t)$ are summed up to $E_D(t)$, where the overbar denotes a temporal average in the following equations.

$$E_S(t) = \bar{E}_S \cdot \cos(2\pi \cdot f_s \cdot t + \varphi_S)$$
$$E_{LO}(t) = \bar{E}_{LO} \cdot \cos(2\pi \cdot f_{LO} \cdot t)$$
$$E_D(t) = E_S(t) + E_{LO}(t) = \bar{E}_S \cdot \cos(2\pi \cdot f_s \cdot t + \varphi_S) + \bar{E}_{LO} \cdot \cos(2\pi \cdot f_{LO} \cdot t)$$

The resulting beam with its electrical field $E_D(t)$ is directed to a detector which is sensitive to light intensities proportional to $|E_D(t)|^2$. As the detector current $I_D(t)$ cannot follow the optical frequencies with THz variations (second eq. below, last term) because those are much higher than the bandwidth of the detector, only the time-averaged signal of $|E_D(t)|^2$ is detected:

$$I_D(t) \propto \overline{|E_D(t)|^2}$$
$$I_D(t) \propto \tfrac{1}{2}\bar{E}_S^2 + \tfrac{1}{2}\bar{E}_{LO}^2 + \eta_{eff}\bar{E}_S\bar{E}_{LO}\left[\overline{\cos(2\pi \cdot (f_S - f_{LO}) \cdot t + \varphi_S)} + \underbrace{\overline{\cos(2\pi \cdot (f_S + f_{LO}) \cdot t + \varphi_S)}}_{=0}\right]$$
$$I_D(t) \propto \tfrac{1}{2}\bar{E}_S^2 + \tfrac{1}{2}\bar{E}_{LO}^2 + \eta_{eff}\bar{E}_S\bar{E}_{LO}\cos(2\pi \cdot (f_S - f_{LO}) \cdot t + \varphi_S)$$

with the efficiency η_{eff} of the optical mixing, which describes the efficiency of the matching of the spatial interference of the atmospheric signal beam and the LO beam (called "mode matching") and their temporal correlation. The efficiency η_{eff} is 1 for coherent beams, which are perfectly correlated in the spatial and temporal domain, whereas it approaches zero for non-coherent signals. In the case of

non-coherent radiation sources, e.g., solar background light, the signal $E_s(t)$ consists of a large number of oscillating waves with a random phase distribution φ_S. Thus the temporal average over the first cos-function (second equation above) would approach zero and the information about the difference frequency $f_S - f_{LO}$ would be lost. This makes the coherent detection method almost insensitive to non-coherent background light.

Due to the temporal averaging of the detector, only the difference frequency $f_S - f_{LO}$ is perceived, as the time average of the second cos-function with the sum frequency $f_S + f_{LO}$ signal is zero (second equation above, last term). The difference frequency is called the beat frequency f_{beat} and is a direct measure of the Doppler frequency shift and can be obtained using heterodyne ($f_{LO} \neq f_0$) or homodyne ($f_{LO} = f_0$) detection:

$$f_{\text{beat,heterodyne}} = |f_S - f_{LO}| = |f_0 + \Delta f - f_{LO}|$$
$$f_{\text{beat,heterodyne}} = |\Delta f|$$

The beat frequency f_{beat} provides only the absolute value of the frequency difference. It cannot be decided which frequency f_S or f_{LO} is higher or lower. For *heterodyne detection*, the frequency f_{LO} is selected such that it is different from the frequency of the outgoing laser pulse by a constant offset ($f_0 - f_{LO}$). Then the Doppler shift frequency Δf, with its positive and negative sign, is spread around this offset frequency, which allows determination of the sign and magnitude of the Doppler shift. In the case of *homodyne detection*, where $f_{LO} = f_0$, it is not possible to determine the sign of the Doppler shift.

The principle optical layout of a heterodyne detection DWL is shown in Fig. 30.2, where the frequency of the emitted laser pulse f_0 is monitored by a further heterodyne optical path with a reference detector. This allows for correcting the temporal variation of the emitted laser pulse frequency f_0. In addition, an optical beam steering device ("scanner") is included to deflect the outgoing laser pulse towards different LOS directions for the measurement of several wind components. The detector current $I_D(t)$ is digitized by an analog-to-digital converter and used to estimate the beat frequency $f_0 - \Delta f - f_{Lo}$. Several frequency estimators can be applied either in the time domain or after transforming the temporal signal by a Fourier transformation in the frequency domain.

In contrast to Doppler radar, where a fixed relation exists between the phase of successive radar pulses, this is not the case for laser pulses, which are generated by independent amplification processes within the laser. In addition, the backscattering particles in the atmosphere are randomly distributed and are not stationary but obey a thermal motion. This results in a randomly fluctuating phase $\varphi_{s,i}(t)$ from different scattering particles i for different times t, which causes a spatial and temporal variation of the signal called the *speckle" effect*. Thus it is not possible to average the time signal $I_D(t)$ of successive laser pulse returns for a coherent detection lidar, as is usually done for other lidar types. A temporal averaging of $I_D(t)$ would result in the loss of the phase information and thus in a loss of the

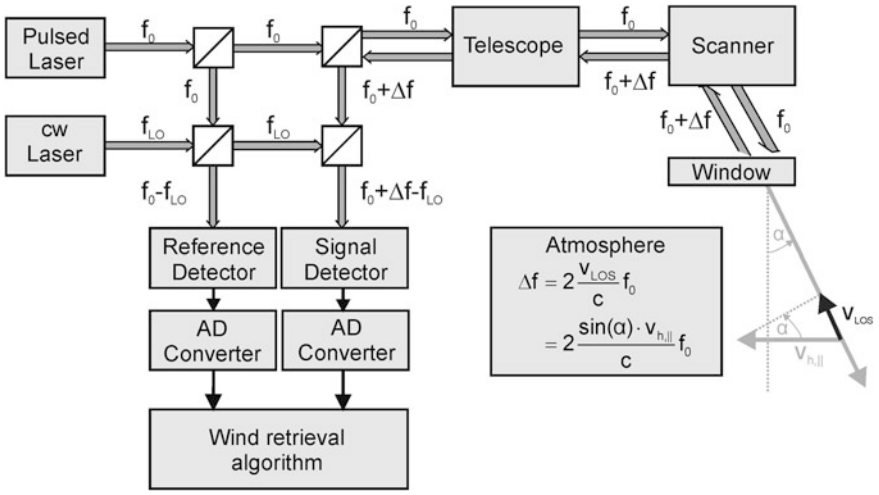

Fig. 30.2 Concept for a coherent, heterodyne DWL with optical mixing of cw laser with f_{LO} and emitted laser pulse with f_0 on the reference detector and Doppler shifted $f_0 + \Delta f$ atmospheric backscatter return on the signal detector; analog–digital (AD) converters provide digital signal; the laser pulses are transmitted through a telescope, scanner, and window to the atmosphere; the scanner is used to deflect the beam in the LOS direction and the v_{LOS} component is determined by the off-nadir angle α and component of the horizontal wind speed $V_{H,II}$ in the plane of the LOS vector

frequency information. The averaging of the signals of several laser pulse returns has to be done in the frequency domain. Several techniques for wind retrieval from coherent Doppler lidar signals are discussed by Smalikho (2003).

Coherent detection is limited to the narrow bandwidth of aerosol backscatter. Thus longer wavelengths in the infrared spectral region are preferred due to the availability of appropriate laser sources for eye-safety reasons and to minimize the broad bandwidth molecular backscatter and solar background. First coherent DWL were developed using gas lasers with CO_2 as active medium, which emit wavelengths of 10.6 μm, and applied to the detection of aircraft wake vortices by Huffaker et al. (1970). DLR activities started in the 1980s with cw-CO_2 lasers (Köpp et al. 1984). Today, solid-state lasers operating at 2.02 μm or fiber lasers with wavelengths of 1.5–1.6 μm are common sources for coherent DWL. The first airborne Doppler lidar measurements were made on board aircraft in the 1980s and were pointed sidewards (Bilbro et al. 1984) or in the flight-direction. Airborne Doppler lidars with conical scanning technique were developed at DLR in the 1990s using cw lasers (Rahm 1995) as true airspeed sensors (Rahm 2001). A pulsed airborne Doppler lidar with a conically scanning technique was developed within the scope of the French-German project WIND (Wind Infrared Doppler Lidar, see Fig. 30.3 and Table 30.2) The WIND instrument was the first airborne Doppler lidar for atmospheric research in 1999 capable of retrieving wind profiles to the ground by using nadir conical scanning (Reitebuch et al. 2001; Werner et al. 2001);. Airborne DWL observations with WIND were used to

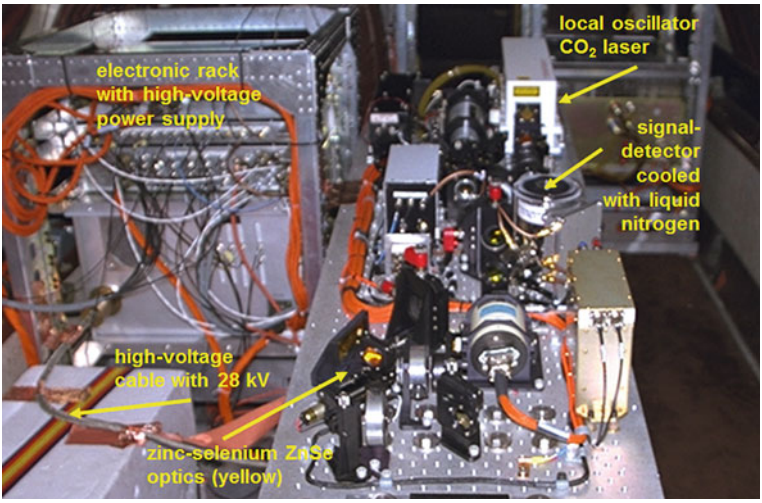

Fig. 30.3 Photo of the optical receiver of the 10.6 μm DWL WIND inside the Falcon aircraft with electronic rack and high-voltage supply for the CO_2 gas laser (*left*) and optical bench (*right*) with local oscillator CO_2 laser, signal detector with liquid nitrogen cooling, and optical elements made of zinc–selenium ZnSe; not visible are the telescope, scanning device and high-power, pulsed CO_2 laser

investigate the flow across the Alps (Reitebuch et al. 2003), the interaction between sea breeze and mistral winds (Drobinski et al. 2005; Bastin et al. 2006) and the *monsoon* flow in West Africa (Bou Karam et al. 2008; Drobinski et al. 2009). Based on the experience with WIND, a more compact airborne DWL was developed using a solid-state 2-μm laser (see Table 30.2). The 2-μm DWL was deployed during field campaigns in Europe for mesoscale studies (Weissmann et al. 2005a; Kiemle et al. 2011), in the North Atlantic for targeted observations (Weissmann et al. 2005b; Weissmann and Cardinali 2007), in Japan for typhoon targeting (Weissmann et al. 2012) and during the eruption of the *Eyjafjallajökull volcano* in 2010 (Schumann et al. 2011).

The detection and characterization of aircraft wake vortices was a major research application at DLR already in the mid-1980s using coherent DWLs with CO_2 lasers at 10.6 μm during several deployments at Frankfurt airport (Köpp 1994). By the end of the 1990s pulsed DWLs with fast scanning capability based on solid-state 2-μm lasers were used (Köpp et al. 2004) and wake vortex detection was demonstrated by DLR on board an aircraft (Rahm et al. 2007; Rahm and Smalikho 2008). Coherent DWLs were used at IPA for deriving turbulent quantities (Banakh et al. 1999; Smalikho et al. 2005), and to estimate turbulent fluxes in combination with trace-gas lidars (Giez et al. 1999; Kiemle et al. 2007; Kiemle et al. 2011) or turbulent flows around wind turbines (Käsler et al. 2010).

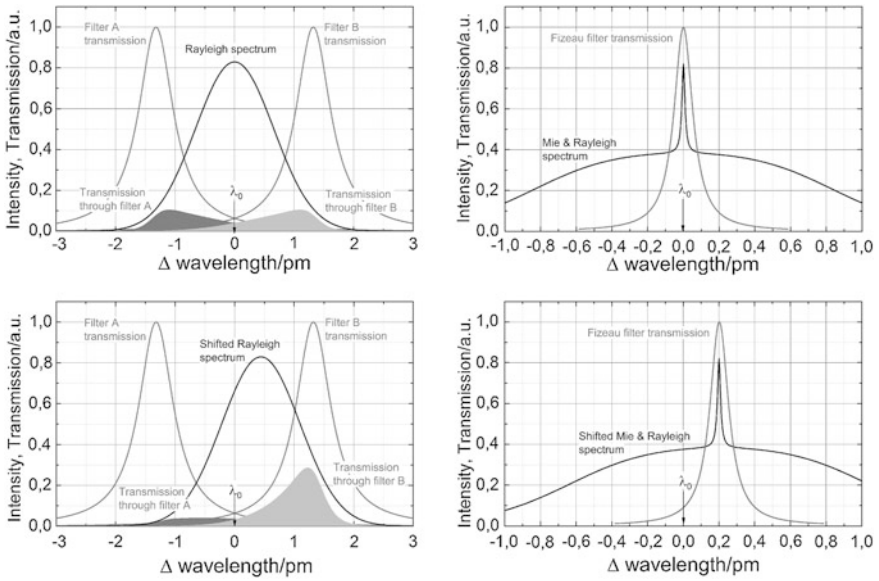

Fig. 30.4 Principle of direct-detection DWL with double-edge technique (*left*) and fringe-imaging technique (*right*) used for the ALADIN airborne demonstrator A2D and the space-borne lidar mission ADM-Aeolus; Rayleigh spectrum from molecular backscatter and double-edge filter A and B transmission (*left*) and additional Mie spectrum from aerosol backscatter and Fizeau filter transmission (*right*) for zero LOS speed (*top*) and Doppler shifted spectrum (*bottom*) for a wavelength of 355 nm; note the different wavelength scales on the left and right figures. The different signal levels transmitted through filter A and B (*left*, *grey areas*) are used to determine the Doppler shift from molecules; the transmitted signal at different spatial locations of the Fizeau filter transmission is used to determine the Doppler shift from aerosols (*right*)

30.3.2 Direct-Detection Doppler Wind Lidar

Compared to coherent DWL, where heterodyne detection by optical mixing is used as the principle, the direct-detection approach uses optical bandpass filters or alternatively an interferometer as spectral analyzer. The direct-detection DWL approach is mainly applied for using the broad bandwidth Rayleigh backscatter return from molecules, although it can also be applied to narrow bandwidth aerosol return. The principle of wind measurement with direct-detection Doppler lidar was first described and realized by Benedetti-Michelangeli et al. (1972) with a frequency-stabilized Ar-ion laser emitting at 488 nm and a spherical *Fabry–Perot interferometer* as spectral analyzer.

30.3.2.1 Filter Technique

The filter method is described using a common approach called the double-edge technique, which is also to be used for the future European space-borne DWL on the *Atmospheric Dynamics Mission* ADM-Aeolus. The double-edge technique uses

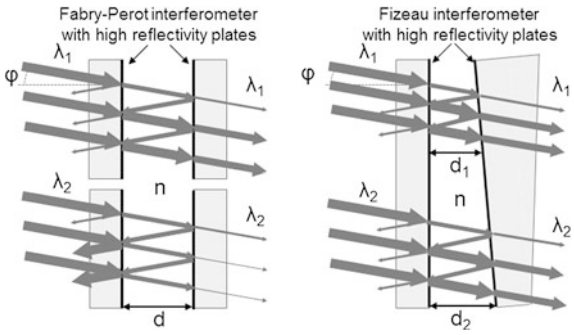

Fig. 30.5 Principle of a plane Fabry–Perot interferometer with distance d (*left*) and a Fizeau interferometer with distance d_1 (*upper part*) and d_2 (*lower part*) of the etalon (*right*) with two highly reflecting plates forming the etalon (refractive index n) and optical beams (*arrows*) for two different wavelengths λ_1 and λ_2; the thickness of the beams indicates their intensity; λ_1 fulfills the criteria for constructive interference for the Fabry–Perot interferometer ($\lambda_1 = 2{\cdot}d{\cdot}n{\cdot}\cos(\varphi)/k$) and for the upper part of the Fizeau interferometer ($\lambda_1 = 2{\cdot}d_1{\cdot}n{\cdot}\cos(\varphi)/k$) and is thus transmitted; λ_2 does not fulfill the interference criteria for the Fabry–Perot interferometer and is thus reflected, while it is transmitted in the lower part of the Fizeau interferometer ($\lambda_2 = 2{\cdot}d_2{\cdot}n{\cdot}\cos(\varphi)/k$)

two bandpass filters which are placed symmetrically around the maximum of the molecular spectral line with their maximum filter transmission at the edges of the line around the inflexion point (Fig. 30.4, left). The transmitted signal through each of these filters is detected, which is proportional to the convolution of the spectral filter transmission and the backscattered line shape from the atmosphere. For the case that both filters A and B have equal transmission characteristic ($T_A(\lambda) = T_B(\lambda + \Delta\lambda_{A,B})$) and are symmetrically placed around the Rayleigh spectral line with a separation of $\Delta\lambda_{A,B}$, then both filters would measure the same signal $I_A = I_B$, for a zero LOS wind speed. In the case of a Doppler shift $\Delta\lambda$, the transmitted signal through one filter would be higher than through the second one (Fig. 30.4, left, bottom). The ratio or contrast of I_A and I_B is used as a measure for the Doppler shift and is called the response R:

$$R = \frac{I_A}{I_B} \text{ or } R = \frac{I_A - I_B}{I_A + I_B}$$
$$R = f(T_A(\lambda), T_B(\lambda), I_{RBS}(\lambda, \Delta\lambda, T, p))$$

The Doppler shift can be derived from the response if the correspondence between both is determined through a calibration procedure. This is necessary because the response R is not only dependent on the Doppler shift $\Delta\lambda$, but also on the actual Rayleigh-Brillouin line shape of the backscattered signal $I_{RBS}(\lambda,\Delta\lambda,T,p)$, and thus depends on atmospheric temperature T and pressure p, as well as on the filter transmissions $T_A(\lambda)$ and $T_B(\lambda)$. Exact knowledge of the actual filter transmissions of the instrument is more critical for wind retrieval than is

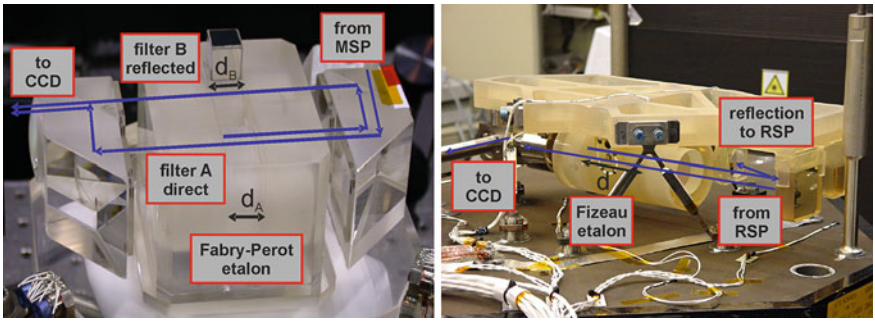

Fig. 30.6 Photo of the optical receiver of the ALADIN airborne demonstrator A2D with Rayleigh spectrometer (RSP, *left*) and its optical paths for filter A and B, and Mie spectrometer (MSP, *left*); optical beam paths towards the charge-coupled device (CCD) detector are indicated with *arrows*, the distances d_A and d_B of the Fabry–Perot etalons (*left*) and d of the Fizeau etalon (*right*) are indicated by double *arrows*

knowledge of sensitivity to the atmospheric temperature and pressure, which can be applied as a first-order correction to the retrieved wind speed.

Certainly the filter A and B bandwidths have to be smaller than the bandwidth of the molecular return signal (FWHM = 1.6 pm at 15 °C for a wavelength of 355 nm), which can be realized by a Fabry–Perot interferometer. As sketched in Fig. 30.5, a plane Fabry–Perot interferometer consists of two plane parallel optical plates (called the "etalon") with high reflectivity and separated by distance d (Vaughan et al. 1996). Multiple-beam interference of the incoming light leads to positive interference patterns (called fringes"), if the wavelength λ is a multiple k of the optical path $\lambda = 2 \cdot d \cdot n \cdot \cos(\varphi)/k$ (k: integer, n: refractive index, φ: incidence angle on plates).

When using a Fabry–Perot interferometer as a filter, then only the central ring ($\varphi = 0$) and not the complete, radial interference pattern is used. The two filters can be realized with the same Fabry–Perot interferometer, but with slightly different distances d_A and d_B at different parts of the Fabry–Perot interferometer (Fig. 30.6, left).

The double-edge technique was pioneered by Chanin et al. (1989) for stratospheric winds using the broadband molecular return with a laser wavelength of 532 nm, and for tropospheric winds using laser wavelengths of 1 064 nm and 532 nm for the narrow band aerosol return and 355 nm for the molecular return by Korb et al. (1992) and Gentry et al. (2000). The first airborne direct-detection Doppler lidar was developed by IPA and EADS-Astrium (Reitebuch et al. (2009), see Table 30.2 for key parameters, and Fig. 30.6) as the airborne prototype A2D (ALADIN airborne demonstrator) for the future space-borne instrument ALADIN (Atmospheric Laser Doppler Instrument) on the ADM-Aeolus mission of the European Space Agency ESA.

Another possibility to realize the filter technique is to use specific, very narrow-band molecular or atomic absorption lines. With absorption lines, only one filter

can be realized ("single edge technique"), which relies on the change in absolute intensity through this filter. The technique is limited to spectral regions where narrow bandwidth absorption lines are occurring at the emission wavelength of lasers. This is realized in the green spectral region at 532 nm and with iodine I_2 absorption lines. The use of the iodine molecular absorption lines as narrow-band filters for both molecular and aerosol backscatter at 532 nm was recently applied to measure sea-surface winds during the 2008 Olympic games in Beijing, China (Liu et al. 2008) and is also common for *high spectral resolution lidars* (HSRL, Esselborn et al. 2008).

30.3.2.2 Fringe Imaging Technique

Another implementation of the direct-detection DWL uses the spatial location of interference patterns ("fringes") and its displacement when a Doppler shift is present. This can be realized by a Fabry–Perot interferometer, where the movement of the radial-angular pattern of interference rings is used to determine the Doppler shift (Hirschberger and Ehret 2011). Another approach is based on a Fizeau interferometer, as for the satellite mission ADM-Aeolus and its airborne prototype A2D for determining the Doppler shift from the narrow bandwidth aerosol return (Fig. 30.6, right). A *Fizeau interferometer* also consists of an etalon, but opposed to a Fabry–Perot interferometer, where both etalon plates are plane-parallel with constant distance d, one of the plates is tilted by a small angle of several μrad. The resulting variation in distance d results in a distinct spatial location of an interference pattern where the condition for positive interference is fulfilled (Fig. 30.5, right). A Doppler shift of the signal will result in spatial displacement of the fringe, which is imaged onto a detector (Fig. 30.4, right). Therefore, a detector with several sensitive areas (e.g., a charge-coupled device CCD, as nowadays used in digital cameras) to provide pixel-wise discrete, spectral channels is needed for the fringe imaging technique.

30.4 Selected Results from Recent Airborne Wind Lidar Observations for Atmospheric Research

Two examples of airborne DWL observations during recent years are briefly discussed here to illustrate their potential for atmospheric research.

30.4.1 Monsoon Observations in West Africa

The monsoon is a seasonal atmospheric circulation induced by asymmetric heating of land and sea, which results in greater annual variation of temperature over large land areas compared with neighboring ocean surfaces. The monsoon onset is of great importance for Western Africa as it is linked to the start of the rain period. Here, the monsoon is a southerly, cool and humid flow and it was resolved on a large horizontal scale and vertically by airborne DWL observations during the

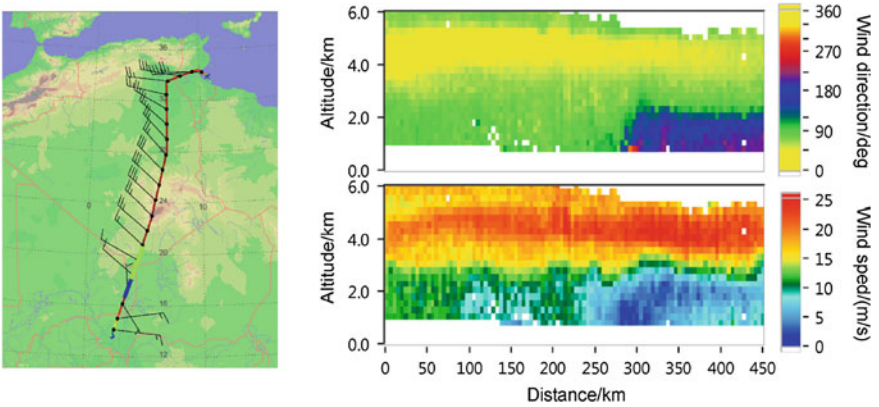

Fig. 30.7 Horizontal wind vector observations up to 6 km altitude above sea level with wind direction (*top*) and wind speed (*bottom*) from the Falcon aircraft from 20.6°N/4.7°E (0 km) to 16.8°N/3.3°E (450 km) on a flight from Djerba (Tunisia) to Niamey (Niger) on 30 June 2006 during the African Monsoon Multidisciplinary Analysis (AMMA) field campaign in West Africa; the Falcon flight track and the horizontal wind vector in flight altitude are shown (*left*) and the location of the observations are indicated on the flight track; the wind direction shows the southerly monsoon flow up to 2 km altitude; the African easterly jet in 4–5 km altitude is visible with higher wind speeds up to 25 m/s; *white* colors indicated non-valid winds after quality control and terrain for low altitudes

African Monsoon Multidisciplinary Analysis (AMMA) field experiment in 2006. The Falcon aircraft flew from Djerba (Tunisia) to Niamey (Niger) on 30 June 2006, and northwesterly winds were observed at a flight altitude of 11 km by the insitu wind sensor (Fig. 30.7, left). The southerly monsoon flow up to 2 km is clearly visible in the wind direction observations from the DWL up to the so-called intertropical discontinuity, where it interacts with the northeasterly, dry wind, the so-called Harmattan (Fig. 30.7, right, top). Above the monsoon flow the dry African easterly jet (a subtropical jet with an easterly wind direction in the troposphere over Africa) is present as a layer of higher wind speeds up to 25 m/s around 4–5 km altitude with a northeasterly direction on this day (Fig. 30.7, right, bottom). Airborne DWL observations were performed for the first time in this region and used for numerous case studies of dynamic processes in combination with mesoscale models, e.g., Bou Karam et al. (2008) and Drobinski et al. (2009).

30.4.2 Targeted Observations in the North Atlantic with 2-μm Wind Lidar

Airborne DWL observations for the purpose of improving weather forecasts were performed over the North Atlantic Ocean in November 2003. The Falcon aircraft with the 2-μm DWL and a dropsonde unit was directed towards so-called sensitive

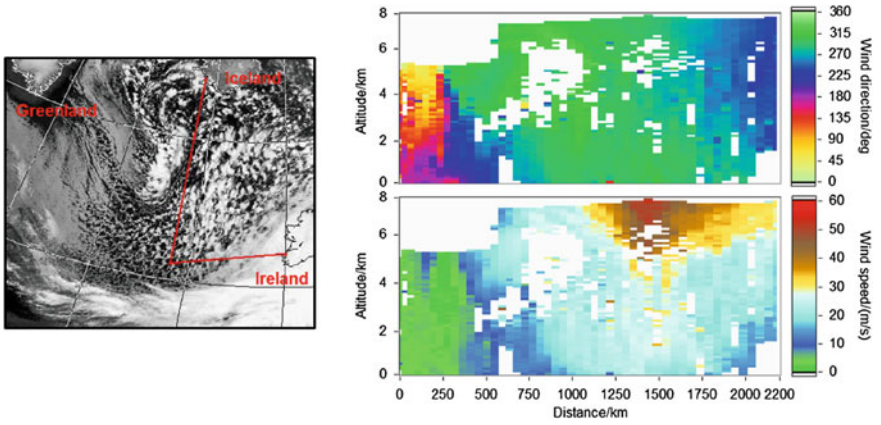

Fig. 30.8 2-μm DWL observations from a Falcon flight from Iceland to Ireland (*left*, superimposed on a visible satellite image) on 25 November 2003 during the Atlantic THORPEX Regional Campaign A-TReC (figure adapted from Weissmann et al. (2005b)); the observations are averaged over 40 km for horizontal wind direction (*top*) and wind speed (*bottom*); a high-altitude jet stream with wind speeds up to 67 m/s was targeted; location of the turning curve from north-south to west-east in the flight track is at a distance of 1 400 km; *white* colors indicate non-valid winds after quality control and different flight altitudes before and after 600 km distance

regions" where observations should be most beneficial for medium-range weather forecasts. These targeted lidar observations were assimilated into the numerical model of the *European Centre for Medium-Range Weather Forecasts* (ECMWF) and had a significant positive impact on the analysis and forecast results (Weissmann and Cardinali 2007). The wind observations from the lidar were averaged along the flight track to about 40 km in order to match the horizontal resolution of the lidar to the numerical model, which allows assigning a low representativeness error. An example is shown for the flight on 25 November 2003 from Iceland to Ireland (Fig. 30.8). Although the atmosphere is rather cloudy, the horizontal wind speed and direction could be derived without large gaps in coverage. The wind speed increased while flying from north to south up to 67 m/s while approaching a high-altitude jet stream. The instrumental error of the lidar observations could be quantified by a comparison with collocated wind measurements from dropsondes and resulted in a random error of 0.75–1.0 m/s (Weissmann et al. 2005b). The results from the impact studies with the assimilated DWL observations underlined the importance of additional wind observations, especially over the oceans, and demonstrated the potential of DWL. For the first time real DWL observations were assimilated into a global model and the findings support the high expectations for the future satellite DWL on ADM-Aeolus.

30.5 Outlook

Today, coherent DWLs with near-infrared wavelength (typically between 1.5–1.6 μm and 2 μm) are commercially available and routinely applied for assessing the wind energy potential and at some airports for wind profiling and wind shear detection. Research and development activities are focused towards further miniaturization for deployment of DWLs in wind turbines for gust detection or on airborne sensors for true-air speed, wake-vortex or *clear-air turbulence* detection. A space-borne direct-detection wind lidar will be deployed on the ADM-Aeolus mission in the near future as a technical demonstrator for future, operational wind lidars on *polar orbiting satellites* for global wind profiling.

References

Banakh, V.A., Smalikho, I.N., Köpp, F., Werner, C.: Measurements of turbulent energy dissipation rate with a CW Doppler lidar in the atmospheric boundary layer. J. Atmos. Oceanic Technol. **16**, 1044–1061 (1999). doi:10.1175/1520-0426(1999)016<1044:MOTEDR>2.0.CO;2

Bastin, S., Drobinski, P., Guenard, V., Caccia, J.-L., Campistron, B., Dabas, A., Delville, P., Reitebuch, O., Werner, C.: On the interaction between sea breeze and summer Mistral at the exit of the Rhone valley. Mon. Weather Rev. **134**, 1647–1668 (2006). doi:10.1175/MWR3116.1

Benedetti-Michelangeli, G., Congeduti, F., Fiocco, G.: Measurement of aerosol motion and wind velocity in the lower troposhere by Doppler optical radar. J. Atmos. Sci. **29**, 906–910 (1972). doi:10.1175/1520-0469(1972)029<0906:MOAMAW>2.0.CO;2

Bilbro, J., Fichtl, G., Fitzjarrald, D., Krause, M., Lee, R.: Airborne Doppler Lidar Wind Field Measurements. Bull. Am. Meteorol. Soc. **65**, 348–359 (1984). doi:10.1175/1520-0477(1984)065<0348:ADLWFM>2.0.CO;2

Bou Karam, D., Flamant, C., Knippertz, P., Reitebuch, O., Pelon, J., Chong, M., Dabas, A.: Dust emissions over Sahel associated with the West African monsoon inter-tropical discontinuity region: a representative case study. Quart. J. Roy. Meteorol. Soc. **134**, 621–634 (2008). doi:10.1002/QJ.244

Chanin, M.L., Garnier, A., Hauchecorne, A., Porteneuve, J.: A Doppler lidar for measuring winds in the middle atmosphere. Geophys. Res. Lett. **16**, 1273–1276 (1989). doi:10.1029/GL016i011p01273

Drobinski, P., Bastin, S., Guenard, V., Caccia, J.-L., Dabas, A.M., Delville, P., Protat, A., Reitebuch, O., Werner, C.: Summer Mistral at the exit of the Rhône valley. Quart. J. Roy. Meteorol. Soc. **131**, 353–375 (2005). doi:10.1256/qj.04.63

Drobinski, P., Bastin, S., Janicot, S., Bock, O., Dabas, A., Delville, P., Reitebuch, O., Sultan, B.: On the late northward propagation of the West African monsoon in summer 2006 in the region of Niger/Mali. J. Geophys. Res. **114**, D09108 (2009). doi:10.1029/2008JD011159

Esselborn, M., Wirth, M., Fix, A., Tesche, M., Ehret, G.: Airborne high spectral resolution lidar for measuring aerosol extinction and backscatter coefficients. Appl. Optics **47**, 346–358 (2008). doi:10.1364/AO.47.000346

Gentry, B.M., Chen, H., Li, S.X.: Wind measurements with 355-nm molecular Doppler lidar. Opt. Lett. **25**, 1231–1233 (2000). doi:10.1364/OL.25.001231

Giez, A., Ehret, G., Schwiesow, R., Davis, K.J., Lenschow, D.H.: Water Vapor Flux Measurements from ground-based vertically-pointed water vapor differential absorption and Doppler lidars. J. Atmos. Oceanic Technol. **16**, 237–250 (1999). doi:10.1175/1520-0426(1999)016<0237:WVFMFG>2.0.CO;2

Henderson, S.W., Gatt, P., Rees, D., Huffaker, R.M.: Wind lidar. In: Fujii, T., Fukuchi, T. (eds.) Laser Remote Sensing, pp. 469–722. Taylor & Francis Group, London (2005)

Hirschberger, M., Ehret, G.: Simulation and high-precision wavelength determination of noisy 2D Fabry–Pérot interferometric rings for direct-detection Doppler lidar and laser spectroscopy. Appl. Phys. B **103**, 207–222 (2011). doi:10.1007/s00340-011-4391-9

Huffaker, R.M., Jelalian, A.V., Thomson, J.A.: Laser-Doppler system for detection of aircraft trailing vortices. Proc. IEEE **58**, 322–326 (1970). doi:10.1109/PROC.1970.7636

Käsler, Y., Rahm, S., Simmet, R., Kühn, M.: Wake measurements of a multi-MW wind turbine with coherent long-range pulsed doppler wind lidar. J. Atmos. Oceanic Technol. **27**, 1529–1532 (2010). doi:10.1175/2010JTECHA1483.1

Kiemle, C., Brewer, W.A., Ehret, G., Hardesty, R.M., Fix, A., Senff, C., Wirth, M., Poberaj, G., LeMone, M.A.: Latent heat flux profiles from collocated airborne water vapor and wind lidars during IHOP_2002. J. Atmos. Oceanic Technol. **24**, 627–639 (2007). http://dx.doi.org/10.1175/JTECH1997.1

Kiemle, C., Wirth, M., Fix, A., Rahm, S., Corsmeier, U., Di Girolamo, P.: Latent heat flux measurements over complex terrain by airborne water vapour and wind lidars. Quart. J. R. Meteorol. Soc. **137**, 190–203 (2011). doi:10.1002/qj.757

Köpp, F., Schwiesow, R.L., Werner, C.: Remote measurements of boundary-layer wind profiles using a CW Doppler lidar. J. Clim. Appl. Meteorol. **23**, 148–154 (1984). doi:10.1175/1520-0450(1984)023<0148:RMOBLW>2.0.CO;2

Köpp, F.: Doppler lidar investigation of wake vortex transport between closely spaced parallel runways. AIAA J. **32**(4), 805–810 (1994). doi:10.2514/3.12057

Köpp, F., Rahm, S., Smalikho, I.: Characterization of aircraft wake vortices by 2-μm pulsed Doppler Lidar. J. Atmos. Oceanic Technol. **21**(2), 194–206 (2004). doi:10.1175/1520-0426(2004)021<0194:COAWVB>2.0.CO;2

Korb, C.L., Gentry, B.M., Weng, C.: The edge technique: theory and application to the lidar measurement of atmospheric winds. Appl. Optics **31**, 4202–4213 (1992). doi:10.1364/AO.31.004202

Liu, Z., Liu, B., Wu, S., Li, Z., Wang, Z.: High spatial and temporal resolution mobile incoherent Doppler lidar for sea surface wind measurements. Opt. Lett. **33**, 1485–1487 (2008). doi:10.1364/OL.33.001485

Rahm, S.: Measurement of a wind field with an airborne continuous-wave Doppler lidar. Opt. Lett. **20**, 216–218 (1995). doi:10.1364/OL.20.000216

Rahm, S.: Precursor experiment for an active true airspeed sensor. Opt. Lett. **26**, 319–321 (2001). doi:10.1364/OL.26.000319

Rahm, S., Smalikho, I., Köpp, F.: Characterization of aircraft wake vortices by airborne coherent Doppler lidar. J. Aircr. **44**, 799–805 (2007). doi:10.2514/1.24401

Rahm, S., Smalikho, I.: Aircraft wake vortex measurement with airborne coherent Doppler lidar. J. Aircr. **45**, 1148–1155 (2008). doi:10.2514/1.32896

Reitebuch, O., Werner, C., Leike, I., Delville, P., Flamant, P.H., Cress, A., Engelbart, D.: Experimental validation of wind profiling performed by the airborne 10 μm-heterodyne Doppler lidar WIND. J. Atmos. Oceanic Technol. **18**, 1331–1344 (2001). doi:10.1175/1520-0426(2001)018<1331:EVOWPP>2.0.CO;2

Reitebuch, O., Volkert, H., Werner, C., Dabas, A., Delville, P., Drobinski, P., Flamant, P.H., Richard, E.: Determination of air flow across the Alpine ridge by a combination of airborne Doppler lidar, routine radio-sounding and numerical simulation. Quart. J. Roy. Meteorol. Soc. **129**, 715–728 (2003). doi:10.1256/qj.02.42

Reitebuch, O., Lemmerz, C., Nagel, E., Paffrath, U., Durand, Y., Endemann, M., Fabre, F., Chaloupy, M.: The airborne demonstrator for the direct-detection Doppler wind lidar ALADIN on ADM-Aeolus. Part I: instrument design and comparison to satellite instrument. J. Atmos. Oceanic Technol. **26**, 2501–2515 (2009). doi:10.1175/2009JTECHA1309.1

Schumann, U., Weinzierl, B., Reitebuch, O., Schlager, H., Minikin, A., Forster, C., Baumann, R., Sailer, T., Graf, K., Mannstein, H., et al.: Airborne observations of the Eyjafjalla volcano ash

cloud over Europe during air space closure in April and May 2010. Atmos. Chem. Phys. **11**, 2245–2279 (2011). doi:10.5194/acp-11-2245-2011

Smalikho, I.: Techniques of wind vector estimation from data measured with a scanning coherent Doppler lidar. J. Atmos. Oceanic Technol. **20**, 276–291 (2003). doi:10.1175/1520-0426(2003)020<0276:TOWVEF>2.0.CO;2

Smalikho, I., Köpp, F., Rahm, S.: Measurement of atmospheric turbulence by 2-μm Doppler lidar. J. Atmos. Oceanic Technol. **22**(11), 1733–1747 (2005). doi:10.1175/JTECH1815.1

Vaughan, J.M., Steinvall, O., Werner, C., Flamant, P.H.: Coherent laser radar in Europe. Proc. IEEE **84**, 205–226 (1996). doi:10.1109/5.482229

Weissmann, M., Braun, F., Gantner, L., Mayr, G., Rahm, S., Reitebuch, O.: The Alpine mountain-plain circulation: airborne Doppler lidar measurements and numerical simulations. Mon. Weather Rev. **133**, 3095–3109 (2005a). http://dx.doi.org/10.1175/MWR3012.1

Weissmann, M., Busen, R., Dörnbrack, A., Rahm, S., Reitebuch, O.: Targeted observations with an airborne wind lidar. J. Atmos. Oceanic Technol. **22**, 1706–1719 (2005b). http://dx.doi.org/10.1175/JTECH1801.1

Weissmann, M., Cardinali, C.: Impact of airborne Doppler lidar observations on ECMWF forecasts. Quart. J. R. Meteorol. Soc. **133**, 107–116 (2007). doi:10.1002/qj.16

Weissmann, M., Langland, R.H., Cardinali, C., Rahm, S.: Influence of airborne Doppler wind lidar profiles near typhoon Sinlaku on ECMWF and NOGAPS forecasts. Quart. J. R. Meteorol. Soc. **138**, 118–130 (2012). doi:10.1002/qj.896

Werner, C., Flamant, P.H., Reitebuch, O., Köpp, F., Streicher, J., Rahm, S., Nagel, E., Klier, M., Herrmann, H., Loth, C., et al.: Wind infrared Doppler lidar instrument. Opt. Eng. **40**, 115–125 (2001). doi:10.1117/1.1335530

Werner, C.: Doppler wind lidar. In: Weitkamp, C. (ed.) Lidar—Range-Resolved Optical Remote Sensing of the Atmosphere, pp. 325–354. Springer, New York (2004)

Witschas, B., Vieitez, M.O., van Duijn, E.-J., Reitebuch, O., van de Water, W., Ubachs, W.: Spontaneous Rayleigh-Brillouin scattering of ultraviolet light in nitrogen, dry air, and moist air. Appl. Optics **49**, 4217–4227 (2010). doi:10.1364/AO.49.004217

Tunable Light Sources for Lidar Applications

31

Andreas Fix

Abstract

Differential absorption lidar measurements require coherent radiation with high peak and average power in the nanosecond time regime at wavelengths coincident with specific absorption features of the atmospheric trace species to be measured. Often these wavelengths do not coincide with readily available laser lines. Optical parametric oscillators and amplifiers can efficiently convert light from fixed frequency sources into broadly tunable laser radiation providing a generic approach to developing versatile lidar transmitters. This technology has been advanced at DLR to a stage of maturity suitable for airborne and spaceborne applications.

31.1 Introduction

If a *lidar* researcher came across a magical wishing well and threw a penny into it to have a wish come true, she or he might possibly ask for an ideal *laser* source for use in *differential absorption lidar* (DIAL) atmospheric *trace gas* investigations. Unfortunately, such a magical lantern does not exist in reality. It would have to be tunable over a significant portion of the electromagnetic spectrum with narrow bandwidth and high *spectral purity* (with negligible broadband spectral components) to ideally match the absorption features of the different trace gases. For pulsed direct detection lidar which is the focus here, it must provide both high peak

A. Fix (✉)
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_31,

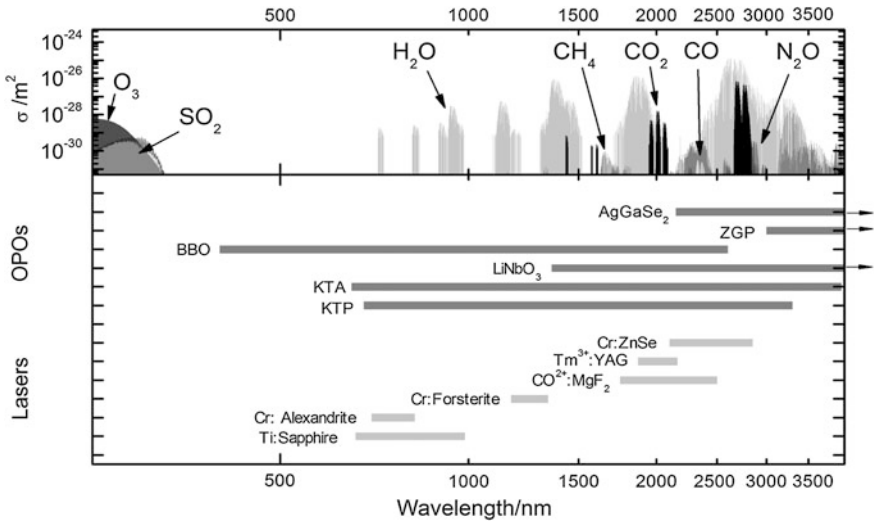

Fig. 31.1 Tuning ranges of selected OPO materials in comparison to those of tunable solid-state *lasers* in the wavelength range from the UV (250 nm) to the infrared (4 000 nm). The upper panel gives the *absorption cross section* of the six most prevalent trace gases plus *sulfur dioxide* (multiplied by their typical abundance at the Earth's surface)

and average power. Its pulses have to be in the nanosecond time regime (corresponding to a geometrical length of the pulse of a few meters) to provide the desired spatial resolution. For operation on *aircraft* or spacecraft, this *coherent* light source should provide a high electrical-to-optical *efficiency*, be maintenance-free and rugged, and have a small volume.

No single laser source fulfills those requirements. Most lasers emit at fixed wavelengths which, only in rare cases, coincide with the relevant absorption lines of atmospheric species. Established laser materials such as Ti:Sapphire or Cr:Alexandrite that provide some tunability are scarce. Others perform poorly or lack heritage.

There is, however, a solution that comes quite close to this wishful thinking: *Nonlinear frequency conversion* using *optical parametric oscillator* (OPOs) and *optical parametric amplifiers* (OPAs) is an effective means to convert the radiation of fixed-frequency ("pump") lasers into broadly tunable, coherent radiation at wavelengths where conventional lasers perform poorly or are unavailable. Although these parametric devices generate coherent radiation like lasers, their principle of operation is based on optically driven *polarization* and thus differs from lasers since no energy is stored in a medium and no stimulated emission takes place. As a necessary condition, OPOs require coherent pump radiation which is converted in a three-wave mixing process. The tuning range of OPOs is basically determined by the transparency of the optical *crystals*, which facilitates this nonlinear optical (NLO) process. This range can be much broader than for tunable lasers (see Fig. 31.1). The nonlinear optical devices open up a generic approach from the *ultraviolet* (UV) to the infrared and beyond for a variety of applications.

The potential of OPOs was rapidly recognized after its first realization in 1965 (Giordmaine and Miller 1965) and few years later for lidar applications (Baumgartner and Byer 1978; Endemann and Byer 1981; Brassington 1982). Still, it was a long way from a mere subject of scientific investigation to useful devices for applications such as lidar. In the early days, OPOs were notorious for their unreliability and insufficient spatial and spectral properties. In the last few decades significant advances in the development of novel, high-quality nonlinear crystals with low absorption and high optical damage threshold, high peak power (all-) solid-state pump lasers, and optical *coatings* have transformed OPOs into attractive sources of broadly tunable coherent radiation. They are applicable not only for lidar but also for spectroscopy, biomedicine, environmental monitoring, imaging and chemical process control.

It should be noted that optical parametric devices can be operated in various temporal regimes from femtosecond pulses to continuous wave. For the current state-of-art the reader is referred to the literature (Orr et al. 2009; Ebrahim-Zadeh and Dunn 2004). The purpose of this contribution is to introduce the basic principles of operation, describe recent developments of pulsed OPOs in the nanosecond time regime, and describe achievements at DLR in applying OPOs on board aircraft or spacecraft for the detection of *water vapor*, *ozone*, *methane* and *carbon dioxide*.

31.2 Basic Principles of Pulsed Optical Parametric Oscillators and Amplifiers

The core of an OPO or OPA is a nonlinear optical crystal. Under the irradiation of electromagnetic fields dielectric crystalline media become polarized. The polarization can be expanded into a power series of the electric field. Once the fields become strong enough, the second order term cannot be neglected. This second order nonlinearity facilitates the interaction between three waves with different frequencies and is responsible for nonlinear optical effects such as sum or difference frequency generation but also for optical parametric amplification and oscillation. The strength of the interaction is characterized by the nonlinear susceptibility tensor $\chi^{(2)}$ or, more conveniently, by the nonlinear coefficient d_{eff}, a scalar that is derived from $\chi^{(2)}$ considering the orientation of the crystal axes with respect to the fields and their planes of polarization.

In classical theory, the interaction of the three involved electrical fields E_p, E_s, and E_i, with the angular frequencies $\omega_{p,s,i}$ fulfilling the condition $\omega_p = \omega_s + \omega_i$, along the direction of propagation z are given by the coupled amplitude equations that are derived from *Maxwell's equations* for nonmagnetic media:

$$\frac{\partial E_s}{\partial z} + \alpha_s E_s = i\frac{\omega_s d_{eff}}{n_s c} E_p E_i^* \cdot e^{i\Delta kz}$$

$$\frac{\partial E_i}{\partial z} + \alpha_i E_i = i\frac{\omega_i d_{eff}}{n_i c} E_p E_s^* \cdot e^{i\Delta kz}$$

$$\frac{\partial E_p}{\partial z} + \alpha_p E_p = i\frac{\omega_p d_{eff}}{n_p c} E_s E_i \cdot e^{-i\Delta kz}$$

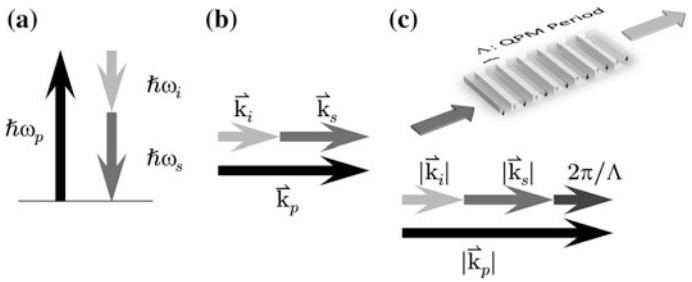

Fig. 31.2 **a** Energy conservation of optical parametric frequency conversion in the photon picture. A pump photon (with angular frequency ω_p) is split into a less energetic signal and idler photons whose photon energies sum up to the pump photon energy. **b** Phase matching of optical parametric frequency conversion. The sum of the wave vectors of signal and idler equals that of the pump. **c** The artificial grating structure in the patterned crystal compensates for the wave vector mismatch

In these equations, $n_{s,i,p}$ are the refractive indices. Losses in the media are considered by the absorption coefficients $\alpha_{s,i,p}$. $\Delta k = k_p - k_s - k_i$ is the difference of the wave vectors, also termed phase mismatch.

Efficient energy transfer in nonlinear frequency conversion processes requires a constant phase relation between the interacting waves during their propagation through the nonlinear media (i.e., $\Delta k = 0$). In the special case of parametric frequency conversion, the amplitudes E_s and E_i, which are termed "signal" and "*idler*" (where conventionally the wave with higher frequency is called the signal), are small compared to the pump field E_p. In this case, the signal and idler fields receive gain at the expense of the pump and grow from a low level. Indeed, the fields emerging from vacuum fluctuations are adequate to start this process.

It is convenient to look at this process from a *photon* point of view (see Fig. 31.2). In parametric frequency conversion, an energetic pump photon is split under adherence to energy conservation into two photons with lower energy or frequency, thereby producing gain at the signal's and idler's frequencies.[1] Which pair of signal/idler photons is generated is governed by another important constraint which is equivalent to the momentum conservation: the phase-matching condition which requires that the sum of the wave vectors of signal and idler equals that of the pump.

In the transparency region of isotropic media, this condition, however, cannot be fulfilled. But in birefringent NLO crystals, the phase-matching (PM) condition $\Delta k \approx 0$ can be achieved by adjusting its ordinary (o−) and extraordinary (e−) ray refractive indices via angle and temperature. Thus, changing the angle or

[1] In a similar way, two photons with the same frequency can add up to yield their double frequency (or half the wavelength). This is termed *second harmonic generation* . *Third harmonic generation* is achieved when second harmonic and fundamental are mixed. The resulting wavelength then equals a third of the original wavelength. For further reading on laser physics and nonlinear optics the reader is referred to standard textbooks; e.g. Saleh and Teich 2007.

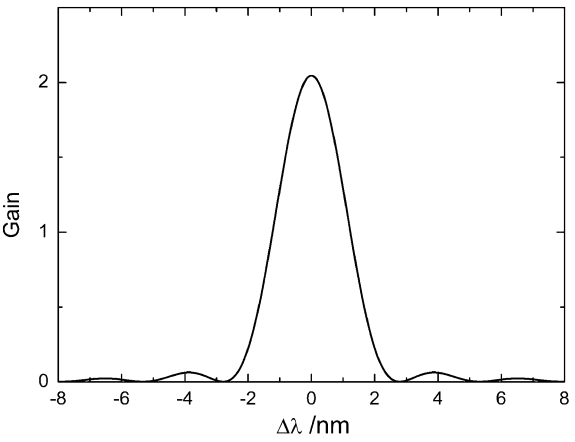

Fig. 31.3 Single-pass signal gain of a typical OPO configuration as function of the wavelength deviation from optimum phase matching. For this calculation, a KTA crystal was chosen at a signal wavelength of 1572 nm, pumped with 100 mJ of a Nd:YAG laser with a flat beam profile of 5 mm diameter and a pulse length of 8 ns

temperature of a suitable crystal enables the selection of the signal/idler pair and is the basis of the broad tunability of OPOs and OPAs. In particular, angular phase matching enables the broadest tuning ranges. As an alternative to birefringent phase matching, which has recently gained significant importance, the PM condition can also be satisfied by so-called quasi-phase matching (QPM). Here, an artificial grating structure in the nonlinear crystal is used to periodically reset the accumulated phase error between the driving and the generated waves (Armstrong et al. 1962). Probably the most significant advantage of QPM over conventional birefringent phase matching is that already established nonlinear crystals can be used at wavelengths outside their normal phase matching region.

Depending on the properties of the crystal, phase matching can be applied in several variations. In Type-I phase matching both signal and idler have the same polarization perpendicular to that of the pump. Conversely, in Type-II phase matching, signal and idler have different polarization directions.

Under the assumption that the idler energy is zero at the input facet of the NLO crystal, the single pass small signal gain G can be calculated from the coupled amplitude equations for a crystal of length l and pump intensity I_p:

$$G(\Delta k) = \frac{|E_s(l)|^2}{|E_s(0)|^2} - 1 = \frac{\Gamma^2 l^2 \cdot \sinh^2 gl}{(gl)^2} \text{ with } \Gamma^2 = \frac{2\omega_s \omega_i \left|d_{eff}\right|^2 I_p}{n_p n_s n_i \varepsilon_0 c^3} \text{ and}$$

$$g = \sqrt{\Gamma^2 - \left(\frac{\Delta k}{2}\right)^2}$$

A result for the parametric gain G of a typical nanosecond OPO configuration is shown in Fig. 31.3. At optimum phase matching the gain is ~ 2. This is by far not enough to efficiently convert pump into signal and idler radiation starting from noise levels. Therefore, it is necessary to build an optical resonator around the crystal to provide optical feedback. Once, the pump energy is increased above a

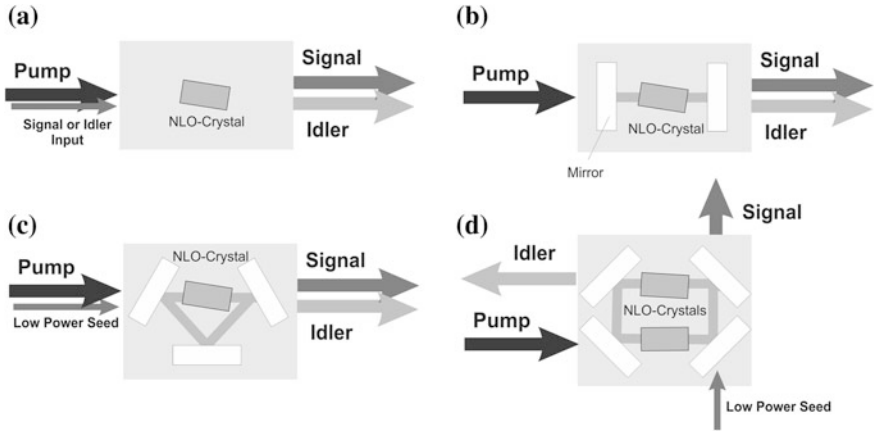

Fig. 31.4 Basic designs of parametric frequency converters: **a** optical parametric amplifier **b** an OPO needs a resonator to provide feedback to the OPA process. **c** For bandwidth control the OPO is injection seeded by low power seed radiation at either signal or idler wavelength. In order to prevent back reflection into the seed laser, ring cavities, for example consisting of three (**c**) or four (**d**) resonator mirrors, are advised

certain threshold where the gain compensates for the losses of the resonator, the systems starts to oscillate. In this case one speaks of an OPO in contrast to an OPA, which lacks a resonator (Fig. 31.4).

As can be seen from Fig. 31.3, the typical gain bandwidth of OPOs is of the order of nanometers. Thus, the resulting signal and idler bandwidth are of the same magnitude. This is often much too large for applications such as DIAL, where bandwidths close to the time-bandwidth limit are required. It is possible to reduce the bandwidth of OPOs by introducing frequency-selective elements into the resonator, but a technique that has proven to very efficiently reduce the bandwidth of OPOs by orders of magnitude (Bjorkholm and Danielmeyer 1969) and which has also shown its reliability in operation on aircraft is injection seeding. This technique requires radiation from an external (low-power) source with a narrow bandwidth to be injected into the OPO cavity (see Figs. 31.4c, d and 31.5). For successful injection seeding, the pump laser has to operate in a single longitudinal mode.

Still, the selection of crystals is the main criterion for successful operation of an OPO. Transparent crystalline materials with non-zero second-order nonlinearity are required. The choice of the most suitable crystal material for a given application is often far from trivial; it should involve the consideration of many aspects, such as transparency for all wavelengths involved, a possibility for phase matching, adequate nonlinear coefficient, high resistance to optical damage, low absorption, good optical homogeneity, and favorable thermo-optical and thermo-mechanical properties. Finally, the availability of crystals of consistently good quality and appropriate size is an issue as well.

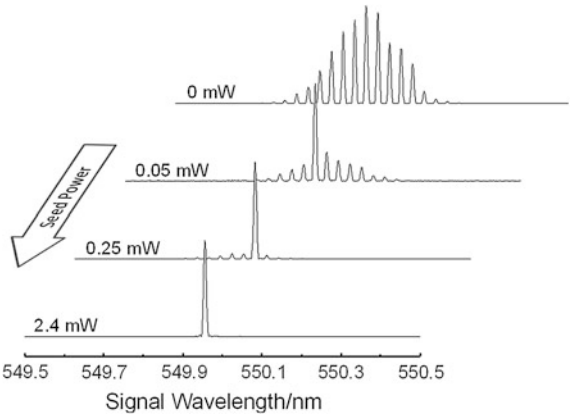

Fig. 31.5 Operation principle of injection seeding. With the help of an OPO with large mode separation the individual modes were experimentally resolved. It is clearly seen how the side modes become suppressed and spectral purity improved once the power of the seed laser is increased; after (Fix and Wallenstein 1996)

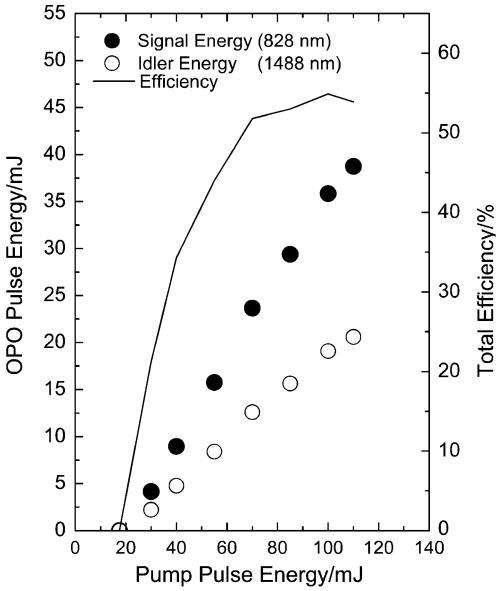

Fig. 31.6 Pulse energy of a KTP OPO in a linear resonator configuration. At a maximum pump energy of the frequency-doubled Nd:YAG pump laser of 110 mJ the signal and idler energies are as high as 38.7 and 20.6 mJ, respectively. The total efficiency exceeds 50 %

From a multitude of nonlinear optical crystals (Nikogosyan 2005) only a handful are useful for practical OPOs/OPAs. Most notably, these are lithium niobate ($LiNbO_3$), potassium titanyl phosphate (KTP) and arsenate (KTA), beta-bariumborate (BBO), lithium triborate (LBO), or, for the mid-infrared spectral range, silver gallium selenide ($AgGaSe_2$) or zinc germanium phosphide (ZGP). In recent years, significant advances have been achieved with QPM materials like periodically poled lithium niobate (PPLN), lithium tantalate (PPLT), or KTP (PPKTP). Since the QPM grating is mostly introduced into the crystals by post-growth techniques like periodic electric field poling, their apertures are relatively

small. Although crystals with aperture sizes as large as 5 mm have been poled (Saikawa et al. 2008), typical commercially available crystals have sizes only up to 1 mm. Therefore, optical damage threshold is an issue and with these crystals it is scarcely possible to achieve OPO energies in the 0.1-J-range at ns pulses.

Due to these advantages, OPOs are now routinely operated when tunable light sources are requested that deliver high peak (>1 MW) and average powers (>1 W) and high conversion efficiency (>30 %) in a robust and compact set-up. As a typical example, Fig. 31.6 shows energy and efficiency of a 532-nm pumped OPO using KTP as the nonlinear crystal. This OPO used a simple, relatively short linear cavity, as depicted in Fig. 31.4b. However, when high-quality beam profiles and excellent spectral properties are required in conjunction with high energy and efficiency, more sophisticated designs are needed. In the past few years this technology has been advanced at DLR to utilize optical parametric frequency converters in lidar applications. In the following, the status of IPA research towards airborne and spaceborne applications is presented.

31.3 Optical Parametric Oscillators at 935 nm for Airborne Differential Absorption Water Vapor Lidar

The development of an OPO at 935 nm in the spectral region of the 3*v* overtone vibrational band of H_2O was stimulated by the need to develop an airborne water vapor DIAL with a high measurement sensitivity at *tropopause* height, particularly in cases of very dry air from the lower *stratosphere*. For this purpose, operation in the 935-nm spectral region is recommended. In this wavelength range, the line strengths of suitable water vapor absorption lines are more than a magnitude higher than in the 4*v* overtone vibrational bands of H_2O near 720 nm or 830 nm, where previous airborne DIAL systems have been operated (at NASA, CNRS, and DLR) using tunable dye, Cr:Alexandrite or Ti:Sapphire lasers. Unfortunately, this spectral region is already at the edge of the tuning range of Ti:Sapphire lasers so OPOs appeared to be the method of choice.

A novel airborne all-solid-state DIAL system for the *Falcon research aircraft* has thus been developed whose transmitter is based on an injection-seeded KTP OPO pumped by the second harmonic of a Q-switched, diode-pumped single-mode Nd:YAG laser at a pulse repetition frequency (PRF) of 100 Hz (Poberaj et al. 2002). The OPO which consists of a three-mirror ring cavity is optimized for operation in the spectral region between 920 and 950 nm with an average output power of 1.2 W. The narrow bandwidth (<150 MHz) and high spectral purity (>99 %) is achieved using an external cavity diode laser as the seed, which is guided by single-mode polarization-maintaining fibers (similar to those used in fiber communication systems) to the OPO cavity. The wavelength of the seed laser is stabilized to the peak of a water vapor line using a multipass absorption cell. In the laboratory, only a few tens of microwatts are sufficient to obtain spectral purity values of 99 %, which is a necessary condition for precise water vapor measurements. In the aircraft, the spectral purity is continuously monitored. When

unseeded, the OPO bandwidth is on the order of 0.26 nm (90 GHz). This is much broader compared to the line width of one particular water vapor absorption line (1–2 GHz). Therefore, the unseeded multimode signal can be used for the non-absorbing off-line or reference measurement. A simple shutter, which was closed for every second pump pulse, therefore enabled alternating on- and off-line measurements. For the first time, daytime measurements of two-dimensional water vapor cross sections with high vertical (500–750 m) and horizontal (6–20 km) resolution were performed from an aircraft in the tropopause region (Poberaj et al. 2002). However, this DIAL has also successfully been applied to investigate water vapor flux in the *atmospheric boundary layer*. This H_2O DIAL was put into service in 1999 and was regularly flown during various scientific research missions in Europe, Northern Africa, the Americas, and Australia until 2006.

The successful operation of the DLR H_2O DIAL led to mission proposals to the *European Space Agency* for a spaceborne H_2O DIAL water vapor measurement with high vertical resolution and high accuracy on a global scale from Earth's surface to the lower stratosphere.

In this context a new, much improved, airborne multiwavelength water vapor DIAL (WALES) has been developed, acting as an airborne *demonstrator* for a space mission and, at the same time, as a state-of-the-art *instrument* for airborne atmospheric studies (Wirth et al. 2009). In order to be able to cover the range from the atmospheric boundary layer to the lower stratosphere in all climate regions, a measurement scheme was developed that uses four different wavelengths in the 935 nm absorption band of H_2O, each one especially adapted to a restricted altitude range of the *atmosphere*.

Therefore the most important new attribute of WALES is multiwavelength operation. The required four wavelengths are generated by using two identical all-solid-state laser systems consisting of an injection-seeded KTP OPO pumped by a diode-pumped Nd:YAG laser at a PRF of 100 Hz each. The pump lasers were developed in-house and feature several unique properties and significant improvements over existing systems (Wirth et al. 2009).

The OPOs, which are pumped at 532 nm at maximum pulse energies of 220 mJ, are set up as planar four-mirror ring oscillators with two KTP crystals cut for Type-II phase matching and resonant for the signal wave, only. Narrow-band operation of the OPOs is again realized using injection seeding. For this purpose, a seed laser system has been designed using four different distributed feedback (DFB) lasers that are stabilized to the respective on- and offline wavelengths (Wirth et al. 2009). Herewith, a *frequency stability* is achieved that is adequate for H_2O measurements (<20 MHz).

Each of the two lasers seed one of the OPOs. The switching between wavelengths is performed on a shot-to-shot basis using fast fiber switches. Thus all H_2O absorption lines can be addressed on a shot-to-shot basis. The whole seed laser system is fully fiber-coupled, resulting in high stability and a modular set-up. Each seed laser provides $\sim$5 mW at the end of its fiber.

The injection seeded OPOs have a pump energy threshold of $\sim$17 mJ/pulse and at full pump energy deliver about 60 mJ of output energy with a decent beam quality ($M^2 < 8$) in 5.5 ns long pulses. The line width is $\leq$150 MHz. For lidar

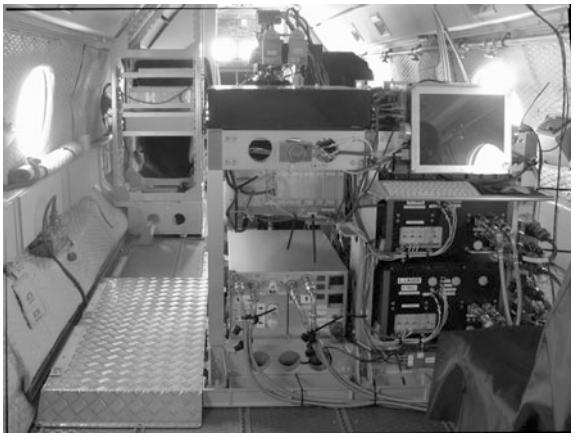

Fig. 31.7 The WALES system as installed on board the G550 *HALO* aircraft

operation, the output energy of the OPOs is usually reduced to ~45 mJ per pulse. Under these conditions, the spectral purity of the OPOs, which is constantly monitored using a side-mode beating technique (Mahnke and Wirth 2010), is reliably better than 99.9 %.

The average power of WALES is ~7.5 times higher than its predecessor system. Together with a larger receiving telescope (1.7 times area increase) and advancements in the detection system, WALES indeed constitutes a tremendous improvement for airborne water vapor DIAL. Next to the water vapor detection capability, WALES features additional detection channels for the fundamental and second harmonic of the pump laser, each with co- and cross-polarized detection and a high spectral resolution channel to retrieve *aerosol extinction* using the HSRL technique.

Assembly of this new instrument was finished in 2007. Since then, the WALES system has confirmed its unique capabilities and reliability in a variety of *airborne campaigns* on board the DLR research aircraft Falcon 20 performed under different climatic conditions over central Europe, the Cape Verde islands, the European *Arctic*, and Japan. Subsequently, WALES has been retrofitted for operation on board HALO and had its maiden flight on this new German research aircraft in 2010 (Fig. 31.7).

While a spaceborne H_2O DIAL mission is yet put on hold due to its lacking technology readiness level, WALES has clearly demonstrated its capabilities both as a state-of-the-art instrument for airborne atmospheric research and as a demonstrator for future space-borne water vapor DIAL instruments.

31.4 Optical Parametric Oscillators in the Ultraviolet Spectral Range for the Detection of Ozone

Another important trace gas that can be measured by DIAL is ozone (O_3), which in the lower atmosphere is known to be an air pollutant with harmful effects on human and animal health, crop production and forest growth. In the stratosphere,

however, where about 90 % of the ozone resides, it is beneficial since it prevents UV radiation from reaching Earth's surface. Ozone is also a tracer for stratosphere-to-*troposphere* exchange.

Thus, the ability to measure ozone both in the lower troposphere as well in the tropopause region is of significant interest. *Satellites* have difficulties in measure tropospheric ozone through the *stratospheric ozone layer*, therefore ozone DIAL systems installed on aircraft have significant advantages for providing a detailed description of the spatial distribution of ozone over large areas.

An O_3 DIAL transmitter needs wavelengths in the UV spectral range. As has been pointed out by many authors, wavelengths in the spectral range of 280–300 nm are appropriate for a DIAL targeting lower tropospheric ozone. Wavelengths above 300 nm do not enable sufficient differential absorption at low concentrations but are appropriate for ozone concentrations that are found in the tropopause region or stratosphere. Wavelengths below 280 nm are strongly absorbed by ozone or attenuated by atmospheric extinction, thus reducing the measurement range significantly. Since O_3 does not exhibit a line spectrum, the stringent line width requirements that must be fulfilled for H_2O, CH_4 or CO_2 do not apply here. Although a variety of lidar systems operate at fixed UV wavelengths it is generally accepted that the preferred O_3 DIAL transmitter should be tunable. Tunability allows optimizing the lidar wavelengths for a selection of optimum absorption cross sections in order to improve the signal-to-noise ratio, to optimize the measurement range, to minimize measurement errors caused by aerosol extinction and gradients, and also to minimize interference with other molecules such as sulfur dioxide (SO_2) or nitric dioxide (NO_2).

OPOs cannot directly generate radiation in this wavelength range, but there are possibilities to convert the radiation of OPOs from the visible or near infrared to the UV using the nonlinear processes of *second harmonic generation* or sum frequency mixing (SFM). Unfortunately, the efficient frequency conversion of OPOs is often hampered due to insufficient spectral and beam propagation characteristics. However, in a series of studies (Fix and Ehret 1998; Fix et al. 2002; Peuser et al. 2009) novel design architectures have been investigated that use OPOs pumped at 532 nm with intracavity sum-frequency mixing to achieve enhanced conversion efficiencies.

By means of different mixing schemes almost the entire wavelength range between 234 and 325 nm could be covered. When pumped with the harmonics of a (flashlamp-pumped) Nd:YAG laser at 10 Hz (fundamental energy: 400 mJ), UV energies of almost 20 mJ can be achieved with a conversion efficiency from the fundamental of the pump laser to the UV as high as 5 %.

Based on this experience, an airborne O_3 DIAL system (TropOLEX), consisting of a near infrared OPO with sum-frequency mixing internal to the cavity was developed. It is continuously tunable between 281 and 300 nm and thus well adapted to the requirements for measuring tropospheric ozone. In order to generate UV wavelengths in the range from 280 to 300 nm using a Nd:YAG with frequency doubling and tripling, there are different possibilities to combine OPO and SFM processes. For TropOLEX it was decided to use the second harmonic of a Nd:YAG laser for pumping

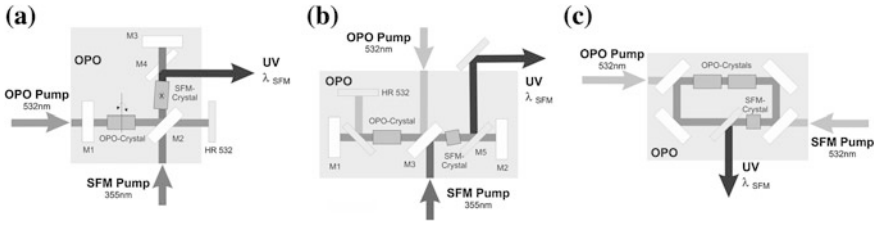

Fig. 31.8 Various OPO set-ups with intracavity sum frequency mixing for the generation of tunable UV radiation

the OPO and the third harmonic of the same laser for sum frequency mixing with the OPO's idler radiation (Fig. 31.8b). Since none of the other possible materials combine as many advantages as KTP and BBO (Nikogosyan 2005), these crystal materials were chosen for the OPO and SFM, respectively. The OPO needs to generate idler radiation in the range of 1498 and 1781 nm. Sum frequency mixing with the third harmonic of the pump converts the radiation to the target wavelengths. The system design follows the earlier concept (Fix and Ehret 1998) except that the cavity mirror design was changed (Fig. 31.8b), which enables easier manufacturing for the wavelengths involved. The pump source for this system is an injection-seeded flashlamp pumped Nd:YAG laser with an output energy of 400 mJ at a PRF of 10 Hz. The maximum UV energy generated is about 14 mJ, yielding an optical conversion efficiency of 3.5 % with respect to the fundamental of the Nd:YAG pump. The spectral bandwidth in the UV is <0.05 nm (<160 GHz), which is much better than required. The OPO system is rather small in size ($20 \times 17 \times 10$ cm^3) and comprises high mechanical stability. In order to change the UV wavelengths from on- to off-line in between two successive pulses, both crystals have to be rotated synchronously with the laser's PRF. Therefore, both crystals are attached to temperature-stabilized galvomotors with an accuracy that allows stable wavelength operation.

The OPO transmitter was merged with a receiving and data acquisition system for operation on DLR's single-engine Cessna Grand Caravan aircraft.

Figure 9 shows the lidar as installed in the aircraft. The entire power consumption is about 1 kW and the overall weight 270 kg. The typical resolution of the two-dimensional ozone measurement is up to 220 m in the vertical and 600 m in the horizontal direction. During first test flights it could be demonstrated that two dimensional ozone cross sections throughout the lower troposphere could be recorded with an overall error of less than 10 %.

A further challenge that has been tackled is the ability to measure water vapor and ozone in the tropopause region with high spatial resolution at the same time with a single instrument. To accomplish this task, a further UV-OPO was devised that could be pumped with the same pump laser system as the WALES H_2O DIAL. Since this pump laser operates at a PRF of 100 Hz its average power is about tenfold compared to the TropOLEX system. To handle this high average power the optical design had to be revised (Fig. 31.8c). As for upper tropospheric and lower

Fig. 31.9 The Ozone DIAL TropOlex as installed in the Cessna Grand Caravan aircraft

stratospheric ozone the wavelengths have to be longer; here the OPO signal radiation (699–784 nm) was mixed with the second harmonic of the pump to generate wavelengths between 302 and 317 nm. In this entire range this frequency converter generates more than 1 W of average power, i.e., pulse energy of more than 10 mJ at 100 Hz. As before, on- and off-line wavelengths are generated using galvomotors to reach the respective phase matching angles on a shot-to-shot basis. The mechanical and optical design was devised such that the UV converter fits into a module identical with the OPO for water vapor detection at 935 nm (Fig. 31.10). Therefore, it is now indeed possible to measure water vapor and ozone at the same time by replacing one 935-nm OPO module by the new UV-OPO module. The maiden flight of this new combination of lidars (called AMALFI) is still pending.

31.5 Infrared Optical Parametric Oscillators and Amplifiers for the Detection of Greenhouse Gases

Even more challenging than the applications described above is the detection of the most important *anthropogenic greenhouse* gases, carbon dioxide (CO_2) and methane (CH_4), using lidar techniques. Since both gases exhibit adequate absorption lines only in the near and mid-infrared spectral range, where the quality of state-of-the-art detectors is still limited, the Integrated Path Differential Absorption (IPDA) technique has advantages compared with DIAL (Milton et al. 1997), even if this method lacks range resolution and provides but column-integrated *mixing ratios* of the greenhouse gases.

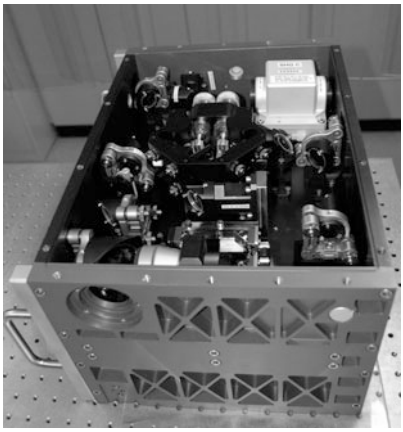

Fig. 31.10 Photograph of the UV frequency converter for ozone DIAL measurements. The outer dimensions of this module are identical to those of the 935-nm module used in the WALES system for water vapor DIAL and thus can be exchanged (courtesy F. Steinebach)

The first IPDA system built at DLR (in co-operation with ADLARES GmbH and under the aegis of E.ON Ruhrgas, now Open Grid Europe) was a system to detect leaks in natural gas pipelines using a small helicopter (CH_4 airborne remote monitoring, CHARM®). In order to fulfill the requirements for routine survey of gas pipelines, such a system has to be capable of identifying small natural gas leaks (leakage flow rates of 0.05–10 m^3/h) detected against the natural background level of $\sim$1.8 μmol/mol within a single pulse pair from a flight altitude of 80–150 m and a helicopter velocity of 70–150 km/h. Careful analysis suggested using the strong absorption at $\sim$3.3 μm of methane, which give useful absorption to achieve the necessary sensitivity. Again, an OPO was chosen to fulfill the stringent requirements to the transmitter. In a few cases, OPOs have been used before for similar applications but none of the known existing airborne methane DIAL systems fulfilled the requirements for pipeline monitoring in terms of sensitivity, compactness, and flexibility. As the lidar transmitter, a single longitudinal mode OPO was developed that uses KTP as the nonlinear crystal in a three-mirror ring resonator (see Fig. 31.11). A custom-designed diode-pumped Nd:YAG laser (Rofin-Sinar, Germany) with a PRF of 100 Hz and an average power of 3 W serves as pump. A double-pulse option enables the emission of two closely separated pulses ($\sim$200 μs) with almost equal energy. The usable idler output energy of the OPO in the 3.3-μm range is of the order of 1 mJ per single pulse. Single longitudinal mode operation is again achieved using the technique of injection seeding, but here at the OPO's signal wavelength (at $\sim$1 580 nm) to provide narrowband output at the idler wavelength in the target spectral range. Two fiber-coupled diode lasers operating at slightly different wavelengths are alternately switched to be injected into the OPO cavity that is actively matched to the seed. This enables the OPO to alter its wavelength from on-line to off-line within a double pulse sequence. This lidar transmitter was integrated together with an optical head for precise beam steering, receiving telescope, detectors as well as data acquisition system and cooling unit into a container sized

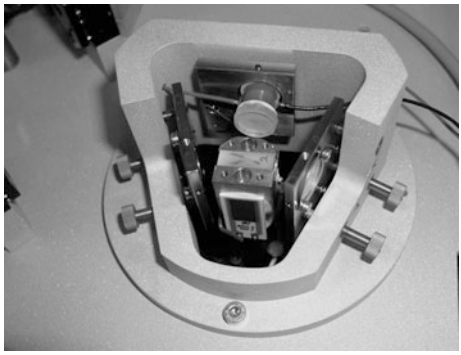

Fig. 31.11 Photograph of the OPO at 3.3 μm used as the transmitter for gas leak detection (courtesy: P. Mahnke)

$2.3 \times 1.1 \times 0.5\ m^3$. Since its development, CHARM® is routinely employed to survey natural gas transportation pipelines and has received the national authority's certification to monitor urban gas grids.

Compared to identifying small "hot spots" such as gas leaks, the challenges to measure CH_4 and CO_2 for climate research from aircraft or spacecraft are much more stringent. While in the former case measurement accuracy is not so essential (for localization it suffices to distinguish the leak from background levels), the measurement of CH_4 and CO_2 for climate research requires extremely high accuracy and precision for flux estimation. On the other hand, horizontal averaging is acceptable, whereas this is not the case for pipeline monitoring.

In order to fulfill the demanding requirements for accuracy and precision, the spectral aspects of the laser transmitter play a major role in the definition of spaceborne and airborne instruments. To support the proposed missions for CO_2 and CH_4 IPDA measurements from space (i.e., Advanced Space Carbon and Climate Observation of Planet Earth A-SCOPE, and Methane Remote Sensing Lidar Mission MERLIN), activities have started on transmitter development and are continued at IPA.

For CO_2 IPDA, two possible wavelength regions at 1.57 and 2.05 μm have been identified as suitable for measurement from space. For CH_4 IPDA, wavelengths around 1.65 μm show favorable properties for achieving the measurement requirements with respect to absorption cross section, temperature insensitivity, and optimum weighting functions. The basic transmitter requirements are listed in Table 31.1.

The experimental investigations within this study concentrated on the wavelength range at 1.57 μm because here the requirements are the most stringent. Since optical parametric frequency conversion is a versatile concept, the results can of course be readily transferred to other wavelengths ranges of interest.

The first OPO to demonstrate the IPDA technique for CO_2 column measurement was based on a KTP OPO in a three-mirror ring configuration, very similar to the earlier approach of the 935-nm OPO for water vapor measurements. This OPO was pumped with a flashlamp pumped Nd:YAG laser at 1 064 nm at a PRF of 10 Hz and generated signal radiation at 1 572 nm. The output energy was on the order of 10 mJ. Injection seeding was performed using DFB diode lasers at that wavelength range.

Table 31.1 Basic transmitter requirements for spaceborne instruments measuring carbon dioxide and methane with IPDA

Parameter	CO_2		CH_4
Wavelength range (μm)	1.57	2.05	1.64
Transmitter pulse energy (mJ)	50	50	10
Bandwidth (MHz)	60	60	60
Mean frequency stability	70 kHz	270 kHz	10 MHz
Spectral purity (%)	99.97	99.74	99.0

With the help of this laboratory set-up, first IPDA measurements were successfully performed (Amediek et al. 2008). In a separate activity aimed at obtaining reliable information about the absolute ground reflectance and its variations at the particular measurement wavelength, which has an impact on the measurement sensitivity of spaceborne and airborne IPDA systems, the OPO was employed in an airborne measurement campaign. For this purpose the UV-OPO of the TropOLEX system was exchanged against this OPO transmitter at 1 572 nm to perform extended measurements on board the Cessna Grand Caravan (Amediek et al. 2009).

The need to meet the stringent requirements for spaceborne and airborne IPDA of CO_2 and CH_4, particularly in terms of the spectral properties, led to the development of an improved transmitter at 1 572 nm (Fix et al. 2011). The OPO closely follows the OPO design of the WALES system. However, to maintain the spectral purity and good beam quality even at high pulse energies the OPO signal was amplified in a subsequent optical parametric amplifier stage. The set-up of this system is depicted in Fig. 31.12.

As pump laser for the OPO and OPA, a diode-pumped Q-switched Nd:YAG laser generating pulse energies of ∼250 mJ at 100 Hz repetition frequency was used. For the OPA experiments, the laser design was altered to use a partially reflecting mirror (PR) between oscillator and amplifier (Fig. 31.12). Approximately 15 mJ of pump energy from the oscillator were extracted and used to pump the OPO. The remaining part is amplified up to ∼200 mJ and can be used to pump the OPA.

Both, OPO and OPA use KTA as the nonlinear material since it has much lower absorption in the idler spectral range (∼3300 nm) compared to, for example, KTP. In the OPO two identical crystals were employed and generated signal pulse energies on the order of 3–4 mJ. In a subsequent OPA stage this was amplified to yield output energies of more than 50 mJ at 1 572 nm (see Fig. 31.13) corresponding to the average power of about 5.5 W.

In order to stably seed the OPO, a stabilization concept has been implemented based on experience with the WALES development (Mahnke and Wirth 2010). The OPO when seeded with a fiber laser showed an absolute stabilization of its center frequency of <70 kHz (for averaging times of more than 3 s). The seed lasers were stabilized to either a CO_2 absorption cell or to a frequency comb.

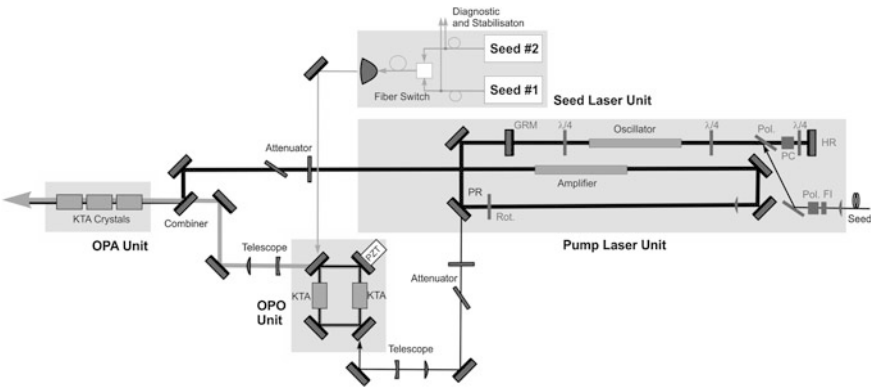

Fig. 31.12 Set-up of the entire narrow band OPO/OPA system consisting of the Nd:YAG pump laser, the OPO unit, a seed laser unit system for narrow band operation of the OPO/OPA, and an OPA stage

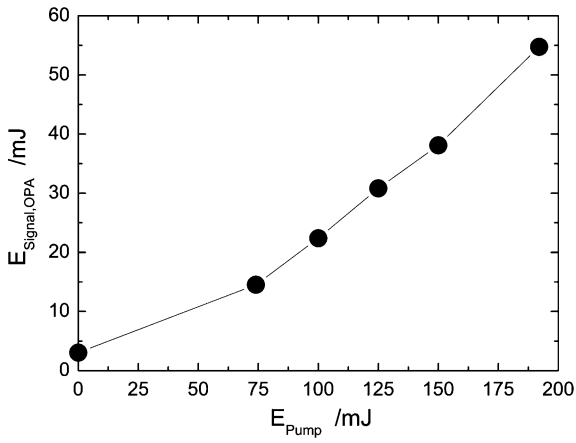

Fig. 31.13 OPA signal energy as a function of pump energy. The pulse energy exceeds 50 mJ per pulse, which corresponds to more than 5 W of average power

The bandwidth of the seeded OPO was measured to be on the order of 90 MHz. Although this is somewhat broader than required, this is purely dependent on the pulse width of the pump laser (8 ns). Longer pulses that were not accessible within the scope of this study will enable further reduction of the spectral width of the OPO, which is already close ($\sim 1.2\ \times$) to the time-bandwidth limit.

The seed power available for the OPO was of the order of 5–9 mW. However, even for seed powers in the range of 150 μW spectral purity better than 99.98 % was achieved, as confirmed by long-pass absorption measurements using a multipass cell with an optical path length of 210 m filled with pure CO_2 (Fix et al. 2011).

Being that promising, this design architecture is now implemented for the airborne demonstrator CHARM-F and also serves as the baseline for MERLIN to be launched in the time frame of 2016.

31.6 Summary and Outlook

Advances in developing reliable and versatile tunable laser sources for lidar applications were summarized. Various transmitters based on optical parametric frequency conversion have successfully been devised for airborne DIAL and IPDA measurements of H_2O, O_3, CH_4 and CO_2. In addition, those devices are also applicable to other lidar applications where tunability is required, e.g., fluorescence lidar or *high spectral resolution lidar*, to match the wavelength to the bandpass of narrowband filters.

The OPO technology has been advanced to a state where it is now considered to be the baseline approach for the German–French methane satellite mission MERLIN. It is therefore likely that OPOs find their way into space to make the visions of science fiction authors (Brin 1980) finally come true.

References

Amediek, A., Fix, A., Wirth, M., Ehret, G.: Development of an OPO system at 1.57 μm for integrated path DIAL measurement of atmospheric carbon dioxide. Appl. Phys. B **92**, 295–302 (2008). doi:10.1007/s00340-008-3075-6

Amediek, A., Fix, A., Ehret, G., Caron, J., Durand, Y.: Airborne lidar reflectance measurements at 1.57 μm in support of the A-SCOPE mission for atmospheric CO2. Atmos. Meas. Tech. **2**, 755–772 (2009). doi:10.5194/amt-2-755-2009

Armstrong, J.A., Bloembergen, N., Ducuing, J., Pershan, P.S.: Interactions between light waves in a nonlinear dielectric. Phys. Rev. **127**, 1918–1939 (1962). doi:10.1103/PhysRev.127.1918

Baumgartner, R.A., Byer, R.L.: Continuously tunable ir lidar with applications to remote measurements of SO_2 and CH_4. Appl. Opt. **17**, 3555–3561 (1978). doi:10.1364/AO.17.003555

Bjorkholm, J.E., Danielmeyer, H.G.: Frequency control of a pulsed parametric oscillator by radiation injection. Appl. Phys. Lett. **15**, 171–173 (1969). doi:10.1063/1.1652954

Brassington, D.J.: Differential absorption lidar measurements of atmospheric water vapor using an optical parametric oscillator source. Appl. Opt. **21**, 4411–4416 (1982). doi:10.1364/AO.21.004411

Brin, D.: Sundiver. Bantam Book, New York (1980)

Ebrahim-Zadeh, M., Dunn, M.H.: Optical parametric oscillators. In Bass, M., Enoch, J.M., Van Stryland, E.W., Wolfe, W. L. (eds.) OSA Handbook of Optics, vol. 4, pp. 22.21–22.72. McGraw-Hill, New York (2004)

Endemann, M., Byer, R.L.: Simultaneous remote measurements of atmospheric temperature and humidity using a continuously tunable IR lidar. Appl. Opt. **20**, 3211–3217 (1981). doi:10.1364/AO.20.003211

Fix, A., Wallenstein, R.: Spectral properties of pulsed nanosecond optical parametric oscillators: experimental investigation and numerical analysis. J. Opt. Soc. Am. **B13**, 2484–2497 (1996). doi:10.1364/JOSAB.13.002484

Fix, A., Ehret, G.: Intracavity frequency mixing in pulsed optical parametric oscillators for the efficient generation of continuously tunable ultraviolet radiation. Appl. Phys. B **67**, 331–338 (1998). doi:10.1007/s003400050512

Fix, A., Wirth, M., Meister, A., Ehret, G., Pesch, M., Weidauer, D.: Tunable ultraviolet optical parametric oscillator for differential absorption lidar measurements of tropospheric ozone. Appl. Phys. B **75**, 153–163 (2002). doi:10.1007/s00340-002-0964-y

Fix, A., Büdenbender, C., Wirth, M., Quatrevalet, M., Amediek, A., Kiemle, C., Ehret, G.: Optical parametric oscillators and amplifiers for airborne and spaceborne active remote sensing of CO_2 and CH_4. Proc. SPIE **8182**, 818206 (2011). doi:10.1117/12.898412

Giordmaine, J.A., Miller, R.C.: Tunable coherent parametric oscillations in LiNbO3 at optical frequencies. Phys. Rev. Lett. **14**, 973–976 (1965). doi:10.1103/PhysRevLett.14.973

Mahnke, P., Wirth, M.: Real-time quantitative measurement of the mode beating of an injection-seeded optical parametric oscillator. Appl. Phys. B **99**, 141–148 (2010). doi:10.1007/s00340-010-3923-z

Milton, M.J.T., Gardiner, T.D., Molero, F., Galech, J.: Injection-seeded optical parametric oscillator for range-resolved DIAL measurements of atmospheric methane. Opt. Commun. **142**, 153–160 (1997). doi:10.1016/S0030-4018(97),00260-5

Nikogosyan, D.N.: Nonlinear optical crystals: a complete survey. Springer, New York (2005)

Orr, B.J., He, Y., White, R.T.: Spectroscopic applications of tunable optical parametric oscillators. In: Duarte 2nd, F. (ed.) Tunable Laser Applications. CRC, New York (2009)

Peuser, P., Platz, W., Fix, A., Ehret, G., Meister, A., Haag, M., Zolichowski, P.: Compact, passively Q-switched, all-solid-state master oscillator-power amplifier-optical parametric oscillator (MOPA-OPO) system pumped by a fiber-coupled diode laser generating high-brightness, tunable, ultraviolet radiation. Appl. Opt. **48**, 3839–3845 (2009). doi:10.1364/AO.48.003839

Poberaj, G., Fix, A., Assion, A., Wirth, M., Kiemle, C., Ehret, G.: Airborne all-solid-state DIAL for water vapour measurements in the tropopause region: system description and assessment of accuracy. Appl. Phys. B **75**, 165–172 (2002). doi:10.1007/s00340-002-0965-x

Saikawa, J., Miyazaki, M., Fujii, M., Ishizuki, H., Taira, T.: High-energy, broadly tunable, narrow-bandwidth mid-infrared optical parametric system pumped by quasi-phase-matched devices. Opt. Lett. **33**, 1699–1701 (2008). doi:10.1364/OL.33.001699

Saleh, B.E.A. Teich, M.C.: Fundamentals of Photonics. Wiley, New York (2007)

Wirth, M., Fix, A., Mahnke, P., Schwarzer, H., Schrandt, F., Ehret, G.: The airborne multi-wavelength water vapor differential absorption lidar WALES: system design and performance. Appl. Phys. B **96**, 201–213 (2009). doi:10.1007/s00340-009-3365-7

Advanced Numerical Modeling of Turbulent Atmospheric Flows

32

Christian Kühnlein, Andreas Dörnbrack and Thomas Gerz

Abstract

The present chapter introduces the method of computational simulation to predict and study turbulent atmospheric flows. This includes a description of the fundamental approach to computational simulation and the practical implementation using the technique of large-eddy simulation. In addition, selected contributions from IPA scientists to computational model development and various examples for applications are given. These examples include homogeneous turbulence, convective boundary layers, heated forest canopy, buoyant thermals, and large-scale flows with baroclinic wave instability.

32.1 Introduction

There are three fundamental approaches to quantifying the properties of atmospheric flows. The first approach is observations using devices such as passive and active remote sensing, instrumented airborne platforms, soundings, or ground-based

C. Kühnlein (✉)
Ludwig-Maximilians-Universität München (LMU),
Meteorological Institute Munich (MIM), Theresienstraße 37,
80333 München, Germany
e-mail: christian.kuehnlein@lmu.de

A. Dörnbrack · T. Gerz
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Andreas.Dornbrack@dlr.de

T. Gerz
e-mail: Thomas.Gerz@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_32,

sensors. Although observations are invaluable for gaining knowledge about the actual state of the atmosphere and are of utmost importance for scientific evidence, they only provide an incomplete picture of the atmospheric state in general. Furthermore, observations can tell only very little or nothing about the future state. Also the second possible approach, linear theory, has a limited range of applicability because of the highly nonlinear nature of atmospheric flows. An appropriate means to describe nonlinear flow system evolution is numerical solution of the fundamental physical laws, generally called *computational physics* (Potter 1973).

With the advent of more and more powerful computing systems over the last decades, computational *simulation* methods have been developed as an indispensable tool for predicting and studying atmospheric flows from microscale *turbulence* to planetary-scale *Rossby waves*. Importantly, computational simulations admit physical studies in a model environment, where specific parameters can be varied purposely. As the resulting influence on the solution can be assessed systematically, computational physics serves as a virtual laboratory for studying atmospheric flows. This capability is especially valuable as observations in the atmosphere cannot be repeated under the same ambient conditions. Even though computational simulation methods are fairly accurate nowadays for many applications, the model predictions contain errors due to numerical approximations, insufficiently known initial and boundary conditions, as well as incomplete model physics *parameterizations*.

Computational flow simulations have been a major pillar of research at IPA for the last $\sim$30 years. Methodological developments as well as numerous, multifaceted applications have been carried out at the institute. Here, the focus is mainly on the methodology of *large-eddy simulation* (LES), which nowadays is the basis for most practical implementations of computational simulation of atmospheric flows. The basic idea of LES is to simulate the large, energy-containing eddies of the atmosphere and to parameterize (i.e., approximate by simplified models) the effect of smaller motion scales. LESs are typically applied to study microscale (<1 km) and meso-γ-scale (2–20 km) atmospheric phenomena. Most of the applications are devoted to the atmospheric *boundary layer*. In particular, urban and canopy flows, the dispersion of constituents, as well as dry and moist *convection* are of special relevance.

The chapter is organized as follows. The subsequent Sect. 32.2 briefly introduces the fundamental physical laws underlying computational simulations in atmospheric flow modeling and hints at their solution. In Sect. 32.3, the practical implementation of the physical laws in computational models is addressed, leading to the idea of large-eddy simulation. Section 32.4 highlights selected contributions of IPA scientists in the field of computational modeling of turbulent flows.

32.2 Fundamental Approach

The fundamental physical laws describing the spatiotemporal evolution of atmospheric flows are given here for incompressible *Boussinesq fluids* with a constant background density ρ_0:

$$\frac{D\mathbf{v}}{Dt} = -\frac{1}{\rho_0}\nabla(p - p_0) - \mathbf{g}\frac{\theta - \theta_0}{\theta_0} + \frac{1}{\rho_0}\nabla \cdot \boldsymbol{\tau} \tag{32.1a}$$

$$\nabla \cdot \mathbf{v} = 0 \tag{32.1b}$$

$$\frac{D\theta}{Dt} = \frac{1}{\rho_0}\nabla \cdot \mathbf{h} + \mathcal{H}, \tag{32.1c}$$

where (32.1a) represents the momentum equation expressing Newton's second law of motion for a fluid under various forces, (32.1b) describes the conservation law of mass for an incompressible fluid, and (32.1c) is a thermodynamic equation related to the conservation of internal energy. For viscous fluids this corresponds to the *Navier-Stokes equations*. The dependent variables in the system (1) are the fluid velocity vector $\mathbf{v}$, and the *potential temperature* $\theta = T(p_0/p)^{R/c_p}$, where T is the fluid temperature, p is the pressure; p_0, θ_0, and ρ_0 refer to constant reference values (the latter for density), R and c_p are thermodynamic constants. The symbol $D/Dt = \partial/\partial t + \mathbf{v} \cdot \nabla$ (where ∇ is the nabla operator) denotes the Lagrangian derivative following the *advection* of material "fluid parcels" with the velocity $\mathbf{v}$, $\mathbf{g} = (0, 0, -g)^T$ in (32.1a) represents the acceleration vector due to Earth's gravity, and $\mathcal{H}$ in (32.1c) a heat source or sink (e.g., from radiative heating or cooling). The terms involving the viscous stress tensor $\boldsymbol{\tau}$ and heat flux vector $\mathbf{h}$ describe dissipation of momentum and diffusion of heat, respectively.[1] Surface heat fluxes enter the normal component of $\mathbf{h}$ at the boundary. Finally, we note that the governing Eq. (32.1a–32.1c) can be expanded with additional equations describing moist physics associated with the formation of water and ice clouds, as well as other processes when these are of interest.

Given well-posed initial and boundary conditions, the governing Eq. (32.1a–32.1c) completely describe all degrees of freedom of the flow in a deterministic manner. Whether one attempts to solve these incompressible Boussinesq equations for atmospheric flows, their nonlinearity precludes analytical solutions. In general, atmospheric flows are turbulent, i.e., they are chaotic, multiscale and dissipative. Nevertheless, treating the evolving atmospheric state as an initial value problem in computational physics, one can obtain numerical solutions to the nonlinear Eq. (32.1a–32.1c) , in order to predict the spatiotemporal evolution of the flow and the attendant thermodynamic fields.

[1] A homogeneous "molecularly well-mixed" fluid is assumed, which is why no diffusion term occurs in the mass conservation; see Tuck (2008).

In essence, *numerical methods* discretize the continuous Eq. (32.1a–32.1c) in space and time using techniques such as finite differences, finite volumes, finite elements, or spectral methods and combinations of them. The resulting algebraic equations are solved computationally. Basically, numerical methods are justified when the numerical solution converges to the solution of the original continuous equations as the discrete representation (or the grid spacing) is refined. However, convergence in a strict mathematical sense is not attainable for turbulent flows (see the discussion in the subsequent section).

32.3 Computational Model Development

There are well-established approaches for the numerical simulation of turbulent flows. *Direct numerical simulation* (DNS) completely solves the governing Eq. (32.1a–32.1c) for the entire spectrum of motion scales. The basic idea is that a turbulent flow consists of a spectrum of eddies of different sizes. The geometry of the domain and the energy content of the larger-scale flow determine the so-called integral length scale l. The energy of the eddies cascades down to smaller scales where, eventually, at the smallest scales random molecular motion dissipates the energy into heat—a concept proposed by Richardson (1922). Theoretically, the energy transfer is best described by the celebrated *Kolmogorov* $k^{-5/3}$ *law* for the *inertial subrange* (Kolmogorov 1941). Therefore, a DNS covering all scales is considered to be the most accurate approach for simulating turbulent flows. On the other hand, DNS is enormously expensive in terms of computational demands.

This is easily seen by estimating the computational demands. The ratio of the integral flow scale l to the dissipative scale—given by the Kolmogorov microscale η—follows the relation $l/\eta \sim Re^{3/4}$ (Geurts 2003). The nondimensional *Reynolds number* $Re = u' l/\nu$ (where u' is a characteristic velocity fluctuation and ν the kinematic fluid viscosity) is a measure of the relative importance of inertial forces over the viscous molecular forces in the fluid. For turbulence to exist $Re \gg 1$. A complete flow representation requires grid increments as least as fine as the order of η, in a computational domain that encompasses the largest flow scales l. Consequently, the number of grid cells in three-dimensional space scales as the 9/4 power of the Reynolds number, $Re^{9/4}$, resulting in enormous requirements for computing power as Reynolds numbers for the free atmosphere are 10^7 and higher. In fact, the situation is even worse because integration schemes normally require reducing the time step increment according to the refinement of the grid increments in space, which therefore yields an essential scaling $\sim Re^3$ for the computational problem. Therefore, DNS of atmospheric flows is for the foreseeable future restricted to small-scale processes and to low Re-numbers. Present-day state-of-the-art computing systems allow the use of $\sim 10^9$ spatial degrees of freedom, and this number is expected to increase approximately by an order of magnitude every 10 years.

The only practicable way to simulate turbulent atmospheric flows numerically is to reduce the complexity of the governing Eq. (32.1a–32.1c), resulting in a feasible size of the associated computations. For this purpose, large-eddy simulations use the important property of turbulent flows that most of the kinetic energy occurs at larger scales (Kolmogorov 1941). Usually, LES employs spatially filtered equations. The basic idea is that a spatial low-pass filter with a width Δ eliminates the high-frequent and short-wave components of the turbulent flow. The common assumption is that the filter width Δ lies in the inertial subrange $l \gg \Delta \gg \eta$. This means, large-eddy simulations—in German called "Grobstruktursimulationen"—explicitly capture the larger turbulent eddies, whereas the net effect of the smaller, less energetic eddies has to be modeled. Pioneering work in developing LES methodology was carried out by Smagorinsky (1963), Lilly (1967), Deardorff (1974), and Schumann (1975).

Formally, the dependent variables $\phi = \mathbf{v}, \theta$ are decomposed as $\phi = \overline{\phi} + \phi'$ into an average $\overline{\phi}$ (representing scales larger than Δ) and fluctuations ϕ' (representing scales smaller than Δ). Commonly, $\overline{\phi}$ and ϕ' are referred to as resolved (grid scale) and unresolved (*subgrid scale*, abbreviated as SGS) components of the turbulent flow field, respectively. Inserting the decomposed components into the governing Eq. (32.1a–32.1c) and averaging them leads to equations which describe the filtered larger-scale turbulent flow.

A typical specification for the filter width is $\Delta \sim \Delta_h$, i.e., the governing Eq. (32.1a–32.1c) are averaged over length scales comparable in size to the grid increments Δ_h of the underlying numerical *discretization*. In fact, discretization of the governing equations on the grid already implicitly filters flow scales smaller than $\sim \Delta_h$ (Schumann 1975). Therefore, explicit filter operations (e.g., convolution with a top-hat or a Gaussian filter) are mostly not applied.

Due to the nonlinearity of the advective terms, however, filtering of the governing Eq. (32.1a–32.1c) generates new terms proportional to the divergence of the SGS fluxes involving the correlations of the unresolved variables ϕ'. These new subgrid terms are formally similar to the divergence of the Fickian fluxes in (32.1a) and the heat flux in (32.1c). However, the new, unknown SGS correlations $\boldsymbol{\tau}_t$ and $\mathbf{h}_t$ define turbulent eddy fluxes instead of the molecular fluxes $\boldsymbol{\tau}$ and $\mathbf{h}$. Hence, these eddy flux divergences must be specified by new assumptions in order to close the governing Eq. (32.1a–32.1c).

Typically, this is achieved by means of special SGS models (also called parameterizations) formulated as functions of the resolved fields. Many different approaches for SGS modeling exist. The most popular one in atmospheric LES models is based on the eddy viscosity/diffusivity assumption where the eddy flux $\boldsymbol{\tau}_t$ is defined to be proportional to a turbulent eddy viscosity and the strain rate tensor of the resolved (or filtered) velocity field. Additionally, the eddy viscosity must be specified by appropriate length and velocity scales. It is impossible to discuss the large variety of available SGS models within the scope of the present chapter. The interested reader is referred to the relevant literature, e.g., Schumann (1991, 1995), Geurts (2003) and references therein.

Although LES allows flows to be simulated with much higher *Re*-numbers compared to a DNS, simulating the three-dimensional time-dependent evolution is still computationally very demanding by today's standards. Therefore, LES is nowadays applied for (quasi-)realistic Reynolds numbers in limited regions of the atmosphere, as the examples provided in the subsequent Sect. 32.4 will show. It is obvious from limitations of the computational model formulation as well as the inherent unpredictability of turbulent flows that LES and DNS cannot provide a representation of the flow in every detail. Therefore, the simulations mostly aim at calculating accurately the statistical properties of the turbulent flow (e.g., average transports of momentum, heat and moisture, spectral distributions, and so forth). Nevertheless, LES and DNS predict explicitly the large-scale structures in the evolving flow field.

32.4 Examples for Applications

In this section, selected contributions in the area of computational simulation of turbulent atmospheric flows by IPA scientists are highlighted.

32.4.1 DNS of Turbulent Shear Flows

In the 1980s, a DNS could only handle flows with *Re*-numbers of about 80. This Reynolds number, although much lower than under typical atmospheric conditions, is high enough to allow some insight into elementary nonlinear interactions in turbulent flows remote from walls. Effects of constant stable *thermal stratification* and vertical shear on the evolution of *homogeneous turbulence* were investigated using DNS for the first time by Gerz et al. (1989). These conditions are relevant to the dynamics of the stably stratified free atmosphere. It was found that the low *Re*-number DNS represented the essential physical processes correctly, in agreement with measurements from water tank and wind tunnel experiments. Analyzing the heat transport, the DNS revealed a persistent counter-gradient flux in the vertical direction, i.e., heat was transported from lower-temperature to higher-temperature regions. Additionally, the DNS indicated a counter-gradient flux of momentum, which was not reported from the experiments.

A follow-on study (Gerz and Schumann 1996) revealed that counter-gradient fluxes of heat and momentum occur at large scales when thermal stratification is strong. However, they always do occur at small scales, independently of thermal stratification. A conceptional model was proposed to explain such persistent counter-gradient fluxes at small scales, see Fig. 32.1. In the simulated shear flow, inclined vortex structures (Ω-loops) evolve by self-organization. In these coherent structures, a head-down eddy transports large lumps of warm fluid downward, a head-up eddy transports lumps of cold fluid upwards. Eventually, these coherent eddies effectively manage the co-gradient, negative flow of momentum and heat at

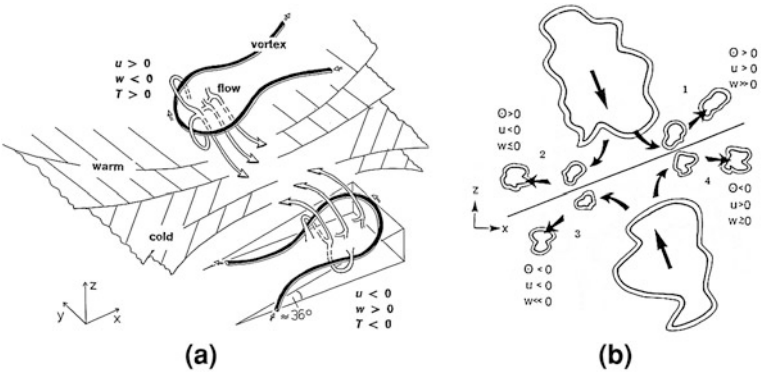

Fig. 32.1 **a** Sketch of a pair of inclined coherent vortices (Ω-loops) in homogeneously sheared turbulence at neutral and weakly stable stratification with the induced motion of large warm and cold fluid lumps (*arrows*). **b** Sketch of the convergence zone between the two vortices where the large fluid lumps collide and break-up in smaller fluid parcels with labeled properties. The fluid properties u, w, T (or θ) read local velocity components in x and z direction and temperature, respectively. Figures and notation taken from Gerz and Schumann (1996)

large scales, as expected. It happens that these eddies predominantly occur in pairs when the forcing is homogeneous: a head-down eddy always lies above a head-up eddy. Hence, the advected large fluid lumps move against each other and collide in the convergence zone between the two eddies. Small fluid parcels result from the collision of the large lumps, as depicted in the four sectors (Fig. 32.1b). The small parcels differ in their properties such that on average small but persistent counter-gradient, positive fluxes of heat and momentum occur at these small scales.

Spectral analysis revealed that the energy of the small-scale motions is provided by a quick transport of kinetic energy from the scales of production to the dissipating scales. The collision mechanism is an inherent property of the turbulence dynamics. Therefore, the persistent counter-gradient fluxes at small scales reflect a universal character of homogeneous turbulence, and they are found over a large range of Reynolds numbers.

32.4.2 LES of the Convective Boundary Layer

To extend the investigations of turbulent structures in the free atmosphere (Sect. 32.4.1) to atmospheric flows impacted by the presence of Earth's surface, large-eddy simulations have been conducted with growing complexity at IPA during the past 20 years. An early example is the simulation of the *convective boundary layer* uniformly heated from below and topped by a layer of uniformly stratified fluid (Schmidt and Schumann 1989). The SGS fluxes were computed from algebraically approximated second-order closure transport equations for which all essential coefficients were determined from the inertial-range theory of

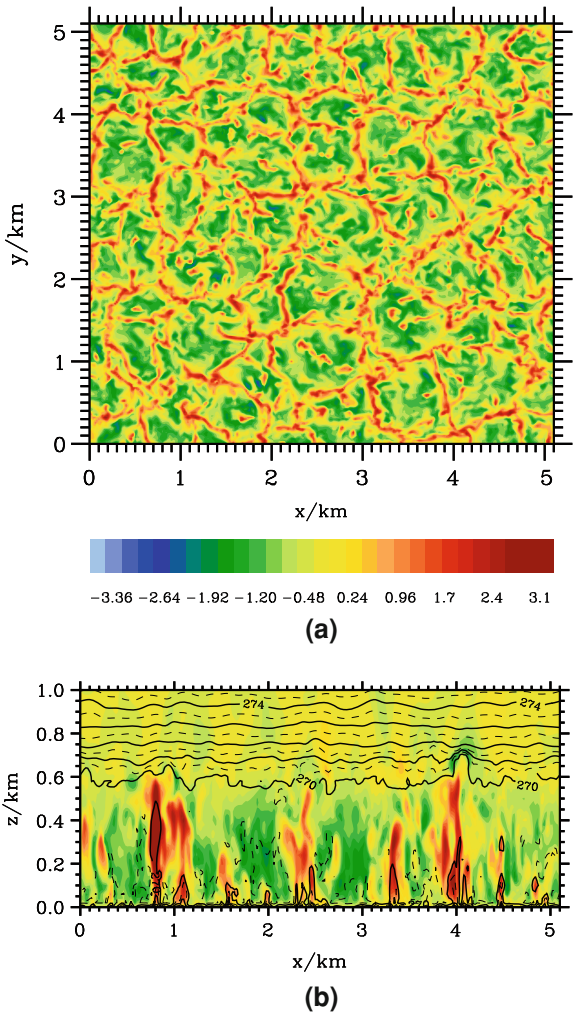

Fig. 32.2 Vertical wind w (ms^{-1}, color shaded) at $z = 100\,m$ altitude **a** and w superimposed by the potential temperature θ (K, black contour lines) in a vertical cross section **b** through the middle of the domain after 120 min simulation time. Numerical simulation performed with the computational model EULAG using $128 \times 128 \times 96$ grid points

homogeneous isotropic turbulence. In Schumann (1991), a first-order turbulence closure model—in which only a prognostic budget equation for the SGS kinetic energy is solved—with adapted SGS length scales in anisotropic, stably-stratified regions was found advantageous over the rather complex second-order closure, which also reveals realizability problems.

Figure 32.2 shows simulated polygonal "spoke" patterns arising in the lower half of the convective boundary layer. In the atmosphere, this cell-type organization of convection often occurs during cold-air outbreaks over warm ocean waters, but it can be also observed over land. Usually, buoyant updrafts with warm air rising in rather narrow columns are surrounded by larger areas which slowly sink as downdrafts. The circulation extends vertically over the whole depth of the

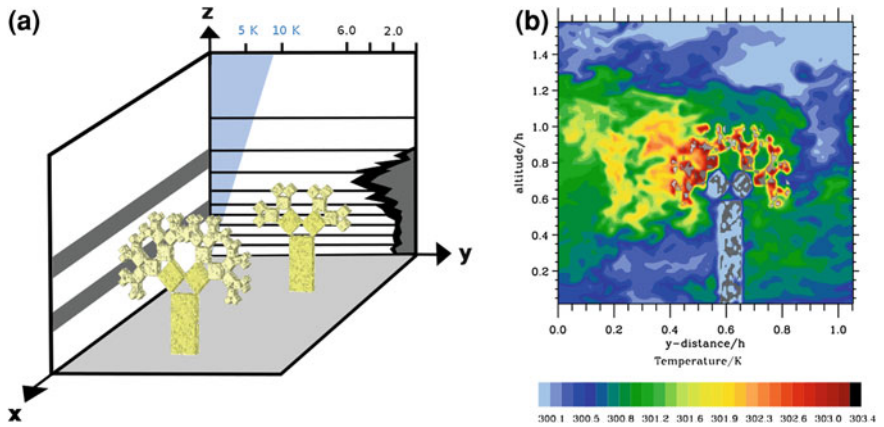

Fig. 32.3 **a** Forest canopy LES setup with two enlarged and numerically resolved exemplary fractal Pythagoras trees (*yellow*). The varying tree height (*upper grey bar*), crown bottom (*lower grey bar*), stretched vertical grid (*black lines*), and leaf area density profile LAD(z) (*black, grey shading*) is shown. The tree crown (*light yellow*) can be diabatically warmed or cooled. The atmospheric background state can be initialized as stable (*blue*) or neutral. **b** A rising thermal is observed as it detaches from the crown space of the fractal Pythagoras tree (*grey*). *h* is the height of the canopy

convective boundary layer. The LES results are in fair agreement with available measurements and simulations as far as mean profiles and spectra of turbulence are concerned. Therefore, the LES, performed with 160 × 160 × 48 grid cells at that time, has been serving as a benchmark test for atmospheric LES models.

32.4.3 LES of Turbulent Flows Through Heated Canopies

Surface heterogeneities introducing rough and solid obstacles for the atmospheric boundary layer flow play a major role in exchange processes of momentum, heat, and constituents. Especially the interpretation of field observations in urban environments and in forests requires accurate representation of the aerodynamic characteristics of rough surfaces in a *numerical model*. The complex geometry of forest stands demands special treatment in the numerical model. For this purpose, LES is an appropriate virtual laboratory for investigating thermodynamic aspects of such flows.

Schröttle and Dörnbrack (2012) investigate the turbulence structure of a heterogeneous and diabatically heated forest stand by high-resolution numerical modeling. The objective is to resolve the flow over a wide range of scales from the plant-scale up to the top of the surface layer (5 cm–100 m). For this purpose, the forest elements are mimicked with immersed boundaries representing fractal Pythagoras trees; see Fig. 32.3. This approach extends the original idea of representing a forest as a porous body of horizontally uniform leaf area density with a constant drag coefficient, as utilized in the pioneering LES conducted by Shaw and Schumann (1992).

In the numerical simulations, vortex shedding behind individual branches or trunks is resolved, as well as the integrated response of the fractal Pythagoras grove on the flow. It could be shown that the reference run for neutral thermal stratification produces physically meaningful results that are quantitatively comparable to former field-scale LES and wind tunnel experiments. The turbulence structure due to different thermal background stratification and heating of the crown space shows the onset and formation of coherent structures; see Fig. 32.3.

32.4.4 Adaptive Meshes

The standard approach for numerical discretization of the governing Eq. (32.1a–32.1c) is to employ rectangular grids with uniform spacing. However, as mentioned before, turbulent geophysical flows are multiscale, i.e., scales covering many orders of magnitude occur simultaneously. Additionally, nonlinear interactions between the motion scales result in sharp gradients and scale contractions, as is obvious at cloud edges or in the formation of atmospheric *fronts*. From a physical perspective, uniform grid spacing may not be the optimal choice for resolving such nonlinear processes.

Nowadays, *numerical weather prediction* and other numerical models simulating atmospheric flows resolve the atmospheric boundary layer with more grid cells. Near the surface the scale of the most energetic eddies is much smaller than those occurring in the free atmosphere. An appropriate strategy to keep computational costs limited is a vertically stretched grid, with very fine grid spacing near the surface that becomes wider upward. Besides reduced computational costs, simulations with stretched vertical grid spacing become less dependent on the parameterization of the subgrid processes in the surface layer.

However, it is a trickier issue to resolve horizontal gradients of atmospheric flows because the ceaseless wind advects, for example, clouds and fronts embedded in the larger-scale ambient flow. Early on, numerical strategies such as *nesting* were proposed (Clark and Farley 1984), and mostly static refinement of the grid has been widely used in atmospheric models until now. Indeed, one has to know the location or the track of the phenomena well in advance to place the better-resolved nest accordingly in the numerical model.

A more effective way of grid refinement is to change the local grid spacing in response to the flow evolution during the course of the simulation. *Solution-adaptive meshes* fulfil this requirement. A numerical solver for atmospheric flows with a solution-adaptive mesh capability was developed recently by IPA scientists (Kühnlein et al. 2012). The particular technique was implemented in the established, state-of-the-art computational model EULAG (Prusa et al. 2008), a work which continued the close collaboration of IPA with the National Center of Atmospheric Research (NCAR) starting in the 1980s (Schumann and Moeng 1991a, b; Moeng and Schumann 1991; Nieuwstadt et al. 1991). Here, the available mesh points are moved in the simulation domain in response to a previously specified monitor function. This monitor function controls the local mesh point density according to a priori stated physical or numerical-error-based refinement criteria.

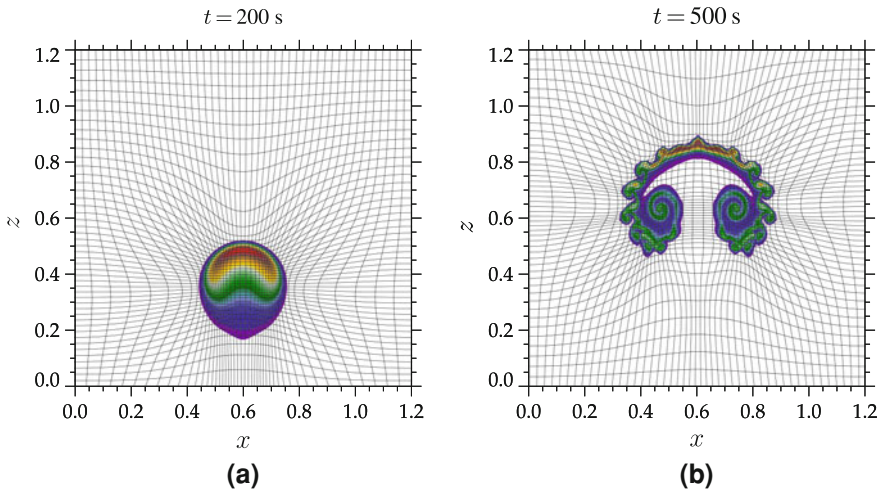

Fig. 32.4 Two snapshots from the adaptive simulation of an idealized buoyant thermal in a neutrally-stratified quiescent atmosphere. Indicated are contours of potential temperature perturbation θ' from the constant background (*shaded*; warmer colors correspond to a higher temperature), overlaid by the adaptive mesh (*solid lines*; only every 2nd mesh line is drawn). Horizontal distance x and height z are given in km

Figure 32.4 presents two snapshots from an idealized, solution-adaptive mesh simulation of a rising buoyant thermal—a highly relevant problem for small-scale atmospheric flows (Kühnlein 2011). For this simulation setup, the monitor function is specified proportional to a combination of vorticity and the gradient of potential temperature. Figure 32.4a, b shows that the refined mesh continuously conforms with the shape of the rising thermal. In particular, fine-scale eddies from shearing instabilities along the edge of the thermal are resolved using only a relatively low total number of mesh cells (Fig. 32.4b).

Large-scale atmospheric flows can benefit equally by the use of a solution-adaptive mesh. For the first time, the performance of solution-adaptive moving mesh methods was investigated for simulating the life cycle of an atmospheric *baroclinic wave instability*—an archetype of the weather patterns (cyclones and anticyclones) in Earth's mid-latitudes (Kühnlein et al. 2012). The nonlinear interactions of this flow problem generate a broad motion spectrum involving propagating and amplifying synoptic-scale baroclinic Rossby waves as well as fine-scale phenomena of fronts and internal *gravity waves*. Figure 32.5 provides two snapshots from the solution-adaptive simulation with a monitor function specified proportional to the gradient of potential temperature at 600 m altitude. Applying this choice of the monitor function, the moving mesh scheme provides finer mesh spacing in the regions of the evolving frontal zones. As demonstrated in Kühnlein et al. (2012), the use of the solution-adaptive moving mesh technique significantly improves the large-scale flow statistics. Moreover, adapting the mesh near the frontal boundaries better resolves the associated large local gradients,

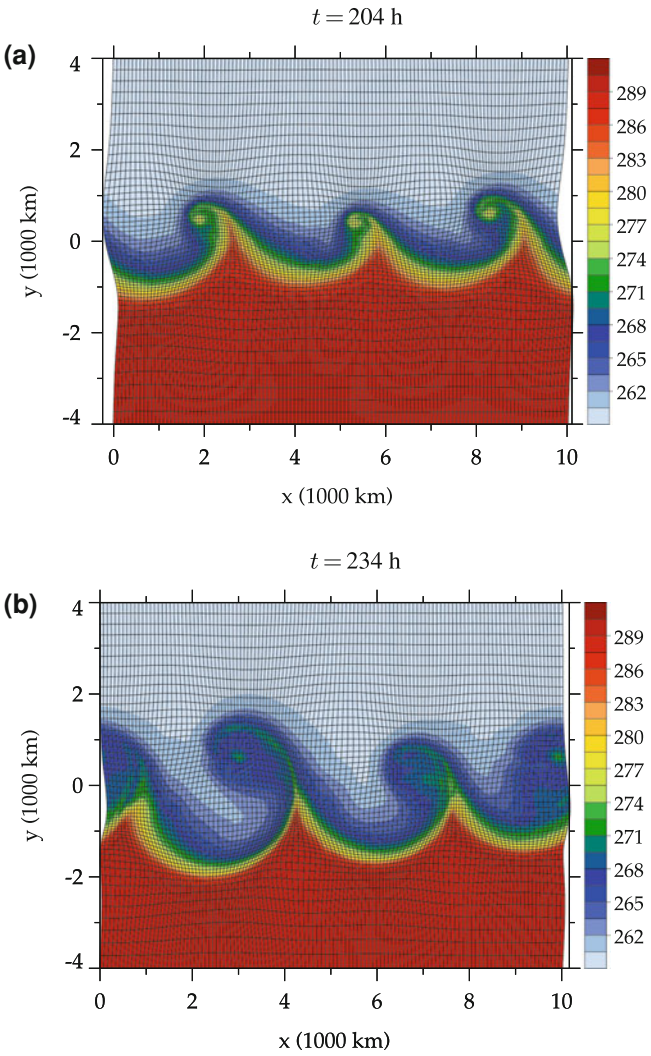

Fig. 32.5 Two snapshots from the adaptive simulation of a baroclinic wave instability. Depicted are contours of potential temperature θ (K, shaded) on a horizontal cross section at height $z = 600\,\mathrm{m}$, overlaid by the solution-adaptive mesh (*solid lines*)

and allows the representation of the intricate mechanisms of internal gravity wave excitation from the baroclinic regions. For the baroclinic instability flow, the adaptive mesh scheme reduces the computing time by a factor of 2–3 compared to a uniform mesh simulation with similar solution quality. Additionally, the solution-adaptive mesh model is able to compute the flow significantly more accurately than a static mesh model using the same computational resources.

32.5 Future Trends

Nowadays, the most important tool for investigating turbulent atmospheric flows is computational simulation, and its relevance for research and operational purposes is likely to strengthen in the future. A main driver behind this trend has been and will be the relentless growth of available computing power. However, the success of computational physics is founded similarly on improvement to the numerical models. Some of the most influential contributions to the computational simulation of turbulent flows are cited in the present chapter.

In the not so distant future, advances in computing power—especially due to massive parallel architectures—will allow atmospheric flow simulations with unprecedented spatial resolutions in *mesoscale* domains. This will benefit, for instance, modeling studies and predictions of cumulus convection (Craig and Dörnbrack 2008), or internal gravity wave breaking (Dörnbrack 1998), which is inter alia relevant for predicting clear-air turbulence to enhance aircraft safety. Another interesting example with reference to research at IPA are the dynamics of aircraft-induced wake vortices that interact with the evolving ambient atmosphere.

Ongoing developments in numerical models of the atmosphere and other branches of computational physics are guided by the use of massive parallelism with thousands of processors. Grid arrangements, numerical algorithms, and physical parameterization schemes that feature high scalability on these parallel computing systems will generally be the preferred choice. In addition, standard serial approaches for analyzing the associated huge amounts of model output data easily reach their limits, and this circumstance challenges new developments. Parallel computing approaches for postprocessing and visualization will certainly become more important in the next few years.

As addressed in Sect. 32.4.4, adaptive mesh methods are promising to more efficiently represent multiscale atmospheric flows in computational models. However, there are still a number of open research issues (Kühnlein et al. 2012). Among others, reliable mesh refinement criteria that work under various atmospheric flow situations and suitable closure models for subgrid scale processes on (solution-)adaptive meshes are not yet available. Regarding the parameterization of turbulence, the special class of numerical algorithms named "high-resolution methods" has the ability to provide an implicit subgrid turbulence model due to the self-adaptive dissipative nature of the numerical truncation terms. This so-called implicit LES approach (Grinstein et al. 2007) is of interest for atmospheric flow simulations with adaptive meshing where the formulation of explicit subgrid models for the turbulent eddy fluxes $\boldsymbol{\tau}_t$ and $\mathbf{h}_t$ is difficult.

References

Clark T.L., Farley R.D.: Severe downslope windstorm calculations in two and three spatial dimensions using anelastic interactive grid nesting: a possible mechanism for gustiness. J. Atmos. Sci. **41**, 329–350 (1984). 10.1175/1520-0469(1984)041<0329:SDWCIT>2.0.CO;2

Craig, G.C., Dörnbrack, A.: Entrainment in cumulus clouds:what resolution is cloud-resolving? J. Atmos. Sci. **65**, 3978–3988 (2008). doi:10.1175/2008JAS2613.1

Deardorff, J.W.: Three-dimensional numerical study of turbulence in an entraining mixed layer. Bound. Layer Meteor. **7**, 199–226 (1974). doi:10.1007/BF00227913

Dörnbrack, A.: Turbulent mixing by breaking gravity waves. J. Fluid Mech. **375**, 113–141 (1998)

Gerz, T., Schumann, U.: A possible explanation of countergradient fluxes in homogeneous turbulence. Theor. Comput. Fluid Dyn. **8**, 169–181 (1996). doi:10.1007/BF00418056

Gerz, T., Schumann, U., Elghobashi, S.E.: Direct numerical simulation of stratified homogeneous turbulent shear flows. J. Fluid Mech. **200**, 563–594 (1989). doi:10.1017/S0022112089000765

Geurts, B.J.: Elements of Direct and Large-Eddy Simulation. R.T. Edwards Inc., Flourtown (2003)

Grinstein, F.F., Margolin, L.G., Rider, W.J. (Eds.): Implicit Large Eddy Simulation: Computing Turbulent Flow Dynamics. Cambridge University Press, New York (2007)

Kolmogorov, A.: The local structure of turbulence in incompressible viscous fluid for very large Reynolds' numbers. Akademiia Nauk SSSR Doklady **30**, 301–305 (1941)

Kühnlein, C.: Solution-adaptive moving mesh solver for geophysical flows. Ph.D. thesis, Ludwig-Maximilians-Universität München (2011)

Kühnlein, C., Smolarkiewicz, P.K., Dörnbrack, A.: Modelling atmospheric flows with adaptive moving meshes. J. Comput. Phys. **231**(7), 2741–2763 (2012). doi:10.1016/j.jcp.2011.12.012

Lilly, D.K.: The representation of small scale turbulence in numerical simulation experiments. IBM Sci. Comput. Symp. Environ. Sci. 195–210 (1967)

Moeng, C.-H., Schumann, U.: Composite structure of plumes in stratus-topped boundary layers. J. Atmos. Sci. **48**, 2280–2292 (1991). doi:10.1175/1520-0469(1991)048<2280:CSOPIS>2.0.CO;2

Nieuwstadt, F.T.M., Mason, P.J., Moeng, C.H. et al.: Large-eddy simulation of the convective boundary layer—A comparison of four computer codes. Thin Solid Films **1**, 1–4 (1991)

Potter, D.: Computational Physics. Wiley, New York (1973)

Prusa, J.M., Smolarkiewicz, P.K., Wyszogrodzki, A.A.: EULAG, a computational model for multiscale flows. Comput. Fluids **37**, 1193–1207 (2008)

Richardson, L.F.: Weather Prediction by Numerical Process. Cambridge University Press, Cambridge (1922)

Schmidt, H., Schumann, U.: Coherent structure of the convective boundary layer derived from large-eddy simulations. J. Fluid Mech. **200**, 511–562 (1989). doi:10.1017/S0022112089000753.

Schröttle, J., Dörnbrack, A.: Turbulence structure in a diabatically heated forest canopy composed of fractal pythagoras trees. submitted to Theor. Comput. Fluid Dyn. Dec 2011 (2012)

Schumann, U.: Subgrid scale model for finite difference simulations of turbulent flows in plane channels and annuli. J. Comput. Phys. **18**, 376–404 (1975). doi:10.1016/0021-9991(75)90093-5

Schumann, U.: Subgrid length-scales for large-eddy simulation of stratified turbulence. Theor. Comput. Fluid Dyn. **2**, 279–290 (1991). doi:10.1007/BF00271468

Schumann, U.: Stochastic backscatter of turbulence energy and scalar variance by random subgrid-scale fluxes. Proc. R. Soc. A. **451**, 293–318 (1995). doi:10.1098/rspa.1995.0126

Schumann, U., Moeng, C.-H.: Plume budgets in clear and cloudy convective boundary layers. J. Atmos. Sci. **48**, 1758–1770 (1991a). doi:10.1175/1520-0469(1991)048<1758:PBICAC>2.0.CO;2

Schumann, U., Moeng, C.-H.: Plume fluxes in clear and cloudy convective boundary layers. J. Atmos. Sci. **48**, 1746–1757 (1991b). doi:10.1175/1520-0469(1991)048<1746:PFICAC>2.0.CO;2

Shaw, R.H., Schumann, U.: Large-eddy simulation of turbulent flow above and within a forest. Bound Layer Meteor **61**, 47–64 (1992). doi:10.1007/BF02033994

Smagorinsky, J.: General circulation experiments with the primitive equations. Mon. Weather Rev. **91**, 99–164 (1963). doi:10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2

Tuck, A.F.: Atmospheric turbulence: a molecular dynamics perspective. Oxford University Press, (2008)

Cloud Resolving Modeling of Contrail Evolution

33

Simon Unterstrasser, Ingo Sölch and Klaus Gierens

Abstract

Contrails are ice clouds that form behind aircraft. As a result of burning kerosene in the engines, water vapor is emitted that rapidly freezes and forms ice crystals. If the atmosphere is sufficiently moist and cold, these contrails expand and persist for many hours. This chapter describes the numerical modeling of contrails on a local scale with cloud resolving simulations. Emphasis is put on the description of microphysical modeling. With this methodological approach valuable information on contrail evolution for a multitude of atmospheric and aircraft parameters can be obtained.

33.1 Introduction

Cloud resolving simulations (CRS) help to better understand how *contrails* evolve in detail and how they are affected by aircraft properties and meteorological conditions. CRS modeling of contrails is needed not only for basic understanding of the processes in contrails and *contrail cirrus*. Simplified global contrail

S. Unterstrasser (✉) · I. Sölch · K. Gierens
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Simon.Unterstrasser@dlr.de

I. Sölch
e-mail: Ingo.Soelch@dlr.de

K. Gierens
e-mail: Klaus.Gierens@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_33,

prediction models covering many contrails (Schumann 2011) can be improved or tuned by using detailed CRS results as benchmarks. Results can be applied to the development of contrail parameterizations in large-scale models (Burkhardt and Kärcher 2011) which give global estimates of climate impact. The climate impact of contrails depends strongly on the variability of contrail optical properties, and statistical approaches (Kärcher et al. 2009a) benefit from CRS results. Finally, understanding is required when discussing so-called mitigation options, i.e., technical or operational means to reduce the *climate impact* of *aviation*. This may include means to avoid contrails or to make them optically thinner and their lifetime shorter. In this chapter, the temporal and spatial dynamics of contrails and the detailed modeling of contrails and contrail cirrus are described.

33.2 Temporal Evolution of Contrails

The temporal evolution of contrails can be divided into three phases in which different physical processes dominate the evolution and in which different phenomena occur. These phases are the *jet phase*, the *vortex phase*, and the *dispersion phase*; they are described in the following sections. Contrails form when the *Schmidt–Appleman criterion* is fulfilled, and they are long-living (persistent) in *ice supersaturated regions*. For the following discussion it is assumed that these conditions are met.

33.2.1 Contrail Formation in the Jet Phase

A jet engine under cruise conditions emits exhaust gases with high velocity of about 200 m/s relative to the ambient air in order to produce the required engine thrust as a consequence of momentum conservation. The exhaust jet has an initial diameter given by the exit diameter of the engine that is on the order of 1 m. The large axial velocity and the corresponding strong velocity gradient between the jet and ambient air generate highly turbulent flows which lead to entrainment of ambient air into the jet. This results in (1) an increase of jet diameter, (2) a decrease of jet speed, (3) mixing and dilution of exhaust gases with ambient air, and (4) cooling of the exhaust gases. The cooling and mixing lead to liquid saturation in case the Schmidt–Appleman criterion is met and consequently a contrail forms in the jet plume within a fraction of a second after the exhaust gases leave the engine. Figure 33.1 illustrates how rapidly exhaust contrails form behind an aircraft as a result of the strong jet dynamics (Schumann 1996).

Figure 33.1 shows a faint iridescent sheet of ice crystals between the thick exhaust contrails. These are formed by aerodynamic effects, namely the strong cooling of the air when it passes over the wings. When air is accelerated while flowing over the wing its pressure and temperature drop and substantial supersaturation can occur that allows liquid aerosol droplets in the air stream to freeze. These ice crystals form a so called aerodynamic contrail. While the formation

Fig. 33.1 Photograph showing a contrail produced from a Boeing B747 with four engines. Two contrail formation modes are visible: exhaust contrails becoming visible shortly behind the engines, and an aerodynamically produced contrail. The latter is the faint iridescent ice sheet in the middle between the thick exhaust contrails. Copyright Josef P. Willems

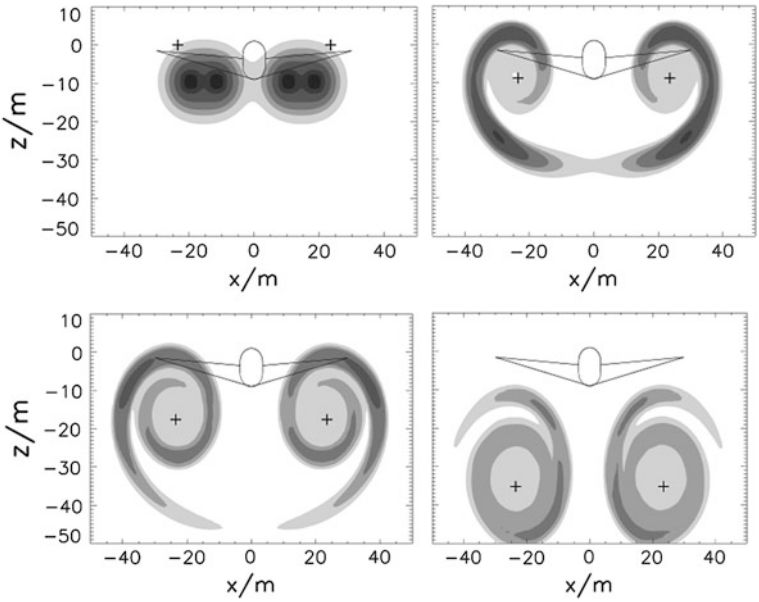

Fig. 33.2 Concentration of exhaust gases from a four-engine aircraft directly behind the aircraft (*top left*) and 5s (*top right*), 10s (*bottom left*) and 20s (*bottom right*) after the passage of the aircraft. This corresponds to distances of 1.2, 2.5 and 5 km behind the airplane. The positions of the downward sinking vortex centers are marked '+'. The cruise altitude corresponds to z = 0 m. The data are obtained from 3D simulations and are averaged along the flight direction

process of aerodynamic contrails has been described recently (Gierens et al. 2009; Kärcher et al. 2009b), the later evolution of such contrails has not yet been investigated, neither numerically nor observationally. *Aerodynamic contrails* are therefore not further discussed in this chapter.

Figure 33.1 also shows that the exhaust contrails, which initially mark the position and shape of the more or less straight jets, are quickly deflected and twisted. This results from the roll-up of a pair of counter-rotating vortices that are generated by the downwash behind an aircraft. The expanding exhaust jet interacts with the vortices and finally most of the exhaust material is trapped inside the vortex system, which inhibits further dilution of the exhaust plume. Wake vortices also pose a hazard to landing aircraft at congested airport runways and are an important research issue of aircraft safety.

Figure 33.2 shows the entrainment of the exhaust material (including the contrail) into the nascent vortex pair in a cross-sectional viewing plane as obtained from a three-dimensional simulation. The position of the aircraft is included for orientation, and the crosses mark the vortex centers. The outer jets, which are closer to the vortex centers, are entrained into the vortices first, and then the inner jets follow. About 20 s after engine exit the exhaust gases and particles and the contrail ice are trapped into a pair of vortex tubes. This marks the end of the jet phase and the beginning of the vortex phase, which is considered next.

33.2.2 Vertical Expansion During the Vortex Phase

The two vortices mutually induce a downward motion with an initial speed of around 1–2 m/s depending on size, weight, and speed of the aircraft. The *vortex system* descends a few hundred meters until it decays after a few minutes and also carries the trapped ice crystals downward. In a stably stratified environment (which is generally the case in flight altitudes) the descending vortex pair induces baroclinically-produced vorticity around the primary vortex. The *baroclinic vorticity* detrains ice crystals from the primary vortex and forms a curtain of ice, the so-called secondary wake, which can extend down from the flight altitude to the primary vortices. Under just-saturated conditions there appears a gap between the primary and secondary wakes, as shown in Fig. 33.3. The upper part is the faint *secondary wake*; the lower one is the *primary wake*, which gets weaker with time due to ice sublimation.

The descent of the vortex tubes leads to *adiabatic* heating, and consequently to an increase in saturation vapor pressure. The *relative humidity* with respect to the ice (RHi) decreases due to deposition of moisture on ice particles and later decreases in the vortex system because of adiabatic heating. The ice crystals start to sublimate as soon as the RHi decreases below 100 %. Ice crystals which are not sufficiently large can sublimate completely, thus ice crystals get lost and the primary wake becomes weaker. Sometimes it disappears completely and only the secondary wake survives, as in Fig. 33.3, because the latter is not subject to adiabatic heating. Analytical calculations show that the relative humidity inside the primary wake drops by 10 % for each 80 m of vertical displacement. The smaller the supersaturation initially, the more ice crystals are lost. In a weakly supersaturated situation more than 99 % of the initially formed ice crystals may get lost in the vortex phase, whereas under highly supersaturated conditions (say, 130 % RHi and higher) most crystals survive.

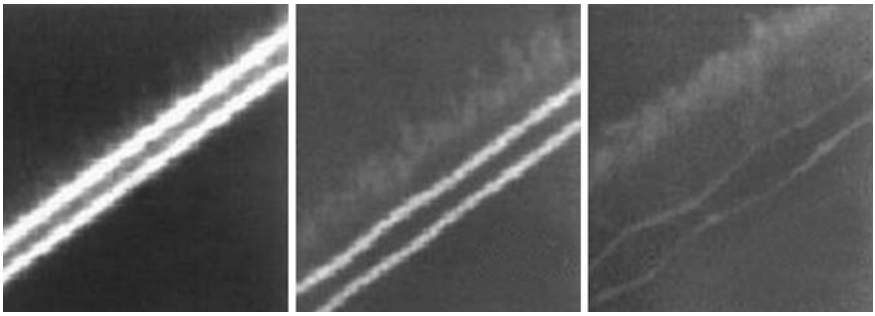

Fig. 33.3 Photographs of a contrail after 25, 61 and 103 s, adapted from Sussmann and Gierens (1999)

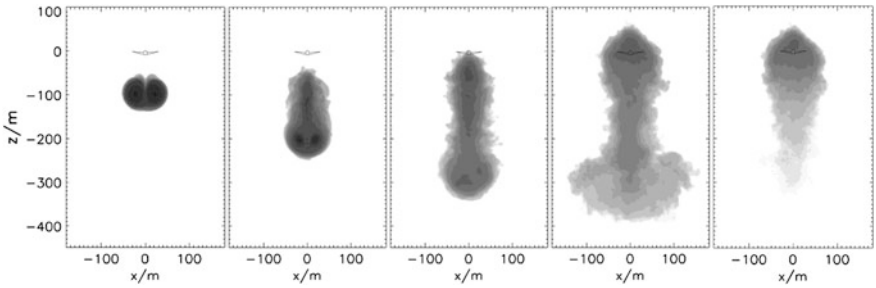

Fig. 33.4 Ice crystal concentration in a contrail during the vortex phase after 1, 2, 3 and 5 min for an ambient relative humidity RHi = 120 % and after 5 min for RHi = 110 % (from *left* to *right*). The *darker* the area, the higher is the concentration. The data are averages along the flight direction from 3D simulations

Besides relative humidity, the extent of crystal loss depends on further meteorological, microphysical and aircraft parameters. These affect either crystal growth and sublimation or the characteristics of wake vortex decay, which in turn determines the extent of adiabatic heating.

Ice crystal concentrations in a plane normal to the flight direction as obtained from three-dimensional simulations are shown in Fig. 33.4. For large airliners like Airbus A340 or Boeing B747 contrails can eventually reach a height of more than 400 m. The first four panels (from left to right) show a contrail evolution at an ambient relative humidity RHi = 120 %. Initially the highest concentrations can be found in the primary wake, and appear in the secondary wake at the end, as explained above. For comparison, the rightmost panel shows a case of RHi = 110 %. Very few crystals survive in the primary wake and the overall vertical extent is smaller, although the final vortex displacement is the same in both cases.

Vertical expansion of the contrail and crystal loss are the most significant processes during the vortex phase since, acting on a time scale of minutes, they affect the evolution and properties of the contrail in the later phase, i.e., the

dispersion phase where it can be transformed into contrail cirrus on time scales of hours.

33.2.3 Spreading During the Dispersion Phase

After the wake vortex system decays via a number of mechanisms, the surviving ice crystals in the primary wake are no longer trapped and get mixed with the ice in the secondary wake. They now evolve further under the influence of local meteorological conditions. The initial conditions for contrail dispersion vary widely because of the vortex phase. Horizontal spreading sets in when the horizontal wind varies along the contrail height (called vertical wind shear). Shear-driven spreading together with shear-enhanced turbulent mixing causes supersaturated ambient air to get mixed into the contrail, allowing the ice crystals to grow, in particular at the outer regions of the contrail. In contrast, ice crystals in the inner regions of the contrail grow little, and the relative humidity is close to 100 %. Another important process for maintaining supersaturation even in the contrail's inner parts is updraft. If the aircraft has produced the contrail in a region where the air is lifted, the resulting adiabatic cooling decreases the saturation vapor pressure, which is a source of condensable vapor. Such persistent contrails always contain much more (orders of magnitude) ice mass than was initially emitted by the aircraft. Once a spreading contrail loses its linear shape or overlaps with naturally formed *cirrus* clouds and, hence, becomes hardly distinguishable from naturally formed cirrus, it is termed contrail cirrus.

Horizontally spreading contrails at low ambient supersaturation tend to become optically thinner, that is, less visible. Simultaneous crystal growth via mixing and updraft increase the optical thickness, so the contrail can remain visible for hours. Minnis et al. (1998) was able to follow an intentionally produced contrail for more than 17 h in satellite images.

Crystal growth leads to sedimentation, that is, falling ice crystals. As explained above, the crystals grow faster in the outer than in the inner zones. Thus, the inner crystals fall very slowly and remain essentially at their altitude, while the outer crystals start to fall earlier and they may even grow further while falling through supersaturated layers, which can accelerate their falling. Thus, contrails can produce fall streaks, as shown in Fig. 33.5. Wind shear below the contrail then leads to curved (comma-like) fall streaks. Once the ice crystals reach subsaturated air on their way down, they start to sublimate. Sedimentation steadily removes ice mass from the contrail. As long as crystal growth due to mixing and updraft balance this loss, the contrail remains visible. However once sedimentation loss can no longer be balanced (for instance when all contrail crystals are eventually large enough to fall quickly) the contrail starts to dissipate. The contrail will dissipate as well once the meteorological conditions enforce a decrease of ambient relative humidity below ice saturation, for instance when the updraft is halted and turned into subsidence.

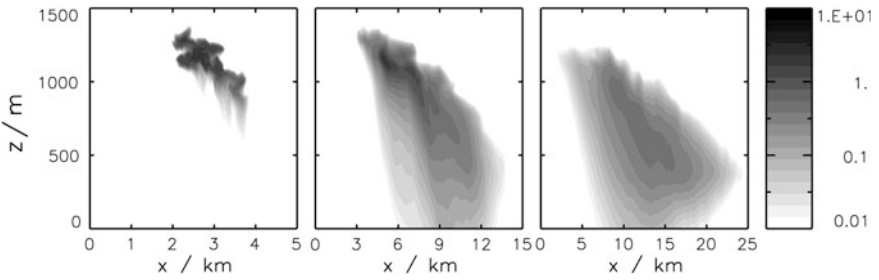

Fig. 33.5 Ice mass concentration IWC (in mg/m^3) from a contrail simulation with temperature 217 K at cruise altitude, relative humidity 110 % and vertical wind shear 2 (m/s)/km. The contrail age is 45 min, 2.5 and 5 h (from *left* to *right*). The displayed plane is normal to the flight direction and the cruise altitude is at the model level z = 1 300 m. Note the different scales on the horizontal axis

Fig. 33.6 Photograph taken at Oberpfaffenhofen of spreading and overlapping contrails. Copyright S. Unterstraßer

Another process that can be important during the dispersion phase is radiation. If the contrail is sufficiently optically thick the *absorption* and emission of radiation can lead to differential heating and cooling, causing buoyancy and thus triggering vertical motion. This in turn affects crystal growth rates again via adiabatic heating and cooling and the corresponding change in saturation vapor pressure (Jensen et al. 1998; Unterstrasser and Gierens 2010a).

Contrail cirrus and natural cirrus differ in some aspects although the physical processes acting on them are the same. Contrail formation is temperature-controlled and can happen at any relative humidity. The formation of natural cirrus by homogeneous nucleation requires a relative humidity above 145 %. Thus, contrails can form and persist in otherwise cloud-free areas (see Fig. 33.6), and increase the fractional coverage of upper tropospheric ice clouds, which is an important aspect of their relevance to *climate change*.

33.3 Modeling Contrails and Contrail Cirrus

33.3.1 Why Do We Need Models?

Linear contrails can be detected with satellites (Mannstein et al. 1999) and thus global properties like mean coverage can be estimated. Aged contrails lose their characteristic linear shape due to vertical wind shear, sedimentation, turbulence and spatial variations in meteorological variables. Such contrails can hardly be distinguished from natural cirrus. The problem of distinction becomes even worse when contrails overlap and interact and form a contrail cluster, which is usual in regions with dense air traffic. Thus it is currently difficult to uniquely determine properties of individual contrail cirrus using measurements. The life cycle and properties of contrail cirrus are not well known and cloud resolving modeling is a promising method to gain information on contrail evolution during a complete life cycle.

The numerical model used at IPA consists of three coupled parts that simulate the dynamic, microphysical and radiative processes of contrails. The description here will focus on *microphysics*.

33.3.2 Dynamics

Atmospheric air flows are described by the *Navier–Stokes equations*, which are valid for many fluid flows. These equations or suitable approximations describe the time evolution of velocity, pressure, density and temperature of the atmosphere and are often solved by a *large-eddy simulation* (LES) approach. For the simulations presented here the EULAG model is employed (Smolarkiewicz and Margolin 1997, 1998). In this chapter the term cloud resolving simulation (CRS) is preferred over large eddy simulation (LES), since LES is a strictly three-dimensional approach and several results presented here were gained with 2D variants of the EULAG model.

33.3.3 Microphysics

33.3.3.1 Microphysical Processes

The microphysical processes involved in the presented simulations are listed below. So far, the jet phase is usually skipped (at IPA) to start directly with the vortex phase. Formation of *ice crystals* by *homogeneous nucleation* or *heterogeneous nucleation* has then already occurred and these processes are not further discussed here.

- *Deposition*: Growth of ice crystals by deposition of water molecules from the vapor phase onto crystal surfaces. This occurs when the surrounding air is ice-supersaturated. The number of ice crystals is not affected by deposition.
- *Sublimation*: The process opposite to deposition. If the surrounding air is ice-subsaturated the ice crystals lose mass. The number of ice crystals is reduced once crystals sublimate completely.

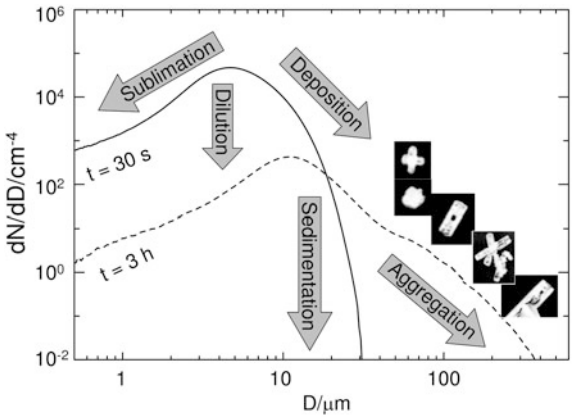

Fig. 33.7 Size distribution (diameters) of ice crystals in a contrail 30 s after formation (*solid line*) and 3 h later, after spreading in ice supersaturated air. Shown are results from a LES simulation. The *arrows* indicate various processes acting on the ice crystals explained in the text. Examples of the variety of shapes and habits of ice crystals common in contrails are shown for illustration only. These are high resolution digital images of crystals collected on a research flight

- *Sedimentation*: In Earth's gravity field ice crystals settle towards the ground. The settling speed depends on ice crystal mass and shape. The shape affects the drag of the ambient air on the falling crystal. Gravity force and drag force are equal once the "terminal" velocity is reached. Terminal velocities range from a few cm/s for small (<10 μm) up to 2 m/s for large (>1 mm) crystals.
- *Aggregation*: Crystals falling with different terminal velocities may collide and stick together. This process leads to large crystals aggregates with complex and variable shapes.
- *Radiation*: Radiation can be absorbed and emitted as well as scattered by ice crystals. Absorption and emission and the corresponding heating and cooling can trigger *convection*.

The effects of these processes (except radiation) can well be illustrated by considering a size spectrum of contrail ice crystals. Figure 33.7 shows how it is altered with time due to these processes. A size spectrum (or *size distribution*) shows how many ice crystals of all sizes are contained in a cloud or a part of it. The crystal size spectrum of a fresh contrail with an age of about 30 s is shown by a solid line in Fig. 33.7 together with a size spectrum of crystals commonly observed in aged contrails (dashed line). The latter may have an age of about 3 h. The specific effects of the microphysical processes listed above are indicated with arrows. The horizontal spreading of contrails leads to *dilution*, that is, to lower total crystal concentrations (crystals of all sizes per unit volume), which affects the spectrum in the indicated way.

33.3.3.2 Levels of Complexity

The microphysical processes can be represented in numerical models in various ways and with different levels of complexity, which may lead to different degrees of details to be simulated. The choice of model complexity depends on available computing resources and on the specific application. Detailed simulations using models with higher complexity are particularly useful if the model input is of high quality and the physical parameters within the model are known with good accuracy.

Generally there are three classes of microphysics models: *bulk models*, *spectral models* and models using a Lagrangian approach. Bulk models represent only gross quantities of the ice population, for instance, total ice mass concentration (single moment models) or both total ice mass and number concentrations (two-moment models). The most simplified bulk models assume a monodisperse crystal size distribution, that is, all crystals within a grid cell are assumed to have the same size. The only advantage of such an assumption is that its numerics can be formulated very easily. The improved way is to implicitly assume a certain type of size distribution, e.g., log-normal or gamma distribution (Spichtinger and Gierens 2009). Such analytical types are close to observed size distributions. The advantage over the simpler models is the addition of a width of the size spectra, that is, the simultaneous existence of small and large ice crystals. Unfortunately, though, there are no prognostic equations for moments of the crystal mass spectrum higher than two (budget equations can only be formulated for the total number, or zeroth moment, and total mass, first moment of the mass spectrum). Thus the temporal evolution of the distribution width has to be connected to the evolution of total mass and number in an ad hoc way. This difficulty shows that it is not desirable to use too elaborate distribution types with more than two parameters. The weak point of bulk models is that the initially chosen distribution type is not allowed to change during cloud evolution. Although the microphysical processes act as indicated in Fig. 33.7, the size spectrum is forced into the predefined shape at the end of each time step. This causes artifacts that sometimes are tolerable, sometimes not. Nevertheless, bulk models are useful for implementation in large-scale models because they are computationally efficient, with only a few additional prognostic equations that have to be solved.

Spectral microphysics codes discretize the overall crystal size range into a number of small subranges (bins) and compute the ice number concentration and the process rates for each bin individually. This approach does not enforce certain shapes of the size spectra, which can therefore freely evolve as in Fig. 33.7; however, they may suffer from numerical diffusion in radius space. The artifacts mentioned above are thus avoided. However, spectral models need much more computing time than bulk models and are therefore not always applicable.

The highest amount of detail can be simulated when single ice crystals are tracked by the model in the sense of a Lagrangian approach. Of course this is not possible for the actual number of crystals in a cloud. But a compromise is possible by tracking only a set of representative ice crystals, called simulation ice particles (SIP), where each SIP represents 1 000–1 000 000 actual ice crystals. Thus it is

possible to deduce the properties of the contrail or contrail cirrus from the evolution of the control sample of SIPs in a statistical sense. The model system EULAG-Lagrangian Cirrus Model (LCM) was developed recently at IPA (Sölch and Kärcher 2010). Gas and aerosol are treated over a fixed Eulerian grid similar to the dynamic and thermodynamic variables, while the ice phase is treated by tracking a large number of SIPs. Typically, more than one hundred million SIPs can be tracked in a simulation, which guarantees statistically meaningful results. This approach became feasible only recently with the advent of supercomputers with sufficient memory and processor performance. The advantages for understanding detailed contrail microphysical properties justify the high numerical effort. First, the crystal size distribution can evolve freely as in the spectral models. Second, individual processes acting on individual crystals can be easily simulated. This even allows investigation of rare events due to statistical fluctuations as they appear in real clouds; no other approach is able to do this. In principle position, *habit*, size, and chemical composition of single crystals in the contrail can be stored and time series can be produced for individual properties to extract the influence of different processes on particle properties in different regions of the ice cloud. Thus with this method not only contrail evolution can be followed, by and large, but also singular events that may happen in a real cloud yet which cannot be simulated with simpler approaches can be looked at.

33.3.4 Radiative Transfer

Radiative transfer codes are employed to calculate the emission/absorption/*scattering* of thermal and *solar radiation* by atmospheric constituents (greenhouse gases, aerosol, water droplets, and ice crystals). Radiation fluxes entering the contrail simulation domain are predetermined. Inside the domain, radiation-induced heating/cooling rates are computed. Growth or sublimation of ice crystals may be modified by radiative heating or cooling of the crystal's surface. Heat exchange between ice crystals and the surrounding air occurs and may induce convective up/downdrafts by differential heating/cooling of the air. The air *heating rates* enter the temperature equation from the dynamics part as a diabatic forcing term.

33.4 Simulations of Contrails and Contrail Cirrus

The EULAG-LCM code described above can be used to simulate ice clouds in general. In the following only contrail simulations are considered. Numerical simulations are the best way to explore the relevant parameter space controlling contrail evolution during the various regimes and to find the most significant parameters. CRS and LES have been employed at several research groups worldwide to study contrails (Jensen et al. 1998; Lewellen and Lewellen 2001; Paoli et al. 2004; Huebsch and Lewellen 2006). The presentation here will focus on recent results gained at IPA.

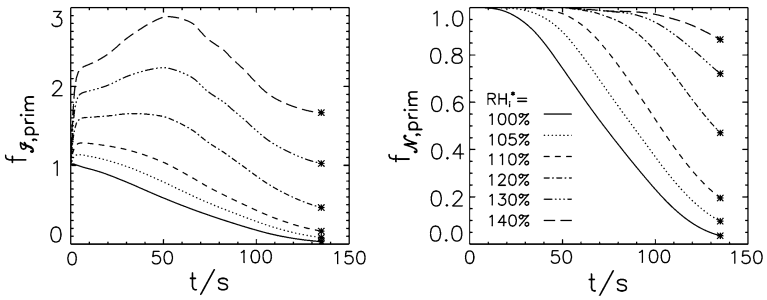

Fig. 33.8 Temporal evolution of ice crystal mass $f_{I,\,prim}$ (*left*) and number $f_{N,\,prim}$ (*right*) concentrations in the primary wake, normalized by their initial values, during the vortex phase. Results are shown for various relative humidity values (100–140 %). The temperature at cruise level was 217 K

33.4.1 Vortex Phase

Two contrail characteristics important for its later evolution into contrail cirrus are determined during the vortex phase, namely its vertical extension (given by the descending distance of the vortex pair) and the number of ice crystals surviving adiabatic heating. These characteristics depend on meteorological and aircraft parameters, and in particular the fraction of surviving ice crystals varies between all and nothing for the same aircraft under different meteorological conditions.

The simulations are initialized with a rolled-up vortex pair, namely with two counter-rotating vortices defined by prescribed initial circulation. The ice crystals initially contain the *water vapor* emitted by the engines, which is an uncritical assumption since they quickly react to the humidity field in the vortex. The number of ice crystals per flight distance is chosen as well; it corresponds to the *emission index* of *soot* particles, which act as *ice nuclei*. The spatial distribution of the ice crystals is uniform within two circles around the vortex cores, with a radius of 20 m representing the approximate plume dimension. Numerical tests have shown that spatially uniform distribution is an uncritical assumption as well.

Figure 33.8 is an example of simulation results which shows what fractions of ice mass and ice number concentration relative to their initial values are present at various times during the vortex phase. Only the values for the primary wake are shown. Simulations have been run for a range of ice supersaturation from zero (saturation) to 40 %. Temperature at cruise altitude was 217 K for the cases shown. During the first 20–40 s the ice mass increases by depositional crystal growth if there is supersaturation. This proceeds until supersaturation is completely consumed and saturation is reached in the plume. This phase can be called the deposition phase. It is short due to the high crystal number concentration and supersaturation is quickly consumed. Subsequently, the ice mass starts to decrease because the adiabatic temperature increase in the descending vortex leads to subsaturated conditions. This may be called the sublimation

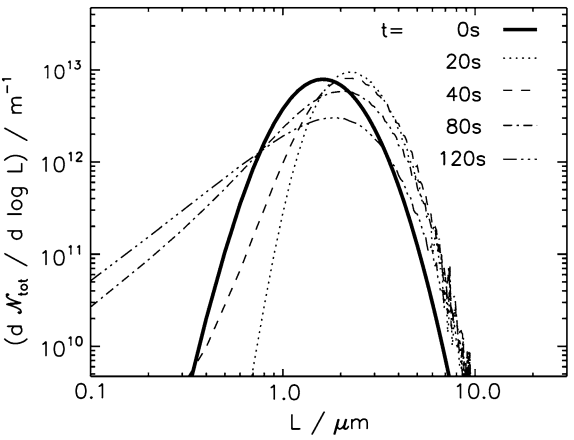

Fig. 33.9 Size distribution of ice crystals in the primary wake for T = 217 K and RHi = 120 % at various times (see legend). The size distribution is integrated over the plane perpendicular to the flight direction

phase. The sublimation rate decreases after about 100 s when the vortex pair's descent gets slower. Contrary to the mass, the normalized crystal number cannot increase and always stays below unity since crystal formation is already completed in the jet phase and no new crystal nucleation occurs during the vortex phase. Only few crystals are lost during the deposition phase, and when the sublimation phase starts only the smallest crystals sublimate completely; crystal loss is still moderate. The crystal loss rate becomes more substantial only later when initially larger crystals also become sufficiently small for complete sublimation.

Figure 33.9 shows the evolution of crystal size distribution with time for RHi = 120 %. The bold curve represents the initial distribution. In the first 20 s the ice crystals grow in the supersaturated environment and the size distribution is shifted to larger crystal dimensions. Sublimation begins after about a minute, the number concentration of small particles increases, and after about 80 s a sublimation tail appears. Crystal loss appears as a downward shift of the curve. A bulk model which was used in an earlier study (Unterstrasser et al. 2008) could not reproduce the formation of a sublimation tail and the results obtained with EULAG-LCM are more plausible in several respects (Unterstrasser and Sölch 2010).

Further simulations showed a strong sensitivity to temperature. In a warmer environment, microphysics proceeds faster and thus the sublimation is stronger and fewer ice crystals survive. Also, the contrail's vertical extent can be smaller, as shown in Fig. 33.4. Further parameters which affect contrail crystal loss and vertical extent are rate of vapor emission (which is proportional to fuel flow rate), rate of soot particle emission (these are needed for contrail formation because they serve as the condensation nuclei), initial shape of crystal size distribution (in particular its width), aircraft type (span, weight, speed, number of engines), *turbulence* level and *stratification* of the ambient atmosphere. Some of these parameters affect only microphysics, the latter two the vortex dynamics. Evidently,

with such a large number of influencing factors, contrail properties can vary widely at the end of the vortex and the beginning of the dispersion phase (Unterstrasser et al. 2008; Unterstrasser and Sölch 2010).

33.4.2 Dispersion Phase

Dispersion phase simulations are best initialized with the final results of vortex phase simulations. Then it is possible to investigate how the transition of a linear contrail into contrail cirrus is affected by various factors that characterize the meteorological situation, especially humidity, temperature, vertical wind shear, stratification, updraft velocity, and radiation. An example of a simulation of a peculiar case is shown in Fig. 33.10. The contrails were measured on 7 November 1990 over southern Germany by a lidar on board the DLR research aircraft Falcon (left panel). The contrails of unknown age extend over 1.5 km vertically and have strong fall streaks. The model is able to reproduce such strong contrails when a relative humidity of 140 % is prescribed. Unfortunately, the humidity data from a nearby *radiosonde* ascent had to be ignored since it reported unrealistically low values (RHi < 20 %). Moreover, the provided data resolution of the radiosonde wind profile turned out to be too coarse to reproduce the peculiar U-shape of the fall streaks and the profile had to be fitted accordingly. With these parameter settings the simulated contrail (right panels) resembles the lidar measurements: strong extinction of radiation in the upper part, less extinction in the fall streak. The overall shape resembles the measurement as well.

The climate impact of a contrail depends mainly on its optical properties and its horizontal extent. IPA simulations have proven relative humidity to be the dominant parameter for climate impact together with wind shear. However, the strong variability at the end of the vortex phase and the large number of meteorological parameters that affect the development, which themselves vary with time, show that the climate impact of single contrails can be quite variable. Results for a constrained number of simulations or measurements cannot be taken as representative for the whole globe. Further results and explanations can be found in Unterstrasser and Gierens (2010a, b).

33.5 Outlook

As shown, cloud resolving simulations of contrails and contrail cirrus can be used to investigate many issues that cannot readily be observed and measured. In particular, it is possible to study which ambient and aircraft/fuel parameters have the strongest influence on contrail development in the three phases of the evolution. However, CRS/LES models must be further developed because academic situations are still being simulated. For principle understanding this is reasonable, but then reality must be added to the simulations. Two ongoing developments at

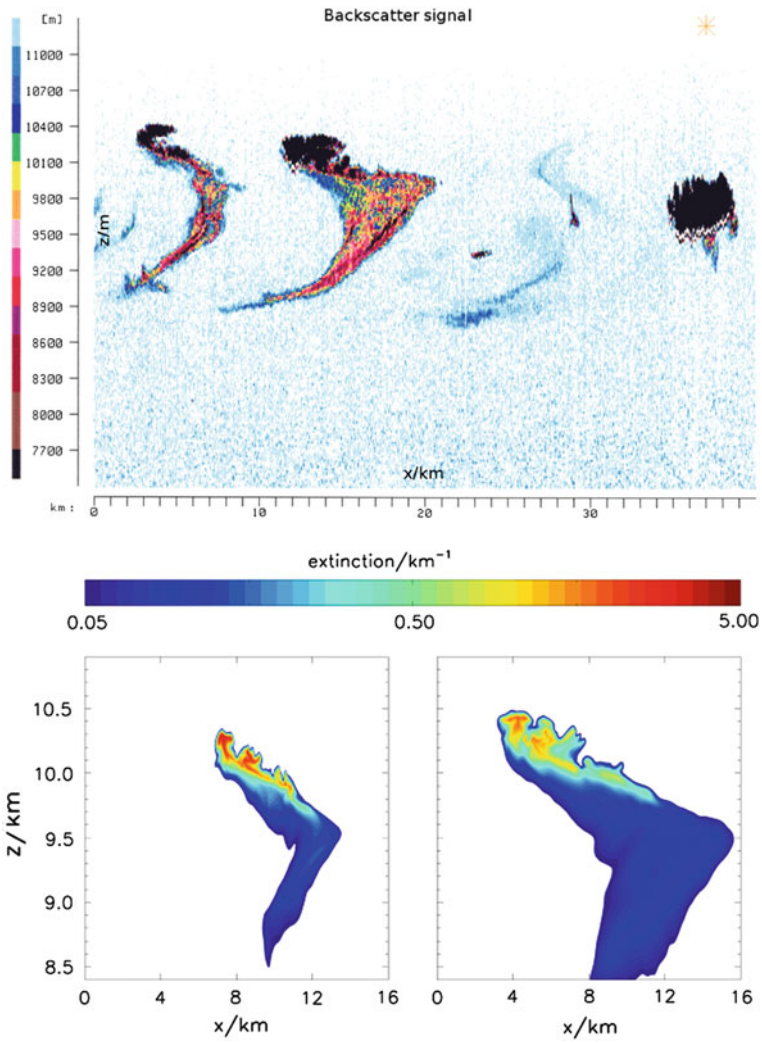

Fig. 33.10 The *top* image shows an IPA-lidar measurement of several aged contrails as observed with the DLR Falcon in 1991 (Schumann 1994). Extinction fields (*bottom*) after 2 and 3 h taken from simulations with similar meteorological conditions are shown on the *bottom*. The value of the outer *contour line* is similar to the lidar detection threshold in the *top panel*

IPA are the simulation of contrail clusters and an improved representation of synoptic situations. In regions like Europe where air traffic is dense, contrails mostly occur in large groups and they start sooner or later to interact with each other. Contrail-to-cirrus transition generally takes some hours and meteorological conditions change over such a long time. These effects need to be simulated more realistically. Contrails often appear in regions where cirrus clouds are already

present. Thus the formation and evolution of contrails within cirrus clouds and the interaction between a contrail and the surrounding cloud is a topic for further research.

References

Burkhardt, U., Kärcher, B.: Global radiative forcing from contrail cirrus. Nature Clim. Change **1**, 54–58 (2011). doi:10.1038/NCLIMATE1068

Gierens, K., Kärcher, B., Mannstein, H., Mayer, B.: Aerodynamic contrails: phenomenology and flow physics. J. Atmos. Sci. **66**, 217–226 (2009). doi:10.1175/2008JAS2767.1

Huebsch, W.W., Lewellen, D.C.: Sensitivity study on contrail evolution. 36th AIAA Fluid Dynamics Conference and Exhibit, AIAA 2006-3749, vol. 14, 2006

Jensen, E.J., Ackermann, A.S., Stevens, D.E., Toon, O.B., Minnis, P.: Spreading and growth of contrails in a sheared environment. J. Geophys. Res. **103**(13), 557–513, 567 (1998). doi:10.1029/98JD02594

Kärcher, B., Burkhardt, U., Unterstrasser, S., Minnis, P.: Factors controlling contrail cirrus optical depth. Atmos. Chem. Phys. **9**, 6229–6254 (2009a). doi:SRef-ID:1680-7324/acp/2009-9-6229

Kärcher, B., Mayer, B., Gierens, K., Burkhardt, U., Mannstein, H., Chatterjee, R.: Aerodynamic contrails: microphysics and optical properties. J. Atmos. Sci. **66**, 227–243 (2009b). doi:10.1175/2008JAS2768.1

Lewellen, D.C., Lewellen, W.S.: The effects of aircraft wake dynamics on contrail development. J. Atmos. Sci. **58**, 390–406 (2001)

Mannstein, H., Meyer, R., Wendling, P.: Operational detection of contrails from NOAA-AVHRR data. Int. J. Rem. Sens. **20**(8), 1641–1660 (1999)

Minnis, P., Young, D.F., Garber, D.P., Nguyen, L., Smith Jr, W.L., Palikonda, R.: Transformation of contrails into cirrus during SUCCESS. Geophys. Res. Lett. **25**, 1157–1160 (1998). doi:10.1029/97GL03314

Paoli, R., Hélie, J., Poinsot, T.: Contrail formation in aircraft wakes. J. Fluid Mech. **502**, 361–373 (2004). doi:10.1017/S0022112003007808

Schumann, U.: On the effect of emissions from aircraft engines on the state of the atmosphere. Ann. Geophys. **12**, 365–384 (1994)

Schumann, U.: On conditions for contrail formation from aircraft exhausts. Meteor. Z. **5**, 4–23 (1996)

Schumann, U.: A contrail cirrus prediction model. Geosci. Model Dev. **5**, 543–580 (2012). doi:10.5194/gmd-5-543-2012

Smolarkiewicz, P., Margolin, L.: On forward-in-time differencing for fluids: an Eulerian/semi-Lagrangian non-hydrostatic model for stratified flows. In: Lin, C.A., Laprise, R., Ritchie, H. (eds.) Numerical methods in atmospheric and oceanic modelling: the Andre J. Robert memorial volume, pp. 127–152. Canadian Meteorological and Oceanographical Society, Ottawa (1997)

Smolarkiewicz, P., Margolin, L.: MPDATA: a finite-difference solver for geophysical flows. J. Comp. Phys. **140**, 459–480 (1998)

Sölch, I., Kärcher, B.: A large-eddy model for cirrus clouds with explicit aerosol and ice microphysics and Lagrangian ice particle tracking. Q. J. Royal Meteorol. Soc. **136B**, 2074–2093 (2010). doi:10.1002/qj.689

Spichtinger, P., Gierens, K.M.: Modelling of cirrus clouds—Part 1a: Model description and validation. Atmos. Chem. Phys. **9**, 685–706 (2009)

Sussmann, R., Gierens, K.: Lidar and numerical studies on the different evolution of vortex pair and secondary wake in young contrails. J. Geophys. Res. **104**, 2131–2142 (1999)

Unterstrasser, S., Gierens, K., Spichtinger, P.: The evolution of contrail microphysics in the vortex phase. Meteorol. Z. **17**, 145–156 (2008). doi:10.1127/0941-2948/2008/0273

Unterstrasser, S., Gierens, K.: Numerical simulations of contrail-to-cirrus transition—Part 2: Impact of initial ice crystal number, radiation, stratification, secondary nucleation and layer depth. Atmos. Chem. Phys. **10**, 2037–2051 (2010a). doi:10.5194/acp-10-2037-2010

Unterstrasser, S., Gierens, K.: Numerical simulations of contrail-to-cirrus transition—Part 1: An extensive parametric study. Atmos. Chem. Phys. **10**, 2017–2036 (2010b). doi:10.5194/acp-10-2017-2010

Unterstrasser, S., Sölch, I.: Study of contrail microphysics in the vortex phase with a Lagrangian particle tracking model. Atmos. Chem. Phys. **10**, 10003–10015 (2010). doi:10.5194/acp-10-10003-2010

Global Atmospheric Aerosol Modeling

34

Johannes Hendricks, Mattia Righi and Valentina Aquila

Abstract

Global aerosol models are used to study the distribution and properties of atmospheric aerosol particles as well as their effects on clouds, atmospheric chemistry, radiation, and climate. The present article provides an overview of the basic concepts of global atmospheric aerosol modeling and shows some examples from a global aerosol simulation. Particular emphasis is placed on the simulation of aerosol particles and their effects within global climate models.

34.1 Introduction

The atmospheric aerosol is defined as the population of small, levitating liquid or solid particles dispersed in the air. The sizes of these particles typically range from a few nanometers to about ten micrometers. The major aerosol constituents are mineral dust, sea salt, black and organic carbon, sulfate, nitrate, ammonium, and

J. Hendricks (✉) · M. Righi
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Johannes.Hendricks@dlr.de

M. Righi
e-mail: Mattia.Righi@dlr.de

V. Aquila
NASA Goddard Space Flight Center, Mail Code: 614,
Greenbelt, MD 20771, USA
e-mail: valentina.aquila@nasa.gov

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_34,

water. Atmospheric aerosol particles result from natural sources such as sea spray, the wind-driven release of desert dust, or volcanoes, as well as anthropogenic sources such as the combustion of fuels. Furthermore, aerosol particles are released by biomass burning, that is, forest and grassland fires, which can be caused by lightning or human activity such as fire clearance. Aerosol particles are either directly emitted into the atmosphere or formed indirectly by the conversion of emitted precursor gases.

Increases in the amount of atmospheric aerosol resulting from anthropogenic emissions can strongly affect air quality, with adverse effects on human health. Anthropogenic aerosol particles are also important agents in climate change since they scatter and absorb solar and terrestrial radiation. They can further affect climate indirectly via the interaction with clouds and the chemistry of some greenhouse gases, particularly ozone.

Many activities in the field of atmospheric research focus on aerosol particles and their roles in air pollution and climate change. In addition to aerosol measurements in the atmosphere or laboratories, numerical models are applied to simulate the chemical and microphysical properties of atmospheric aerosol particles as well as their spatial and temporal distributions. They are used to analyze aerosol budgets and life cycles and to study specific aerosol-related effects such as aerosol-induced changes in clouds, radiation, and atmospheric chemistry, as well as the resulting climate impacts. Also the roles of specific aerosol types or individual aerosol sources in the climate system can be studied. Models allow simulations of the impact of anthropogenic aerosol on air quality and climate not only for past and present-day conditions but also for future scenarios, enabling studies of the efficiency of mitigation strategies to reduce future anthropogenic aerosol emissions and the related atmospheric impact. Models can provide consistent, temporally and spatially resolved information about atmospheric aerosol properties and aerosol-related effects in an entirety that, in many cases, is not accessible by observations. Measurements, however, cannot be replaced by simulations. They are essential to provide information from the real atmosphere which is crucial to evaluate the quality of models and to define model parameters that cannot be calculated solely from theory.

Atmospheric aerosol models are available for different spatial scales. Global models resolve aerosol properties on large scales and typically have spatial resolutions between about ten and several 100 kilometers. Low resolutions are typically applied in global models that are used to simulate the climate impacts of aerosols, since such simulations can be very expensive and low resolutions reduce the computational costs. With the resolutions typical for global models only large-scale mean aerosol properties can be simulated, and local phenomena, for example, high particle concentrations occurring close to highways or within the plumes of large ocean going ships, cannot be resolved. Higher resolutions are achieved, for instance, with regional, urban-scale, and street canyon models, which are major tools in air pollution research. Another type of frequently applied aerosol models are aerosol box models which simulate aerosol transformations within an

individual air parcel. Derivatives of such models can be used, for instance, to study the aerosol within expanding exhaust plumes.

Many recent studies focus on modeling the global atmospheric aerosol, its anthropogenic perturbations, and the resulting effects on climate. At the DLR Institute of Atmospheric Physics many efforts have been directed at the development and application of atmospheric aerosol models and their integration in global climate models (Lauer et al. 2005, 2007; Aquila et al. 2011). Besides general characterization of the global atmospheric aerosol (Lauer and Hendricks 2006), a particular focus of the institute has been modeling global aerosol perturbations induced by emissions of the transport sector (aviation, land-based transport, and shipping) as well as simulating the resulting changes in clouds, radiation, and climate (Köhler et al. 2001; Hendricks et al. 2004, 2005, 2011; Lauer et al. 2007; Righi et al. 2011).

The present chapter introduces technical aspects of global aerosol models (Sects. 34.2–34.4) and discusses an example of a global aerosol simulation performed at the Institute of Atmospheric Physics (Sect. 34.5). It further suggests possibilities for future model improvements (Sect. 34.6). Particular emphasis is placed on the simulation of aerosol particles and their atmospheric effects within global climate models. The major focus is modeling tropospheric aerosol since aerosol particles in the troposphere are particularly relevant for air pollution and climate. The intention of the chapter is to provide a brief overview of global aerosol modeling concepts. For more details on the subject, the reader is referred to Ghan and Schwartz (2007) and references therein.

34.2 Representation of Aerosol in Global Models

34.2.1 Aerosol Types and Size Distribution

Global atmospheric aerosol modeling requirements are governed by several aspects. The models should provide information about the atmospheric mass concentration (mass per volume of air) of different aerosol constituents as well as the aerosol particle number concentration (number of particles per volume of air). Information on the particle size distribution is also important, since many processes controlling aerosol life cycles and climate impacts are size dependent. The aerosol constituents can show different degrees of mixing. For instance, black carbon particles can occur as pure particles or they may have acquired a coating composed of other constituents. The degree of mixing can affect the atmospheric impact of the particles, particularly their effects on clouds and radiation. Therefore, it is also desirable to resolve different mixing states in different size ranges. Global aerosol modeling also requires a proper description of all processes that control the formation and life cycles of aerosol particles. If aerosol effects on climate are to be quantified, effects of aerosol particles on clouds, radiation, and atmospheric chemistry have to be considered as well. In addition, for climate applications the

model needs to be capable of simulating climatically relevant time periods, lasting from a few decades to centuries or even longer. A common way to analyze the effects of different aerosol types (e.g., anthropogenic aerosol) on air quality and climate is to conduct sets of different simulations including a reference simulation and, for comparison, model runs where specific aerosol types or processes are neglected. Hence, a possibly large number of model runs is necessary. These requirements imply the need for huge computing resources.

Long-term global simulations with comprehensive aerosol-climate models mostly require the use of very expensive supercomputing facilities. In order to reduce the computational burden, simplifications in the model representation of aerosol particles are necessary. The first generation of global aerosol models considered only the mass concentration of some aerosol constituents. With increasing computer power, more comprehensive schemes have been implemented. Many current global aerosol models also enable the simulation of particle number concentration and size distribution, in addition to particle composition. State-of-the-art size resolving box or process models consider large numbers of individual size bins (bin schemes) or even track the size and composition of individual particles (particle-resolved models). Application of such approaches in global models, however, would require very large computing resources, particularly when simulations on climatically relevant time scales are performed. Hence, large-scale aerosol models, particularly those embedded in climate and Earth system models, apply simplified size distribution representations.

As a common simplification, the aerosol size distribution is approximated by a superposition of several so-called size modes. In many global models, each mode is represented by a log-normal distribution, that is, a bell-shaped function in a coordinate system with a logarithmic abscissa:

$$n(\ln D) = \frac{dN}{d\ln D} = \frac{N_t}{\sqrt{2\pi}\ln\sigma}\exp\left[-\frac{(\ln D - \ln D_{med})^2}{2\ln^2\sigma}\right]. \qquad (34.1)$$

Here, N_t is the total particle number concentration of the mode, D is the particle diameter, σ is the geometric standard deviation describing the width of the mode, and D_{med} is the median diameter, which for a log-normal distribution coincides with the geometric mean diameter (whose logarithm corresponds to the arithmetic mean of ln D) and the diameter where the function shows its maximum value. The log-normal function in the form of Eq. 34.1 represents a number size distribution $n(\ln D)$ which expresses the differential number concentration dN of particles present within a differential size range corresponding to the coordinate interval between ln D and ln $D + d \ln D$. Comprehensive numerical schemes to simulate aerosol microphysical processes with the log-normal approach have been developed, for instance by Whitby and McMurry (1997).

The applicability of the modal approach and the log-normal approximation to atmospheric aerosol has been demonstrated by Whitby (1978), who showed that the size distribution of tropospheric aerosol particles can be approximated by a

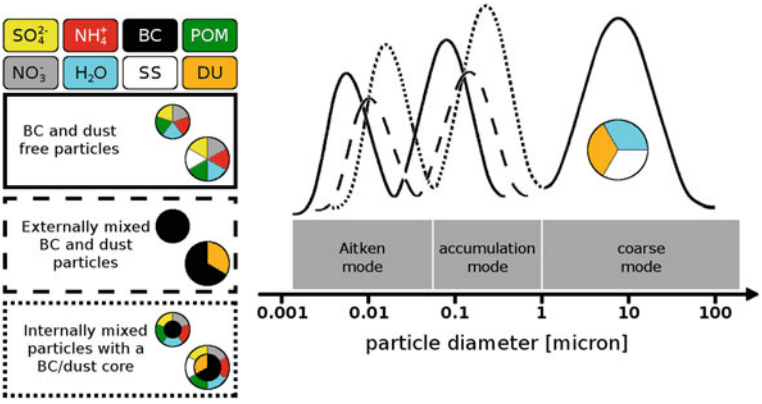

Fig. 34.1 Schematic overview of the aerosol modes considered in the EMAC/MADE-IN global aerosol model (Aquila et al. 2011). The model includes the aerosol constituents sulfate (SO_4^{2-}), nitrate (NO_3^-), ammonium (NH_4^+), water, black carbon (BC), particulate organic matter (POM), sea salt (SS), and mineral dust (DU). Different particle types are distinguished by three submodes in the Aitken and accumulation mode size ranges. The composition and size distributions of these submodes are indicated by the three different line styles of the box frames (*left panel*) and size distribution functions (*right panel*). See text for more details

superposition of three log-normal size modes covering the particle size ranges of a few nanometers (nm) to about 100 nm (Aitken mode), about 100 nm to 1 μm (accumulation mode), and about 1 μm to several 10 μm (coarse mode). Most current global aerosol models that apply the log-normal approach consider more than three modes. According to measured size distributions, some models include an additional mode, called the nucleation mode, to separate very fine particles in the size range of a few nanometers from the Aitken mode. Additional modes are also included to distinguish between different particle types in the same size range. Figure 34.1 exemplarily shows the modal concept of the global aerosol model EMAC/MADE-IN (Aquila et al. (2011); Sect. 34.5) developed at the Institute of Atmospheric Physics. In the size ranges of the Aitken and accumulation modes, the model considers separate modes for insoluble particles (black carbon and mineral dust) with and without a liquid coating as well as particles without insoluble inclusions.

As a further simplification, the standard deviations σ of the modes are kept constant in most aerosol models that apply the log-normal approach. This results in only two free parameters being necessary for describing the size distribution of each mode: N_t and D_{med} (Eq. 34.1). With these parameters, the total volume concentration of the mode (total aerosol volume per volume of air) can be derived assuming spherical particles:

$$V_t = \frac{\pi}{6}\int_{-\infty}^{+\infty} D^3 n(\ln D) d(\ln D) = N_t \frac{\pi}{6} D_{med}^3 \exp\left[\frac{9}{2}\ln^2 \sigma\right]. \qquad (34.2)$$

The total volume concentration V_t can also be derived from the mass concentrations of the different aerosol constituents present in the mode, taking into account the constituents' specific densities. Hence, if the mass concentrations of the individual species and the total particle number concentration N_t are known for each mode, the median diameters D_{med} of the modes can be derived (Eq. 34.2) and the particle size distribution is fully specified. Therefore, the models simulate the constituents' mass concentrations as well as the particle number concentration for each individual mode as basic prognostic quantities.

34.2.2 Aerosol Life Cycle

Atmospheric aerosol is influenced by particle sources, transport due to atmospheric motion, chemical and microphysical transformations, as well as particle sinks. Global aerosol models usually include a number of submodels to represent the different processes affecting atmospheric aerosol. Microphysical processes which affect the aerosol size distribution on the microscale, particularly aerosol–aerosol and aerosol–gas interactions, are frequently subsumed within a specific submodel. In the following, such submodels are referred to as aerosol modules. Particle emissions, chemical transformations, atmospheric transport and the deposition of particles on the Earth's surface are often simulated by other submodels. Processes usually considered in global aerosol models as well as their technical representations in modal schemes are described in the following.

34.2.2.1 Aerosol Microphysical Processes

Nucleation: Under specific conditions, gas molecules aggregate to form molecule clusters. Further growth of these clusters leads to the formation of very fine aerosol particles. This nucleation process is commonly simulated in large-scale atmospheric models by means of empirical parameterizations describing the particle nucleation rate dependent on a set of key parameters such as temperature, humidity, and the concentrations of the involved trace gases. Parameterizations applied in global models usually describe the nucleation of particles containing sulfuric acid (H_2SO_4) and water. Nucleation results in increases in particle number and mass concentration which, in modal models, are assigned to the finest-sized mode.

Condensation/evaporation: The surface of liquid aerosol particles experiences a permanent exchange of molecules with the gas phase. After condensation, compounds with low volatility, such as sulfuric acid, predominantly remain in the particulate phase. Large amounts of condensed compounds with higher volatility, such as nitric acid (HNO_3), can reevaporate to the gas phase. In global models, it is usually assumed that evaporation of low-volatile compounds is negligible. The condensation of these species is mostly simulated by applying particle-size-dependent physical descriptions of the gas fluxes to the particle surfaces. For sulfuric acid, the competition between condensation and nucleation is also

considered. In the case of constituents with higher volatility, the partitioning between gas and particles is often simulated by assuming that condensation and evaporation reach equilibrium within the time step used in the model to discretize time. This has the advantage that only the equilibrium partitioning has to be determined and an explicit physical description of the fluxes of condensing/evaporating mass and their temporal change can be avoided. In many cases, however, such fluxes are explicitly calculated. This is required, for instance, if the time scales to reach equilibrium are larger than the model time step. Global atmospheric models operate with time steps typically in the range of a few minutes to an hour. Especially for larger particles in the accumulation or coarse mode size range, the equilibrium time scales can be considerably larger.

Coagulation: The process of collision and sticking of aerosol particles, which results in the formation of larger particles, is called coagulation. Coagulation of atmospheric aerosol is mainly driven by the Brownian motion of the particles. Modal models consider coagulation of particles from the same mode (intramodal coagulation) as well as coagulation of particles from different modes (intermodal coagulation). The total coagulation rates are usually calculated by integration over the whole size distribution considering coagulation efficiencies (coagulation kernels) dependent on the sizes of the interacting particles. Intramodal coagulation leads to a reduction of the particle number concentration in the respective mode. Intermodal coagulation results in a reduction of the particle number concentration in the smaller-sized mode and a corresponding transfer of mass of the involved aerosol constituents from the smaller-sized to the larger-sized mode.

The uptake of mass due to condensation and coagulation can lead to changes in particle size and composition. Such changes can also be induced by the uptake of aerosol particles in cloud droplets (Sects. 34.2.2.2 and 34.2.2.3). If a cloud droplet evaporates, an aggregate of all aerosol particles taken up by the droplet remains, including also additional aerosol mass generated within the droplet. These changes can necessitate a reassignment of particles to other modes of the model. Hence, aerosol models rearrange particle mass and number accordingly at each time step.

34.2.2.2 Emissions and Chemical Transformations

Emissions: Aerosol particles can be emitted directly, like in the case of the wind-driven release of dust at the Earth's surface or the generation of black carbon particles during combustion processes. Aerosol matter can also be formed in the atmosphere by the transformation of emitted trace gases, called aerosol precursors (see below). The representation of emissions of aerosol and its precursors in global models depends on the specific emission process. Emissions are mostly implemented in the form of prescribed spatial and temporal distributions of emission fluxes. Emissions that depend on weather conditions, such as wind-driven emissions of mineral dust or sea salt, can either be prescribed or calculated during the simulation by taking into account meteorological parameters such as wind speed. The latter approach results in a higher consistency of emissions and meteorology.

Chemical transformation of aerosol precursors: Chemical processes in the atmosphere lead to the conversion of highly volatile species into condensable vapors which contribute to the nucleation of new particles or to the condensation of aerosol mass on pre-existing aerosol. As an example, photochemical processes can convert gaseous sulfur dioxide (SO_2) into gaseous H_2SO_4, which is then transformed into aerosol sulfate. SO_2 can also be converted into sulfate by liquid phase chemistry within cloud droplets. If the cloud droplet evaporates, sulfate is released in the form of aerosol residues. These conversion processes can be simulated with the aid of gas and liquid phase chemistry schemes. Since these schemes are computationally expensive, many global aerosol models consider only simplified aerosol precursor chemistry based on a limited number of species and reactions. Some models consider pre-calculated data sets of the concentrations of relevant trace gases. Important conversion processes to be taken into account are, for instance, the oxidation cycles of SO_2 into H_2SO_4 and of NO_x ($NO + NO_2$) into HNO_3.

34.2.2.3 Transport and Deposition

Atmospheric transport: To simulate the transport of aerosol, its precursor gases, and other quantities relevant for the aerosol simulation, global models include transport schemes. Different schemes are applied to represent transport by large-scale wind systems (advection), small-scale vertical motions driven by local heating of air masses (convection), and small-scale turbulent motion (diffusion). The winds and other parameters relevant for these calculations are either prescribed or calculated by the model. While an online calculation of these parameters enables effects of aerosols on weather and climate to be quantified in complex aerosol-climate simulations, prescribed meteorology simplifies the model system but enables only passive aerosol transport.

Wet deposition: Aerosol particles can be included in cloud droplets or ice particles either by acting as nuclei for cloud particle formation or by colliding with pre-existing cloud particles. When cloud particles are transformed into precipitation (rain or snow), the included aerosol mass is either "washed out" of the atmosphere or released at lower altitudes in case of evaporation or sublimation of rain or snow, respectively. Falling rain droplets or snowflakes collect additional aerosol particles, increasing the amount of aerosol that is removed from the atmosphere. Global models simulate these important aerosol sinks, called wet deposition, by coupling aerosol modules to cloud and precipitation schemes. The aerosol modules provide information about the relevant aerosol properties, and the cloud and precipitation schemes calculate the cloud coverage and cloud microphysical properties as well as the precipitation formation rates and precipitation fluxes. By combining this information, global models calculate the amount of aerosol removed from the atmosphere through wet deposition.

Dry deposition: Aerosol particles can also be deposited on the Earth's surface in the absence of precipitation (dry deposition). Collisions of particles with the surface and possible sticking can occur due to diffusive transport by turbulent air

motion in combination with the Brownian motion of fine particles or gravitational settling of larger particles. In global models, dry deposition is commonly expressed in terms of a deposition velocity which is calculated as a function of particle size and composition as well as local aerodynamic and surface properties.

34.3 Simulation of Aerosol Effects on Atmosphere and Climate

Effects on radiation: Anthropogenic or natural aerosol perturbations can change the Earth's radiation budget and consequently affect weather and climate. Global climate models include submodels to simulate the transfer of shortwave (solar) and longwave (terrestrial) radiation through the atmosphere. Among other processes, these submodels consider the scattering and absorption of radiation by aerosol particles. This requires information about the optical properties of the aerosol, that is, its specific characteristics controlling the interaction with radiation at different wavelengths. Such information can be derived from the microphysical and chemical aerosol properties (number concentration, size distribution, and composition) calculated by the aerosol module.

Effects on clouds: Aerosol particles have important effects on clouds. On the one hand, they serve as nuclei in the formation of cloud droplets and ice crystals. On the other hand, aerosol effects on radiation can induce dynamical and thermodynamical changes that can affect cloud properties and lifetime. The latter effects are implicitly simulated by the model, provided aerosol-radiation interactions are considered. Aerosol-induced cloud particle formation is commonly described by means of parameterizations driven by the simulated aerosol parameters which characterize the particle size distribution and composition, as well as dynamical and thermodynamical parameters such as temperature, relative humidity, and the cooling rate or updraft velocity. Since cloud formation frequently occurs on spatial scales which cannot be resolved by global models, many of these parameterizations include assumptions on the small-scale variation of the input parameters within the grid-boxes of the global model. The parameterizations return information about the properties of the formed cloud, particularly the cloud particle number concentration, which is used as a basis for further calculations of cloud development in the cloud schemes of the models. This enables simulations of the impact of aerosol changes on clouds and the resulting radiative and climate effects. At the Institute of Atmospheric Physics, particular efforts have been directed at developing parameterizations to describe aerosol-induced cirrus cloud formation (Kärcher et al. 2006).

Effects on atmospheric chemistry: Aerosol particles have important effects on atmospheric gas phase chemistry affecting, for instance, the ozone concentration. Gas phase chemistry is influenced by direct interactions of aerosol particles with gas molecules. Aerosol particles can take up condensable species like HNO_3,

which reduces their gas phase concentrations. This affects the concentrations and chemistry of other important trace gases. For instance, a reduction in gas phase HNO_3 can change the concentration of NO_x and HO_x ($OH + HO_2$), which play key roles in ozone chemistry. In addition, several gas phase compounds are converted to other gases by heterogeneous reactions on aerosol particle surfaces. For instance, heterogeneous reactions can lead to an efficient activation of chlorine radicals and subsequent ozone destruction in the stratosphere. Such effects are simulated in the gas phase chemistry schemes of the models by considering the loss and possible transformation of gases due to interactions with aerosol particles. In addition to their direct chemical impacts, aerosol particles also have indirect effects on atmospheric chemistry since their effects on radiation and clouds, which are simulated as described above, influence the photolysis, transport, and deposition of gaseous compounds.

34.4 Model Evaluation

To evaluate a model's capability to simulate atmospheric aerosol, the simulation results need to be compared to observations. Measurements from different kinds of observation platforms, such as ground based stations, aircraft, or satellites, provide a vast range of information about the mass concentrations of different aerosol constituents, particle number concentrations for different size ranges, composition of individual particles, optical aerosol properties, and other quantities. An example of a comparison of model results with observations is discussed below.

34.5 Examples of Results from Aerosol Model Simulations

34.5.1 Model Description

Different versions of global atmospheric aerosol models have been developed and applied at the Institute of Atmospheric Physics (Lauer et al. 2005; Lauer et al. 2007; Lauer and Hendricks 2006; Righi et al. 2011). The latest version is the EMAC/MADE-IN model system (Aquila et al. 2011) developed by coupling the aerosol module MADE-IN (modal aerosol dynamics model including insoluble modes) with the global climate-chemistry-model system EMAC (ECHAM/MESSy atmospheric chemistry model; Jöckel et al. 2005; Jöckel et al. 2006). MADE-IN follows the log-normal approach to simulate the aerosol size distribution (Sect. 34.2). The different modes as well as the aerosol constituents considered by MADE-IN are shown in Fig. 34.1 (see also Sect. 34.2.1). In the following, some examples of results are presented from a simulation performed by Aquila et al. (2011) with EMAC/MADE-IN considering emission data for the year 2000.

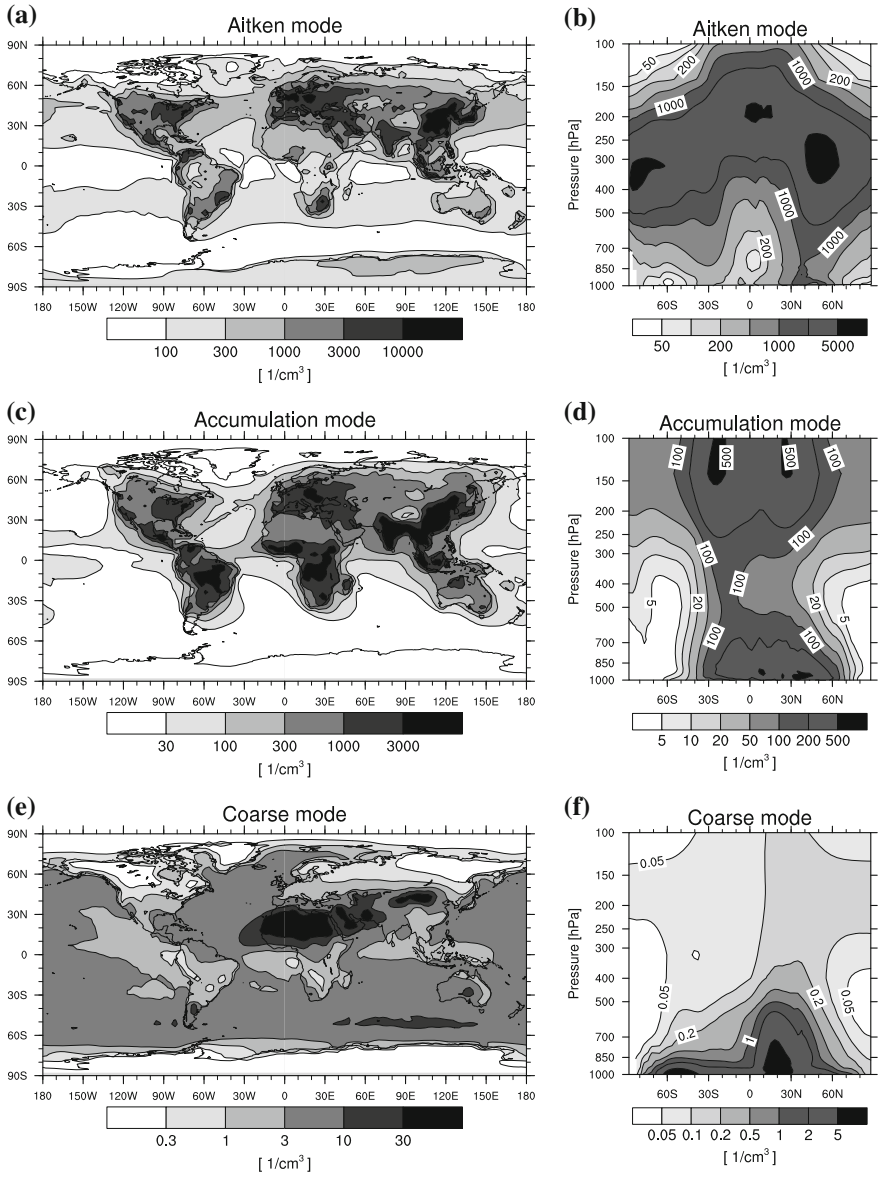

Fig. 34.2 Simulated annual mean horizontal global distribution of the aerosol particle number concentration in the atmosphere at the Earth's surface (*left*) and its zonal mean (average along circle of latitude) vertical distribution (*right*). A distinction is made between the concentrations of particles in the Aitken (**a**, **b**), accumulation (**c**, **d**), and coarse mode (**e**, **f**) size ranges. Pressure is used as the vertical coordinate in **b**, **d**, and **f**. Pressures of 500, 250, and 100 hPa, for instance, approximately correspond to altitudes of 5, 10, and 16 km, respectively. The simulation was performed with the EMAC/MADE-IN global aerosol model (Aquila et al. 2011)

34.5.2 Results

Figure 34.2 shows the annual mean number concentrations of particles in the size-ranges of the Aitken, accumulation, and coarse modes (Sect. 34.2.1), as simulated with EMAC/MADE-IN. The number concentrations of the different particle types shown in Fig. 34.1 were added within the respective size range. The horizontal distributions of the Aitken and accumulation modes (Figs. 34.2a and 34.2c) show the largest particle number concentrations over highly populated continental areas (Central Europe, Southeast Asia, and Eastern USA) where particles and their precursors are generated by human activities such as the combustion of fossil fuels. Particles in the accumulation mode size range also show large number concentrations over Central Africa and South America where frequent forest and grassland fires generate a large amount of aerosols of this size. Deserts such as the Sahara are a very important source of mineral dust aerosols. Many mineral dust particles are in the size range of the coarse mode, which, accordingly, presents its largest number concentrations over deserts (Fig. 34.2e). Comparatively large concentrations of coarse mode particles also occur over the oceans, where sea spray injects large sea salt particles into the atmosphere. For the Earth's surface, the simulation reveals typical particle number concentrations of about 100–10 000/cm^3 in the Aitken mode, 10 to several 1 000/cm^3 in the accumulation mode, and 1 to several 10/cm^3 in the coarse mode size range. These values are in good agreement with the concentrations measured in the atmosphere.

At low levels over the main emission areas, the simulated particle number concentrations decrease with altitude (Figs. 34.2b, 34.2d, and 34.2f). This decrease is mainly caused by washout due to rain and snow, spatial dispersion, agglomeration of particles due to coagulation, and gravitational settling of large, preferentially coarse mode particles. Above a specific altitude level, the concentrations of Aitken and accumulation mode particles increase again with height. Particularly in the Aitken mode size range, very large number concentrations of several 1 000/cm^3, on average, occur at pressure levels between about 400 and 200 hPa (about 7 and 12 km) at mid-latitudes and about 300 and 100 hPa (about 9 and 16 km) in the tropics. Here, large numbers of small sulfate aerosol particles nucleate after SO_2 is oxidized into gaseous sulfuric acid. The model assigns these particles to the Aitken mode. Due to condensation of additional sulfuric acid and coagulation, some of the newly formed particles grow to comparatively large sizes. This also increases the number concentration of accumulation mode particles. The modeled aerosol mass concentrations are considerably smaller at these altitudes than at the Earth's surface (Fig. 34.3) and the mean size of the particles is smaller than at the ground.

Figure 34.3 presents the simulated global average composition and total mass concentration of particles in the three different size ranges, at the surface and at 250 hPa (about 10 km). Aerosol water is not considered in this analysis since its contribution to aerosol mass is highly variable due to changes in the atmosphere's relative humidity. Particulate organic matter (POM) and sulfate (SO_4^{2-}) provide the largest contributions to the Aitken and accumulation mode mass at the Earth's

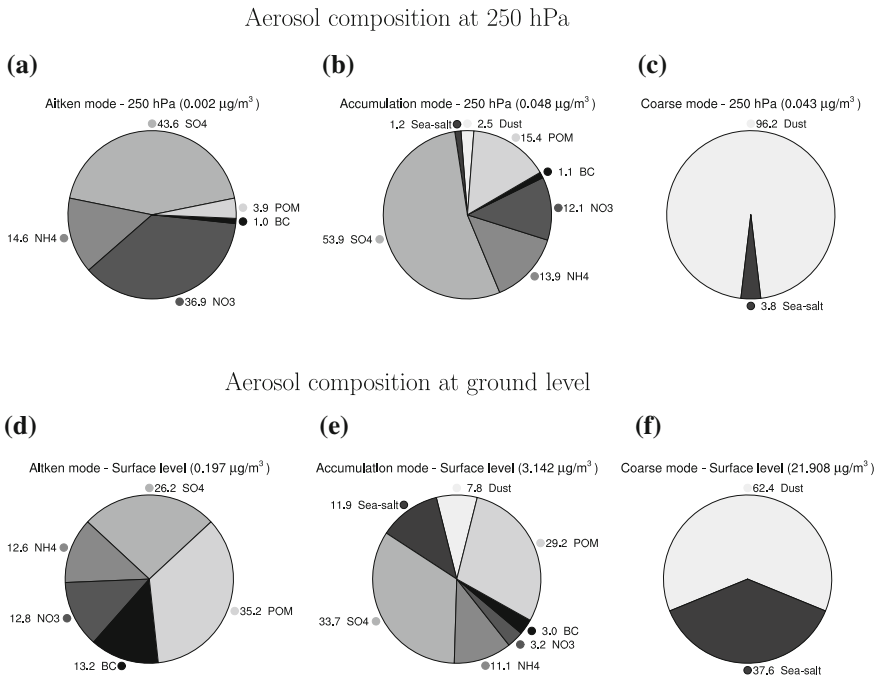

Fig. 34.3 Average composition (contribution of different constituents in % of mass) and total dry mass concentration (water content not considered) of the global aerosol in the Aitken (**a**, **d**), accumulation (**b**, **e**), and coarse mode (**c**, **f**) size ranges. Results are shown for the lowest model layer located just above the Earth's surface (*bottom row*) as well as for the 250 hPa pressure level, which corresponds to an altitude of about 10 km (*top row*). The simulation was performed with the EMAC/MADE-IN global aerosol model (Aquila et al. 2011). See Fig. 34.1 for an explanation of the aerosol constituents considered

surface. Other relevant constituents are black carbon (BC), nitrate (NO_3^-), and ammonium (NH_4^+). The accumulation mode also includes important contributions by fine sea salt and dust particles. The only non-water constituents of the coarse mode are mineral dust and sea salt, since the model neglects interactions of the coarse mode with other species, a simplification which should be abolished in the future. The global averages of the total mass concentrations of the different size modes, simulated for the lowest model layer, are 21.9, 3.1, and only 0.2 $\mu g/m^3$ in the coarse, accumulation, and Aitken mode, respectively. The relative contributions of sulfate and nitrate to the aerosol mass present in the Aitken and accumulation modes are much larger at 250 hPa than at the surface. This is caused by the formation of these compounds from precursor gases during vertical transport of air masses. In contrast, the contribution of directly emitted aerosol constituents, such as BC, decreases with altitude. Also, the composition of the coarse mode shows altitudinal changes. At 250 hPa, it is mainly composed of mineral dust since coarse sea salt particles are assumed to be larger than dust particles and, consequently, their removal due to

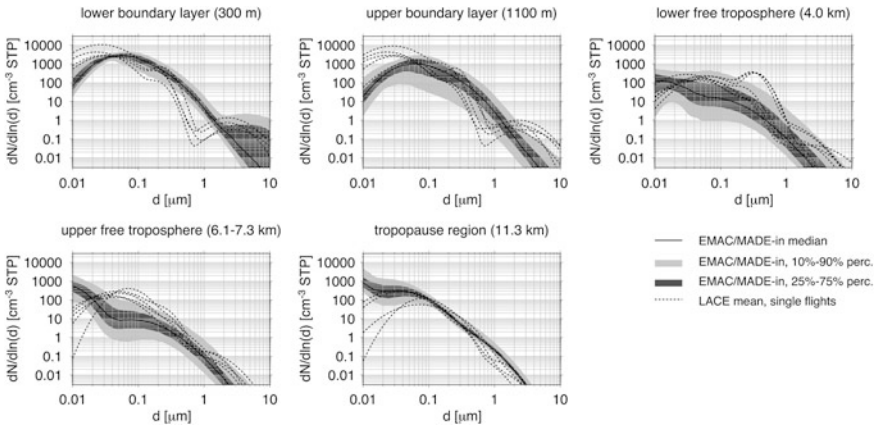

Fig. 34.4 Aerosol size distribution, i.e., number concentration per logarithmic size interval versus particle diameter, at different altitudes measured by Petzold et al. (2002) during the LACE campaign (Lindenberg aerosol characterization experiment) over Central Europe in August 1998 and simulated with the EMAC/MADE-IN global aerosol model. The number concentrations were converted to standard temperature and pressure (STP) conditions (273 K, 1013 hPa). Particle diameters correspond to dry aerosol (water content not considered). The dotted lines indicate size distributions measured during different flights. The solid line and shaded areas indicate the median, the 25th and 75th, and the 10th and 90th percentiles of the model results. The figure is adapted from Aquila et al. (2011)

gravitational settling is more efficient. The total mass concentrations are strongly reduced at 250 hPa. They amount to only 0.04, 0.05, and 0.002 $\mu g/m^3$ in the coarse, accumulation, and Aitken mode, respectively. Due to efficient gravitational settling of large particles, the coarse mode concentration experiences the largest reduction.

In order to evaluate the model's capability to simulate atmospheric aerosol correctly, the simulation results have been compared to observational data (Aquila et al. 2011). Figure 34.4 shows an example of a comparison of model results with data from aircraft measurements performed by the Institute of Atmospheric Physics. Aerosol size distributions at different altitudes are presented. The measurements were performed over the area of Berlin (Germany) in August 1998 (Petzold et al. 2002). The comparison reveals that the model is mostly able to reproduce the shape of the aerosol size distribution as well as the particle number concentrations at all altitudes. Discrepancies can be caused, for instance, by deficiencies in the size distribution representation of the aerosol module or by specific meteorological conditions during the campaign which cannot be reproduced by the simulation.

34.6 Outlook

A number of model improvements should be implemented to further enhance the quality of global aerosol-climate model simulations. These include the application of more detailed aerosol modules that, on the one hand, enable higher resolution of the particle size distribution, for instance, by the application of bin schemes (Sect. 34.2.1). On the other hand, the number of different particle types, mixing states, and aerosol constituents considered in the modules can be increased. For instance, a more detailed representation of organic aerosol constituents would be an important improvement, since most current global aerosol models subsume the whole suite of different types of particulate organic matter within one species. Particle-resolved schemes would dramatically increase the level of detail in the description of aerosol, but they still require too high computational expense to be applied in global climate simulations. Some global aerosol models still apply simplified representations of atmospheric chemistry. Hence, a coupling of aerosol modules to more comprehensive gas and liquid phase chemistry schemes is desirable. Also, the representation of aerosol effects on clouds needs to be further improved. Particular progress is required, for instance, in simulating aerosol interactions with ice and mixed-phase clouds. In addition to these modifications, also increases in the spatial and temporal resolution of the models can contribute to enhancing the level of detail in future global aerosol simulations.

Aerosol simulations would also benefit from further research on still existing gaps in fundamental knowledge about several aerosol-related processes. Further in situ and laboratory measurements supported by theoretical analysis are desirable to increase understanding of these processes. To give some prominent examples, further research should focus on the formation and properties of organic aerosols, the interactions of aerosols with liquid, mixed-phase and ice clouds, and on chemical transformations within or on the surfaces of aerosols.

References

Aquila, V., Hendricks, J., Lauer, A., Riemer, N., Vogel, H., Baumgardner, D., Minikin, A., Petzold, A., Schwarz, J.P., Spackman, J.R., et al.: MADE-in: a new aerosol microphysics submodel for global simulation of insoluble particles and their mixing state. Geosci. Model Dev. **4**, 325–355 (2011). doi:10.5194/gmd-4-325-2011

Ghan, J.S., Schwartz, E.S.: Aerosol properties and processes—a path from field and laboratory measurements to global climate models. Bull. Am. Meteor. Soc. **88**, 1059–1083 (2007)

Hendricks, J., Kärcher, B., Döpelheuer, A., Feichter, J., Lohmann, U., Baumgardner, D.: Simulating the global atmospheric black carbon cycle: a revisit to the contribution of aircraft emissions. Atmos. Chem. Phys. **4**, 2521–2541 (2004)

Hendricks, J., Kärcher, B., Lohmann, U., Ponater, M.: Do aircraft black carbon emissions affect cirrus clouds on the global scale? Geophys. Res. Lett. **32**, L12814 (2005). doi:10.1029/2005GL022740

Hendricks, J., Kärcher, B., Lohmann, U.: Effects of ice nuclei on cirrus clouds in a global climate model. J. Geophys. Res. **116**(D18206), 1–24 (2011). doi:10.1029/2010JD015302

Jöckel, P., Sander, R., Kerkweg, A., Tost, H., Lelieveld, J.: Technical note: the modular earth submodel system (MESSy)—a new approach towards earth system modeling. Atmos. Chem. Phys. **5**, 433–444 (2005). doi:10.5194/acp-5-433-2005

Jöckel, P., Tost, H., Pozzer, A., Brühl, C., Buchholz, J., Ganzeveld, L., Hoor, P., Kerkweg, A., Lawrence, M.G., Sander, R., et al.: The atmospheric chemistry general circulation model ECHAM5/MESSy1: consistent simulation of ozone from the surface to the mesosphere. Atmos. Chem. Phys. **6**, 5067–5104 (2006). doi:10.5194/acp-6-5067-2006

Kärcher, B., Hendricks, J., Lohmann, U.: Physically based parameterization of cirrus cloud formation for use in global atmospheric models. J. Geophys. Res. **111**, D01205 (2006). doi:10.1029/2005JD006219

Köhler, I., Dameris, M., Ackermann, I., Hass, H.: Contribution of road traffic emissions to the atmospheric black carbon burden in the mid-1990s. J. Geophys. Res. **106**, 17997–18014 (2001)

Lauer, A., Hendricks, J., Ackermann, I., Schell, B., Hass, H., Metzger, S.: Simulating aerosol microphysics with the ECHAM/MADE GCM—Part I: model description and comparison with observations. Atmos. Chem. Phys. **5**, 3251–3276 (2005). doi:10.5194/acp-5-3251-2005

Lauer, A., Hendricks, J.: Simulating aerosol microphysics with the ECHAM4/MADE GCM—Part II: Results from a first multiannual simulation of the submicrometer aerosol. Atmos. Chem. Phys. **6**, 5495–5513 (2006). doi:10.5194/acp-6-5495-2006

Lauer, A., Eyring, V., Hendricks, J., Jöckel, P., Lohmann, U.: Global model simulations of the impact of ocean-going ships on aerosols, clouds, and the radiation budget. Atmos. Chem. Phys. **7**, 5061–5079 (2007). doi:10.5194/acp-7-5061-2007

Petzold, A., Fiebig, M., Flentje, H., Keil, A., Leiterer, U., Schröder, F., Stifter, A., Wendisch, M., Wendling, P.: Vertical variability of aerosol properties observed at a continental site during the Lindenberg aerosol characterization experiment (LACE 98). J. Geophys. Res. **107**, 8128 (2002). doi:10.1029/2001JD001043

Righi, M., Klinger, K., Eyring, V., Hendricks, J., Lauer, A., Petzold, A.: Climate impact of biofuels in shipping: global model studies of the aerosol indirect effect. Environ. Sci. Technol. **45**, 3519–3525 (2011). doi:dx.doi.org/10.1021/es1036157

Whitby, E., McMurry, P.: Modal aerosol dynamics modelling. Aerosol. Sci. Techol. **27**, 673–688 (1997)

Whitby, K.T.: The physical characteristics of sulfur aerosols. Atmos. Environ. **12**, 135–159 (1978)

Earth System Modeling

35

Patrick Jöckel

Abstract

Earth system models are important research tools for improving understanding of the climate system and for simulating climate projections. This chapter is devoted to the basic construction principles and challenges of such models, whereas application examples are provided in companion chapters. Since they still do not incorporate the full complexity of the real climate system (and maybe never will), Earth system models nowadays typically focus on specific aspects, for instance on the role of chemically active substances in the climate system.

35.1 Introduction

Earth system models (ESMs) are computer programs for simulating the terrestrial *climate* system. They serve as virtual laboratories, since no real planetary scale systems for conducting *experiments* are available. As such, ESMs depict a condensate of our knowledge about the "real world *climate system*" and with increasing knowledge they are continuously further developed and become more and more complex. Besides constituting tools for fundamental research on processes in the Earth system, ESMs are used for assessing the *anthropogenic* influence on climate and for future climate projections. Lately, ESMs are being

P. Jöckel (✉)
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Patrick.Joeckel@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_35,

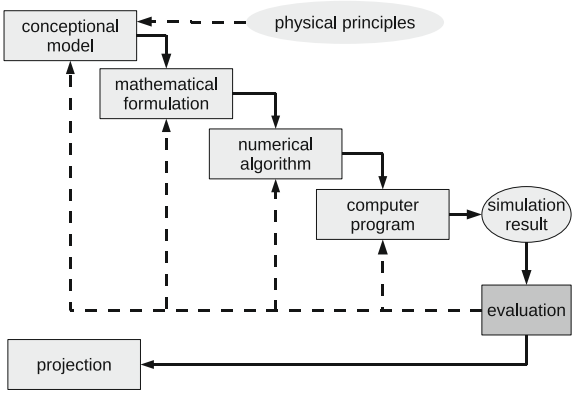

Fig. 35.1 Construction principle of Earth system models

more and more used to test climate change *mitigation* strategies and to examine potential climate engineering measures and their side effects.

At DLRs Institute of Atmospheric Physics (IPA) the most important applications are the *assessment* of the impact of land-, ship- and air-*traffic* on air quality and climate, including potential mitigation efforts; the role of natural cirrus *clouds* and *contrails*; analysis of *satellite* based *remote sensing* observations of radiatively active substances (e.g., *methane*); the role of *aerosol*; and assessment of the future development of the stratospheric *ozone layer*. Moreover, *simulation* results of complex ESMs are the basis for the development of simplified *climate models*.

35.2 Earth System Models: A Construction Guideline

Any meaningful application of ESMs requires a description of all relevant processes of the real system in sufficient detail. However, these prerequisites pose a fundamental dilemma in computational Earth system science, because a priori it is neither clear what the relevant processes are, nor what "sufficient detail" implies. As a consequence, computational Earth system science and the construction of ESMs are continuously evolving processes, of which the most important steps are outlined in Fig. 35.1.

The first step is to build a conceptional model based on fundamental physical principles (e.g., conservation of energy and momentum, etc.). In a second step, the mathematical formulation of the model usually leads to a set of nonlinear, coupled differential equations. Except for some rare idealized cases these cannot usually be solved analytically, so a numerical solution method appropriate for the system of equations needs to be found. This *numerical method* needs to be coded as computer program, usually in a so-called high-level language (Fortran, C, C++, etc.), and translated by a compiler into an executable command sequence. The numerical solution (i.e., the simulation result), which is always an approximate solution, needs to be evaluated with real world data, e.g., with direct, in situ observations,

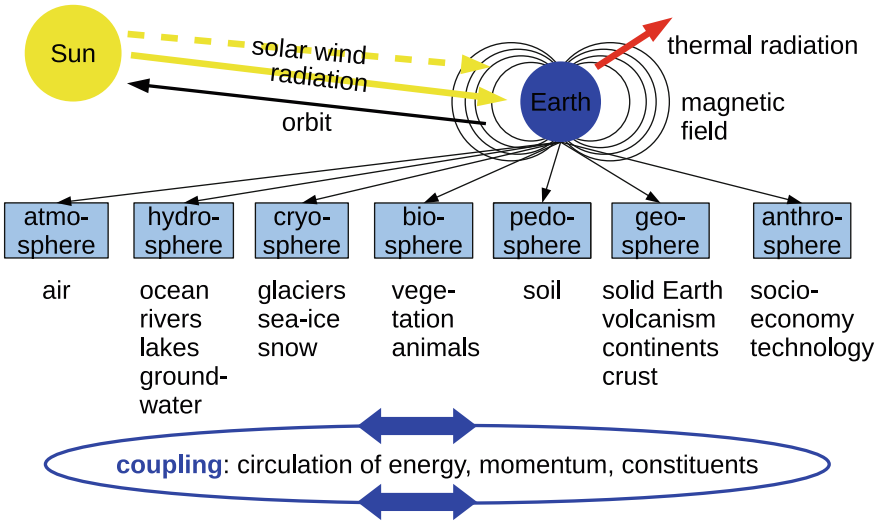

Fig. 35.2 Conceptional view of the Earth system

with indirect observations such as those provided by remote sensing *instruments* on satellites, or with data from climate reconstructions (e.g., from ice cores, sediment cores, tree rings, etc.). This last step, which is a tremendous task on its own, is an absolute prerequisite for meaningful ESM-based projections. This implies that each ESM needs to represent the past and present state of the terrestrial climate system adequately, before it can be applied for climate projections, mitigation strategies or geo-engineering assessments. As the dashed lines indicate, each step bears the potential necessity to go one or more steps back for further refinement, if the constructed model fails in one way or another.

To start with, Fig. 35.2 shows a conceptional model of the Earth system (more precisely, the terrestrial climate system), which as such forms the basis for a comprehensive ESM as ultimate aim.

Earth is in radiative balance with the *Sun*; incoming shortwave *radiation* is partly reflected and partly absorbed; the *absorption* leads to heating and consequently to the *emission* of thermal radiation. This balance is under continuous variation, since Earth's orbit around the Sun and solar activity (irradiance, magnetic field, solar wind) vary in time. Further variations arise from changing Earth characteristics like changes in *albedo* (*reflectivity*) and absorption potential (*greenhouse effect*). Earth itself can conceptionally be further decomposed into subspheres, as listed in the figure. These subspheres are more or less strongly coupled through the *circulation* of energy, momentum and constituents (e.g., *water*, carbon, nitrogen, sulphur, etc.). In addition, the upper *atmosphere* (ionosphere, plasmasphere) is electromagnetically coupled with the solar and Earth's magnetic fields and influenced by the solar wind (a stream of charged particles).

A mathematical description of this system requires the identification of so-called state-variables which vary in space and time through the action of numerous physical, chemical, biological and, last but not least, socioeconomical processes. For instance, the fluid dynamic part of the subsphere "atmosphere" (see Fig. 35.2) can be formulated as compressible flow (of air) in a shell around a rotating sphere by four fundamental equations:

35.1. Conservation of momentum (Newton's second law)
35.2. Conservation of mass (*continuity equation*)
35.3. Conservation of energy (first law of *thermodynamics*)
35.4. Equation of state (ideal gas for air in good approximation)

These four equations constitute a coupled set of partial differential equations (PDE) for the four state variables *pressure*, *temperature*, density and velocity of air. Through the partial derivatives, variations in space and time are not independent. Moreover the equations are nonlinear, i.e., they contain squares of the state variables which lead to deterministically chaotic behavior of the solutions. A further complication arises because the system is in a rotational frame of reference and the equations therefore need to consider the fictitious *Coriolis* force. But most important, the system is subject to external forcing, since the equations contain a dissipation term (friction, Eq. 35.1) and require heat input and release (Eq. 35.3), including heat *uptake* and release by the phase transitions of water. An analytical solution for this system does not exist.

Assuming the forcing terms (dissipation and heat input/release) are given, the PDE system of the "dynamics" can be solved numerically with a variety of different algorithms, e.g., finite differencing or spectral methods. Common to all numerical methods is that the continuous analytical equations need to be approximated by discrete algebraic equations that can be solved numerically by computers. This translation is called *discretization* and is one of the key aspects in numerical mathematics. Discretization defines the spatial resolution of the model; the equations are solved algebraically for discrete points in space and time which define the model grid and time step length. An important limitation of such discretization is the *Courant–Friedrichs–Lewy* (CFL) criterion, which basically says that information in one time step (which is given by the discretization in time) must not travel farther than one unit of the spatial grid (given by the spatial discretization). If this criterion is violated, the numerical algorithm becomes unstable and the solution unphysical and therefore meaningless. This has an important implication for the construction of models, since according to this criterion, the spatial grid size and the discrete time step length for time integration cannot be chosen independently. In other words, if the spatial resolution is increased (finer grid size) to achieve more accurate results, the time step length needs to be reduced appropriately. Increasing the number of grid-points and reducing the time step length, however, both increase the computational demands, i.e., the number of operations the computer must calculate to simulate a given time period. This is limited by the available computational power and requires a compromise between model resolution and simulated time span. Nowadays, global models of the atmosphere on the most powerful computers can simulate time

periods of days with a spatial resolution of O(10) km, e.g., for *numerical weather prediction*, or time periods of centuries up to a millennium with a spatial resolution of O(100) km.

Depending on the chosen discretization method, the *transport* of constituents must be implemented as a separate process (see Sect. 35.4). For instance, the spectral transform technique used for solving the dynamic system (Eqs. 35.1–35.4) is not suitable for calculating the transport of constituents due to the *Gibbs phenomenon*, which is common to all spectral methods: small negative (and therefore unphysical) overshoots of constituent *mixing ratios* are calculated. The separate implementation of the transport process on the corresponding Eulerian ("quadratic Gaussian") grid, however, also has some drawbacks: Inconsistencies with respect to the implicit air mass transport arise which lead to a violation of mass conservation (Jöckel et al. 2001). Furthermore, transport algorithms on Eulerian grids suffer from numerical diffusion.

For these reasons, IPA developed a constituent transport algorithm in the Lagrangian frame of reference, i.e., instead of discretizing the atmosphere into stationary grid boxes, the atmosphere is divided into air parcels which are transported with the air flow (Reithmeier and Sausen 2002). This approach turned out to be very successful for the *atmospheric* transport of *water vapor* (Stenke et al. 2008) and other chemically active substances (Stenke et al. 2009). The currently ongoing development targets on replacing the complete spectral dynamics by a Lagrangian approach.

35.3 Forcing Terms: Parameterizations and Feedback

The solution for the realistic case requires, as indicated above, also knowledge about the forcing terms (sometimes called the "*physics*" of the model). The distribution of the heat input in space and time depends on the distribution of light (radiation) and the distribution of the absorbers, i.e., the surface albedo, clouds, aerosol and the *greenhouse gases* (water vapor, *carbon dioxide*, methane, *nitrous oxide*, *ozone*, and others). As a consequence, the mathematical model of the atmosphere as described by the four equations above needs to be further expanded by additional equations, most important by equations for the *hydrological cycle*, since water vapor is the most important greenhouse gas and energy transport is strongly coupled to water vapor transport through the absorption and release of *latent heat* during phase transitions. Given the limitations of the numerical discretization above, this has another important consequence: some processes happen on spatiotemporal scales which are smaller than the typical grid size and are therefore not resolved. A demonstrative example are clouds, which are important players in the climate system. Their spatial scale ranges from O(1 000) km down to <O(1) km, and at least the smaller clouds literally fall through the grid of a typical climate model. Their effect must therefore be described at least approximately in each grid box, based on physical principles, available information (e.g., temperature and water content) for the grid-box, and

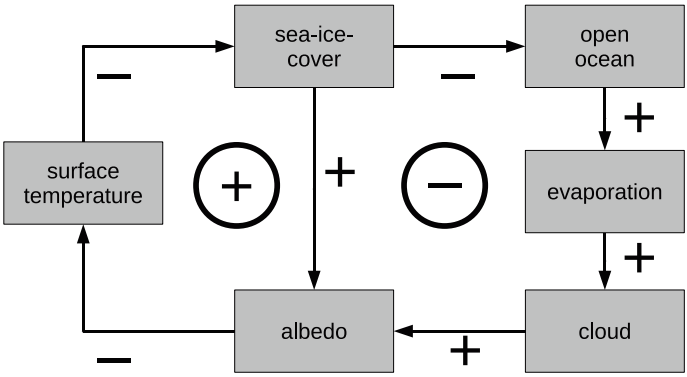

Fig. 35.3 Ice-albedo feedback as an exemplary feedback mechanism between forcing and state variables (positive and negative responses are indicated by *plus* and *minus* signs, respectively)

assumptions about the processes involved. In other words, these clouds are not directly simulated, but their effect is parameterized. Since there is naturally a certain degree of freedom in the formulation of such *parameterizations*, there are related uncertainties and differences among existing ESMs.

Adding parameterizations for the forcing terms has another important consequence: they introduce *feedback* mechanisms between forcing terms and state variables, implying that they are no longer independent of each other. This is illustrated in Fig. 35.3 for ice-albedo feedback. The *sea-surface temperature* is a state variable and the heat input, which determines it, is modified by the albedo and the clouds.

An increasing sea-surface temperature leads to a decrease in sea-ice *coverage*, since more sea ice melts. This in turn leads to an overall decline of surface albedo, because the dark *ocean* absorbs more light, whereas ice reflects it. Increasing absorption by the ocean, however, leads to increased sea-surface temperature, and as such a positive feedback loop (left) is created. Less sea-ice cover implies, on the other hand, more open ocean surface, where more water can evaporate. With more water vapor in the atmosphere, more clouds can potentially form, which will increase the albedo, i.e., reflect more light and decrease the absorption by the ocean. Thus a negative feedback loop (right) is created. Which of the two loops is more effective is a priori unclear.

35.4 Operator Splitting: The Fundamental Concept of Earth System Model Implementations

As an intermediate summary, ESMs are computer programs which combine different algorithms. For instance, the description of the subsphere "atmosphere" (see Fig. 35.2) requires at least one algorithm for the solution of the basic equations describing the fluid dynamic system and several others for calculating the forcing

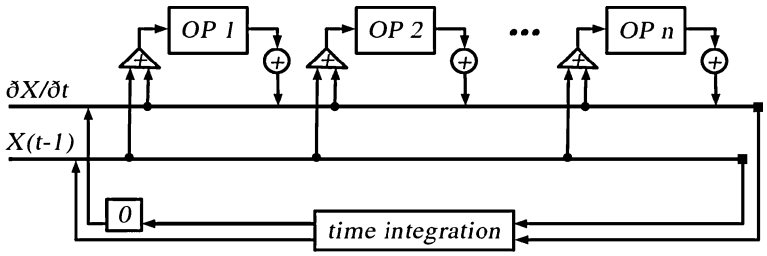

Fig. 35.4 Operator splitting concept

terms (e.g., absorption, reflection and *scattering* of light, the orbit of Earth around the Sun) and for the parameterizations of processes that are not resolved by the chosen spatiotemporal discretization (clouds, *convection*, diffusion, *gravity waves*). An extension by other subspheres (see Fig. 35.2), e.g., to construct a coupled atmosphere–ocean model, or a more detailed description of processes, increases the complexity and the number of different algorithms that need to be combined. The fundamental concept which enables the coupling of different model components into one simulation system is the *operator splitting* concept (see Fig. 35.4).

Based on the state variable X (e.g., air temperature) at the initial condition or the recent time step [X(t − 1)], the first operator (OP 1, e.g., the radiation algorithm) calculates a tendency (partial time derivative, here a temperature change). Subsequently, the second operator (OP 2, e.g., the cloud algorithm) calculates an additional tendency (indicated by a plus sign in a circle, e.g., the temperature increase due to latent heat release through *condensation*) based on X(t − 1) and the tendencies of all prior operators (indicated by the plus sign in a triangle). This continues for all operators modifying the temperature. The total temperature change during the model time step is then determined by the sum of all partial contributions. From this the time integration scheme calculates the initial value for the next time step and resets the total tendency to zero (box with zero).

Mathematically, this concept approximates the total derivative of a state variable with respect to time by a sum of partial derivatives. In other words, the solution for advancing the state variables in time is not calculated in one step, but rather sequentially, where each step is described by an operator attributed to a specific process.

35.5 Handling Complexity: The Coupling Issue

The complexity of the models, the fact that most of them grew historically over decades, and the increasing number of different operators pose the challenge of how to couple the algorithms technically. Figure 35.5 gives an overview of the different concepts.

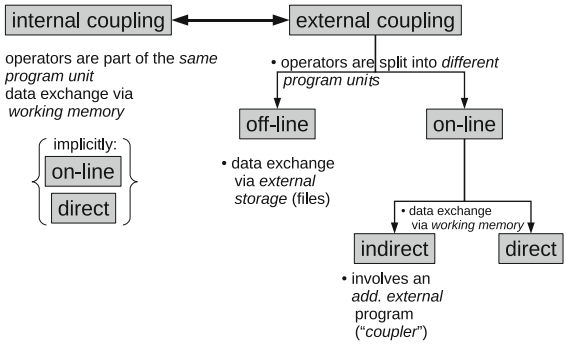

Fig. 35.5 Different concepts for coupling operators in ESMs as outlined in the text

Technically, the simplest operator is formed by a subroutine which calculates a specific output for the provided input. The granularity of this decomposition into operators can be formalized in a way that a set of subroutines describes a specific natural process of the system [which in the Modular Earth Submodel System (MESSy) is called a submodel; Jöckel et al. (2005)]. Such a process can still be regarded as an operator. Even further, a set of processes describing a specific domain (or subsphere) of the Earth system, such as the atmosphere or the ocean subsystem, can still be regarded as an operator. The coupling describes how these operators communicate which each other. Two operators are internally coupled if they are part of the same program unit, i.e., the same executable, and exchange information on-line, i.e., via the working memory of the computer. In contrast, operators are externally coupled if they are split into different program units, i.e., different executables. External coupling implies that information to be exchanged between the operators needs to leave the program unit bounds and is either exchanged off-line via storage to and import from an external (e.g., disk-based) file system, or on-line via the working memory. Whereas off-line external coupling is computationally inefficient due to the low input/output bandwidths and hardly allows a two-way data exchange, on-line external coupling requires special software or a model infrastructure for communication between two (or more) executables (such as for instance the message-passing interface, MPI). The external on-line coupling approach can be further differentiated into indirect external on-line coupling, in which two different program units communicate via an additional intermediate program unit (meaning that another executable is required, which is not identical to the message passing environment process) and direct external online coupling, where the two program units exchange their information without such an additional intermediate program unit (e.g., within the same message-passing environment). The intermediate program unit performing the data exchange in the indirect external on-line coupling approach is commonly called a *coupler*. In most implementations, the coupler does, however, not only perform the data exchange (including transpositions) but also further services, such as for instance rediscretizations, i.e., transformations of data from one domain model grid to another.

The choice of the most suitable coupling method depends on many aspects, not least on the desired application. In summary, a proper choice usually reflects a compromise which minimizes the computational and communication overheads and the implementation effort [which depends mostly on the code structure of the involved legacy model(s)] and maximizes the desired degree of flexibility and sustainability. Internal coupling is the most widely used approach, since it is inherent to all ESMs.

35.6 The Modular Earth Submodel System: Towards a Comprehensive Earth System Model

The internal coupling approach has been formalized (Jöckel et al. 2005, 2010) by specifying submodel implementations for the operators (processes or diagnostics) to be coupled. Furthermore, an interface structure was defined which allows the strict separation of process implementations from a shared model infrastructure (such as memory management, input/output, metadata handling, call sequence, etc.). This approach, explained here in more detail, is being applied and co-developed further at IPA together with a consortium of other institutions. The basic concept of the Modular Earth Submodel System (MESSy) is a four-layer structure: the base model layer will ultimately consist only of a central clock and run-time control. Currently however, it is typically a *general circulation model* (GCM) of the atmosphere. The base model interface layer comprises the base-model-specific implementation of the MESSy infrastructure; it can be regarded as a multiple socket outlet for communication between the base model and the submodels. The submodel interface layer represents the connector of a specific process to the infrastructure. And last but not least, the submodel core layer comprises the base-model-independent implementation of a specific process in the Earth system, or of a diagnostic tool of the model system. It can be regarded as an operator using the data exchanged via its submodel interface layer with other submodels and/or the base model. The MESSy user interface is based on the namelist concept of the Fortran95 standard (ISO/IEC-1539-1). Each submodel is controlled by (at least) two namelists containing parameters and switches: one affecting the internal complexity and flow control of a specific submodel, and the other for the coupling of a specific submodel to the base model and to other submodels via the MESSy infrastructure. The base model interface layer of MESSy is based on so-called generic (or infrastructure) submodels, where the term generic indicates that these submodels are also coded as base-model-independent submodels, i.e., organized into an interface and a core layer (e.g., Jöckel et al. 2008). Each generic submodel serves a specific, superordinate purpose, such as memory management, input/output, flow control, etc.

The Modular Earth Submodel System is designed to facilitate tailor-made model setups for a wide variety of scientific applications, without the requirement of source code modifications or recompilation. Each submodel can be switched on

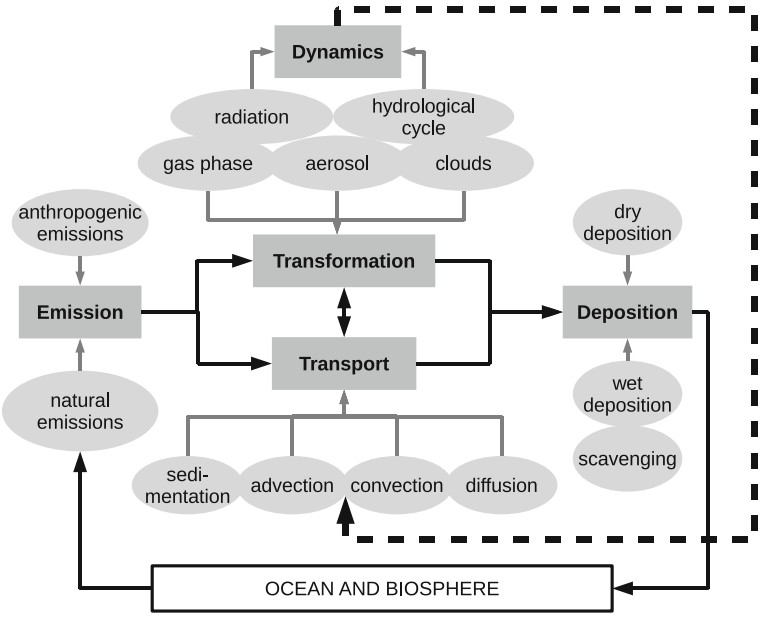

Fig. 35.6 A conceptual model of atmospheric chemistry

or off individually via the user interface at run-time. This is particularly useful for the analysis of feedback mechanisms. Furthermore, different formulations of the same process can coexist, allowing *uncertainty* estimates for different parameterizations or implementations. The highly *modular* structure enables parallel developments; nevertheless, a state-of-the-art model configuration is always available. Due to the continuously increasing complexity, a modular approach is unavoidable for the overall maintenance of such a research tool, to which several scientists at different institutions contribute.

35.7 The Modular Earth Submodel System: Atmospheric Chemistry and Beyond

Originally, MESSy was designed to enable the expansion of a general circulation model of the atmosphere by atmospheric *chemistry* processes. Conceptionally, this is a further zooming-in into the "atmosphere" box of Fig. 35.2, as shown in Fig. 35.6.

Chemically active substances are released by natural processes from the ocean, *soil* and biosphere, from volcanoes and also through anthropogenic activity, such as human-induced *biomass burning*, industry, traffic, animal husbandry, agriculture, etc. These primary released substances undergo photochemical reactions and form secondary, intermediate substances in the gas phase, in cloud droplets and aerosol, on ice and aerosol surfaces, etc. Substances are transported with the air

stream on different spatio-temporal scales, depending on their typical photochemical *lifetime* (i.e., the time before they react further). *Long-range transport* (large-scale *advection*) can cause distant effects; convective transport modifies the vertical distribution of tracers on short *time scales*, etc. Ultimately, substances are deposited by dry *deposition*, or by wet deposition caused by uptake in cloud droplets, scavenging by *rain*, and thus feedback to other subspheres such as the ocean, the biosphere or the pedosphere (soil).

The "Dynamics" box in Fig. 35.6 is mathematically described by the four equations listed above, including the forcing terms (radiative heating/cooling and friction). The distributions of radiatively active substances which are also chemically active (such as methane and ozone) determine the local absorption and emission of *electromagnetic radiation* and therefore the differential heating and cooling of the atmosphere, and thus contribute to the forcing terms. The forcing, in turn, affects the dynamics (i.e., the air flow), and therefore the transport pathways of the chemically active constituents (tracers). This feedback between the "chemistry" and the "dynamics" is the fundamental characteristic of *chemistry climate models* (CCMs). The most prominent application of CCMs is assessment of the stratospheric ozone layer.

Due to the modular structure of MESSy, the feedback between "chemistry" and "dynamics" can nicely be switched off (conceptually this implies the cutting of the dashed arrow in Fig. 35.6) and thus the model can be run in a quasi-*chemistry transport model* mode (Deckert et al. 2011). This is important if small perturbations of the chemical subsystem are to be examined, for instance, the question of the impact of nitrous oxide emissions of *air traffic* on climate. Small perturbations of nitrous oxide emissions of *aircraft* in a model configuration with feedback (of chemistry to dynamics) require two (with and without perturbation) long (several decades) simulations to allow the detection of *statistically significant* differences. The reason is the nonlinear, deterministic chaotic nature of the system, also known as the *butterfly effect*. Any small perturbation leads to a different realization (*weather* situation) of the same climate, thus the two simulations cannot be compared directly but only statistically, since the meteorological situations at a given time between both simulations differ. As a consequence, the signal to be detected becomes blurred by the internal variability (noise) of the system, i.e., by a too low signal-to-noise ratio. Switching off the feedback allows detection of the chemical signal by direct comparison of two shorter simulations: without the feedback, the meteorological situations between the chemically perturbed and chemically unperturbed simulations do not diverge, presupposing that the initial meteorological condition is the same in both simulations.

After the first version of MESSy was implemented (Jöckel et al. 2006) in connection with an atmospheric GCM, it turned out that the overall concept was much more powerful than the original goal of expanding a GCM into a CCM. Based on a revision and after the release of the second version of MESSy (Jöckel et al. 2010) with an enhanced and even more flexible model infrastructure, several new developments were triggered. An existing model of the physical ocean has been coupled successfully to the atmospheric component by the internal coupling

approach of MESSy (Pozzer et al. 2011). It could be shown that this approach is competitive, though with advantages and disadvantages, with a previously implemented indirect external on-line coupling approach involving additional coupler software.

Moreover, MESSy has been successfully coupled to a limited-area atmospheric model (Kerkweg and Jöckel 2012a) as the basis for a new regional atmospheric chemistry model. In addition, this regional model has been coupled on-line to the global model following a direct external on-line approach implemented as client–server architecture, the so-called multimodel driver (Kerkweg and Jöckel 2012b). This coupling approach allows for the concurrent *nesting* of an arbitrary number of regional model instances into the global model or into each other. The system enables on-line dynamic (Hofmann et al. 2012) and chemical downscaling, thus providing an alternative to globally highly resolved models by offering the possibility to zoom into specific regions of interest. The most important application will be accompanying simulations with increased resolution for measurement campaigns, which will improve the observational data analysis and re-integration of the findings into improved model components.

35.8 Earth System Models: Computational Challenges

ESMs have, from the beginning of computer systems, been among the most demanding applications in computational science. The first numerical weather prediction (Charney et al. 1950) was performed on the most powerful *supercomputer* of that time (ENIAC, Electronic Numerical Integrator and Computer, which operated in the decimal system) and took 24 h for a 24-hour forecast. In 2008, the same simulation with the same simplified two-dimensional model and the same initial condition was possible on a common cell-phone in less than a second (Lynch and Lynch 2008). This nicely illustrates Moore's (empirical) Law, which states (simplified) that the computational power (expressed in floating point operations per second (FLOPS/s) per unit cost doubles approximately every 18 months. As a consequence, the achievable FLOPS/s of *high performance computing* (HPC) systems have increased since von Neumann's time by 12 orders of magnitude and the development is still continuing, approaching systems that are capable of 1 Exa-(=10^{18})-FLOPS/s. Despite this tremendous increase in the power of HPC systems, ESMs are still exploiting them up to the limits, mainly for two reasons: increasing grid resolution and increasing complexity through the incorporation of more and more processes. However, based on currently available chip technology, the retention of *Moore's Law* is approaching its limit, mainly because of the huge power consumption. A further increase, this is at least the current view, can only be achieved by an increased degree of parallelization, i.e., with more parallel tasks on many more (but less powerful and less power consuming) computational cores. Whereas nowadays HPC systems have typically O(10^3)–O(10^5) computational cores, the next "exa-scale" generation is expected to have

$O(10^6)$ cores or even more. However, each core is expected to be equipped with much less memory and to have much lower computational power than today's Central Processing Units (CPUs). In addition, the applied chip-architecture is expected to change. Currently the debate on the application of Graphical Processing Units (GPUs), which also have intrinsically a high degree of parallelism, is ongoing. These developments, which in any case constitute a change of paradigm in HPC, currently pose the biggest challenge for the future of ESMs. The reason is that such a high degree of parallelism cannot be achieved with today's models. Completely new numerical algorithms, even new programming models and languages which take the specifics of the upcoming hardware architectures into account, are required. A simple porting of existing ESMs to the next, more powerful, architecture is presumably not straightforward and requires specific knowledge about the applied HPC systems.

References

Charney, J.G., Fjörtoft, R., von Neumann, J.: Numerical integration of the barotropic vorticity equation. Tellus **2**, 237–254 (1950)

Deckert, R., Jöckel, P., Grewe, V., Gottschaldt, K.-D., Hoor, P.: A quasi chemistry-transport model mode for EMAC. Geosci. Model Dev. **4**, 195–206 (2011). doi:10.5194/gmd-4-195-2011

Hofmann, C., Kerkweg, A., Wernli, H., Jöckel, P.: The 1-way on-line coupled atmospheric chemistry model system MECO(n) Part 3: Meteorological evaluation of the on-line coupled system. Geosci. Model Dev. **5**, 129–147 (2012). doi:10.5194/gmd-5-129-2012

Jöckel, P., von Kuhlmann, R., Lawrence, M.G., Steil, B., Brenninkmeijer, C.A.M., Crutzen, P.J., Rasch, P.J., Eaton, B.: On a fundamental problem in implementing flux-form advection schemes for tracer transport in 3-dimensional general circulation and chemistry transport models. Q. J. Roy. Meteor. Soc. **127**, 1035–1052 (2001)

Jöckel, P., Sander, R., Kerkweg, A., Tost, H., Lelieveld, J.: Technical note: the modular Earth submodel system (MESSy)—a new approach towards Earth system modeling. Atmos. Chem. Phys. **5**, 433–444 (2005). doi:10.5194/acp-5-433-2005

Jöckel, P., Tost, H., Pozzer, A., Brühl, C., Buchholz, J., Ganzeveld, L., Hoor, P., Kerkweg, A., Lawrence, M., Sander, R., et al.: The atmospheric chemistry general circulation model ECHAM5/MESSy1: consistent simulation of ozone from the surface to the mesosphere. Atmos. Chem. Phys. **6**, 5067–5104 (2006). doi:10.5194/acp-6-5067-2006

Jöckel, P., Kerkweg, A., Buchholz-Dietsch, J., Tost, H., Sander, R., Pozzer, A.: Technical note: coupling of chemical processes with the modular Earth submodel system (MESSy) submodel TRACER. Atmos. Chem. Phys. **8**, 1677–1687 (2008). doi:10.5194/acp-8-1677-2008

Jöckel, P., Kerkweg, A., Pozzer, A., Sander, R., Tost, H., Riede, H., Baumgaertner, A., Gromov, S., Kern, B.: Development cycle 2 of the modular Earth submodel system (MESSy2). Geosci. Model Dev. **3**, 717–752 (2010). doi:10.5194/gmd-3-717-2010

Kerkweg, A., Jöckel, P.: The 1-way on-line coupled atmospheric chemistry model system MECO(n) Part 1: Description of the limited-area atmospheric chemistry model COSMO/MESSy. Geosci. Model Dev. **5**, 87–110 (2012a). doi:10.5194/gmd-5-87-2012

Kerkweg, A., Jöckel, P.: The 1-way on-line coupled atmospheric chemistry model system MECO(n) Part 2: on-line coupling with the Multi-Model-Driver (MMD). Geosci. Model Dev. **5**, 111–128 (2012b). doi:10.5194/gmd-5-111-2012

Lynch, P., Lynch, O.: Forecasts by PHONIAC. Weather **63**, 324–326 (2008)

Pozzer, A., Jöckel, P., Kern, B., Haak, H.: The atmosphere–ocean general circulation model EMAC-MPIOM. Geosci. Model Dev. **4**, 771–784 (2011). doi:10.5194/gmd-4-771-2011

Reithmeier, C., Sausen, R.: ATTILA—atmospheric tracer transport in a Lagrangian model. Tellus B **54**(3), 278–299 (2002)

Stenke, A., Grewe, V., Ponater, M.: Lagrangian transport of water vapor and cloud water in the ECHAM4 GCM and its impact on the cold bias. Clim. Dyn. **31**, 491–506 (2008). doi: 10.1007/s00382-007-0347-5

Stenke, A., Dameris, M., Grewe, V., Garny, H.: Implications of Lagrangian transport for simulations with a coupled chemistry-climate model. Atmos. Chem. Phys. **9**, 5489–5504 (2009)

Evaluating Climate-Chemistry Response and Mitigation Options with AirClim

36

Volker Grewe and Katrin Dahlmann

Abstract

The evaluation of climate change mitigation options addresses the whole air traffic system. Any optimization with respect to climate change requires a representation of this system and hence a simplification of the individual components and models. AirClim is such a model for simplified evaluation of the approximate chemistry-climate impact of air traffic emissions. The model represents the major responses of the atmosphere to emissions in terms of composition and climate change. The model is used to evaluate both the mean response and the uncertainty range of the climate impact of any change in the air traffic system. The uncertainty range is derived by a Monte-Carlo simulation using random variations of uncertain model input parameters. This uncertainty range is found to be much smaller than the uncertainties in knowledge of the air traffic climate impact in general.

36.1 Introduction

Climate change is a challenge to society and mitigation of climate change plays an important role. Mitigation options cover a wide range from technical to economic measures. They include fossil fuel replacements, like fuel production from algae or

V. Grewe (✉) · K. Dahlmann
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Volker.Grewe@dlr.de

K. Dahlmann
e-mail: Katrin.Dahlmann@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_36,

hydrogen production from wind energy parks, more fuel efficient aircraft, less climate sensitive air traffic routings, and emission trading schemes. But how is the effectiveness of such a measure evaluated? Actually, the interrelation between human activity and the response in terms of climate change is extremely complex. One important aspect of human activity is the emission of species which act as greenhouse gases or have the potential to alter the composition of the atmosphere and thereby change Earth's radiation budget. Directly emitted greenhouse gases like carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), or chlorofluorocarbons (CFCs), have mostly a lifetime in the order of decades, since the loss processes are slow. These species are quite well-mixed since the atmosphere is turbulent and one important implication is that the impact of an emitted species on climate is independent of its location, making it easier to assess its climate impact.

Besides direct emissions of greenhouse gases, there are a number of indirect impacts from anthropogenic emissions, like emissions of nitrogen oxides ($NO_x = NO + NO_2$), which affect ozone and methane, and emissions of water vapor, which lead to contrail and contrail cirrus clouds. Emissions of nitrogen oxides are a consequence of many anthropogenic activities, since much of daily life is based on combustion processes, e.g., transport, heating, and electricity generation. During combustion atmospheric oxygen (O_2) is dissociated and reacts with otherwise stable atmospheric nitrogen to form the radicals NO and NO_2. Nitrogen oxides (NO_x, the sum of NO and NO_2) are a prerequisite for tropospheric ozone production, and ozone is again an important greenhouse gas. Compared to the direct emission of greenhouse gases, this indirect alteration of ozone by nitrogen oxide emissions is very complex. The lifetime of such an alteration depends on both the lifetime of the primary emission (NO_x) and the lifetime of ozone itself. The lifetime of NO_x depends on chemical transformation to HNO_3 and subsequent rain, washing it out of the atmosphere and leading to acid rain or to direct deposition on the ground. The lifetime of ozone is mainly determined by chemical destruction and varies between hours and months. Chemical regimes are very different, with fast and slow conversions between summer and winter, day and night, and between the surface and tropopause altitude, respectively. Besides nitrogen oxide emissions, contrails play a major role, and the formation of contrails in ice supersaturated regions, their lifetime, and their impact on climate are highly variable, both in space and time. When we look at the contribution of air traffic on climate change, we see that a large part arises from these indirect effects. This is one of the reasons why assessments of the contribution of air traffic to climate change are characterized by large uncertainties.

This high spatial and temporal variability in the contribution of air traffic emissions to climate creates difficulties and challenges, but also opportunities to mitigate the climate impact of air traffic. The variability gives the possibility to avoid times and regions where emissions have a large climate impact by re-routing air traffic horizontally and vertically. But it is also a major obstacle to the precise assessment of the climate impact of air traffic and it is a challenge to evaluate mitigation options because of large uncertainties. Here the model AirClim is introduced, with which air traffic climate change mitigation options can be

assessed and which takes into account these uncertainties. An example for an application of AirClim is presented in Sect. 36.7.

36.2 From an Emission to Global Warming

36.2.1 Emissions of Long-lived Species: Carbon Dioxide

Any species emitted into the atmosphere remains in the atmosphere until it is either photochemically converted, washed out by rain, or deposited at the ground, biosphere or in the ocean. Some species, e.g., carbon dioxide, are however not permanently removed from the atmosphere, but stored in a reservoir and re-emitted to the atmosphere after some time. All these processes have to be taken into account when simulating the impact of an emission based on its concentration. Calculations with detailed models are used to deduce the lifetimes of these processes. This is combined with a general approach which describes the fate of an emission: the response x(t) of a system at a time t caused by sufficiently small perturbations e(t′), e.g., an emission at time t′, is approximated with linear perturbation theory using the impulse-response (Green's) function G_x. The function G_x basically includes the characteristic lifetimes of the system and describes the change in x at time t caused by a unit emission at time $t' = 0$ (Hasselmann et al. 1997; Sausen and Schumann 2000):

$$x(t) = \int_{t_0}^{t} G_x(t - t')\, e(t')\, dt'. \tag{36.1}$$

Climate simulations showed that the climate response can be sufficiently described by a sum of five exponential terms with different lifetimes τ_j and respective weighting factors α_j *(Sausen and Schumann* 2000). Hence, by using.

$$G_C(t) = \sum_{j=1}^{5} \alpha_j e^{-t/\tau_j}, \tag{36.2}$$

a CO_2 concentration change $\Delta CO_2(t)$ of an emission scenario E can be calculated:

$$\Delta CO_2(t) = \int_{t_0}^{t} G_C(t - t')\, E(t')\, dt'. \tag{36.3}$$

The lifetimes vary between 1.7 years and infinity. The latter refers to an irreversible climate change caused by changes in the oceanic carbon uptake in a warmer climate.

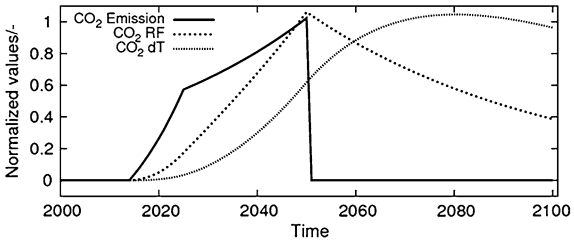

Fig. 36.1 Nondimensional temporal development of CO_2 emissions (solid), radiative forcing (*dashed*) and temperature (*dotted*) for an emission scenario which shows an academic aircraft fleet with an entry into service in 2015, a strong increase in fleet size until 2025, a further, but slower, increase until 2050, and total termination of fleet use in 2050 [from Grewe and Stenke (2008)]

The change in CO_2 concentration alters the radiative balance of the atmosphere (radiative forcing, RF). The RF of this concentration change is estimated on the basis of climate simulations used to determine the responses of the climate system, e.g., to a doubling of the CO_2 concentration. The relation between concentration changes and its RF normally includes the background CO_2 concentration and saturation effects, e.g., the RF is less and less effective the higher the CO_2 concentration. For a carbon dioxide emission scenario this procedure provides a time series of annual mean radiative forcing values.

The respective climate response in terms of near-surface temperature changes can be again derived with impulse-response functions (see Eq. 36.1) with a lifetime of around 30 to 40 years, which describes the inertia of the ocean–atmosphere system (Sausen and Schumann 2000):

$$\Delta T(t) = \int_{t_0}^{t} G_T(t - t')\, RF^*(t')\, dt', \tag{36.4}$$

where RF* is the RF normalized to a doubling of CO_2 and G_T the respective Green's function for the atmosphere–ocean system.

The general response of the atmosphere–ocean system to an emission of CO_2 is shown in Fig. 36.1. The dashed lines show an arbitrary academic emission scenario which describes an aircraft fleet with an entry into service in 2015, a strong increase in fleet size until 2025, a further, but slower, increase until 2050, and total termination of fleet use in 2050. For small CO_2 emissions, as assumed here, the RF is linearly proportional to the CO_2 concentration change, increases from 2015 to 2050, reaches its peak in 2050, and only slowly decreases afterwards due to the long lifetime of the perturbation. The emission peak coincides with the peak in RF in 2050, whereas the inertia of the atmosphere–ocean system delays the peak in temperature change by around 30 years. This has an important implication for any political measures to, e.g., reach a 2 °C goal in 2050, since any measure of CO_2

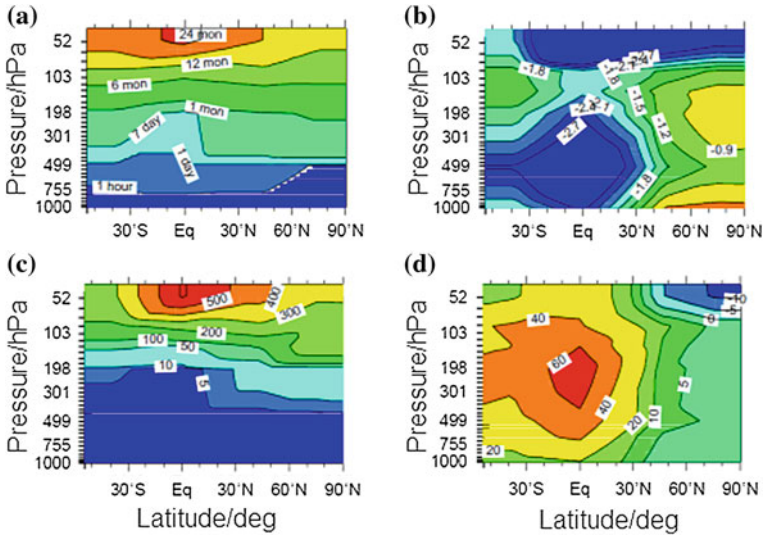

Fig. 36.2 Water vapor perturbation lifetime (**a**), percentage methane lifetime change (**b**), and radiative forcing at the tropopause for water vapor (**c**) and ozone perturbations (**d**) normalized to the same total annual emission of 1 Pg water vapor and 1 Tg of NO_x as N mass in mW/m^2 [from (Grewe and Stenke 2008)]

requires a minimum 30 years advance, and their effect will still be strongly controlled by past emissions. Hence to reach the 2 °C goal by 2050 with measures aiming at CO_2, they have to go into effect well before 2020.

36.2.2 Short-Term Air Traffic Impacts

The description of the impact of a CO_2 emission (Sect. 36.2.1) shows that both the perturbation lifetime of a species and the response time of the atmosphere–ocean system control the temporal evolution of the temperature response to an emission. Air traffic not only emits CO_2, but also water vapor, NO_x, CO, unburned hydrocarbons and soot, and thereby alters the concentration of water vapor, ozone and methane and influences cloudiness by contrails, contrail cirrus and soot contrails. The lifetimes of these perturbations are highly variable and largely depend on their location.

Figure 36.2 shows a summary of results from 24 climate-chemistry model simulations. In each simulation constant NO_x and water vapor emission was prescribed in a specific altitude range and latitudinal band, covering the northern hemisphere polar regions, the mid-latitudes tropics and southern hemisphere mid-latitudes, and regions between the surface and 20 km in altitude. Water vapor emitted near the surface has a lifetime of less than an hour (Fig. 36.2a) and hence does not have the potential to significantly impact climate change, whereas at altitudes above 12 km ($\sim$200 hPa) the water vapor perturbation lifetime exceeds a

month and reaches 2 years at 20 km (~50 hPa) in the tropics. The radiative forcing from emitted water vapor in the respective latitude band (Fig. 36.2c) shows an insignificant impact in the troposphere and an increase in the stratosphere from around 10 to more than 500 mW/(Pg(H_2O)/year) (1 Pg = 10^{15} g).

Under most conditions, NO_x emissions lead to an increase in ozone and a decrease in methane, since the most relevant chemical reactions are Reaction 1

$$NO + HO_2 \rightarrow NO_2 + OH,$$

followed by photolysis of NO_2 and a recombination of the produced atomic oxygen (O) with molecular oxygen (O_2) to form ozone (O_3) and Reaction 2

$$CH_4 + OH \rightarrow CH_3 + H_2O,$$

where the produced OH from reaction (36.1) leads to a loss of methane. However, the reactivity of the reactions and the competition with other reactions leads to regionally significantly different effects on the ozone and methane prompted by an emitted NO_x molecule. Furthermore, the lifetime of ozone and methane and consequently the effect on radiative forcing varies regionally. This leads to different patterns of the effect of emitted NO_x on the change of methane and ozone (Figs. 36.2b and d): in the troposphere, the negative RF due to methane destruction is strongest in the tropics at around 5 km (~500 hPa), whereas the positive ozone RF is strongest at around 12 km (~200 hPa).

Simplified climate-chemistry response models which are applied to air traffic investigations on climate change have to account for non-CO_2 emissions, since these effects are far too relevant to be ignored. Ponater et al. (2006) used detailed model calculations to obtain the RF of individual components (ozone, contrails) and inserted those results in Eq. 36.4 to obtain the associated temperature changes. This provides a good overview of the importance of individual components. However, this methodology is limited in applicability, since no changes in routing, aircraft design, etc. can be evaluated without additional detailed simulations with climate-chemistry models. Hence the climate-chemistry response model AirClim (Sect. 36.3) was developed to close this gap.

36.3 The Concept of the Climate-chemistry Response Model AirClim

The path from an emission to climate change includes many complex processes (see Sect. 36.1). Evaluating climate change mitigation options or optimizing air traffic on a system-wide level demands a fast, but reliable assessment model. Climate-chemistry models are far too time-consuming with respect to computing time. That led to the idea to develop and apply a response model which reproduces the response of a climate-chemistry model without actually calculating all the physical and chemical effects.

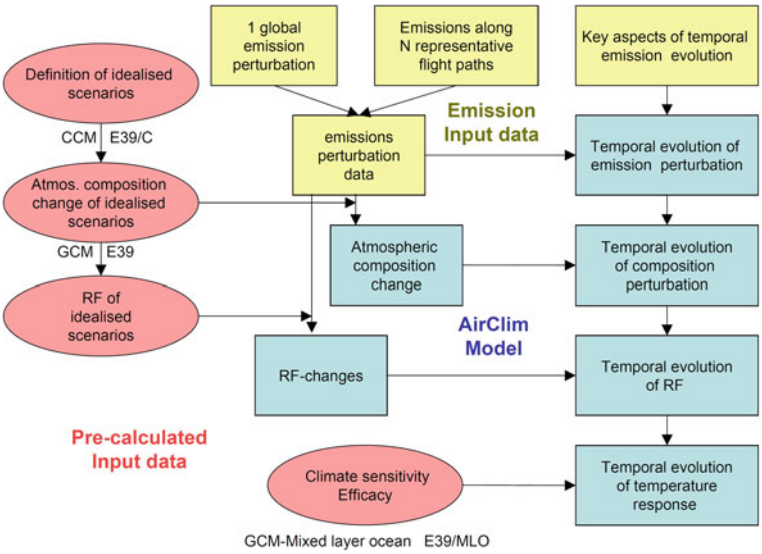

Fig. 36.3 Overview of the AirClim model (Grewe and Stenke 2008)

An overview on the AirClim model is given in Fig. 36.3. The model combines precalculated data (red), partly described in Sect. 36.2.2, with emission data (input to AirClim, yellow) and provides atmospheric composition changes, RF and finally the temporal evolution of temperature changes (Sect. 36.2.1, blue). Figure 36.4 shows the basic mechanism applied in AirClim for those species. The detailed climate-chemistry model E39/C is applied to a number of different point emissions which differ in altitude and latitude (left part). The relation between the emission location and the response of the model (here in terms of radiative forcing) forms the basis for the AirClim response model. These relations are calculated for H_2O, NO_x (leading to ozone production and methane destruction), and contrails. Hence for any emission $E^{species}(t')$ of an individual species at time t' [and a varying location x(t')] in the time period $[T_1,T_2]$ the response ΔC is calculated by

$$\Delta C^{species} = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} E^{species}(t') \sum_{k=1}^{4} \varepsilon_k(t') \frac{\Delta C_{id}^{species}}{E_{id}^{species}} dt', \tag{36.5}$$

where the indices id refer to the concentration change and emission of the pre-calculated simulations of the idealized emission scenario and ε_k are the weighting factors for the interpolation of the precalculated values onto the actual emission location (see Fig. 36.4). The time interval $[T_1,T_2]$ can either represent the duration of an individual flight or annual emission. This means that the response of a specific emission scenario is then the respective combination of the responses of the point emissions. The respective RF can be derived accordingly by replacing

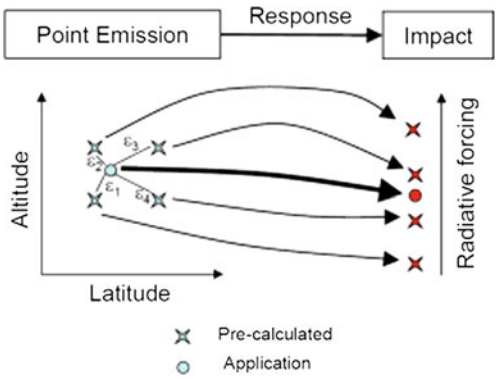

Fig. 36.4 Principle mechanism of the response model AirClim. Stars indicate precalculated response relations derived with a detailed climate-chemistry model, i.e., the RF (=response, impact) caused by an emission at a certain latitude and altitude. Any given emission at an arbitrary latitude and altitude (*circle*) can be represented by a linear combination (ε) of the predefined emissions, and the respective response is then the same linear combination of the precalculated responses

ΔC with RF. This describes the response of short-lived species (lifetime up to a year), where the location highly affects the response. The temporal development of the response is calculated according to the temporal development of the respective emission. AirClim is designed to take into account two types of information: a general air traffic emission scenario E_{ann} with annual mean global values over a period of decades (Fig. 36.3 "Key aspects") and a detailed 3D emission data set $E^{species}$ for one specific year (=t_{norm}) (Fig. 36.3, "emission perturbation"). For very long-lived species like CO_2, the location is unimportant, but the response is spread over years (see also Fig. 36.1). Other perturbations, like the perturbation of stratospheric water vapor or changes in stratospheric NO_y and hence ozone, may have lifetimes exceeding a year, but much shorter than a decade. For those species the lifetime is taken into account by a linear differential equation

$$\frac{d}{dt}\Delta C^{species} = P(t) - \tau_{species}^{-1}\Delta C^{species}, \tag{36.6}$$

where the lifetime τ is derived in analogy to Eq. 36.5. The production term $P(t) = s\ E_{ann}(t)$ of the concentration change is proportional to the emission scenario $E_{ann}(t)$ with a proportionality factor s, which guarantees that at time $t_{norm} + \tau_{species}$ the concentration calculated in Eq. 36.5 is identical to the concentration in Eq. 36.6. That means that with Eq. 36.5 a snapshot of the concentration change and the respective RF at a certain time is calculated, whereas Eq. 36.6 provides the temporal development.

Methane is a species which has a lifetime of around 8 to 9 years. Hence, it does not fit into one of the categories, neither short-term effects, like contrails and ozone nor long-term effects like carbon dioxide. The lifetime is short enough that the location of

a NO_x emission significantly impacts the methane depletion, i.e., methane lifetime change (Fig. 36.2b). On the other hand the lifetime is too long to calculate a large number of response functions for the methane change, since the detailed simulations would be too CPU-time-consuming. Hence, not the methane concentration change caused by a NO_x emission, but the change in its lifetime δ (percentage change) is calculated, which can be derived with only short-term (e.g., 5 year) simulations. The temporal evolution of the lifetime change $\delta(t)$ is derived by scaling the lifetime for the regarded snapshot with the emission scenario. In AirClim the temporal evolution of the change in methane caused by a change in its lifetime is calculated with a differential equation Eq. 36.9, which is derived as the difference of two differential equations, for the unperturbed (Eq. 36.7) and perturbed situation (Eq. 36.8).

$$\frac{d}{dt} C^{CH_4}(t) = P(t) - (\tau_{CH_4})^{-1} \, C^{CH_4}(t) \tag{36.7}$$

$$\frac{d}{dt}\left(C^{CH_4}(t) + \Delta C^{CH_4}(t)\right) = P(t) - ((1+\delta(t)) \, \tau_{CH_4})^{-1} \left(C^{CH_4}(t) + \Delta C^{CH_4}(t)\right) \tag{36.8}$$

$$\begin{aligned} \frac{d}{dt}\Delta C^{CH_4}(t) &= \left(\frac{\delta(t)}{1+\delta(t)}\right) \tau_{CH_4}^{-1} \, C^{CH_4}(t) - \left(\frac{1}{1+\delta(t)}\right) \tau_{CH_4}^{-1} \, \Delta C^{CH_4}(t) \\ &= \left(\frac{1}{1+\delta(t)}\right) \tau_{CH_4}^{-1} \left(\delta(t) \, C^{CH_4}(t) - \Delta C^{CH_4}(t)\right) \end{aligned} \tag{36.9}$$

With this approach the temporal evolution of concentration changes caused by an emission scenario, and further the temporal evolution of the respective RF which then causes a change in near surface temperatures, are calculated, see Sect. 36.2.1 and especially Eq. (36.4).

36.4 Verification of the AirClim Model

It is important to stress that there is a difference in the concept of the detailed models and their response models. Detailed models are based on physical laws and resolve physical processes. Response models do not resolve any physical processes; instead, the response is extracted. This limits the applicability of the response model to those set-ups which were chosen in the precalculation of the basic response functions for AirClim. For example, a steady emission throughout the year was chosen for deriving the response relations, i.e., the climate-chemistry model resolves individual weather situations throughout the years, and the impact of the targeted emissions for all these weather situations is calculated as an annual mean response. Hence, the response model AirClim implicitly considers the variability of the weather situation. However, there is no possibility to resolve any response at a time-scale smaller than a year, i.e., for an individual weather pattern.

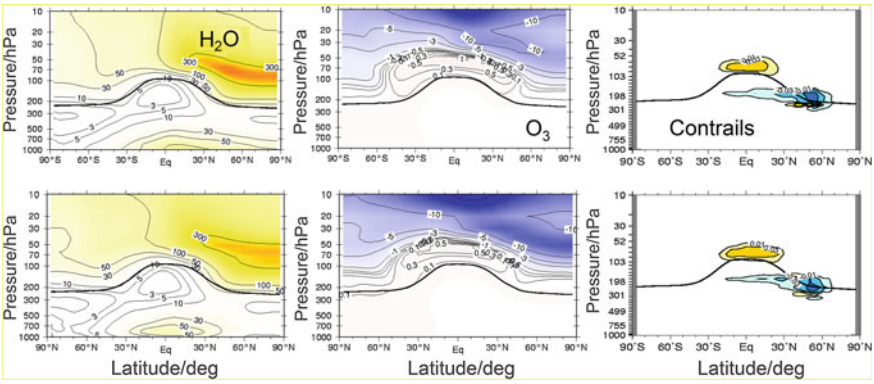

Fig. 36.5 Simulated distributions of water vapor (nmol/mol), ozone (nmol/mol) and contrail coverage (%) for a partial substitution of a subsonic fleet by supersonic aircraft. *Top*: climate-chemistry model; *bottom*: AirClim (for details see Grewe and Stenke 2008)

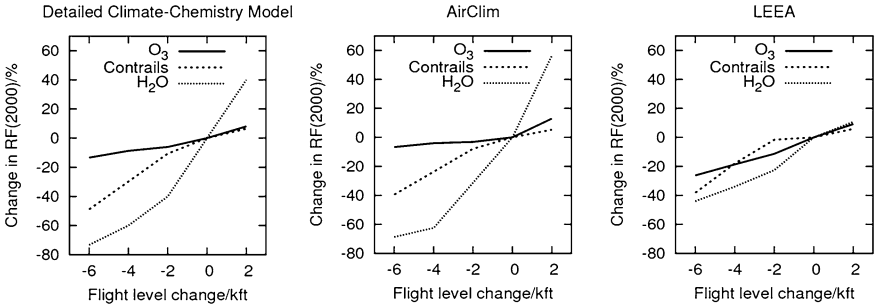

Fig. 36.6 Sensitivity of RF for the year 2000 for a general change in flight altitude based on TRADEOFF data for a detailed climate-chemistry simulation (*left*), AirClim (*mid*), and LEEA model (*right*) for the components ozone (*solid*), linear contrails (*dashed*) and water vapor (*dotted*) (Fichter 2009)

A necessary step is a verification of this modeling approach. This implies testing whether for a given emission scenario the response simulated with the climate-chemistry model is reasonably reproduced with AirClim. Naturally, this emission scenario has to be independent of the point emissions used to derive the response function. This was performed with a supersonic emission scenario (Fig. 36.5), showing impressively the possibilities of response modeling to reproduce the more complex results of the chemistry-climate model.

Recently, an updated version of AirClim (AirClim 2.0) has been developed (Dahlmann 2011) which includes, beside the effect of CO_2, the effects of water vapor emissions, contrail formation and its transformation into cirrus, as well as NO_x impacts on ozone production and methane destruction, and the longer-lived ozone reduction arising from methane depletion, as well as higher resolution in the

subsonic cruise levels (Fichter, 2009). Figure 36.6 illustrates that AirClim (center) well reproduces the sensitivity of the RF for general changes in flight altitudes (varying between −6 000 feet and +2 000 feet) simulated by the detailed model (left). An intercomparison with the method developed within the low emissions effect aircraft project (LEEA) (Köhler et al. 2008) shows that the sensitivity of the results to the underlying base model is small for ozone and contrails, but largely differs for water vapor, though the absolute water vapor effect is small.

36.5 Evaluating Mitigation Options for Air Traffic

AirClim is designed to be used for evaluating climate change mitigation options for air traffic. But what exactly do we understand by the wording "climate change"? Is there a single parameter describing climate change? No, clearly climate change manifests itself in a multitude of parameters, like sea level rise, precipitation changes, changes in extreme weather events and increase in global mean near-surface temperature, to only name some prominent features. However, it is widely accepted that the global mean near-surface temperature change is a valid, stable and well established proxy for climate change. The basic causal relationship of a climate relevant emission can be briefly summarized as follows (Fuglestvedt et al. 2010): An emission changes the atmospheric composition, which might include cloudiness and cloud properties, which in turn changes the way solar radiation travels through the atmosphere warming Earth's surface and how Earth radiates to space. The changed radiation budget leads to changes in climate, e.g., to increasing temperatures. This chain of processes is represented in climate and chemistry-climate models. Physical climate metrics are a short-cut to this extremely complicated process and directly relate emissions to climate change.

The simplest metric is the emitted mass of a species, e.g., CO_2. However, this includes no information on climate impacts, though it is widely used. Instead, the metric radiative forcing (RF) is basically the radiation change caused by a concentration change due to by past emissions, relative to pre-industrial times. This metric is linked much closer to the actual atmospheric response. However, it does not yet consider the effect on, e.g., temperature. The *global warming potential* (GWP, absolute in units W/m^2 or relative to CO_2 as a dimensionless factor) sums up future impacts on radiation changes from today's concentration change to a given time horizon. The choice of the time horizon is not based on physical considerations and in literature has been subjectively chosen between 20, 50 and 100 years. The global temperature potential (GTP) (again absolute or relative to that of CO_2) translates the radiation changes caused by a concentration change in a temperature change at a certain point in time by using response functions. GTP results in a physical quantity which is directly associated with climate change. The GTP depends critically on the chosen time horizon. The dependency on the time horizon is largely reduced by using the average temperature response (ATR),

which is the mean future temperature development. This overview of climate metrics reveals an important and often underestimated aspect in the evaluation of mitigation options and in the evaluation of the impact of emissions on climate: all these metrics somehow target "climate change", but they provide different physical quantities for measuring climate change and hence they provide answers to different questions. The underlying emission scenario is equally important. Addressing the question "What is the climate impact of air traffic," requires an emission scenario from 1750 to 2020 and investigating the contribution to temperature changes, whereas the question "What is the climate impact of an emission today?" requires a pulse emission and investigation of the contribution to future climate change.

Hence it is important to first clearly formulate the objective. This very much defines what kind of climate metric and emission scenario has to be chosen. Not the climate metrics are ambiguous, but the ill-posed underlying questions often lead to confusion. Consider the question: "what is the current contribution of air traffic to anthropogenic climate change?" This implies that we have to consider all past air traffic emissions and evaluate today's temperature change (absolute GTP) caused by these emissions. All other combinations of emission scenarios and above listed metrics would not address the question appropriately.

The ultimate objective of evaluating mitigation options is to reduce the climate change due to aviation emissions in the future. It might not be important if the climate change takes place at the beginning or at the end of a time horizon. Therefore, the ATR is an adequate metric for evaluating mitigation options as it is the mean temperature change over a time horizon but less dependent on this time horizon than, e.g., GTP. The underlying emission scenarios are then a base case representing past emissions and a best estimate for future emissions, and respective mitigation scenarios.

36.6 Why Even Large Uncertainties in Atmospheric Processes Are Not Limiting Application of Climate Optimization in Aircraft Design

There are many uncertainties in calculating the climate change of aviation emissions due to parameterization and partly insufficient understanding of atmospheric processes which result in different model results. An overview of the uncertainties is given, e.g., in Grewe and Stenke (2008); Lee et al. (2009); and Lee et al. (2010). The main sources of uncertainties are related to the calculation of lifetime (τ), radiative forcing (RF) and *efficacy* (λ), i.e., to the relation between RF and the temperature response of different climate agents. To account for these uncertainties, AirClim can be run inside a Monte Carlo simulation, which then propagates the uncertainties in the input parameters to an uncertainty in the outcome, i.e., temperature change. A Monte Carlo simulation is a stochastic method to approximate numerical problems by numerous repetitions of a random experiment (e.g., 10 000 times). For each random experiment, a random number within a given

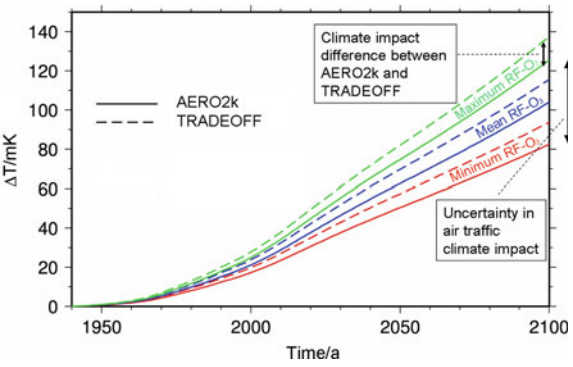

Fig. 36.7 Temporal development of temperature changes due to two different emission inventories and three different assumptions for the strength in radiative forcing of ozone

uncertainty range and distribution function taken from literature is chosen for each of the uncertainty parameters (τ, RF and λ) and AirClim starts with this parameter setting. Thereby, a broad range of possible parameter settings are calculated. The spread in possible results is very wide due to the large uncertainties. However, it should be noted that the uncertainties are not statistical uncertainties which arise from atmospheric variability, but uncertainties arising from different model studies and processes which are less understood. That means that, for instance, the efficacy is 30 % lower than the one used in AirClim, but this applies to every calculated scenario. This does not matter for analyzing the uncertainty of absolute climate change for one scenario, but it is important for analyzing the differences in climate change for different scenarios.

To illustrate this concept, Fig. 36.7 shows the temporal development of temperature change due to two different emission inventories (TRADEOFF and AERO2k) and three different assumptions for the strength of radiative forcing of ozone. The uncertainty in absolute contribution to climate change in 2100 is large and the difference between the absolute climate changes of both emission inventories is not significant. Analyzing the difference in climate change between both emission inventories for each assumption separately shows that TRADEOFF causes a higher climate change than AERO2k for each assumption, leading to a statistically significant difference between both emission inventories. For quantifying the mitigation gain, the important information is whether scenario A or scenario B has lower climate impact, and not the exact absolute difference. Thus the difference between both scenarios is calculated for each parameter setting separately, i.e., for each step of the *Monte Carlo* simulation which leads to lower uncertainties.

To gain statistically significant results, it is important to analyze the relative difference between the climate impact of a scenario and a reference scenario for each step of the Monte Carlo simulation. The reference scenario can either be one of the calculated scenarios (for example the scenario of current aircraft or routing) or the mean of all calculated scenarios, depending on the question to be answered.

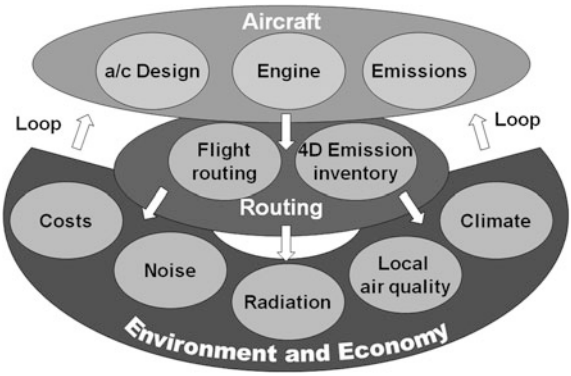

Fig. 36.8 Sketch of a multi-objective optimization loop for air traffic

36.7 An Example for an Application of AirClim: Optimizing Supersonic Business Jets

36.7.1 Basic Components of Climate Optimization Systems for Aircraft Design

The optimization of an aircraft with respect to its climate impact requires a set of numerical models representing three different research areas: the aircraft itself, the routing or air traffic management, and the environment and economy (Fig. 36.8). At the aircraft design level, a proper description of the aircraft and engine flight performance, e.g., the aerodynamics, fuel consumption and specific emissions such as nitrogen oxides, soot, etc., are important issues. The aircraft design process includes, among other aspects, the definition of the aircraft's outer geometry, cabin layout, mass breakdown, thrust requirements and flight performance. Every aircraft is designed for a set of performance requirements that include among others the flight speed, operating range and payload capacity. By varying the free design parameters, e.g., engine thrust, wing area, etc., the best aircraft configuration (e.g., minimum cost or minimum fuel) is derived for the given design task. The flight performance of an aircraft is thus the result of an optimization process involving the relevant design parameters, which ultimately has a direct influence on the environmental impact of the aircraft through its fuel consumption and emitted exhaust products. For the assessment of climate impact and operational mitigation options, descriptions of aircraft mass, aerodynamic and propulsion characteristics are sufficient for the first step.

In the next step, air traffic (routing) is simulated with a set of flights. Based on regional passenger demand, a global route network is developed including data concerning city pairs, flight schedules, annual frequencies and aircraft types. Each flight in the network is simulated with a trajectory calculation module based on given vertical and lateral flight profiles. Free design parameters include cruise speed and altitude. The resulting 4D (space and time) trajectories are assembled to

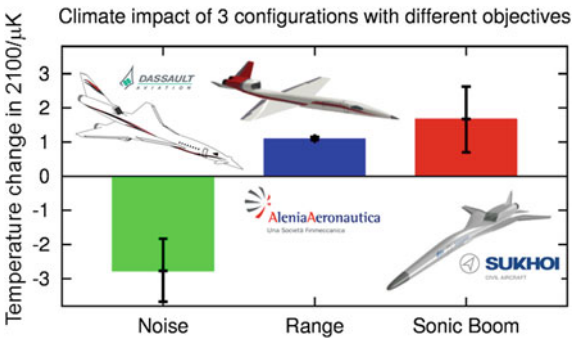

Fig. 36.9 Climate impact of three supersonic aircraft configurations. The colored bars represent the climate impact in terms of temperature change in the year 2100. The error bars indicate the uncertainty range of the temperature difference between the individual configuration and the mean temperature response of all three configurations (see Sect. 36.7)

a global fleet emission inventory which serves as input for calculating the environmental impact. The economic impacts such as costs for flight crew, fuel and overhaul are determined based on the flight time and fuel consumption for each trajectory. In general, environmental impacts are manifold; here the focus is on the climate impact simulated with the response model AirClim.

36.7.2 Optimized Aircraft Design

The challenge of multi-objective optimization was addressed in the European project HISAC (2005–2009), which aimed at a green supersonic small-scale aircraft, i.e., a supersonic business jet. Besides minimal climate impact, three different objectives were to be met: noise reduction during landing and take-off, optimal aircraft range, and sonic boom avoidance. Three different aircraft manufactures designed an aircraft configuration to meet one objective. All three designs were based on common specifications like number of passengers (=8, option up to 19), subsonic cruise mach number (MN) of 0.95, maximum speed of 1.6 MN, maximum altitude FL 510 ($\sim$16 km), height seating of 1.785 m, entry into service by 2015, and a 20-year lifetime (Grewe et al. 2010). The common specification was designed in three loops to obtain a low climate impact, which resulted in a relatively low flight altitude compared to earlier concepts as the impact of water vapor becomes more important the deeper the aircraft flies in the stratosphere, since it accumulates and finally controls the climate impact of a supersonic aircraft. For each loop the individual aircraft designs were then optimized with respect to their additional objective, i.e., noise reduction, range, and sonic boom mitigation. This led to three different designs (Fig. 36.9). Only minor differences exist in climate impact, since the common specification limits variability.

36.8 The Future: Modeling Climate Response to Traffic Emissions

The climate response model AirClim was initially designed for emissions from supersonic aircraft. It is now updated to also cover subsonic aircraft emissions. As a next step of development it is planned to adapt the model to be applicable to emissions of surface transport (cars, trucks, ships, etc.). Hence, the development of AirClim is continuing and will evolve into a new model, "TransClim." This adds further scientific challenges. The contribution of surface traffic emissions to the concentration of species like NO_x, CO, and hydrocarbons are much larger than for the aircraft sector and the chemical system is highly nonlinear, so more complex response functions have to be considered.

References

Dahlmann, K.: Eine Methode zur effizienten Bewertung von Maßnahmen zur Klimaoptimierung des Luftverkehrs. Ph.D., Ludwig-Maximilians-Universität, Munich (2011)

Fichter, C.: Climate impact of air traffic emissions in dependency of the emission location. PhD thesis, Manchester Metropolitan University (2009)

Fuglestvedt, J.S., Shine, K.P., Berntsen, T., Cook, J., Lee, D.S., Stenke, A., Skeie, R.B., Velders, G.J.M., Waitz, I.A.: Transport impacts on atmosphere and climate: metrics. Atmos. Environ. **44**, 4648–4677 (2010). doi:10.1016/j.atmosenv.2009.04.044

Grewe, V., Stenke, A.: AirClim: an efficient climate impact assessment tool. Atmos. Chem. Phys. **8**, 4621–4639 (2008)

Grewe, V., Plohr, M., Cerino, G., Di Muzio, M., Deremaux, Y., Galerneau, M., de Saint Martin, P., Chaika, T., Hasselrot, A., Tengzelius, U., et al.: Estimates of the climate impact of future small-scale supersonic transport aircraft—Results from the HISAC EU-project. Aeron. J. **114**, 199–206 (2010)

Hasselmann, K., Hasselmann, S., Giering, R., Ocana, V., Storch, H.V.: Sensitivity study of optimal CO2 emission paths using a simplified structural integrated assessment model (SIAM). Clim. Change **37**, 345–386 (1997). doi:10.1023/A:1005339625015

Köhler, M. O., Rädel, G., Dessens, O., Shine, K. P., Rogers, H. L., Wild, O., Pyle, J. A.: Impact of perturbations to nitrogen oxide emissions from global aviation. J. Geophys. Res. **113**, D11305 (2008). doi:10.1029/2007JD009140

Lee, D.S., Fahey, D.W., Forster, P.M., Newton, P.J., Wit, R.C.N., Lim, L.L., Owen, B., Sausen, R.: Aviation and global climate change in the 21st century. Atmos. Environ. **43**, 3520–3537, (2009)

Lee, D.S., Pitari, G., Grewe, V., Gierens, K., Penner, J.E., Petzold, A., Prather, M.J., Schumann, U., Bais, A., Berntsen, T., et al.: Transport impacts on atmosphere and climate: Aviation. Atmos. Environ. **44**, 4678–4734 (2010)

Ponater, M., Pechtl, S., Sausen, R., Schumann, U., Hüttig, G.: Potential of the cryoplane technology to reduce aircraft climate impact: a state-of the-art assessment. Atmos. Environ. 6928–6944 (2006). doi:10.1016/j.atmosenv.2006.06.036

Sausen, R., Schumann, U.: Estimates of the climate response to aircraft CO2 and NOx emission scenarios. Clim. Change **44**(1–2), 27–58 (2000)

Part III

Research Trends

The Transition From FALCON to HALO Era Airborne Atmospheric Research

37

Monika Krautstrunk and Andreas Giez

Abstract

DLR's FALCON aircraft has been successfully conducting airborne atmospheric research for 35 years. HALO, a German research aircraft with extended and enhanced capabilities, is preparing to become the new European airborne research flagship for the next few decades. This chapter briefly describes the history behind the HALO project. It introduces the technical and operational features of this new research aircraft in comparison to the FALCON and presents the scientific research planned on HALO for the near future.

37.1 Introduction

In the mid-1970s the Institute of Atmospheric Physics (IPA) initiated a process to acquire an aircraft for high-altitude atmospheric research. Airborne meteorological research at that time was concerned with cloud physics and atmospheric dynamics. Scientists at IPA were using military aircraft such as the Lockheed T33 or the Canberra to investigate the upper *troposphere*. However, researchers were in need of a fast aircraft with a higher ceiling and long-range capability to perform measurements in the vicinity of the *tropopause*, a region sparse of observations, and only a corporate-type *jet* could fulfill these requirements. In October 1974 a committee of representatives from various major research institutions charged

M. Krautstrunk (✉) · A. Giez
DLR, Flight Experiments (FX), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Monika.Krautstrunk@dlr.de

A. Giez
e-mail: Andreas.Giez@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_37,

DLR's parent organization, the "Deutsche Forschungs- und Versuchsanstalt für Luft und Raumfahrt e.V.", or DFVLR, with the task of procuring a suitable aircraft. As a result the French Dassault Mystère *Falcon* 20 E twin-jet (FALCON) was selected. The FALCON is a structurally and aerodynamically robust, midsize aircraft that can be flown even in extreme atmospheric conditions such as areas of high *turbulence* levels and *wind shear* (e.g., in the vicinity of *thunderstorms* or in *contrails* behind airliners). The purchase contract was signed with Avions Marcel Dassault in December 1974. The price for the modified aircraft was 9.8 million Deutsch Mark. The *green aircraft* was delivered, that is without paint job, furniture and other customer-specific installations, in Bordeaux on 29 September 1975. After modification in Paris and a flight test period in Istres, France, the new and heavily modified aircraft arrived at DFVLR in Oberpfaffenhofen on 16 July 1976 to begin its successful career under the call sign D-CMET. The FALCON became one of the most famous *research aircraft* in Europe.

During its lifetime, the FALCON was continually adapted to the changing needs of the research community. In 1993 it was fitted with four hard points under its wings (two on each side) for the installation of *instruments* like particle measurement probes. In 1995 the aircraft engines were replaced and two additional electrical generators (2 × 300 A @ 28 V) were installed in order to provide more power to the *experiments*. The new *engines* increased the aircraft's range from 2 800 to 3 700 km and its ceiling by one kilometer to 12.7 km (42 000 ft), which allowed the FALCON to fly above the typical airways of most global *air traffic* and even reach the *lower stratosphere* at higher latitudes. In 1997 a third large viewport was added on the top of the fuselage, directly above one of the two large openings on the bottom of the fuselage, and in 1998 another universal window interface (220 mm diameter) was added on the starboard side.

But even with all these modernizations the FALCON turned out to be too small and too limited in range for many of the new challenges in atmospheric research. One reason is that the investigation of processes in the *atmosphere*, its interactions with the *biosphere* and *oceans*, and the impact of human activities had become more and more complex.

The need to carry more instrumentation to measure a broader range of parameters and at the same time to fly with more instrument operators required a higher payload capability. Today's environmental research projects focus on regions where dramatic *climate changes* can be observed. Examples are the *polar regions* or the boreal zone, which call for a longer-range research aircraft. The processes near the tropopause, e.g., the exchange of trace gases between troposphere and stratosphere, are not yet fully understood, and respective studies require an aircraft with a ceiling of at least 15 km. The demand for a new atmospheric research aircraft were initially postulated in 1999 by climate experts of the German *Helmholtz* Association. In 2 000 representatives of German atmospheric and climate research institutions met to develop a concept for future airborne research in this field. As a result, on behalf of more than 30 German research institutions, the Max Planck Society (MPG) and the German Aerospace Center (DLR) submitted a proposal for a new German research aircraft capable of flying up to altitudes of

Fig. 37.1 *HALO* during modification at RUAG Aerospace. Due to the number of apertures, the *fuselage* structure must be triple-reinforced around the openings to retain its original stiffness

15.5 km, covering a range of more than 8 000 km, and carrying a payload of up to 3 tons. After a successful peer review by a German science council commissioned by the Federal Ministry of Education and Research (BMBF), project High Altitude and Long Range Research Aircraft (*HALO*) was selected as "eligible without any restrictions" in November 2002. On 16 February 2005, DLR signed a sales contract with Gulfstream Aerospace (GAC) at the Federal Ministry of Education and Research. After a short construction time at GAC in Savannah, Georgia (USA) the green aircraft was transferred to Oberpfaffenhofen in April 2006 where it was heavily modified during a 19 month period by RUAG Aerospace in close cooperation with DLR (see Fig. 37.1). Afterwards, the modified aircraft flew back to GAC on 8 December 2007 for further modifications, paint job and certification. HALO was finally delivered to DLR on 24 January 2009 and is now located in a hangar built exclusively for this new aircraft and its user community. The investment in the modified aircraft, a total of 72 million euro, twice the price of the base aircraft, was financed by the Federal Ministry of Education and Research, the Max Planck Society, Jülich Research Center (FZJ), Karlsruhe Institute of

Fig. 37.2 The DLR jets *Falcon* 20 E (*left*) and the new GAC 550 HALO (*right*) on the apron after a flight test of HALO accompanied by the FALCON after coordinated flight testing

Technology (KIT), and the German Aerospace Center, with FZJ, KIT and DLR being part of the Helmholtz Society. Finally, the Free State of Bavaria contributed to the new HALO hangar. Figure 37.2 shows both aircraft.

This article describes the development of airborne atmospheric research within DLR over the last years. The concept of operation and the range of services provided by the DLR Research Flight Facility are explained before important upcoming research projects on this new platform are presented. The article closes with a description of the technical features of the HALO aircraft in comparison with its predecessor.

37.2 From National to International to Global Operation

Over the course of 35 years the FALCON had flown almost 8 500 h, an annual average of 240 h. In the beginning the research focus was mainly national or European, but soon the range of operations expanded to intercontinental destinations. Figure 37.3 shows a map with sites of important FALCON field experiments along with a listing of about 150 major *field campaigns* which have been flown on this aircraft over the years. The list reflects the diversity of research activities with the FALCON, which started with a field experiment in Scotland in 1977.

In addition to these user-driven projects, DLR's Research Flight Facility in Oberpfaffenhofen successfully applied in 1996 for EC funding of FALCON flight hours under the *European Commission* Framework Program 4 (FP4, Training and

Fig. 37.3 *Campaign* bases and campaign acronyms of large international FALCON missions since 1987

Mobility of Researchers) and began to attract new, inexperienced scientists from Europe to use the aircraft. This project, Scientific Training and Access to Aircraft to Atmospheric Research Throughout Europe (STAAARTE), had the major aim of offering access to three large-scale facilities in Europe (the C 130 Hercules of the Meteorological Research Flight in the UK; the ARAT Fokker 27 from INSU, France; and the FALCON from DLR, Germany) to small atmospheric research projects (8–10 h each) by user groups working within the European community but with no comparable infrastructure in their home countries. This transnational access project was so successful that it was extended and expanded until today, operating under different names (STAAARTE, CAATER, EUFAR-TA). The cooperation with two other European operators during these initial transnational access projects eventually initiated a periodic round-table meeting which resulted in the concerted action project European Research Aircraft and Sensors for Environmental Research (EURASER) aiming at closer cooperation between European aircraft operators and the development of a European strategy for research aircraft facilities and sensor applications. This network was the precursor of the EUFAR project (formerly: European Fleet for Airborne Research, then European Facility for Airborne Research), continuously funded by the EC from Framework Program 6 and now with three main pillars: networking, transnational access and joint research activities.

The year 2009 saw a significant decrease in the activities of the FALCON because scientists and engineers inside and outside DLR were concentrating on the preparation of HALO projects and the associated instrumentation. However, in April 2010 during the *volcanic* eruption of *Eyjafjallajökull* in Iceland, the FALCON experienced the highest public attention of its career. DLR's IPA and Research Flight Facility undertook 17 flights above and in areas of increased ash concentrations and took *remote sensing* and in situ *measurements*. The measurements of *ash* and other volcanic emissions (SO_2, CO) in the eruption plume supported the responsible national agencies in their decisions on closure and re-opening of airspace impacted by ash.

37.3 Research Flight Operations

All aircraft operated at DLR are used as "multipurpose platforms"; the scientific equipment and instrumentation is individually assembled for each specific research project. This quick-change configuration allows for constant, state-of-the-art instrumentation. Such a concept of flexibility, proven in many international FALCON campaigns, will also apply to the larger HALO aircraft. To facilitate this concept a general HALO cabin layout has been certified, which displays the maximum weights and loads in each sector of the cabin. A maximum of 15 racks can be placed in the cabin. Typically, a rack can carry up to 150 kg of instrumentation, has its own secondary power distribution box (SPDB), and is fixed on seat rails on the cabin floor and on additional rails on the cabin ceiling for more stability (Fig. 37.4). HALO research projects will require complex payloads for remote sensing and in situ measurements. These instruments and sensors are operated by various groups from different research institutions. Each instrument must fulfill all airworthiness requirements before it can be certified. Due to the high workload involved in these *certification issues* (more than 100 new instruments are planned for HALO in the next few years) DLR's own Approved Design Organization (DOA) shares the time-consuming certification process with other German DOAs. The DLR DOA is responsible for the compatibility of the entire payload. The DOA and the Research Flight Facilities in Braunschweig and Oberpfaffenhofen are integrated in one DLR facility, "Flight Experiments".

The structure of the Research Flight Facility in Oberpfaffenhofen reflects the spectrum of services associated with the aircraft. The Sensor and Data Group is responsible for the development and operation of the basic sensors and for data acquisition, and they also support users in linking their instrumentation to the data acquisition system. The group provides a standard set of post-processed, quality checked data to users promptly after each flight (e.g., flight path, altitude, temperature, *humidity* and more). The engineers and mechanics of the Engineering Group support users in integrating their instruments into the cabin or the external pods. The Engineering Group manufactures interfaces (e.g., fasteners and mounting plates) and also advises the scientists before they begin instrument development as

Fig. 37.4 Scientific instruments in the cabin are usually integrated into standard racks mounted between seat rails on the cabin floor and attached to additional rails above the cabin windows

to aircraft interfaces and airworthiness requirements. The dispatchers of the Flight Operation and Acquisition Group prepare the flights together with the pilots and apply for all permits needed to operate the aircraft with various sensor systems worldwide. They assure that all equipment is on site, in time, and is customs-cleared if the aircraft is to be operated away from its home base. They are the managers on duty during international aircraft campaigns, especially if other aircraft or ground stations are involved and have to be coordinated. All pilots have an academic background in engineering or natural sciences. Moreover, they are qualified test pilots, which enables them to perform the flight testing as part of the certification process. The Maintenance Group is authorized by the German national aviation authority (LBA Luftfahrtbundesamt) to maintain and repair the aircraft of its own fleet. The maintenance staff closely monitors technical issues for the aircraft during worldwide operations. They are authorized to perform maintenance and repair work whenever possible, even at remote sites. They are on board during flights to assist the pilots in case of emergency.

37.4 Preparation of HALO Science Missions

The basic operations of HALO, with 220 flight hours annually, is financed by six partners, most of whom have also contributed to the procurement cost of the aircraft. The details regarding access, responsibilities, and liabilities are included in

individual contracts between operator and individual partners or in a collaboration contract between all partners. The financing partners send one leading scientist to a scientific steering committee which meets four times a year to plan and schedule scientific projects. In addition, a board of trustees representing board members or directors of the financing institutions meets once a year to decide operational matters and consider the financial and political impact of the steering and finance committees' recommendations. When deemed necessary, external experts from aviation and other industries can be brought in to evaluate high-risk decisions.

In November 2009 the BAsic HALO Measurement And Sensor system (BAHAMAS), a completely new basic sensor and data acquisition system developed by the Sensor and Data Group of DLR's Research Flight Facility, was integrated into the aircraft and successfully tested in flight. *Calibration* of the sensors and the entire system with trailing cone flights and tower fly-bys took place in spring 2010. State-of-the-art instrumentation, new sensor concepts and innovative measurement strategies helped to demonstrate calibration accuracies, which were significantly improved over existing DLR standards. Since 2010 BAHAMAS has been used to support all flight tests performed by DLR.

In 2011 work on HALO focused on certification and initial flight-testing of the *big belly instrumentation pod* together with the ventral fin, and of the new standard *trace gas inlet* (TGI). A *ventral fin*, mounted below the tail of the aircraft, compensates for reduced directional stability potentially caused by the pod. *Inlets* provide in situ measurement instruments inside the cabin with air from outside the aircraft. The flight testing proved the ability of the HALO aircraft to reach 51 000 ft even with these modifications. Flight testing will continue through early 2012. After completion of the certification work, scientific missions can begin.

The external modifications to HALO have a significant influence on the *aerodynamic* behavior of the aircraft. Thus, the certification process which has to prove the airworthiness of the modified aircraft was extremely laborious. For example, DLR experts developed a *model* to simulate and predict the vibrational behavior of the aircraft wing with external stores after mechanically exciting the wing on the ground and subsequently measuring the induced vibrations. Fig. 37.5 shows the jacked up HALO surrounded by scaffolding and plastered with vibration sensors, which are then linked by cables to the computers of the aeroelastic experts. These measurements with full and empty fuel tanks took several weeks; the development of a vibration model took many months.

In October/November 2010, the high potential of HALO was first demonstrated with a "Techno-Mission" with mainly technical but also some scientific objectives. Four different science instruments from Jülich Research Center; Leipzig, Mainz and Wuppertal Universities; and IPA were installed aboard HALO. IPA provided a novel ion trap *mass spectrometer* and a new High Spectral Resolution Lidar for height-resolved *aerosol backscatter*, *extinction* and *polarization* measurements and for *water vapor* measurements.

Already the first HALO mission was very successful. An *anthropogenic pollution* layer in the *lower stratosphere* originating from East Asia could be detected over southern Germany during the HALO flight on 3 November 2010.

Fig. 37.5 HALO jacked-up and with *scaffolding* during *ground vibration tests* in spring 2009. The yellow installations below the aircraft wing are a carrier for particle measurement probes (*left*) and the under-*wing instrumentation pod* (*right*)

The pollution layer contained increased SO_2, HNO_3, H_2O and particle concentrations as measured by the mass spectrometer and the lidar. According to air mass transport and *dispersion model* simulations, this pollution layer was injected into the lower stratosphere by a *warm conveyor belt* associated with an East-Asian *cyclone*. Accompanying aerosol model simulations indicated that *sulfate aerosol droplets* formed in the SO_2-rich plume and grew to sizes that explained the observed increased lidar backscatter. This was the first observation of a direct East-Asian SO_2 transport into the lower stratosphere by a warm conveyor belt. This observation was possible only because of HALO's higher maximum cruising altitude.

The lidar made extensive combined measurements of humidity and cloud optical depth in the vicinity and for the first time within fully developed *cirrus* clouds. A comparison with joint in situ measurements carried out on synchronous flights with DLR's FALCON aircraft showed that the main variability in *relative humidity* over *ice* was caused by fluctuations of water vapor concentration and not by temperature fluctuations. This validated the use of *numerical weather prediction* (NWP) temperature data to convert the absolute humidity as measured by the lidar to relative humidity. Relative humidity over ice was derived from lidar-measured water vapor concentration and NWP-computed temperature. Hence, for the first time cloud optical properties were related to humidity within cirrus. For this purpose, the extended vertical range of HALO was indispensable as the lidar system required to stay high enough above the cloud tops reaching up to 13 km.

Spectrometer measurements of upward *radiance* were performed by the University of Leipzig together with measurements of *actinic radiation* by Jülich Research Center. The data are used to investigate the influence of the heterogeneity of the surface *albedo* on the *retrieval* of cloud properties (optical depth, *effective radius*). This is performed with a statistic approach using a common retrieval method based on radiative transfer calculations in combination with data gathered by lidar during the flights.

The Techno Mission provided a first test of HALO's capabilities, its user interfaces and the performance of new on-board research installations. BAHAMAS, the new sensor and data acquisition system, was an important and complex component. The system performed without problem during all flights and demonstrated the increased capabilities of this modern and powerful data and sensor system.

For the future, these preparations open a new era in which HALO will serve as the new flagship of atmospheric research, with capabilities far beyond those of the FALCON. HALO's scientific user community is enthusiastically looking forward to at least 30 years of successful operation.

For 2012, the first so-called demonstration missions of HALO are planned, including a first atmospheric science mission, TACTS. TACTS (Transport and composition in the upper troposphere and lowermost stratosphere) will investigate air mass transport between the upper troposphere and lowermost stratosphere. This region plays a key role in the energy budget and the transport of trace gases across the tropopause. Instruments from eight German research institutions will be integrated in HALO. The mission is coordinated by the Johann-Wolfgang-Goethe University in Frankfurt. IPA will contribute with in-situ trace gas measurements (e.g., *nitrogen oxides*, SO_2) using *chemiluminescence* detectors and mass spectrometers. TACTS will be continued with a second atmospheric HALO mission Earth System Model Validation (ESMVal) coordinated by IPA. Trace gas and aerosol data will be sampled from 80°N to 80°S in order to obtain an improved understanding of processes parameterized in *climate–chemistry*–aerosol–cloud models.

Another important HALO campaign planned for the following year is coordinated by the Max Planck Society in Mainz with the participation of several HALO partners, including IPA. The project Oxidation Mechanisms Observations (OMO) has the aim of investigating oxidation processes and large-scale air pollution transport in the free troposphere and lower stratosphere. The lifetime of natural and man-made pollutants in the atmosphere is dependent on the oxidation capacity of the atmosphere, which is to a great extent controlled by chemically reactive radicals (OH and HO_2). At present only few measurements exist of these short-lived radicals in the upper troposphere. The OMO measurements will identify the major sources and sinks of HO_x and help to improve chemistry-transport models. Again, IPA is contributing with measurements of reactive nitrogen species (NO, NOy, PAN).

The ML-Cirrus project (Mission on Mid Latitude Cirrus Clouds) is planned to be performed when the under-wing stations are certified. ML-Cirrus, led by IPA, is designed to investigate both natural- and aviation-induced *cirrus* cloud formation. Cirrus clouds influence the *radiation budget* of Earth and therefore have an impact on, for example, the temperature of the atmosphere below the clouds. Many

project partners are participating in a series of measurements to investigate the radiative and microphysical properties of cirrus and changes in these properties during cloud *lifetime*. The results of these measurements will be used to refine *forecast models*, e.g., of *aviation-induced cloudiness*.

Acknowledgement: Hans Schlager, Martin Wirth and Clemens Fricke contributed to the report on the Techno Mission. Helmut Ziereis contributed to the scientific mission preparation over the years.

37.5 Technical Features of FALCON and HALO

Table 37.1 Technical features of the FALCON 20 E and Gulfstream G550 HALO aircraft

37.6 Major Aircraft Modifications

37.6.1 FALCON's major modifications:

- Two large coverable openings in the bottom of the fuselage (515 mm diameter)
- One large opening on top of the fuselage (515 mm diameter)
- One large opening at the left side of the fuselage (250 × 570 mm)
- Four small openings on top of the fuselage (80 mm diameter)
- One opening for dumping *dropsondes* (75 mm diameter)
- One hard point under the *fuselage* (e.g., for the installation of *particle measurement probes*)
- Four hard points under the wings to carry up to 4 particle measurement probes
- One hard point in the tail area for *radiometer* installation
- One universal window interface at the right hand side of the fuselage (220 mm diameter)
- One *noseboom*, 1.8 m long, with a *five-hole pressure sensor* at the tip and pressure lines to a sensor package to determine total and *static pressure* of the undisturbed air around the aircraft as well as the air flow direction
- Multiple Total Air Temperature housings for *temperature* sensors and inlets for sensor packages
- Sensor package (temperature, pressure, humidity, wind), including a temperature-controlled box behind the nose
- Experimental power: 2 × 300 A @ 28 V, 16.8 kW

37.6.2 HALO's major modifications

The specifications for the HALO modifications are based both on experience with the FALCON or other research platforms at DLR and on a survey among the HALO user community in October 2002 (Fig. 37.6). Besides the necessary increase in the number

Table 37.1 Technical features of both aircraft

	FALCON	HALO
Engines	Garett TFE 731-5BR-2C (2 × 4 750 lbs)	Rolls Royce BR710 (2 × 15 385 lbs)
Aircraft performance		
Maximum range	3 700 km	12 500 km
Maximum cruise altitude	42 000 ft/12 800 m	51 000 ft/15 540 m
Takeoff distance (MTOW, ISA; MSL)	2 000 m	1 801 m
Maximum speed	0.865 Mach	0.885 Mach
Fuel consumption	≈ 1 000 l/h	≈ 1 725 l/h
Weights		
Maximum takeoff weight	13 755 kg	41 277 kg
Maximum landing weight	13 100 kg	34 156 kg
Maximum zero fuel weight	10 200 kg	24 721 kg
Maximum fuel	4 006 kg	18 734 kg
Maximum payload	1 t	3 t
Exterior		
Length	18.75 m	29.4 m
Height	5.45 m	7.9 m
Wingspan	16.46 m	28.5 m
Interior		
Cabin length	4.60–5.10 m	15.3 m
Cabin height	1.64–1.72 m	1.88 m
Cabin width	1.50 m	2.24 m
Cabin volume	20 m^3	47.3 m^3

of user interfaces due to the larger aircraft size, many new aspects like standardization of interfaces, universal interface documentation and communality with existing international standards played a significant role in the HALO design process.

- 19 standard apertures (180 mm × 250 mm, 10 on top, 9 on the bottom of the fuselage)
- Two double size (360 m × 250 mm) openings on top of the fuselage
- One dropsonde dispenser aperture in the baggage compartment
- Two upward looking viewports for optical windows (515 mm Ø, fuselage top)
- Two downward looking viewports for optical windows (515 mm Ø, fuselage bottom)

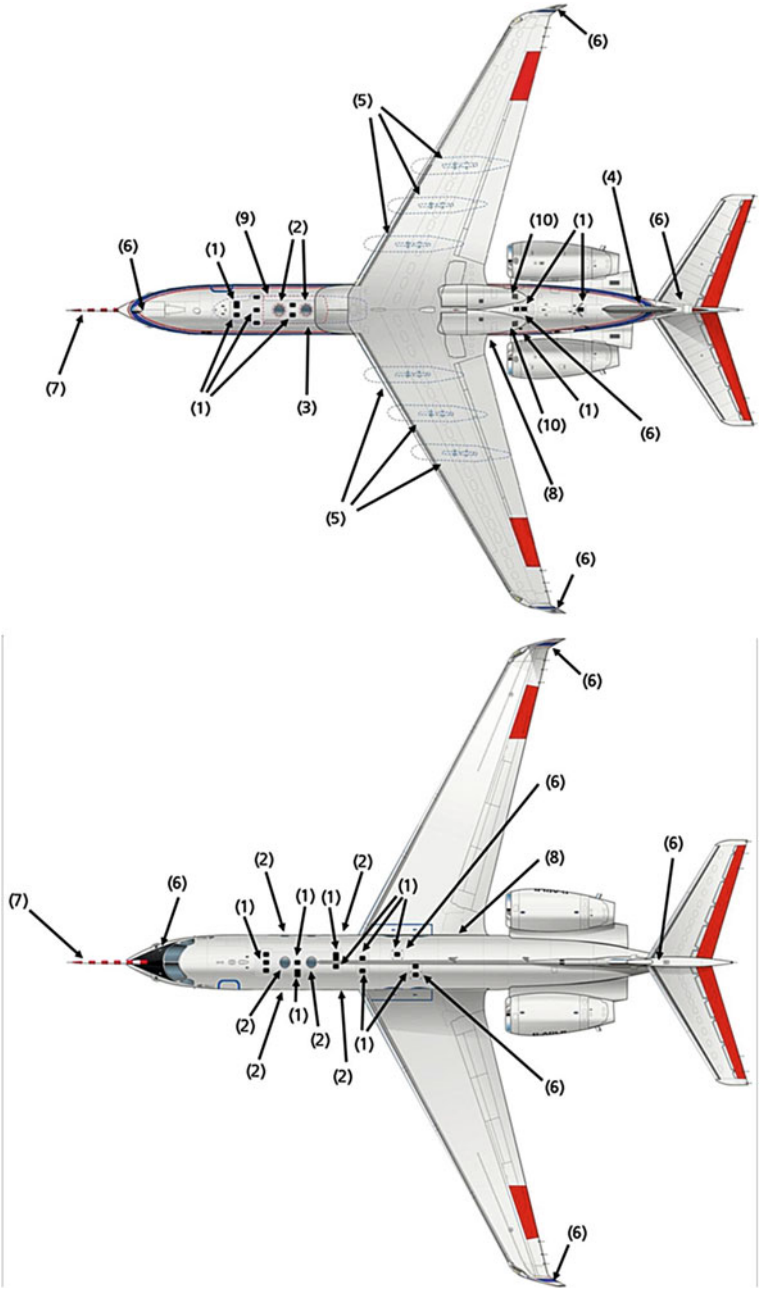

Fig. 37.6 Overview of HALO modifications from below and above: (*1*) *apertures*, (*2*) optical *viewports*, (*3*) *belly instrumentation pod*, (*4*) *ventral fin*, (*5*) wing hard points, (*6*) other hard points, (*7*) *noseboom*, (*8*) Liquid Cooling System, (*9*) shutter door assembly, (10) chemical air exhaust

Fig. 37.7 HALO during *belly pod* and *ventral fin* flight tests photographed from the FALCON, which chased HALO to observe the behavior of the modifications. The belly pod is below the *fuselage* in front of the wings. The ventral fin is below the tail of the fuselage. Both structures can be removed

- Sideward looking optical window mounts (430 mm Ø, or an adapter, can be fitted into each standard cabin window opening)
- All openings can handle significant inertial and aerodynamic loads and are separated as far as possible from each other to minimize interactions. Upward, downward and sideward looking optical viewports have a common axis to allow for scanning applications. A blower system supplies dry hot air to all viewports in order to avoid condensation effects at high *flight levels*. Two chemical exhaust tubes allow dumping of undesired gas flows (pump gases) into the atmosphere
- A modular shutter door assembly prevents downward looking windows from damage during take-off and landing
- Belly (instrumentation) pod (see Fig. 37.7), (7 m long fairing, volume 2 m^3, hard points for 350 kg mounted directly to the aircraft fuselage). The Belly pod consists of seven segments which can be individually adapted to specific campaign needs
- Ventral fin (see Fig. 37.7) to compensate for reduced directional stability caused by the Belly pod
- Three hard points under each wing, capable of carrying up to 400 kg each. Possible installations are:
  - Under-wing pod, 4.5 m long, 0.5 m Ø, with a central instrumental carrier and numerous openings (see Fig. 37.5)

  - Under-wing carriers for up to three particle measurement probes (see Fig. 37.5)
- Three hard points on the fuselage (2 on top, 1 on the bottom, max 50 kg)
- One hard point in the unpressurized nose compartment (max 30 kg) for the basic measurement equipment
- One hard point in the unpressurized tail compartment (max 100 kg)
- One hard point in the tail cone (max 40 kg), an optional larger tail cone is available
- One hard point on the vertical stabilizer fin cap (max 20 kg)
- One hard point at each winglet (max 5 kg)
- Experimental power is provided by a mission power rack that is fed by the two main engines and the auxiliary unit; the mission rack, which can be disconnected at any time from the primary aircraft system, provides at least 42 kW, and is available as:
  - 3 phase, 115 V AC, 400 Hz
  - 3 phase, 115 V AC, 400 Hz anti-ice (only available in flight)
  - 1 phase, 230 V AC, 50 Hz
  - 28 V DC
- Noseboom, 3.8 m long with flow angle sensor; an instrument tray inside the noseboom tip carries the pressure sensor instrumentation (max 3 kg)
- Pressure bulkheads are modified to facilitate pneumatic or electrical connections into the unpressurized areas; therefore, pressure-sealed metal plates can be removed and modified
- Various ductworks for cables are within the cabin and additional ones are linked to the nose/tail/tail cone/vertical stabilizer fin cap/wing stations/winglets
- Six Total Air Temperature housings beneath the aircraft nose for temperature measurements and as inlet system for humidity measurements
- A Liquid Cooling System has been designed to remove the heat produced by instruments in the cabin via a heat exchanger located in the right hand wing to body fairing
- The basic sensor and data acquisition system was developed, manufactured and installed by the DLR Research Flight Facility, Sensor and Data Group. The system includes:
  - Flow and pressure measurement instrumentation for the analysis of wind and aerodynamic parameters
  - Temperature and humidity instrumentation using different sensors
  - A data acquisition system for analog and digital data with up to 1 kHz data rate; users can feed their instrument data into the system via multiple interfaces on a special interface panel
  - Monitoring system for, e.g., a power supply
  - Aircraft reference time standard
  - High precision inertial reference unit
  - Real-time data processing and quicklook unit aircraft data interfaces (e.g., from Air Data Computer, Inertial Reference Units, TCAS, *Lightning* Sensor Systems, GPS, Flight Management Parameters, weather radar,

forward, downward and upward looking cameras, camera of the enhanced vision system and head up displays, interface to the core data bus (ASCB-D) of the aircraft)
  - Prompt provision of post-processed, quality-checked final data after the flight;
- Satellite communication system with a data transmission capacity of 256 kBytes/s
- Telephone communication via *Iridium satellite* systems

The Eyjafjalla Eruption in 2010 and the Volcanic Impact on Aviation

38

Bernadett Weinzierl, Thomas Sailer, Daniel Sauer, Andreas Minikin, Oliver Reitebuch, Bernhard Mayer and Ulrich Schumann

Abstract

This paper gives an overview of the Eyjafjallajökull (short: Eyjafjalla) eruption and its impact on aviation from the perspective of DLR research. The visual detectability of airborne volcanic ash at various distances from the volcano is discussed based on observations and simulations with a radiative transfer model. Ash loads above threshold conditions for safe aviation (2 mg m^{-3}) as in spring 2010 should not be considered exceptional. A future research strategy to better prepare the aviation sector for future volcanic eruptions is outlined.

B. Weinzierl (✉) · T. Sailer · D. Sauer · A. Minikin
O. Reitebuch · B. Mayer · U. Schumann
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: bernadett.weinzierl@dlr.de

T. Sailer
e-mail: sailerthomas@googlemail.com

D. Sauer
e-mail: daniel.sauer@lmu.de

A. Minikin
e-mail: andreas.minikin@dlr.de

O. Reitebuch
e-mail: oliver.reitebuch@dlr.de

B. Mayer
e-mail: bernhard.mayer@lmu.de

U. Schumann
e-mail: ulrich.schumann@dlr.de

B. Weinzierl · D. Sauer · B. Mayer
Ludwig-Maximilians-Universität München (LMU),
Meteorological Institute Munich (MIM), Theresienstraße 37,
80333 München, Germany

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_38, © Springer-Verlag Berlin Heidelberg 2012

38.1 Introduction

"We say hello to all our guests who boarded in Munich, including those who had preferred to go by airplane." This announcement in an ICE train running from Munich to Hamburg was a typical happening after the eruption of Iceland's *Eyjafjalla volcano* in April 2010.

The Eyjafjalla *volcanic ash* caused the most extensive restrictions of the airspace over Europe since the end of World War II. More than 100 000 flights were cancelled between 14 and 20 April 2010. In the course of the "ash crisis" more than 10 million passengers were affected and industry had to cope with reduced transport capacities leading to a slowdown of production in some companies.

Although some scientists had long requested detailed *engine* tests to assess maximum tolerable ash concentrations and doses (accumulated ash mass over time) for "*safe flying*" no such *threshold* values were available in April 2010. While volcanic *eruptions* occur quite frequently around the globe, so far *air traffic management* had to consider flight restrictions only locally or regionally around volcanoes. That volcanic contamination of airspace could affect a very large region with dense air traffic (like the whole of central Europe) at the same time was apparently not expected, and civil *aviation* in Europe was obviously not prepared for such an "unlikely" event. The airspace closures in spring 2010 were critically discussed not only in the media, but also by airline companies and aviation industry. There were even doubts raised about the actual presence of any volcanic ash over central Europe. One of the major criticisms was that actions taken by authorities were arbitrary and that *airspace closures* were based only on model predictions while no measurements were available.

38.2 Disturbance of Air Traffic by Particle Emissions

Atmospheric aerosols, i.e., tiny particles suspended in the air, not only impact *weather*, *climate* and *air quality*, they may also disrupt *aviation*. Among the different aerosol types, *volcanic ash* and *mineral dust* are known to have potentially severe impact. Strong mineral dust storms mainly affect the take-off and landing of an aircraft by causing very poor *visibility*, whereas volcanic ash can pose a *hazard* to aircraft in flight due to adverse effects on the aircraft's engines or navigation equipment. The reason volcanic ash is considered to be more dangerous to *aircraft engines* than mineral dust is that volcanic ash generally consists of predominantly *glassy*-phase material with some crystalline silicate minerals, whereas mineral dust is composed of crystalline minerals only. While the melting temperature for pure quartz, the major constituent of mineral dust, is around 1 650 °C, glassy materials may undergo a state transition from solid to liquid at temperatures far below the typical melting temperatures of crystals or rocks (Casadevall 1993; Kueppers et al. 2010). In a laboratory study with artificial volcanic ash and mineral dust particles, Mechnich et al. (2011) investigated the effects of particle deposition and high-

Table 38.1 Severity index for ash encounters (adapted from ICAO 2001, 2007)

Class 0 encounter	Acrid odor (e.g., sulfur gas) noted in cabin; electrostatic discharge (St. Elmo's fire) on windshield, nose, engine cowls; no notable damage to exterior or interior
Class 1 encounter	Light dust in cabin (no oxygen used); exhaust gas temperature (EGT) fluctuations with return to normal values
Class 2 encounter	Heavy cabin dust ("dark as night" in cabin); contamination of air handling and air conditioning systems requiring use of oxygen; some abrasion damage to exterior surface of aircraft, engine inlet, and compressor fan blades; frosting or breaking of windows due to impact of ash; minor plugging of pitot-static system (insufficient to affect instrument readings); deposition of ash in engine
Class 3 encounter	Vibration of engines owing to mismatch, surging; plugging of pitot-static system to give erroneous instrument readings; contamination of engine oil hydraulic system fluids; damage to electrical system; engine damage
Class 4 encounter	Temporary engine failure requiring in-flight restart of engine
Class 5 encounter	Engine failure or other damage leading to crash

temperature interactions on the thermal barrier coatings of high-pressure turbine airfoils. They concluded that the glassy constituents of volcanic ash can become liquid enough to infiltrate thermal-barrier coatings and cause destabilizing chemical reactions at temperatures which are generally exceeded in jet engine turbines. As a result, damage to the thermal barrier coatings of turbine blades after long-term exposure to volcanic ash cannot be ruled out. Aside from the adverse effects on turbine airfoils other damage, such as the erosion and blinding of windows or clogging of aeronautic sensors, can occur.

In the past, most incidents involving aircraft damage related to volcanic emissions occurred within the first day after the onset of the volcanic eruption, and at distances less than 1 000 km downwind of the volcano, though there have been exceptions with incidents at even larger distances (Casadevall 1994; Guffanti et al. 2010). The severity of effects observed among the different volcanic ash incidents is highly variable and can be classified on the basis of a *severity index* established by the International Civil Aviation Organization (ICAO, see Table 38.1). In some cases only an acrid smell was noted; in other cases the aircraft experienced in-flight degradation of engines or failure of navigational instruments and engines. Some volcanic ash damage is immediately apparent, like engine failures, but some of the effects which are primarily related to the acidic gas sulfur dioxide may take longer to manifest themselves.

Between 1953 and 2009, at least 129 planes inadvertently flew through the volcanic ash plumes originating from *explosive volcanic eruptions* of the Galunggung volcano (1982, Java), Mount Redoubt (1989/90, Alaska) or Mount Pinatubo (1991, Philippines). Of those encounters, 79 caused various degrees of airframe and engine damage, and in nine cases one or more engines failed due to the

melting and resolidification of ash in the jet engine turbine (Guffanti et al. 2010). In all cases of *engine failure*, it was possible to restart the engines in flight, and no Class 5 encounters and no deaths have been reported so far. In response to aircraft incidents with volcanic ash, nine *Volcanic Ash Advisory Centers* (VAAC) were set up in the 1990s all over the world with the aim to coordinate and distribute information about potentially dangerous volcanic ash in the atmosphere. The VAACs are run as part of national weather services and are responsible for ensuring *flight safety* by providing ash plume movement reports and model forecasts. The London VAAC run by the British Met Office is the VAAC responsible for flight safety on the northeast Atlantic routes.

Although aircraft have been capable of flying safely in areas with low ash concentrations, critical threshold values could not be easily defined and the conclusion from the accidental ash encounters in the 1980s and 1990s was that complete avoidance of ash-laden air is the only way to ensure maximum flight safety (Casadevall 1993; ICAO 2001). Therefore, the procedure recommended by ICAO (2001, 2007) and applied after the Eyjafjalla eruption was "(…) regardless of ash concentration—AVOID AVOID AVOID". In particular, it was recommended to avoid any "visible ash".

This "*zero-tolerance rule*" led to the closure of large parts of the European air space for several days in April 2010. To maximize the nonrestricted airspace in which aircraft could operate under a limited presence of volcanic ash, aviation experts agreed on preliminary threshold values for volcanic ash based on empirical assumptions at the end of April 2010 (Schumann et al. 2011). As a result, less frequent closures of airspaces were necessary in the later eruption phase of the Eyjafjalla volcano. On 23 May 2011 after the eruption of Iceland's Grimsvötn volcano, the German government (http://www.bmvbs.de) ruled that aircraft may fly in volcanic ash without special attention below a concentration limit of 0.2 mg m^{-3}. Between 0.2 and 2 mg m^{-3} enhanced maintenance procedures apply, and it is not permitted to fly in areas with concentrations higher than 2 mg m^{-3}. Under certain conditions exceptions to this limit are possible at concentrations between 2 and 4 mg m^{-3} for special flights like "search and rescue" or research flights. It should be noted that the limits for dangerous mass concentrations, and even more for dangerous dose values, are very uncertain. Although accepted threshold values are now available, the hazard of volcanic ash to planes likely depends on more parameters than just the airborne ash mass concentration, including the time spent in an ash layer, the engine type and the power setting. It may also vary between different eruptions (or even within a single eruption) because ash properties are not uniform. Maximum ash concentrations in the vicinity of past Class 4 aircraft encounters were analyzed and found to exceed at least 4 mg m^{-3}, with far higher concentration in the immediate vicinity (Witham et al. 2012).

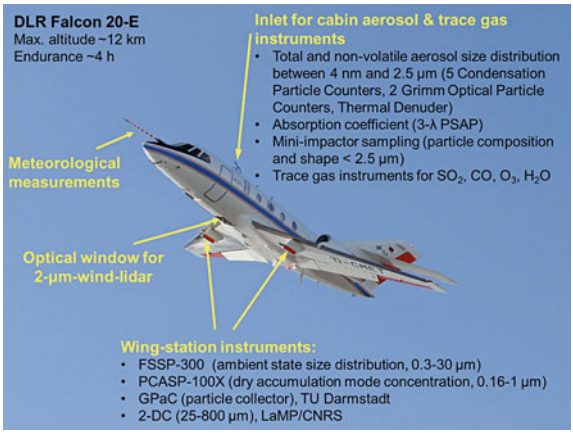

Fig. 38.1 Photograph listing the Falcon instrumentation during the volcanic ash measurements

38.3 The Eyjafjalla Eruption in 2010 and Volcanic Ash Measurements by DLR

Although the widespread effects of the Eyjafjalla eruption in 2010 surprised people in Europe, the eruption itself was not unexpected for volcanologists. Seismic tremors started to increase in April 2009, and a large number of earthquakes were recorded in February and March 2010. After a first eruption on 20 March 2010, the eruption affecting central Europe started on 14 April 2011. The prevailing northwesterly winds very quickly transported airborne ash towards central Europe. Beginning on 15 April 2010, widespread regions of the European *airspace* were closed for about one week. The hindrance of public air travel combined with the substantial economic loss experienced by airlines and airports quickly led to a public awareness of the lack of observational data on the actual location and extent of the volcanic ash layers and their severity in terms of ash mass concentration.

At DLR, a decision to quickly prepare the Falcon aircraft for measurements was made on late Friday, 16 April, shortly after the first airports were closed in northern Europe. To obtain measurements of the airborne ash, the DLR *Falcon* 20-E twin-engine jet aircraft was instrumented with a 2-µm *Doppler lidar* system, several in situ instruments for the measurement of aerosol microphysical and optical properties, meteorological parameters and trace gases (Fig. 38.1). Furthermore, two-stage impactor devices for sampling ash particles for post-flight laboratory analysis of particle composition were installed. The pilots paid particular attention to engine parameters during flights in order to detect any possible engine damages early. Satellite (Iridium) telephone connections enabled important real-time communication of the results from the observers on board the Falcon to the decision-making agencies.

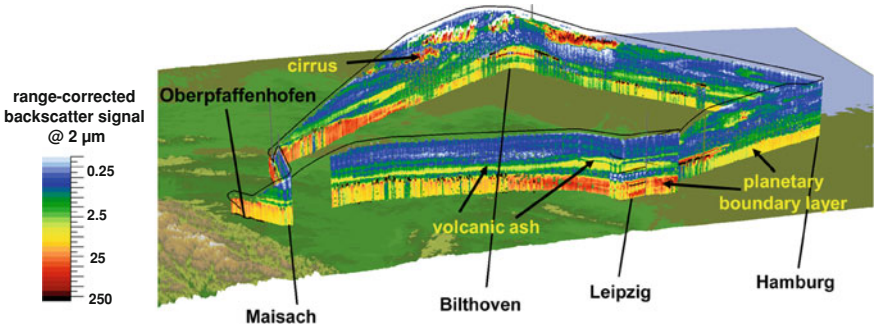

Fig. 38.2 Lidar image showing the range corrected backscatter signal (color-coded) along the flight path of the first Falcon volcanic ash flight on 19 April 2010. Blue colors denote regions with low aerosol concentrations, *yellow* and *red* colors show areas with high aerosol concentrations. Between 4 and 6 km, an elevated volcanic ash layer with a vertical thickness of 0.5–2 km is visible (*yellow*). The viewing direction is from south (*left*) to north (*right*). The flight altitude above Hamburg was 8.2 km. Figure modified from Schumann et al. (2011)

The flight strategy deployed during the volcanic ash measurements was adapted from the experience gained in earlier experiments with the Falcon, including the Saharan Mineral Dust Experiment (SAMUM, 2004–2011) where extended layers of Saharan dust aerosol were investigated (Esselborn et al. 2009; Petzold et al. 2009; Weinzierl et al. 2009; Petzold et al. 2011; Weinzierl et al. 2011). The difficulty in such studies is to fly the aircraft into the appropriate altitudes where the particle layers are present. Here the lidar is an invaluable tool because it is a remote sensing instrument able to detect the entire particle profile in the air column under the aircraft. Thus, it can be used as a pathfinder to direct the aircraft quickly into the aerosol layers for in situ *measurements* inside these layers. Furthermore, the aircraft was often flown over ground stations with upward looking lidar systems (e.g., Munich, Leipzig, Hamburg, Stuttgart, Cabauw) to obtain complimentary measurements of the ash plume. Analysis of such complimentary data sets of the same aerosol layer gives insights into data quality and allows the assessment of measurement uncertainties.

The most important operational question during the first Falcon flight on 19 April 2010 was whether the closure of airspace following the ICAO procedure was justified, i.e., whether volcanic ash was present over Germany or not. Furthermore, airborne data were critically needed to evaluate the quality of the model forecasts. The scientific aims of these measurements included the characterization of particle number, size and mass concentration, the change of those properties as a function of time after the eruption, the chemical plume composition, and the response of lidar and satellite instruments to the volcanic ash plumes. Figure 38.2 shows that an elevated volcanic ash layer situated between 4 and 6 km altitude almost entirely covered Germany.

In situ measurements of ash mass concentrations in the elevated ash layer revealed maximum values of 0.05 [0.03-0.11] mg m^{-3} over Leipzig and Munich on that day (Schumann et al. 2011). Peak ash mass concentrations derived on the basis

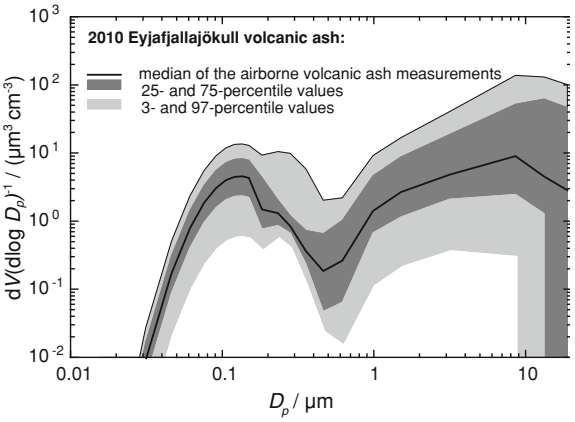

Fig. 38.3 Composite *particle* volume *size distribution* of the Eyjafjalla volcanic ash measurements derived from a set of Condensation Particle Counters (CPC), a Passive Cavity Aerosol Spectrometer Probe (PCASP-100X), and a Forward Scattering Probe (FSSP-300). The dark and light grey shaded areas indicate the 3, 25, 75 and 97 percentile values for the size distributions. Data taken from Weinzierl et al. (2012)

of ground-based lidar measurements were reached on 16 April over Leipzig with values of 1.0 [0.65–1.35] mg m^{-3} (Ansmann et al. 2010), while the maximum of the ash mass concentration of 1.1 [0.65-1.8] mg m^{-3} over Munich was reached on 17 April 2010 (Gasteiger et al. 2011). The median size distribution measured during the Falcon flights is shown in Fig. 38.3. The results compare reasonably well with the measurements performed with the British research aircraft Facility for Airborne Atmospheric Measurements (FAAM) described in Turnbull et al. (2012).

For safety reasons, the flight routes were planned such that the in situ measurements in the volcanic ash plume were only performed close to airports during the first few flights. Depending on the lidar signal, the Falcon crew decided whether it was safe to fly into an ash cloud or whether the ash cloud was to be avoided. After the first flights were trouble-free, profiles in volcanic ash over the Atlantic Ocean between Scotland and Iceland were also made. On 17 May 2010, the Falcon even spent about one hour in a volcanic ash plume with maximum mass concentrations of 0.5 [0.28−0.83] mg m^{-3} without experiencing any notable engine problems.

Figure 38.4a shows a map with all volcanic ash aircraft flight tracks. The highlight of the flights performed in Iceland between 29 April and 3 May was the derivation of the ash mass flux from measurements. Furthermore, important data on ash plume dynamics were gathered during those flights. Figure 38.4b shows a picture of the volcano and the ash cloud on 1 May 2010 (viewing direction from north to south).

In summary, the Falcon flew in *ash plumes* with concentrations up to about 0.8 mg m^{-3} for a few minutes and in an ash plume with approximately 0.2 mg m^{-3} mean-concentration for about one hour without engine damage (Schumann et al. 2011). The Falcon never entered ash plumes exceeding 2 mg m^{-3}, although ash layers with such high mass concentrations were clearly present close to the source. The Falcon results were used in numerous modeling and *intercomparison* studies (Heinold et al. 2011; Stohl et al. 2011; Kristiansen et al. 2012; Matthias et al. 2012; Rix 2012; Turnbull et al. 2012; Webster et al. 2012) and more investigations are ongoing.

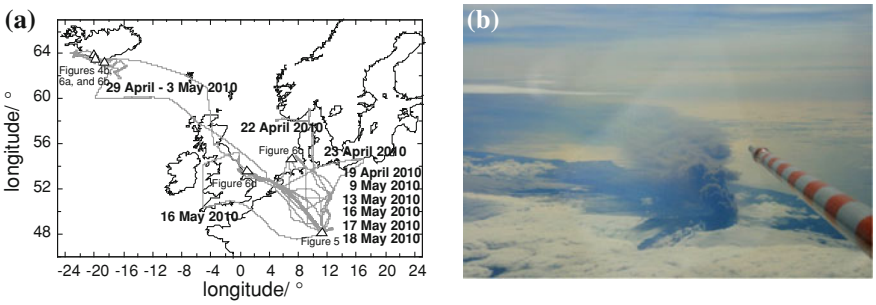

Fig. 38.4 **a** Falcon flight tracks. The triangles indicate the areas where the photographs shown in this article were taken. **b** Photograph showing the ash cloud on 1 May 2010

Fig. 38.5 Photograph taken at Oberpfaffenhofen, Germany on 19 April 2010. Measured ash concentrations in the Munich area ranged between 0.03 and 0.11 mg m^{-3} on that day

38.4 Is There a Volcanic "Ash Cloud"?

For 19 April 2010, the day of the first Falcon flight and day 4 after airspace closure in Germany, blue sky and sunny weather—which would turn out to be typical in April 2010—was forecasted. Figure 38.5 shows a photograph taken on that day. At first glance, no volcanic ash cloud appears to be present. However, the slightly hazy blue sky color indicates a possible presence of large (volcanic ash) particles which scatter the incident solar radiation. The question is whether the term "volcanic ash cloud" is appropriate for a volcanic ash plume observed far from the emission source.

Figure 38.6 shows four photographs of volcanic ash plumes taken at different locations and distances downwind from the volcano. As can be seen from Fig. 38.6, the appearance of airborne volcanic ash changes drastically with distance from the emission source, and therefore *plume age*. Close to the volcano, a well-defined grayish "ash cloud" with sharp edges is visible. Subsequently, 100–1 000 km downwind from the volcano, *dilution* and *sedimentation* processes

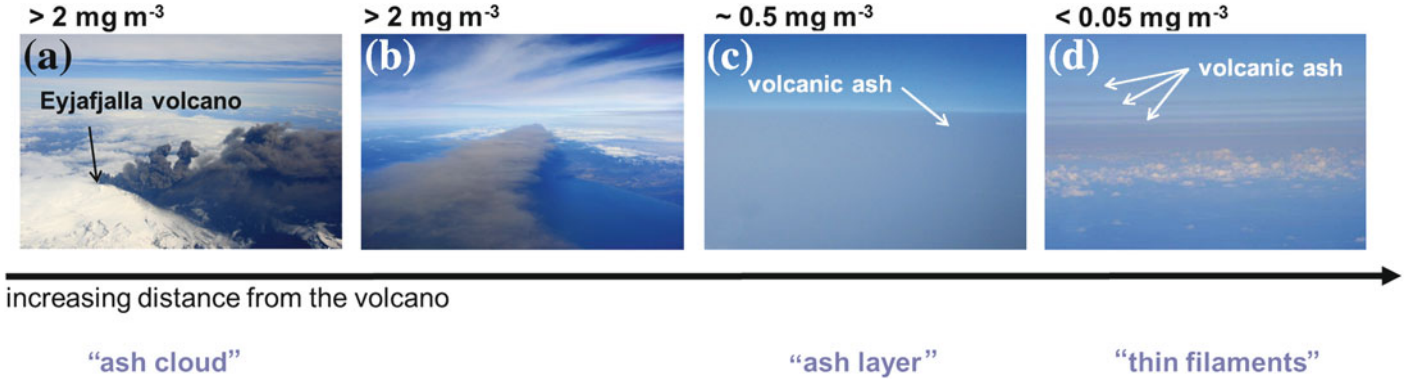

Fig. 38.6 The appearance of the airborne volcanic ash of the Eyjafjalla eruption changed drastically with increasing distance from the volcano. The photographs on the *left* (**a**, **b**) were taken close to the volcano, whereas those on the *right* were taken over the North Sea (**c**) and over Great Britain (**d**), more than 1 500 km downstream from the volcano. The locations where the photographs were taken are indicated in Fig. 38.4a

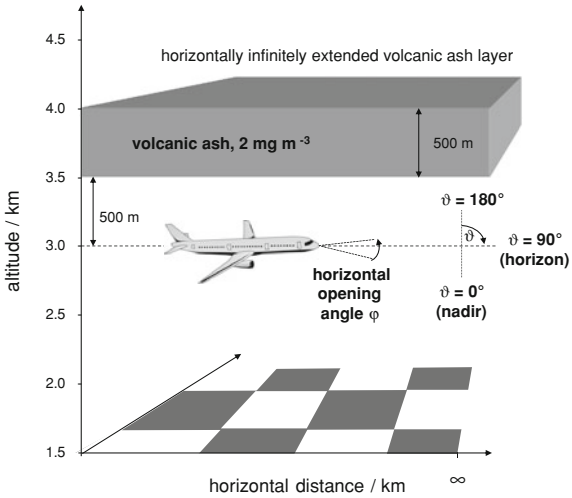

Fig. 38.7 Setup of the idealized simulations for assessing the visibility of volcanic ash. Aerosols were located only in the distinct and horizontally infinitely extended layer between 3.5 and 4.0 km altitude, and clouds were not present. The simulations were performed for three scenarios: the aircraft is *below*, in, and *above* the aerosol layer. To add some artificial contrast to the simulation, the surface albedo was set to a chess board pattern with 1×1 km^2 squares with albedo 0.2 and 0, respectively. Figure modified from Weinzierl et al. (2012)

transform the ash cloud into a grayish-brownish, wide-stretched and often inhomogeneous ash layer with rather blurry edges (Fig. 38.6c). Sometimes the emitted ash does not even form an extended layer, but rather fragments into numerous very thin filaments of about 100 m geometrical thickness (Fig. 38.6d). The smaller the distance between the ash plume and the observer, the less well defined the edges of the plume appear, similar to a pilot's perception of cirrus clouds.

The presence of a grayish-brown ash cloud or ash layer raises the question whether potentially dangerous volcanic ash (mass concentration larger than 2 mg m^{-3}) can be detected just by visual inspection of the sky from the flight deck of an aircraft.

38.5 Is it Possible to Visually Detect Volcanic Ash and Distinguish it from Other Aerosols?

Having in mind the safety criterion of "no flight in visible ash," a key question is whether a pilot has the means to avoid flying through potentially dangerous ash layers just by visual observation of the atmospheric situation from the flight deck. Here we summarize the results of a recent study (Weinzierl et al. 2012).

Assessment of the visual *detectability* of airborne volcanic ash or other aerosols is a very complex task because detectability depends on many parameters, including the size/distance ratio, the *brightness* and color contrast between the

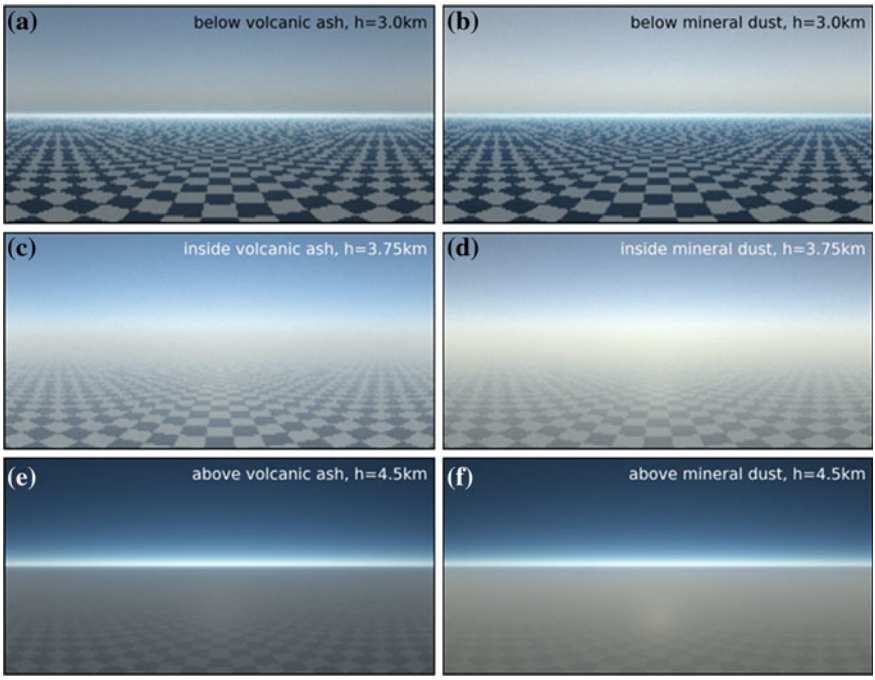

Fig. 38.8 Volcanic ash and mineral dust layers with a mass concentration of 2 mg m^{-3} and a geometrical thickness of 500 m as seen from the cockpit of an aircraft which is *below* (**a**, **b**), *inside* (**c**, **d**), and *above* (**e**, **f**) the aerosol plume. In these images the observer's view is away from the sun

airborne volcanic ash and the background, the *illumination*, the size distribution and mass concentration of the ash, *wavelength*-dependent *light scattering* and *absorption* by the ash, human perception, etc. To characterize visibility or detectability of volcanic ash, the *optical depth* along the *line of sight* through an *ash layer* is more important than just the (vertical) optical depth, which is measured, for example, by *sun photometers*.

The question of the visibility of volcanic ash was assessed in detail after performing a series of highly idealized *radiation transfer* simulations for a variety of selected viewing geometries using the *Monte Carlo* code for the physically correct tracing of photons in cloudy atmospheres MYSTIC (Mayer 2009; Buras and Mayer 2011), which is part of the *libRadtran* package (Mayer and Kylling 2005): Fig. 38.7 shows the setup of the simulations. They resemble the observation of volcanic ash layers several 100 km downwind from the volcano and for meteorological conditions dominated by the presence of a high pressure system, as was observed in April/May 2010.

An important scientific aspect addressed by the model runs was the minimum concentration that can be identified by an observer located 500 m either above or below the aerosol layer: from above, the layer is vaguely identifiable at

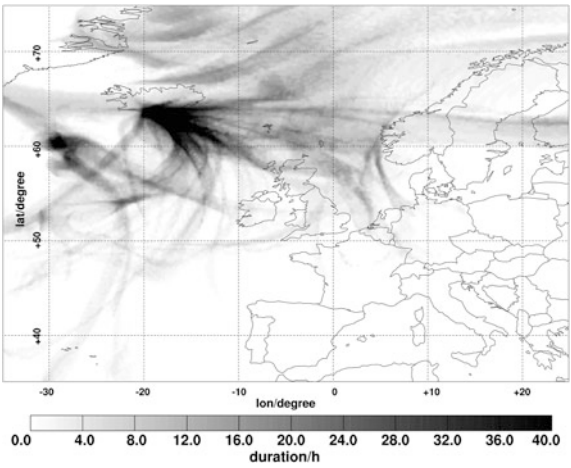

Fig. 38.9 Integrated duration of the occurrence of volcanic ash concentrations above the limit for "*safe flying*" (2 mg m^{-3}) anywhere in the tropospheric column in the period between 14 April and 24 May 2010

concentrations between about 0.2 and 0.5 mg m^{-3} depending on the viewing angle (for more detail see Weinzierl et al. 2012). However, the actual *visibility* of the aerosol layers depends strongly on background contrast and light conditions in a given situation. The actual thickness of the aerosol layer and the position of the observer relative to the layer are also important parameters influencing the result. Therefore, these numbers indicate the lower limit of ash concentration for ash visibility under idealized conditions without clouds during the day.

Aside from volcanic ash, *mineral dust* and *biomass burning aerosol* layers also have a grayish appearance. Analyses of microphysical ash particles showed that volcanic particulate emissions exhibit a bi-modal volume size distribution consisting of a mixture of larger ash particles (about 0.7–10 μm particle diameter) and *sulfuric acid* droplets/particles (smaller than 0.7 μm), whereas mineral dust layers showed only a very distinct *coarse mode* above 0.5 μm (Weinzierl et al. 2012). Figure 38.8 depicts the visual difference between a layer of volcanic ash (left panels) and mineral dust (right panels) from a pilot's view. Although microphysical differences between volcanic ash and mineral dust can be detected by appropriate instruments, it is obvious from Fig. 38.8 that these differences are not detectable by the human eye.

Summarizing the aircraft measurements and the *radiative transfer simulations* leads to the following main conclusions on the visibility of volcanic ash: under clear sky conditions volcanic ash is visible already at mass concentrations far below what is currently considered dangerous for an aircraft engine (2 mg m^{-3}), depending on the thickness of the ash layer and the optical depth along the line of sight. However, it is not possible to visually determine whether a volcanic ash layer is potentially dangerous (mass concentration larger or smaller than 2 mg m^{-3}). The presence of a grayish-brown layer in the atmosphere does not unambiguously indicate the presence of volcanic ash. An uninformed observer is unlikely to recognize an aged volcanic ash layer in his field of view without further

Table 38.2 Volcanic explosivity index (VEI) and recurrence times of volcanic eruptions in Iceland as a function of the VEI (adapted and combined from Newhall and Self 1982; Gudmundsson et al. 2008)

VEI	0	1	2	3	4	5	6	7	8
Eruption type	Non-explosive	Small	Moderate	Moderate-large	Large	Very large	Very large	Very large	Very large
Volume of ejected tephra/km^{-3}	<0.00001	0.00001–0.001	0.001–0.01	0.01–0.1	0.1–1	1–10	10–100	100–1 000	>1 000
Column height/km	<0.1	0.1–1	1–5	3–15	10–25	>25	>25	>25	>25
Duration/h	<1		>12	>12	>12	>12	>12	>12	>12
Stratospheric injection	None	None	None	Possible	Definite	Significant	Significant	Significant	Significant
Examples	Reykjanes, 1926	Poás, 1991	Stromboli, ongoing	Puyehue, 2011	Eyjafjalla, 2010	St. Helens, 1980	Pinatubo, 1991	Tambora, 1815	Yellowstone Caldera
Recurrence time of volcanic eruptions in Iceland/a		10	10	10–20	30–50	100–200	500–1 000		

information. The presence of clouds would make it even more complicated to detect volcanic ash visually because clouds can either directly block the direct line of sight to the ash layer or reduce the contrast if present in the background. Therefore, other means are needed to establish where there is safe airspace after a volcanic eruption. Such means should include aircraft measurements and satellite and ground observations, as well as model predictions which are constrained by observations.

38.6 How Often Were Dangerous Ash Concentration Levels Reached over Germany Based on the Newly Accepted Threshold Values for Safe Aviation?

As mentioned, at the beginning of the Eyjafjalla "ash crisis," the operational rules for aircraft required the complete avoidance of ash-laden air. Assuming that the newly accepted threshold concentrations of 2 mg m^{-3} for safe operation had been applied from the very beginning of the eruption period, one question interesting in retrospect is how often dangerous ash concentrations (above threshold) would have been observed over Germany in spring 2010. This was investigated with the help of the *Lagrangian trajectory model FLEXPART* for the active period of the Eyjafjalla eruption between 14 April and 24 May. More detail on FLEXPART is given in Stohl et al. (2005). The *emission* height of the volcano was taken from *radar* observations in Iceland and information from the VAAC, and an output grid of a horizontal resolution of 0.25° and a vertical resolution of 250 m was used in FLEXPART. The emission strength of the volcano was assumed to be 18.8 Tg of *particulate matter*, because for this value the source term agreed best with the measurements made by the Falcon research aircraft. Figure 38.9 depicts the results of the simulations. Near Iceland, *mass concentrations* of 2 mg m^{-3} were regularly exceeded, as expected. In Germany, high ash mass concentrations exceeding the threshold of 2 mg m^{-3} were simulated over the northwestern and western part of Germany and only for short periods. Over southeastern Germany the threshold of 2 mg m^{-3} was most likely never exceeded in April/May 2010. If the times are added in which the 2 mg m^{-3} limit was exceeded anywhere in the tropospheric column above Germany, this country was covered about 1.5 days with "dangerous" ash concentrations according to the emission assumptions made in FLEXPART. The results of the IPA simulation are in agreement with simulations by Stohl et al. (2011) who showed that the prevailing winds spread the ash emission over the northern hemisphere such that up to 14 % of the European area (10°W–30°E and 36–60°N), mainly near Iceland, exceeded a volcanic ash concentration of 0.2 mg m^{-3}, and 1.5 % of the area exceeded 2 mg m^{-3} at some altitude in the atmosphere.

Table 38.3 Assumptions made for the climatological simulations with FLEXPART

Parameter	Value	References
Emission strength	18.8 Tg (in 35 days)	2.5–50 Tg, best estimate 10 Tg (Schumann et al. 2011) 6.0–17.8 Tg, best estimate 11.9 Tg (Stohl et al. 2011)
Emission height	Four heights between 3 and 11 km above sea level	Radar (Arason et al. 2011) and VAAC (2010)
Emission duration	35 days	37 days (Arason et al. 2011) and 39 days (Petersen et al. 2012)
Removal processes	Wet and dry deposition	(Stohl et al. 2005)
Model parameters	$0.5 \times 0.5°$ ($\sim$50 km) horizontal resolution, 250 m vertical resolution, 3 h temporal resolution, $\sim 10^9$ particles	
Meteorological input data	ERA Interim T255 from ECMWF	

38.7 The Eyjafjalla Eruption in 2010: Was it an Unusual Event?

Volcanic eruptions are a regular phenomenon. During the past 40 years, volcanism has been almost constant with globally 50–70 eruptions per year (Siebert et al. 2010), which implies on average one volcanic eruption per week. The largest concentration of active volcanoes is located on the so-called "*ring of fire*" which extends around the Pacific Ocean and includes volcanoes such as Mount Pinatubo (Philippines), Krakatao and Mount Tambora (both in Indonesia), Mount St. Helens (United States) and El Chichón (Mexico). Within Europe, Iceland has the highest density of volcanoes with currently more than 30 active volcanoes. The explosiveness of volcanic eruptions is measured by the *Volcanic Explosivity Index* (VEI, Newhall and Self 1982). Table 38.2 shows the VEI classification scheme (Newhall and Self 1982) and summarizes the characteristics of individual eruptions. Furthermore, the recurrence times for volcanic eruptions of a certain VEI are summarized following Gudmundsson et al. (2008).

Classified on the basis of VEI, the Eyjafjalla eruption was a medium-size eruption (VEI = 4) with a recurrence time of 30–50 a. The volcanic plume from Eyjafjalla only rarely rose above 8 km altitude (Stohl et al. 2011). However, not only the *explosivity* but also the eruption duration influences the potential volcanic impact on air traffic. The length of a volcanic eruption can vary drastically from hours up to weeks. In the case of the Eyjafjalla volcano, the total eruption (including days with almost no activity) lasted 39 days with two explosive phases

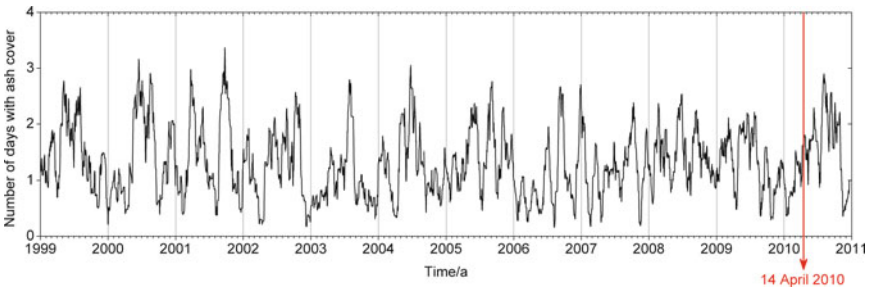

Fig. 38.10 Results from the climatological ash transport simulations for the time between 1999 and 2011. The *black line* shows the number of days on which the ash mass concentration at some altitude in the atmosphere exceeded 2 mg m^{-3} above Germany assuming an emission period of 35 days. The *arrow marks* 14 April 2010, the onset day of the Eyjafjalla eruption, leading to an ash coverage of 38 h according to this study

of 4 and 25 days, respectively, which by far exceeded the duration of any explosive eruption phase in Iceland in the past 30 years (Petersen et al. 2012).

During the eruptive days of the Eyjafjalla volcano, prevailing meteorological conditions led to the fast transport of volcanic ash south-bound into central Europe. A frequency analysis of wind patterns at the 500 hPa level during the past 18 years by Petersen et al. (2012) indicated a large deviation from the climatological mean circulation with a rather unusual *weather pattern* in spring 2010: the observed frequency of the northerly wind component at the 500 hPa level which advected the ash layers towards Europe was 71 % compared to 49 % on average. Mean wind directions at certain altitude levels provide a measure of where volcanic ash is transported, but account neither for the release heights of the volcano nor the amount of ash emitted, nor do they consider complex meteorological situations. To investigate whether the frequency of volcanic ash coverage over central Europe was unusual or not for such an eruption, a *climatological* study of volcanic ash transport was conducted at DLR. Again, the Lagrangian transport model FLEXPART was used and potential volcanic ash transport from 1999 to 2010 was simulated using ERA Interim analysis data from ECMWF. The assumptions made for the climatological study are summarized in Table 38.3.

To estimate the climatological ash transport, eruptions with a duration of 35 days were simulated continuously over 12 years (resembling a hypothetical, idealized permanent volcanic eruption with constant emission characteristics) and the duration was calculated during which the volcanic ash concentration exceeded the threshold of 2 mg m^{-3} at some altitude in the atmosphere over Germany from each eruption. Figure 38.10 shows the climatological mean for Germany from this simulation. The lowest occurrence of volcanic *ash concentrations* above 2 mg m^{-3} over Germany was 3.8 h following an eruption which started on 2 January 2010. The maximum coverage time of critical volcanic ash concentration over Germany was 81 h (for an eruption starting on 18 September 2001). The average coverage time over all eruptions simulated in this 12 year climatological

study was 32 h. The arrow in Fig. 38.10 indicates the onset of the Eyjafjalla eruption on 14 April 2010. The corresponding eruption in the climatology simulation leads to an ash coverage with concentrations above 2 mg m^{-3} for 38 h. It should be emphasized that due to the coarse horizontal resolution of the model (50 × 50 km), the frequency of 2 mg m^{-3} exceedances tends to be underestimated because a mean concentration of 2 mg m^{-3} in the whole grid box is needed for an event to be counted as a threshold exceedance. Falcon measurements showed, however, that the structure and mass concentration of volcanic ash layers can be quite variable within an area of 50 × 50 km. On the other hand, the results of the simulations are sensitive to the input parameters, especially the emission strength. Therefore, the assumption of lower values for the emission strength could result in a less frequent exceedance of the 2 mg m^{-3} limit. As already mentioned, the emission strength was chosen such that the FLEXPART simulations showed the best agreement with the results from the individual Falcon flights.

In summary, the statistical simulations show that the transport of Eyjafjalla ash towards Europe was not an unusual event given the anomalously long duration of the explosive eruption. Also on geological time scales, the transport of volcanic ash from Iceland towards northern and central Europe is not unprecedented. Swindles et al. (2011) investigated the frequency of volcanic ash clouds reaching northern Europe in the past 7 000 years on the basis of peat and lake sedimentary archives. They found a mean return interval of 56 ± 9 a (the range of return intervals of 6–115 a) for the past 1 000 years and suggested that for a 10-a period there is at least a 16 % probability of a tephra fallout event in northern Europe.

38.8 Conclusions

The closure of the airspace following the 2010 eruption of the Eyjafjalla volcano was the logical consequence of the regulations based on the ICAO "*zero-tolerance rule*" based on the information available at the time. Subsequently, to avoid further massive disruption of air traffic, a new regulation with a volcanic ash mass concentration threshold value for "*safe flying*" of 2 mg m^{-3} was established in Germany. In agreement with *ground-based lidar* observations, *satellite* observations, and *model simulations*, *measurements* performed with the Falcon suggest that this new threshold value was only briefly exceeded over the northwestern and western part of Germany and most likely never exceeded over the southeastern part of Germany during the entire Eyjafjalla eruption period in April/May 2010.

With the safety criterion "no flight in visible ash" in mind, we simulated with a radiative transfer model whether a pilot on board an aircraft can visually detect regions with potentially dangerous volcanic ash concentrations and avoid them. These idealized simulations show that ash clouds with mass concentrations above 0.2–0.5 mg m^{-3} are visible under suitable daytime conditions if not obstructed by other clouds, but cannot safely be distinguished from other aerosol layers.

Although the duration of the Eyjafjalla eruption was exceptionally long, ash transport to central Europe should not be considered as exceptional. Future volcanic events may well cover Germany with ash clouds of higher mass concentrations and for even longer periods. Because of the high variability of ash mass concentrations, effective regulations would require not only specification of critical concentration thresholds but also ash dose limits. With the more specific regulation of a concentration threshold value, the prediction of the extent of future ash clouds will be more challenging: now, not only the location of an ash plume but also the mass concentration has to be forecasted accurately. However, the concentration in the ash plume is critically dependent on the emission strength of the volcano, which may vary even within minutes.

Safe aviation requires reliable tools to predict and detect regions with ash loads exceeding certain thresholds. For efficient aviation, one also needs reliable tools to predict and identify regions free of dangerous ash loads. To avoid future economic loss due to unnecessary air traffic closures, such prognostic and diagnostic tools have to be certified and implemented in the decision processes. At this institute, development work to reach these goals has been started based on satellite-based mineral dust detection methods using *Meteosat* in combination with weather prediction data. Geostationary satellite observations are particularly useful for detecting quickly moving ash clouds despite limited spatial resolution. Furthermore, the partially-European-funded project Weather Hazards for Aviation (WEZARD) has been started to identify gaps in the understanding of volcanic ash. One difficulty in testing new airborne instrumentation for volcanic ash is that volcanic eruptions cannot be predicted. However, the microphysical similarities between volcanic ash and mineral dust can be used to test volcanic ash tools under controlled conditions in airborne field studies of mineral dust. In addition, it must be assured that newly developed onboard detectors are capable of distinguishing volcanic ash from mineral dust or water and ice clouds. The development of such methods and experimental validation with existing and further measurements, such as the planned Saharan Aerosol Long-range Transport and Aerosol-Cloud-Interaction Experiment (SALTRACE) campaign, form the core of the future DLR project Volcanic Ash Impact on the Air Transport System (VolcATS) focusing on volcanic research.

References

Ansmann, A., Tesche, M., Gross, S., Freudenthaler, V., Seifert, P., Hiebsch, A., Schmidt, J., Wandinger, U., Mattis, I., Muller, D., et al.: The 16 April 2010 major volcanic ash plume over central Europe: EARLINET lidar and AERONET photometer observations at Leipzig and Munich. Germany. Geophys. Res. Lett. **37**, L13810 (2010). doi:10.1029/2010gl043809

Arason, P., Petersen, G.N., Bjornsson, H.: Observations of the altitude of the volcanic plume during the eruption of Eyjafjallajökull, April–May 2010. Earth Syst. Sci. Data **3**, 9–17 (2011). doi:10.5194/essd-3-9-2011

Buras, R., Mayer, B.: Efficient unbiased variance reduction techniques for Monte Carlo simulations of radiative transfer in cloudy atmospheres: the solution. J. Quant. Spectrosc. Radiat. Transf. **112**, 434–447 (2011). doi:10.1016/j.jqsrt.2010.10.005

Casadevall, T.J.: Volcanic hazards and aviation safety: lessons of the past decade. Flight Safety Foundation—Flight Safety Digest, 1–9 (1993)

Casadevall, T.J.: The 1989–1990 eruption of Redoubt Volcano, Alaska: impacts on aircraft operations. J. Volcanol. Geoth. Res. **62**, 301–316 (1994). doi:10.1016/0377-0273(94)90038-8

Esselborn, M., Wirth, M., Fix, A., Weinzierl, B., Rasp, K., Tesche, M., Petzold, A.: Spatial distribution and optical properties of Saharan dust observed by airborne high spectral resolution lidar during SAMUM 2006. Tellus **61B**, 131–143 (2009). doi:10.1111/j.1600-0889.2008.00394.x

Gasteiger, J., Gross, S., Freudenthaler, V., Wiegner, M.: Volcanic ash from Iceland over Munich: mass concentration retrieved from ground-based remote sensing measurements. Atmos. Chem. Phys. **11**, 2209–2223 (2011). doi:10.5194/acp-11-2209-2011

Gudmundsson, M.T., Larsen, G., Höskuldsson, A., Gylfason, A.G.: Volcanic hazards in Iceland. Jökull **58**, 18 (2008)

Guffanti, M., Casadevall, T.J., Budding, K.: 1953–2009: encounters of aircraft with volcanic ash clouds; a compilation of known incidents. U.S. Geological Survey Data Series 545, ver. 1.0, 12 p., plus 14 appendixes including the compilation database [http://pubs.usgs.gov/ds/545/] (2010)

Heinold, B., Tegen, I., Wolke, R., Ansmann, A., Mattis, I., Minikin, A., Schumann, U., Weinzierl, B.: Simulations of the 2010 Eyjafjallajökull volcanic ash dispersal over Europe using COSMO-MUSCAT. Atmos. Environ. (2011). doi:10.1016/j.atmosenv.2011.05.021

ICAO: Manual on Volcanic Ash, Radioactive Material and Toxic Chemical Clouds (First Edition). Doc 9691-AN/954 variously paged (2001)

ICAO: Manual on Volcanic Ash, Radioactive Material and Toxic Chemical Clouds (Second Edition). Doc 9691-AN/954, variously paged, accessed March 19, 2012, at http://www.paris.icao.int/news/pdf/9691.pdf (2007)

Kristiansen, N.I., Stohl, A., Prata, A.J., Bukowiecki, N., Dacre, H., Eckhardt, S., Henne, S., Hort, M.C., Johnson, B.T., Marenco, F., et al.: Performance assessment of a volcanic ash transport model mini-ensemble used for inverse modeling of the 2010 Eyjafjallajokull eruption. J. Geophys. Res. **117**, (2012). doi:10.1029/2011jd016844

Kueppers, U., Cimarelli, C., Hess, K.U., Dingwell, D.B., Rickerby, D.S., Madden, P.: Thermal stability of Volcanic ash versus turbine ingestion test sands: an experimental investigation. In: Atlantic Conference on Eyjafjallajökull and Aviation (2010)

Matthias, V., Aulinger, A., Bieser, J., Cuesta, J., Geyer, B., Langmann, B., Serikov, I., Mattis, I., Minikin, A., Mona, L., et al.: The ash dispersion over Europe during the Eyjafjallajökull eruption—comparison of CMAQ simulations to remote sensing and air-borne in situ observations. Atmos. Env. **48**, 184–194 (2012). doi:10.1016/j.atmosenv.2011.06.077

Mayer, B., Kylling, A.: Technical note: the libRadtran software package for radiative transfer calculations—description and examples of use. Atmos. Chem. Phys. **5**, 1855–1877 (2005). doi:10.5194/acp-5-1855-2005

Mayer, B.: Radiative Transfer in the Cloudy Atmosphere, vol. 1, EDP Sciences, 75–99 pp (2009). doi:10.1140/epjconf/e2009-00912-1

Mechnich, P., Braue, W., Schulz, U.: High-temperature corrosion of EB-PVD Yttria partially stabilized zirconia thermal barrier coatings with an artificial volcanic ash overlay. J. Am. Ceram. Soc. **94**, 925–931 (2011). doi:10.1111/j.1551-2916.2010.04166.x

Newhall, C.G., Self, S.: The volcanic explosivity index (VEI)—an estimate of the explosive magnitude for historical volcanism. J. Geophys. Res. **87**, 1231–1238 (1982). doi:10.1029/JC087iC02p01231

Petersen, G.N., Bjornsson, H., Arason, P.: The impact of the atmosphere on the Eyjafjallajokull 2010 eruption plume. J. Geophys. Res. **117**, (2012). doi:10.1029/2011jd016762

Petzold, A., Rasp, K., Weinzierl, B., Esselborn, M., Hamburger, T., Dörnbrack, A., Kandler, K., Schütz, L., Knippertz, P., Fiebig, M., et al.: Saharan dust absorption and refractive index from aircraft-based observations during SAMUM 2006. Tellus **61B**, 118–130 (2009). doi:10.1111/j.1600-0889.2008.00383.x

Petzold, A., Veira, A., Mund, S., Esselborn, M., Kiemle, C., Weinzierl, B., Hamburger, T., Ehret, G., Lieke, K., Kandler, K.: Mixing of mineral dust with urban pollution aerosol over Dakar (Senegal): impact on dust physico-chemical and radiative properties. Tellus **63B**, 619–634 (2011). doi:10.1111/j.1600-0889.2011.00547.x

Rix, M., Valks, P., Hao, N., Loyola, D.G., Schlager, H., Huntrieser, H.H., Flemming, J., Koehler, U., Schumann, U., Inness, A.: Volcanic SO_2, BrO and plume height estimations using GOME-2 satellite measurements during the eruption of Eyjafjallajökull in May 2010. J. Geophys. Res. (2012, in press). doi:10.1029/2011JD016718

Schumann, U., Weinzierl, B., Reitebuch, O., Schlager, H., Minikin, A., Forster, C., Baumann, R., Sailer, T., Graf, K., Mannstein, H., et al.: Airborne observations of the Eyjafjalla volcano ash cloud over Europe during air space closure in April and May 2010. Atmos. Chem. Phys. **11**, 2245–2279 (2011). doi:10.5194/acp-11-2245-2011

Siebert, L., Simkin, T., Kimberly, P.: Volcanoes of the World, 3rd edn. University of California Press, Berkeley (2010). ISBN:9780520268777

Stohl, A., Forster, C., Frank, A., Seibert, P., Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. Atmos. Chem. Phys. **5**, 2461–2474 (2005). doi:10.5194/acp-5-2461-2005

Stohl, A., Prata, A.J., Eckhardt, S., Clarisse, L., Durant, A., Henne, S., Kristiansen, N.I., Minikin, A., Schumann, U., Seibert, P., et al.: Determination of time- and height-resolved volcanic ash emissions and their use for quantitative ash dispersion modeling: the 2010 Eyjafjallajokull eruption. Atmos. Chem. Phys. **11**, 4333–4351 (2011). doi:10.5194/acp-11-4333-2011

Swindles, G.T., Lawson, I.T., Savov, I.P., Connor, C.B., Plunkett, G.: A 7000 yr perspective on volcanic ash clouds affecting northern Europe. Geology **39**, 887–890 (2011). doi:10.1130/g32146.1

Turnbull, K., Johnson, B., Marenco, F., Haywood, J., Minikin, A., Weinzierl, B., Schlager, H., Schumann, U., Leadbetter, S., Woolley, A.: A case study of observations of volcanic ash from the Eyjafjallajökull eruption: 1. In situ airborne observations. J. Geophys. Res. **117**, D00U12 (2012). doi:10.1029/2011jd016688

VAAC: Volcanic Ash Advisory from London—Issued Graphics. Accessed March 19, 2012, at http://www.metoffice.gov.uk/aviation/vacc/vaacukvag.html (2012)

Webster, H.N., Thomson, D.J., Johnson, B.T., Heard, I.P.C., Turnbull, K., Marenco, F., Kristiansen, N.I., Dorsey, J., Minikin, A., Weinzierl, B., et al.: Operational prediction of ash concentrations in the distal volcanic cloud from the 2010 Eyjafjallajokull eruption. J. Geophys. Res. **117**, (2012). doi:10.1029/2011jd016790

Weinzierl, B., Petzold, A., Esselborn, M., Wirth, M., Rasp, K., Kandler, K., Schütz, L., Koepke, P., Fiebig, M.: Airborne measurements of dust layer properties, particle size distribution and mixing state of Saharan dust during SAMUM 2006. Tellus **61B**, 96–117 (2009). doi:10.1111/j.1600-0889.2008.00392.x

Weinzierl, B., Sauer, D., Esselborn, M., Petzold, A., Veira, A., Rose, M., Mund, S., Wirth, M., Ansmann, A., Tesche, M., et al.: Microphysical and optical properties of dust and tropical biomass burning aerosol layers in the Cape Verde region—an overview of the airborne in situ and lidar measurements during SAMUM-2. Tellus **63B**, 589–618 (2011). doi:10.1111/j.1600-0889.2011.00566.x

Weinzierl, B., Sauer, D., Minikin, A., Reitebuch, O., Dahlkötter, F., Mayer, B., Emde, C., Tegen, I., Gasteiger, J., Petzold, A. et al.: On the visibility of airborne volcanic ash and mineral dust from the pilot's perspective in flight. Phys. Chem. Earth (2012). doi:10.1016/j.pce.2012.04.003

Witham, C., Webster, H., Hort, M., Jones, A., Thomson, D.: Modelling concentrations of volcanic ash encountered by aircraft in past eruptions. Atmos. Env. **48**, 219–229 (2012). doi:10.1016/j.atmosenv.2011.06.073

Mitigating the Impact of Adverse Weather on Aviation

39

Thomas Gerz, Caroline Forster and Arnold Tafferner

Abstract

Weather has a significant impact on the safety and efficiency of air traffic during all phases of flight. Especially information on adverse weather must be tailored to the user's needs, easy to understand, self-explaining and clear in its message. DLR-IPA has developed a concept and tools to detect, track and predict hazardous weather elements and provide this information in simple unambiguous form to controllers and pilots. It has been demonstrated that these products make a significant contribution to raising the safety and efficiency of the air transport system.

39.1 Weather Impacts the Safety, Efficiency, and Sustainability of Aviation

The weather's impact on aviation is threefold: The safety, economic efficiency and ecologic sustainability of aviation depend, sometimes crucially, sometimes marginally, on the state and motion of the air masses through which an aircraft cruises.

T. Gerz (✉) · C. Forster · A. Tafferner
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Thomas.Gerz@dlr.de

C. Forster
e-mail: Caroline.Forster@dlr.de

A. Tafferner
e-mail: Arnold.Tafferner@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_39,

Vision 2020 (GoP 2001) and follow-on initiatives by the European Commission have set ambitious goals to increase aviation safety, punctuality and sustainability. The goals have to be met under all weather conditions, "flying safely in all weather, aircraft are running on schedule 99 % of the time." However, weather is not a technical problem which can totally be controlled: "we cannot control it but we need to learn to live with the elements and steadily eliminate the service disruption that they may cause." Users have to bear in mind that weather, by nature, is difficult to predict and can be done so only within limits.

Adverse weather is hardly the single source of accidents in aviation, but it is very often a contributing factor. Between 1995 and 2004 weather was the primary cause in 13 % of all aircraft losses; in 33 % of all accidents between 2004 and 2007 weather was involved as a primary or secondary cause. Weather is, on the other hand, the largest single source of delays and inconvenience in air traffic. Airport hubs and flight corridors which operate close to their capacity limits will be disrupted already by small weather events leading to nevertheless high costs and delays. Thus, weather is responsible for 40–50 % of all flight delays in Europe already today. In 2011 weather was the main reason for delays at Munich airport in more than 80 % of all delay cases. Also in the USA, weather is the primary reason for disruptions in the air transport system. For example, up to 90 % of all delays in the airspace over the USA during summer months are due to thunderstorm activity (Leighton 2006).

A Eurocontrol Performance Review Commission (2011) study indicates that the weather phenomena with the largest impact on delays in European air traffic flow management are thunderstorms (30.9 % of the cases) and snow/ice (22.1 %), followed by low visibility/ceiling (21.5 %) and wind (14.5 %). All phases of flight are exposed to weather, but flights in the vicinity of airports, the so-called terminal maneuvering area (TMA), are mainly hindered by weather conditions like adverse winds, strong precipitation, lightning, etc. Today, aviation is one of a few business sectors with large annual growth rates; the number of flights in Europe increased about 21 % from 2003 to 2008. Since growth in aviation basically means a growing number of scheduled flights and not so much a growth in the infrastructure, aviation will become even more vulnerable to weather in the future.

These numbers and trends indicate that there is a clear need for improved weather information that will help all parties to mitigate weather impact well before it happens. The present meteorological information for aviation like aerodrome forecasts (TAF), trend-type landing forecasts (TREND), and aerodrome or en-route warnings for airmen (AIRMET/SIGMET) do not meet the needs of today's aircraft operators and pilots. They are not specific enough, do not cover all phases of flight nor contain all important meteorological parameters.

There is a growing community in aviation who believes that the disruptive impact of weather on aviation can be mitigated by using and integrating dedicated and tailored observations and forecasts of disruptive weather events. Instead of reacting to an event that is already happening (as is the practise today), anticipating the weather to come and taking measures proactively is the growing

consensus among stakeholders. Related studies show that a reduction of weather-induced costs by 40 % and more seems feasible.

39.2 Weather and Air Traffic Management

Future air traffic management (ATM), as envisaged by the European and US-American ATM modernization programs SESAR and NextGen, will focus on trajectory planning in space and time (i.e., in four dimensions, 4d) to facilitate flexible airspace management. Conflicts between flight trajectories must be recognized and resolved quickly, especially in congested airspace, by analyzing and assessing a huge amount of data from as many different types and sources as there are different kinds of aircraft performance, trajectory profiles, aeronautical and geographical data, to name just a few of the factors involved. Furthermore, the operational environment of demand and capacity, the state of use and constraints of airports and airspace, must be considered at the same time.

In order to run a 4d trajectory management system operationally and efficiently it is perspicuous that the aviation-meteorological information (MET) along flight trajectories, at destinations and alternative goals is provided adequately and in a timely manner. Especially information on adverse weather must be tailored to the user's needs, easy to understand, self-explaining and clear in its message. MET must become a lively part of the decision making process in aviation. Essential elements to complete that process include:

- A sufficiently dense weather observation network, weather diagnosis, short-term forecasting ("nowcasting") of disruptive weather, and detailed probabilistic weather forecasts;
- The translation of disruptive weather into aviation constraints ("no-fly" areas) and the conversion of these constraints to aviation impact (demand versus reduced capacity);
- The integration of tailored and timely MET and its impact on aviation in operational processes for collaborative and flexible decision making.

The situation has become so pressing that the World Meteorological Organization (WMO) established an expert team on aeronautical meteorology in 2006. This team is in charge of setting up meteorological services for the terminal area to provide the new information required today to supplement classical terminal area forecasts. This WMO team is also establishing links to the respective working groups of the International Civil Aviation Organization (ICAO) to foster the necessary rule making.

39.3 Weather and Aircraft

Adverse weather conditions due to turbulence, lightning, wind shear, heavy precipitation, hail or icing influence the operating qualities of an aircraft. They reduce passenger comfort or even jeopardize the safety of a flight. Adverse weather may occur at all flight levels everywhere in the atmosphere. Such blockage always increases the workload of the flight crew, which may negatively impact safety. Thunderstorms are top-ranked by pilots as weather situations compromising flight safety since most of the mentioned hazards come together in such an event. Strong turbulence and wind shear lead to accelerations of the aircraft which can result in loss of flight altitude or a rotation of the aircraft; turbulence also gives rise to vibrations of the airframe and wings such that an engine may fall off. Icing, i.e., cloud and rain droplets which freeze on the cold body of an aircraft, can block sensors and flight control devices resulting in false speed indications and reduced maneuverability. A lightning stroke, although of no harm to crew and passengers, may destroy electrical circuits. Hail and heavy precipitation can destroy the cockpit window and the radar dome or result in an engine flame-out. Other hazardous phenomena linked to weather are wake vortices, volcanic ash and desert dust storms; see Hauf et al. (2004) for more details.

Today, information for pilots about such adverse weather is, if at all, based on significant weather charts and the like. This information is handed to pilots before the flight when they check the aircraft for take-off. Such services, however, do not give the necessary, tailored and reliable information required for a particular flight and are often outdated when needed. Significant weather charts, for example, are deduced from coarse numerical weather forecasting models many hours ahead. Moreover, neither do sensors exist today to recognize phenomena like clear-air turbulence or wake vortices from on-board an aircraft, nor are aircraft flight control mechanisms mature or available to mitigate respective encounters.

Hence, what is required today and even more in light of upcoming dynamic airspace management, where autonomous flights will have an increasing share, is, besides a modern ATM system, also an adequately equipped aircraft which

- Recognizes hazardous situations caused by meteorological phenomena,
- Informs pilots and crew in an appropriate way and suggests alternative routing, and
- Is capable of mitigating encountered atmospheric disruptions automatically by control mechanisms.

This information has to originate from both on-board data as well as data uplinked from ground systems.

Examples are given in the following sections of how tailored information on thunderstorms and winter weather conditions can help aircraft pilots and air traffic controllers to master critical situations. More details can be found in Gerz and Schwarz (2012).

39.4 Integrated Observing and Forecasting System

In order to fulfill the needs of aviation stakeholders, both on ground and in the air, DLR-IPA spawned the idea and concept of an integrated weather observing and forecast system named WxFUSION. The system combines data from observations, nowcasting tools, and numerical models in order to detect, track, nowcast up to about 6 h, and forecast beyond 6 h hazardous weather phenomena for aviation purposes (Tafferner et al. 2008a; Forster and Tafferner 2009). WxFUSION is currently under development. The aim of this system is to present a timely, dedicated and adequate description of the weather phenomenon of interest and its future development to all aviation stakeholders both at the ground and in the air. In order to enable an efficient and quick collaborative decision making process on the users' side, the information on weather hazards is presented in a simple, easy to read and unambiguous way which does not need any interpretation. That means the complexity of a weather hazard is reduced to a simple object represented by, e.g., contours encompassing hazardous areas for air traffic. Examples are given below.

When the WxFUSION system is fully active it will combine different data from observation and prediction sources to develop dedicated MET products. Two important components of WxFUSION with regard to thunderstorm hazards have been developed and already successfully demonstrated and validated: the tools Cb-TRAM (Cumulonimbus Tracking and Monitoring, Zinner et al. (2008)) and Radar Tracking and Monitoring (Rad-TRAM, Kober and Tafferner (2009)). The algorithms detect, track, and nowcast up to one hour thunderstorm cells based on satellite and radar data. Data from these tools have been tested both at the ground for use at airports and by air navigation services as well as in the air on board aircraft.

39.5 Enhancing Situational Awareness of Thunderstorms for Flight Crews on Board Aircraft

Within FLYSAFE (http://www.eu-flysafe.org/Project.html), a project of the 6th EU framework program, the focus was on the development and testing of on-board systems providing the pilot with information on the three threats weather, traffic collision, and ground collision. DLR-IPA led the development of a weather information and management system for thunderstorms. The system combines thunderstorm detection and nowcast from Cb-TRAM with ground-based detection and nowcast of heavy precipitation cells and lightning (Tafferner et al. 2008b; Pradier et al. 2009; Zinner and Betz 2009). It could be shown that the data, which can be uplinked from ground to the cockpit, compare well to the weather radar depiction on board an aircraft. Most importantly, the ground data complete the picture of the weather hazard on board the aircraft as they survey a much larger area than the on-board radar and fuse data from multiple observational sources. How the situational awareness of the pilots could be significantly enhanced is outlined in Fig. 39.1.

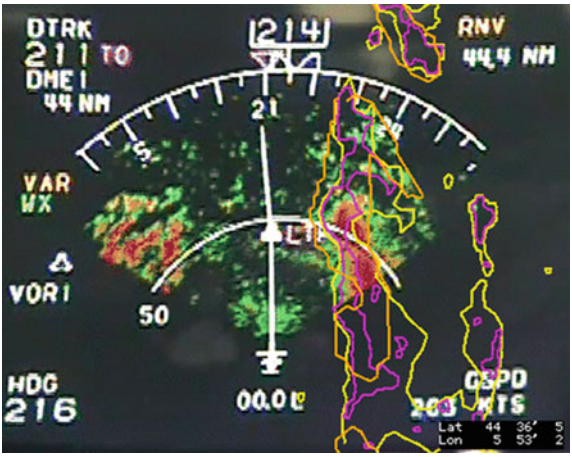

Fig. 39.1 Snapshot of a video of the on-board weather radar image from the French research aircraft SAFIRE ATR-42 close to Grenoble for 14:05 UTC on August 19, 2008 with subsequently superimposed weather object contours from the ground system. The colored shading represents the radar returns from precipitation or ground clutter with *green* indicating *light*, *orange moderate*, and *red strong* radar returns. *Orange* contours indicate Cb-TRAM objects; *yellow* and *pink* contours indicate heavy precipitation cells for two different intensities representing moderate and severe precipitation, respectively. (Figure adapted from Tafferner et al. (2010), © American Meteorological Society)

Figure 39.1shows a snapshot of the radar display recorded during an experimental flight with the French research aircraft SAFIRE ATR-42 (Sénési et al. 2009). Cb-TRAM data and heavy precipitation cells for two different precipitation intensities are indicated as colored contours. The spatial distribution of the Cb cells agrees well with the close range radar depiction to the forward right of the aircraft position (near the 50 nautical mile range circle). However, beyond the close range, Cb cells indicate additional thunderstorm activity and also another cell further away (beyond the 100 nautical mile range circle). Both these Cb cells cannot be seen by the on-board radar at this instant of time as the radar beam is attenuated by the closer Cb cell. After 10 and 20 min these cells are confirmed by the on-board radar as the flight continues. Note that the radar returns on the left side beyond 50 nautical miles are due to ground clutter from the Alps and therefore not confirmed as thunderstorm activity by the ground system.

DLR-IPA has also demonstrated some cases where thunderstorm information from satellite and ground-radar observing systems would most probably have helped to raise the pilot's situational awareness with respect to a weather hazard that may be encountered (Tafferner et al. 2010). One example is the tragic accident of Air France flight AF447 over the South Atlantic on June 1, 2009 where 216 passengers and 12 crew members lost their life. Figure 39.2 shows a Meteosat-9 infrared image with Cb-TRAM thunderstorm objects (contours) superimposed at 01:30 UTC, i.e., 40 min before the last automated message from AF447 was

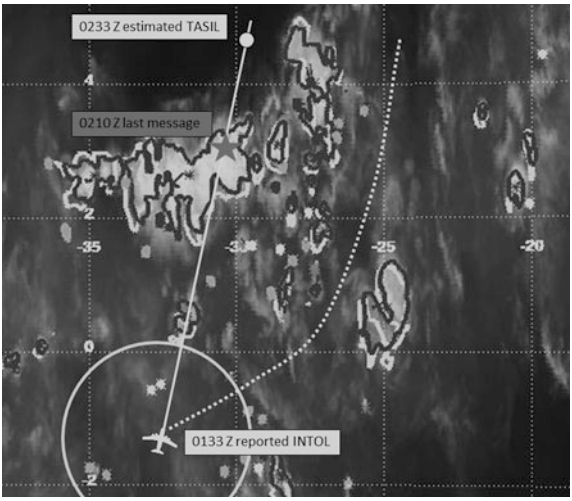

Fig. 39.2 METEOSAT-9 infrared image over the Atlantic east of Brazil with Cb-TRAM analysis (*black* contours) superimposed on June 1, 2009 at 01:30 UTC. *White* shaded areas indicate high thunderstorm clouds. Also marked is the flight route of AF447 between the way points INTOL and TASIL (*white line*), the location of the last automated message received from AF447 (*gray* star), and a possible alternative flight route (*white dotted line*). The *white* circle indicates an approximate on-board radar range of about 80 nautical miles. Geographical latitudes and longitudes are indicated as *thin white dotted lines*. (Figure adapted from Tafferner et al. (2010), © American Meteorological Society)

received by ATC. The white line indicates the flight route between the waypoints INTOL (reported at 01:33 UTC) and TASIL (expected 02:33 UTC); the gray star indicates the location when the last message was received at 02:10 UTC, and the white circle the approximate range of the on-board radar beam (about 80 nautical miles). With information from the on-board radar only, the pilots could not foresee the strong convective activity on their future track. In addition, just by looking out the window it was probably impossible for them to recognize the thunderstorms in the far distance due to the darkness. Furthermore, there were few or no lightning discharges (as later analyzed from the networks for this region) at this time which could have warned the pilots.

Figure 39.2 elucidates that the Cb-TRAM information is able to give a timely overview of the situation. Clearly, if such information is uplinked to the cockpit, it would help pilots to make decisions based on better surveillance. Having considered re-routing constraints like other traffic and fuel consumption, they could then choose another flight track that is not impacted by thunderstorms, e.g., the one indicated by the dotted white line in Fig. 39.2. The AF447 case and the FLYSAFE example tell us that the observational data from radar and satellite processed by Rad-TRAM and Cb-TRAM can provide a good overview of a situation and could help to provide the necessary warnings. However, another example illustrates that

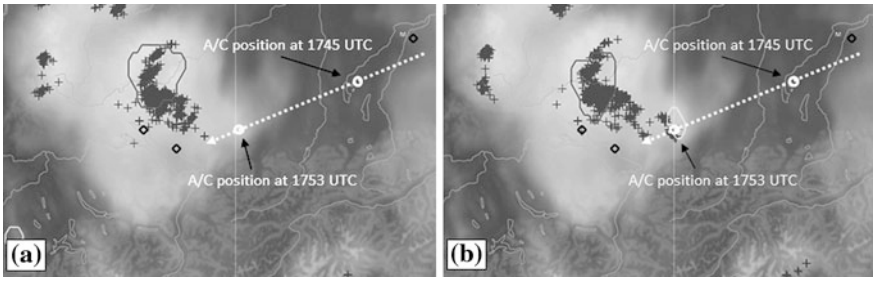

Fig. 39.3 Thunderstorm analysis for southern Germany between Lake Constance and Munich on May 9, 2009 at (**a**) 17:55 UTC and (**b**) 18:00 UTC. METEOSAT-9 infrared image with Cb-TRAM detections (*gray* contours) and lightning incidents (crosses) superimposed. The *white* shaded areas indicate high thunderstorm clouds. The flight track is indicated as a *white dotted line*. (Figure adapted from Tafferner et al. (2010), © American Meteorological Society)

there is sometimes a need for additional data from other sources, which need to be combined with radar and satellite data. On May 9, 2009 at 17:53 UTC, a Lufthansa flight from Munich to Lisbon was hit by severe turbulence southeast of a mature thunderstorm (gray contours in Fig. 39.3) near Lake Constance. Two flight attendants and thirteen passengers were injured, and the aircraft had to stop over in Geneva in order to take care of them.

The METEOSAT-9 infrared satellite image, the Cb-TRAM contours, and lightning discharges at 17:55 UTC, i.e., 2 min after the incident, do not indicate convective activity at the position of the aircraft (Fig. 39.3a). Five minutes later, however, lightning discharges occur and convection initiation is detected by Cb-TRAM at the position of the incident (new contour and crosses in Fig. 39.3b). A thorough analysis revealed that the aircraft was hit by a newly developing cell that could not be seen by Cb-TRAM before 18:00 UTC, and was most likely not recognized by the pilots with the on-board radar, as it was not yet related to heavy precipitation. Such situations are difficult to predict from observations only. It is not uncommon that downdrafts and the associated outflows from mature thunderstorms lift air masses and thereby trigger new cells ahead of the existing storms. To detect these processes and enable timely warnings a combination of all available data from different observations together with numerical model forecast data is necessary, as will be provided in WxFUSION.

39.6 Informing Air Navigation Services and Airport Authorities on Thunderstorms

To date, decision makers at the ground from air navigation services and airport authorities have access to weather observations like weather radar data and lightning information which can be used to gain an overview of the current

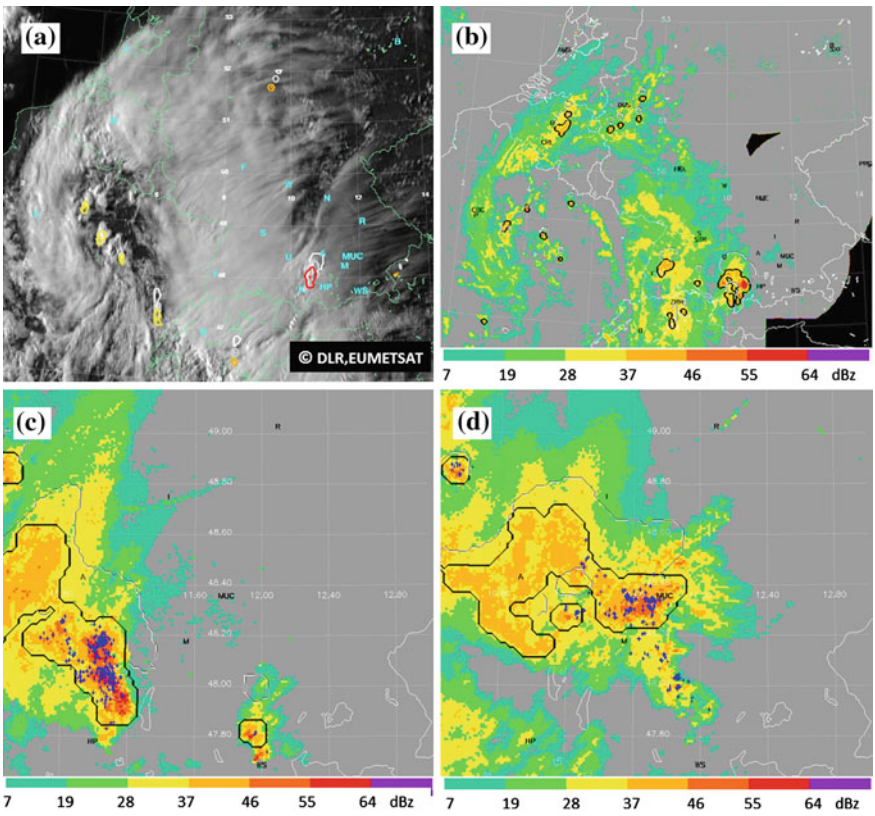

Fig. 39.4 Thunderstorm situation on July 19, 2011. (**a**) METEOSAT-8 HRV image over eastern France, Benelux, Germany, and the Alpine region at 16:15 UTC with Cb-TRAM contours super-imposed. *Yellow*, *orange*, and *red* contours indicate the development stages convection initiation, rapid development, and mature, respectively. (**b**) Radar composite from the German Weather Service (DWD) over the same area with Rad-TRAM contours super-imposed (*black*) at 16:15 UTC. (**c**) as in (**b**), but for the TMA of MUC at 18:15 UTC, and (**d**) as in (**c**), but for 19:45 UTC. The *grey* and *white* polygons in (**a**) are the 30 and 60 min nowcasts, respectively. The *thin black* contours in (**b**), (**c**) and (**d**) are the 60 min nowcasts. *Blue* crosses in (**c**) and (**d**) represent lightning detections. (Figure adapted from Forster and Tafferner (2012))

thunderstorm situation. However, information on the short-term future development, moving direction and moving speed of thunderstorms is not yet available in a satisfactory manner. In order to demonstrate the added value of tailored and timely detection and nowcasting of thunderstorms also for airports and air navigation services, DLR-IPA conducted two campaigns in the summers of 2010 and 2011 at and around Munich airport (MUC) in close collaboration with the German Meteorological Service, DWD. The main goal of these campaigns was real time demonstration of Cb-TRAM and Rad-TRAM running quasi-operationally and providing the products to users at airport operations centers, FMG, the Lufthansa

Hub Control Center, DLH-HCC, and the German Air Navigation Service Provider, DFS. The Cb-TRAM and Rad-TRAM products were presented on web pages as well as on displays that were already well established at the facilities of the FMG, DLH-HCC, and DFS; i.e., the users could test the DLR products on displays they were familiar with. In addition, data from the DLR polarimetric radar POLDIRAD and products from DWD's numerical weather prediction model COSMO-DE were graphically prepared and presented on these web pages. Thunderstorm warnings were issued by DLR-IPA via e-mail to all users whenever thunderstorms were detected and forecast in the TMA of MUC. Feedback was expected from the users on the benefits and disadvantages of the new products for thunderstorm identification, tracking and prediction.

As an example of how the DLR nowcasting products were presented to users, a thunderstorm case from July 19, 2011 is shown in Fig. 39.4. Ahead of a low pressure system, warm and humid air masses were advected to Bavaria. At 16:15 UTC, the vortex of the low pressure system over France can clearly be seen in the high-resolution visible satellite image with Cb-TRAM contours superimposed (Fig. 39.4a). Convection initiation (yellow contours) is detected near the center of the low pressure system, and a mature thunderstorm (red contour) has already developed between Kempten and Hohenpeissenberg (marked by K and HP in Fig. 39.4a). Both Cb-TRAM and Rad-TRAM predict that the cell will move in a northeasterly direction (grey and white contours in Fig. 39.4a and b). It is also characterized by lightning activity, heavy precipitation, and hail (reflectivity >55 dBZ, reddish colors in Fig. 39.4b). A first thunderstorm warning was sent to the users at this time, with Fig. 39.4b attached. The thunderstorm intensified and moved towards MUC. At 18:15 UTC the thunderstorm had crossed the river Lech and reached lake Ammersee (Fig. 39.4c). A second warning was sent to the users at this time with Fig. 39.4c attached. One hour later Cb-TRAM and Rad-TRAM predicted the arrival of the thunderstorm at MUC at about 19:45 UTC, where it was indeed observed at that time (Fig. 39.4d).

The MUC TMA was affected by thunderstorms on 40 and 39 days between June and September in 2010 and 2011, respectively. For these days, a quantitative assessment of the nowcasting quality of Rad-TRAM and Cb-TRAM has been performed. To calculate forecast skill scores an object-based method was applied to the Rad-TRAM data. The observed cells in Rad-TRAM are compared to their nowcast cells at different lead times up to 1 h. If a nowcast object overlaps with an observed one, it is counted as a hit, nonoverlapping observed objects are counted as misses, and nonoverlapping nowcast objects are counted as false alarms. Then the skill score probability of detection (POD) with POD = hits/(hits + misses) and the false alarm ratio (FAR) with FAR = false alarms/(hits + false alarms) are calculated and interpreted. The perfect nowcast would have a POD = 1 and a FAR = 0. As Rad-TRAM cannot predict the formation of a new cell, only those observed objects that also have a nowcast were accounted for in the evaluation. The results are presented in Table 39.1. Up to 30 min, the scores are close to a perfect nowcast, but downgrade quickly for longer forecast lead times. The FARs for lead times longer than 30 min are not yet satisfactory. For instance, traffic

Table 39.1 Mean values of POD and FAR and the respective standard deviations (STD) for all thunderstorm days for the 15, 30, 45, and 60 min nowcasts

	15 min	30 min	45 min	60 min
POD/FAR	0.98/0.01	0.95/0.03	0.89/0.09	0.75/0.23
POD STD/FAR STD	0.06/0.004	0.06/0.007	0.06/0.01	0.05/0.03

managers from the FMG need more accuracy up to 1 h in order to decide whether operations on the movement area of the airport have to be shut down or not.

At the end we received very positive feedback from users including numerous comments and suggestions from personal conversations with supervisors at DFS Tower and Center MUC, directors of operations at DLH HCC, and traffic managers from FMG. The nowcasting up to one hour was found to be accurate enough to enable reasonable planning, e.g., the deployment of additional staff or the ordering of overtime, and it was useful to estimate how long the airport would be affected by a thunderstorm. The display and availability of data also from neighboring countries was of great value especially for the DFS. Thunderstorms moving towards Germany could then be recognized early and action could be triggered under less time pressure. Here, it was however noted that forecasts up to 2 h would be of great advantage, as this is the timeframe for actions with regard to the coordination of European flights at the Central Flow Management Unit of EUROCONTROL.

The thunderstorm warnings issued by DLR-IPA turned out to be very helpful for collaborative decision making at the airport. As all users received a thunderstorm warning at the same time and could get an overview of the current weather situation by inspecting the DLR nowcasting products, they had the same situational awareness and could make decisions and trigger actions with less time pressure than without the warnings and products. Users from DFS stated that the decision making process would be even more efficient if the products could also be uplinked to the cockpit of aircraft.

39.7 Warning Aviation Stakeholders of Winter Weather Conditions

In order to mitigate the impact of winter weather conditions on airport operations more efficiently, all stakeholders at Munich airport agree that substantial improvement in nowcasting these conditions is highly desirable. This comprises onset, duration and type of precipitation, icing at the surface, freezing fog, aircraft icing at ground, and visibility. DLR-IPA is developing a nowcasting system that provides users with 0–2 h forecasts of these winter weather conditions. The task "nowcasting winter weather at Munich airport" can be approached by a stepwise procedure:

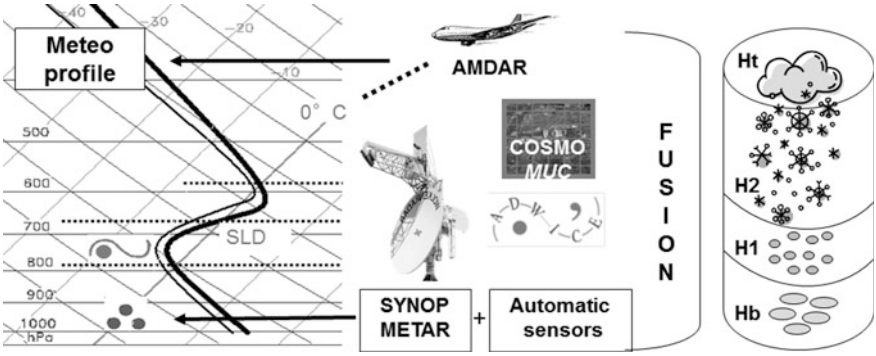

Fig. 39.5 Fusion of various weather observation and forecast data into a winter weather object WWO, symbolized by the cylinder on the *right*. (Figure adapted from Tafferner and Keis (2012))

- Survey—collocate and qualify all MET data at the airport and in the airport environment;
- Analysis—combine the MET data to determine winter weather conditions that hamper operations and jeopardizes safety; represent these conditions with "winter weather objects (WWO)" which reduce the complexity of the weather data into useful information for users, as for example "freezing rain from surface to flight level 50"
- Nowcast—use calculated trends at the airport and surrounding observation sites to determine changes for the WWO
- Forecast—use numerical weather prediction data for early warning.

In Fig. 39.5 various data sources are shown symbolically from which the WWO could be derived. SYNOP and automatic sensors allow surface conditions to be determined, in this example rain with temperature above zero. The temperature/humidity sounding can be provided from the COSMO-DE model or aircraft-measured data (AMDAR), or constructed from both depending on data availability. POLDIRAD observes the precipitation height and is able to determine the hydrometeors within the cloud through its polarimetric capability and related algorithms. ADWICE, the Advanced Diagnosis and Warning System for Icing Environments (Tafferner et al. 2003; Leifeld 2004) uses information on reported weather at the ground together with soundings of temperature and humidity and radar measurements to determine the icing threat to aircraft in flight.

Taken together, the derived analysis can be compacted into the WWO, which is shown schematically as the cylinder at right in the figure. It is obvious that the object can show several different hazard layers in the vertical. For the given case there would be a near surface layer with temperatures above freezing up to height H1 which contains raindrops, a second layer from H1 to H2 which contains super-cooled droplets with corresponding icing threat, and a precipitating cloud layer on top.

For nowcasting icing and snow conditions at the airport, one has to consider weather changes due to advection of air with different characteristics and, especially demanding, possible changes resulting from precipitation and cloud physics processes which can occur within short time spans at the observation site. For capturing both of these effects an approach is followed where WWOs are determined at the various observation sites around Munich airport where both SYNOP and polarimetric radar data are available. Changes in WWOs around the airport can then provide guidance for the expected change at the airport.

For forecasting winter weather conditions beyond the nowcasting range up to about 24 h or more, one can rely on operational forecast models like COSMO-DE. Numerical models have made remarkable progress during the last few years in forecasting the overall weather state, e.g., the surface pressure distribution or whether it will rain or not. The forecast of winter weather phenomena, however, like freezing rain or drizzle, or light or heavy snowfall, is still a demanding task. These phenomena result from the subtle interplay of various factors like the vertical distribution of temperature and humidity, cloud cover and type, snow cover, soil moisture and the composition of the atmosphere with aerosols, which again influence cloud and precipitation processes. The situation gets even more complicated as these processes result from instabilities which are triggered by small changes in the atmospheric parameters, e.g., whether the temperature at the ground or through a certain depth of the atmosphere is slightly above or below 0 °C. In order to better estimate the future atmospheric state, ensemble models give better guidance than a single model run. Combined quantities like ensemble mean, spread, and others allow probabilities to be estimated which can be used for advanced planning. Here, output of the KENDA ensemble model from DWD (Schraff et al. 2011) can be used in future to provide this probability information.

With oncoming winter weather events it is expected that the concept will prove to be useful and applicable to the needs of aviation users. The experience gained from such winter weather cases will also enable the build-up of a fuzzy logic procedure which can improve the nowcasting of winter weather and thus provide a reliable source of information for decision making at Munich airport and elsewhere.

39.8 Conclusion and Next Steps

It can be concluded that the fusion of different data stemming from observations and numerical models is essential for a consistent and reliable description and forecast of weather hazards, a concept that will be realized by DLR-IPA's WxFUSION. The positive feedback from users demonstrated that the nowcasting products provide useful information that is appropriate for use at airports and by air navigation services as well as on board aircraft. They make an important contribution to increasing the safety and efficiency of the air transport system by

providing the same situational awareness to all stakeholders and serving as basis for collaborative and dynamic decision making processes.

Currently, upon request of DFS, preparations are being made for another campaign in summer 2012 which will be supported by DFS, DLH, and FMG in close collaboration with DWD. As a new service it is planned to automate thunderstorm warnings and include them in DFS displays. Moreover, it is planned to up-link the DLR products to the cockpit of an aircraft. Taken together, a first-time demonstration of simultaneously providing users at the ground and pilots in the air with the same view of a thunderstorm situation is envisaged.

The European Space Agency supports research and development for a certified airborne collaborative network to exchange real-time atmospheric data and meteorological conditions from/to business and regional aircraft. The goal is to have in-flight information updates on weather conditions and hazards, and at the same time contribute to global weather observations by providing additional atmospheric measurements complementary to existing data collections. The European Commission is funding a project to develop new cockpit information systems applicable to multiple types of aircraft and helicopters and enabling robust worldwide operations in all weather conditions.

WxFUSION will be further developed for both thunderstorm and wintry weather events, and finally demonstrated and validated at the ground and in the air in close cooperation with decision makers. The aim is to bring the system to an international standard and to apply it Europe-wide as a MET tool serving aviation needs in a Single European Sky.

References

Eurocontrol Performance Review Commission (2011) Performance Review Report 2010. Available at http://www.eurocontrol.int/prc/public/standard_page/doc_prr.html

Forster C, Tafferner A (2009) An integrated user-oriented weather forecast system for air traffic using real-time observations and model data. In: Proceedings of the European Air and Space Conference (CEAS), Manchester, UK, 26–29 October 2009

Forster, C., and A. Tafferner, 2012: Nowcasting thunderstorms for Munich airport. The DLR Project Wetter & Fliegen. In: Gerz T, Schwarz C (eds.) Final research report DLR-FB 2012-02, pp 32–45

Gerz T, Schwarz C (eds.) (2012) The DLR Project Wetter & Fliegen. Final research report DLR-FB, vol 2012-02, p 280

GoP (2001) European Aeronautics: a vision for 2020–Meeting society's needs and winning global leadership. Report of the group of personalities, European Commission, January 2001, p 26

Hauf T, Leykauf H, Schumann U (eds) (2004) Luftverkehr und Wetter, Statuspapier, DLR-Mitteilungen, vol 2004-02, p 57

Kober, K., Tafferner, A.: Tracking and nowcasting of convective cells using remote sensing data from radar and satellite. Meteorol. Z. **1**(18), 75–84 (2009). doi:10.1127/0941-2948/2009/359

Leifeld C (2004) Weiterentwicklung des nowcastingsystems ADWICE zur Erkennung vereisungsgefährdeter Lufträume, Berichte des Deutschen Wetterdienstes, Offenbach am Main, vol 224, p 118

Leighton Q (2006) Modeling and simulation needs for next generation air transportation system research. AIAA modeling and simulation technologies conference and exhibit, Keystone, Colorado, 21–24 August 2006. AIAA 2006-6109, pp 1–8

Pradier, S., Forster, C., Heesbeen, W.W.M., Pagé, C., Sénési, S., Tafferner, A., Bernard-Bouissières, I., Caumont, O., Drouin, A., Ducroq, V., et al.: Description of convective-scale numerical weather simulation use in a flight simulator within the Flysafe project. Meteor Atmos Phys (2009). doi:10.1007/s00703-008-0317-4

Schraff C, Reich H, Rhodin A, Potthast R, Blahak U, Stephan K, Zeng Y, Epperlein D, Leuenberger D, Weusthoff T et al. (2011) COSMO priority project KENDA for Km-scale ensemble-based data assimilation. 9th SRNWP workshop on non-hydrostatic modelling, Bad Orb, 16–18 May 2011

Sénési S, Guillou Y, Tafferner A, Forster C (2009) Cb nowcasting in FLYSAFE: improving flight safety regarding thunderstorm hazards. WMO symposium on nowcasting, Whistler, BC, Canada, 30 August–4 September 2009

Tafferner, A., Hauf, T., Leifeld, C., Hafner, T., Leykauf, H., Voigt, U.: ADWICE–advanced diagnosis and warning system for aircraft icing environments. Weather Forecast **18**(2), 184–203 (2003)

Tafferner, A., Forster, C., Hagen, M., Keil, C., Zinner, T., Volkert, H.: Development and propagation of severe thunderstorms in the upper Danube catchment area: towards an integrated nowcasting and forecasting system using real-time data and high-resolution simulations. Meteorol. Atmos. Phys. **101**, 211–227 (2008a). doi:10.1007/s00703-008-0322-7

Tafferner A, Forster C, Sénési S, Guillou Y, Tabary P, Laroche P, Delannoy A, Lunnon B, Turb D, Hauf T et al. (2008b) Nowcasting thunderstorm hazards for flight operations: the CB WIMS approach in FLYSAFE. 26th Congress of the International Council of the Aeronautical Sciences, Anchorge, AK (USA), 14 September 2008, pp 1–10

Tafferner A, Forster C, Hagen M, Hauf T, Lunnon B, Mirza A, Guillou Y, Zinner T (2010) Improved thunderstorm weather information for pilots through ground and satellite based observing systems. 14th conference on aviation, range, and aerospace meteorology, 90th AMS annual meeting, 17–21 January 2010, Atlanta

Tafferner A, Keis F (2012) Nowcasting winter weather at Munich airport. In: The DLR Project Wetter & Fliegen, Final research report DLR-FB 2012-02, pp 46–57

Zinner, T., Mannstein, H., Tafferner, A.: Cb-TRAM: tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data. Meteorol. Atmos. Phys. **101**, (2008). doi:10.1007/s00703-008-0290-y

Zinner T, Betz H-D (2009) Validation of Meteosat storm detection and nowcasting based on lightning network data. In: Proceedings of the EUMETSAT meteorological satellite conference 2009, Bath, United Kingdom, EUMETSAT P. vol 55, p 8

Probabilistic Weather Forecasting

40

George C. Craig

Abstract

Weather forecasts are approaching the physical limits of predictability. A prediction of a *cyclone* more than a week in advance, or a thunderstorm a few hours ahead, will have a large degree of uncertainty. New techniques using ensembles of forecasts to predict probabilities of *weather* events are being developed to increase the skill of weather forecasting and the value of forecasts to users when exact predictions are impossible.

40.1 Introduction

40.1.1 Improvements in Weather Forecasting

The quality of *weather forecasts* has improved dramatically in recent decades. Major storms rarely come as a surprise and are often predicted many days in advance. The steady improvement in forecast *skill* is monitored by operational weather services using graphs such as Fig. 40.1 for the *European Centre for*

G. C. Craig (✉)
Ludwig-Maximilians-Universität München (LMU),
Meteorological Institute Munich (MIM), Theresienstraße 37,
80333 München, Germany
e-mail: George.Craig@lmu.de

G. C. Craig
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_40,

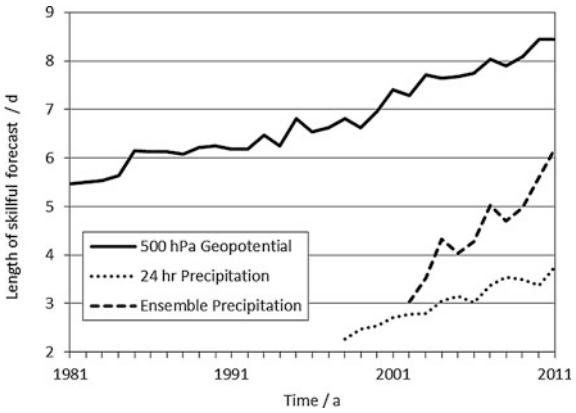

Fig. 40.1 Number of days for which forecasts of 500 hPa geopotential height have useful skill for the northern hemisphere (anomaly correlation > 60 %), plotted as a function of time over the past 30 years (*solid line*). Similarly for forecasts of 24 h accumulated precipitation (stable equitable error in probability space (SEEPS) score > 45 %) (*dotted line*). Similarly for probabilistic forecasts of 24 h accumulated precipitation (conditional ranked probability skill score (CRPSS) > 10 %) (*dashed line*). Adapted from Richardson et al. (2011)

Medium-Range Weather Forecasts (ECMWF), which shows how many days in advance the weather can be predicted with useful skill. Forecasts of large-scale atmospheric flow, the high and low *pressure* systems that are responsible for much of our weather in the middle latitudes, have gained an additional day of lead time every ten years. Forecasts for lead times of 6 days are as good as 3 days forecasts 30 years ago, and weather events more than a week away can often be predicted with useful skill.

This steady improvement has two main sources. The first is the maturing of *numerical weather prediction* models, with ever more detailed descriptions of atmospheric processes running on ever faster *supercomputers*. The second is improved global observations of the atmosphere, particularly from *satellites*, and the development of sophisticated inverse methods to construct the initial model state (*data assimilation*), that now form as large a part of the code and computer time as the forecast *model* itself.

But it seems unlikely that this rate of improvement can go on forever, and the skill of *forecasting* many important weather phenomena has lagged behind. For example quantitative *precipitation* forecasts become essentially useless after fewer than 4 days (Fig. 40.1). Continued progress in *atmospheric* observations and numerical *simulation* give hope that weather forecasting will continue to improve, and IPA is contributing to these developments in many areas, including with new *remote sensing* methods such as lidar, and high resolution modeling techniques. However, it must also be recognized that there are fundamental physical limits to our ability to predict complex systems. Forecasts will always have uncertainty, and the purpose of this article is to describe how progress in weather prediction can

continue in the face of these limits. By expressing a forecast in terms of probabilities, uncertainty in forecasts can be quantified, and the useful information in an uncertain forecast can be extracted. This change in perspective has begun to influence the meteorological community, and there is already evidence that probabilistic forecasts are more valuable than deterministic forecasts, and open more opportunities for rapid improvement (Fig. 40.1). This article will review the factors that limit *predictability* in the atmosphere, and provide examples of how probabilistic forecasts are created and evaluated.

40.1.2 Theoretical Limits: Chaos and Scale Interaction

Earth's *atmosphere* is chaotic, in the mathematical sense of sensitive dependence on initial conditions. This means that arbitrarily small initial errors grow to finite size in a finite time. In other words, even with a perfect forecast model, it is not possible to keep extending the lead time of useful forecasts simply by measuring the initial state more accurately, and the predictability is limited. But the situation is even more complicated since the growth rates of errors depend on their spatial scale, leading to so-called spatio-temporal chaos.

Errors on the scale of a *mid-latitude storm* double in magnitude in a couple of days, but errors on the scale of a thunderstorm double in only about half an hour. A doubling time of 2 days implies an error in surface pressure of 1 hPa (about the accuracy of the current observing system) and would in 12 days grow to an error of 2^6 hPa = 64 hPa, equivalent to having the cyclone in completely the wrong place—a forecast with no skill.

This estimate suggests that a longer forecast would be possible with more accurate pressure measurements, but this is not as simple as deploying more accurate barometers. The problem is not the accuracy of the *instruments*, but rather the gaps in their coverage. The important errors take the form of fluctuations associated with small scale atmospheric flows such as *thunderstorms* or *inertia-gravity waves*, and it would be necessary to increase the spatial resolution of the observing network to include these phenomena.

To determine which scales of motion would need to be observed, however, we must take into account that errors grow faster on smaller scales. *Turbulence* theory relates error growth to kinetic energy for each scale of motion and allows a quantitative estimate of how quickly errors at small scales affect larger scales. The remarkable finding is that even motions on scales smaller than a meter (the famous flap of a butterfly's wing) will grow so rapidly that errors of 1 hPa over distances of 100° km would occur within hours, which would affect the large scale weather patterns within days, as estimated above, and eventually destroy any predictive skill even at the planetary scale in a few weeks.

Chaos implies that the accuracy of observations can never be good enough for an accurate forecast at long range, but spatio-temporal chaos implies that the

accuracy of observations is less important than their coverage. In practice, it is the upscale growth of local errors that limits the lead time of useful forecasts.

40.1.3 Sources of Longer-Range Predictability

While a precise forecast of a weather event longer than a few weeks in the future is not possible, it is still possible to make useful predictions. Edward Lorenz, the father of chaos theory, liked to distinguish between two types of prediction (Lorenz 1996). The first kind uses the equations of motion (also called *primitive equations*) to predict a future state of the system from a set of initial conditions, and is limited by chaos and scale interaction, as described in the preceding section. The second kind of prediction relies on predictable boundary conditions to constrain some properties or statistics of the system. A trivial example is the prediction that average *temperatures* for Germany in July next year will be warmer than in January. The orbital geometry of the *Earth* makes this a near certainty, for although errors in the initial conditions will grow rapidly, they cannot grow forever, and the amplitude at which the growth saturates is constrained by the energy available in the system.

Less trivially, the possible evolution of weather systems is constrained by interactions with other parts of the Earth system that vary more slowly. The *oceans*, with their large thermal inertia, are a good example, but relatively homogeneous regions of the atmosphere such as the *tropics* or the *stratosphere* also influence the growth and decay of mid-latitude storms.

Even less trivially, similar considerations apply to the prediction of smaller-scale weather systems like thunderstorms. Although errors grow rapidly, the evolution of the storm is constrained by its environment. Sometimes the large scale weather pattern makes it impossible for thunderstorms to form, and other times inevitable, although it may not be possible to predict exactly where and when the individual storms will happen. Other factors like *orography* and fluxes of heat and moisture from the earth's surface can also have a decisive influence.

Originally, Lorenz introduced the two types of predictability to explain why *climate* prediction is possible even in the presence of chaos on much shorter *time scales*, but in fact the interplay of rapid error growth with environmental constraints is important for prediction at all lead times, from minutes to millennia. This article will use examples from short-range weather prediction, but the same concepts also apply to seasonal and climate prediction.

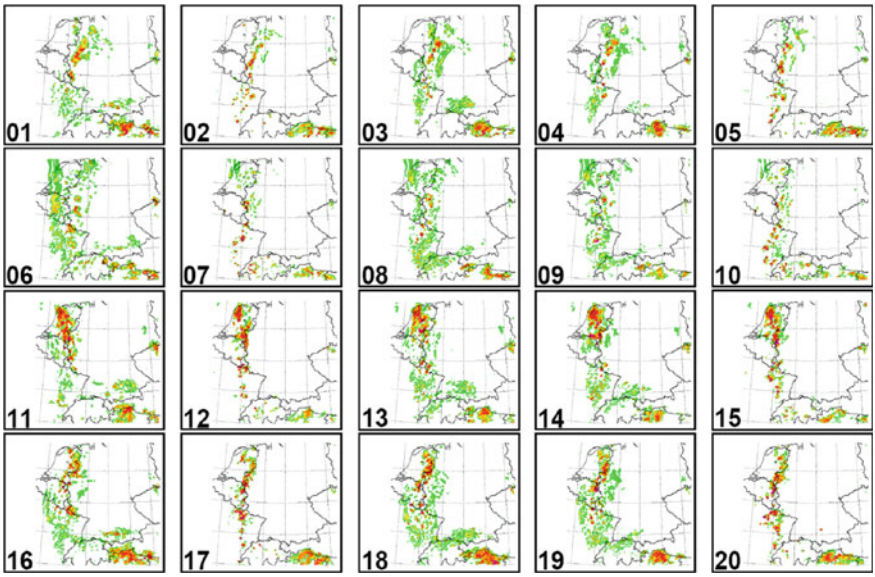

Fig. 40.2 Twenty forecasts of synthetic radar imagery (related to precipitation rate) from the COSMO-DE-EPS for 23:15 UTC August 2007, initialized at 00 UTC, over a domain covering southwestern Germany, northeastern France and northern Switzerland. Radar reflectivities 7, 19, 28, 37, 46 and 55 dBZ are color-coded

40.2 Representing Uncertainty

40.2.1 Ensemble Forecasting

Limited predictability does not prevent making forecasts that have useful and measurable skill. *Uncertainty* in forecasts can be quantified using probabilities. It is standard practice in *physics* and engineering to "dress" each measured or predicted value with an error estimate that typically represents the standard deviation of a Gaussian error distribution. However, this is not adequate for nonlinear systems like the atmosphere, where a Gaussian distribution of initial conditions will be transformed to a more complex, even multimodal, form after a short time. In such cases a *Monte Carlo* approach is more powerful, that is to make a random sampling of initial conditions that are consistent with our current knowledge of the state of the atmosphere and forecast each one with a *numerical model*. The result is an *ensemble* of forecasts, that should sample the possible future states of the atmosphere (Leutbecher and Palmer 2008). An example from the COSMO-DE ensemble prediction system (EPS) currently under development at the Deutscher Wetterdienst (DWD) is shown in Fig. 40.2.

The twenty forecasts of precipitation show a wide range of possible outcomes, but similarities between some of the forecasts indicate that some outcomes may be

more likely than others. To make use of this information, two questions must be addressed. First, does the range of *scenarios* in the ensemble sample the range of outcomes that could occur in nature, and second, how can the forecasts be combined to produce a probability of a particular weather event occurring?

40.2.2 Sampling Sources of Uncertainty

If an ensemble of forecasts randomly samples the states of the atmosphere that are possible, given the available observations, and accurately represents the constraints and small-scale errors that influence the evolution of the system, it can be used to estimate the probability of occurrence of weather situations of interest. However, pure Monte Carlo forecasting with initial conditions that randomly sample all potential states of the atmosphere would require thousands, if not millions, of forecasts, and in practice it is only possible to do ensembles of 20–50 forecasts. With such limited ensemble size, it is crucial to carefully design the members of the ensemble to represent the various sources of uncertainty as efficiently as possible. At present, it is not at all clear how to do this, and current ensemble prediction systems use a wide variety of methods. To illustrate the issues, we show how uncertainty is represented in a particular existing ensemble system, and present examples illustrating two issues that are being actively researched: dependence of predictability on weather regime, and stochastic representation of small-scale processes.

The ensemble of *precipitation forecasts* shown in Fig. 40.2 consists of 20 simulations with a *limited-area model* of the atmosphere and represents two major sources of uncertainty (Gebhardt et al. 2010). Uncertainty in the initial conditions and coming from parts of the world outside the simulation domain are sampled by using forecasts from global models of four different weather services to provide boundary conditions for different members of the ensemble. Uncertainty in small-scale processes within the domain is sampled by creating five different versions of the forecast model by using different values of selected constants in the parameterization schemes that represent the effects of unresolved processes. Since errors at these small scales grow and saturate quickly in time, they are represented by changes to the model evolution equations, rather than by differences at just the initial time. Combining four choices of boundary conditions with five choices of model formulation results in a twenty-member ensemble. Forecasts in the same row in Fig. 40.2 share the same boundary conditions, while those in the same column have the same constants in their *parameterizations*. For this particular weather event, consisting of a line of thunderstorms over western Germany, it appears that the forecasts within each column are relatively similar, showing the importance of growth of small-scale errors for this weather phenomenon.

In general, it might be expected that different sources of uncertainty are dominant in different weather regimes, leading to differing degrees of predictability. For precipitation forecasts like those shown in Fig. 40.2, an important distinction is the degree to which rainfall is controlled by large-scale weather

Table 40.1 Frequency bias index (FBI) and ensemble spread of FBI for precipitation forecasts for two regimes defined by a convective time scale (see text)

	$\log_{10}$(FBI)	$\log_{10}$(spread)
$\tau_c < 6$ h	-0.1 ± 0.6	-0.6 ± 0.5
$\tau_c > 6$ h	-1.0 ± 0.6	-0.1 ± 0.3

From the COSMO-DE EPS for the period 8–16 August 2007. Adapted from Keil and Craig (2011)

features such as cold *front*s. To see the impact of this control, Keil and Craig (2011) took COSMO-DE EPS over a period of 9 days, and divided them into two groups using the convective time scale introduced by Done et al. (2006) as a measure of how quickly *convection* responds to its environment. Two measures of ensemble performance were calculated. First is the log bias score (frequency bias index, FBI), where values less than zero indicate that the total rainfall produced by the ensemble members was less than observed. Second is the spread, defined here as the standard deviation of log bias score among the members of a particular ensemble forecast, where a value of zero corresponds to a spread equal to the root mean square error. Table 40.1 shows that the regime with stronger environmental control (short time scale) has only bias and low spread, indicating a relatively accurate and confident forecast of this quantity. In the other regime precipitation is strongly under-forecast, and the large spread indicates that forecasts in this regime are less certain.

The large forecast uncertainty that occurs when small scale features such as *cumulus* clouds are not under the control of the larger scales suggests that some care should be taken to represent small scale variability accurately. As a result there is a growing interest in stochastic parameterizations of small scale processes that provide a continuous source of small-scale perturbations. For example, the stochastic convection scheme of Plant and Craig (2008) uses a statistical representation of the spatial distribution of convective storms to represent the uncertainty associated with the inability to predict the locations of individual storms. An evaluation of this method in an ensemble forecasting system was carried out by Groenemeijer and Craig (2011) to assess the impact of convective-scale uncertainty compared to that of initial and boundary conditions. Figure 40.3 shows variance in precipitation at individual locations (rather than in total area rainfall, as in Table 40.1) as a function of forecast lead time for a case study where the large-scale environment is expected to exert a significant degree of control. The figure shows that stochastic parameterization provides a large fraction of the variance. Significantly, the variance is roughly proportional to the total precipitation over the region, showing that the statistics of the small scales are coupled to the larger scales of motion.

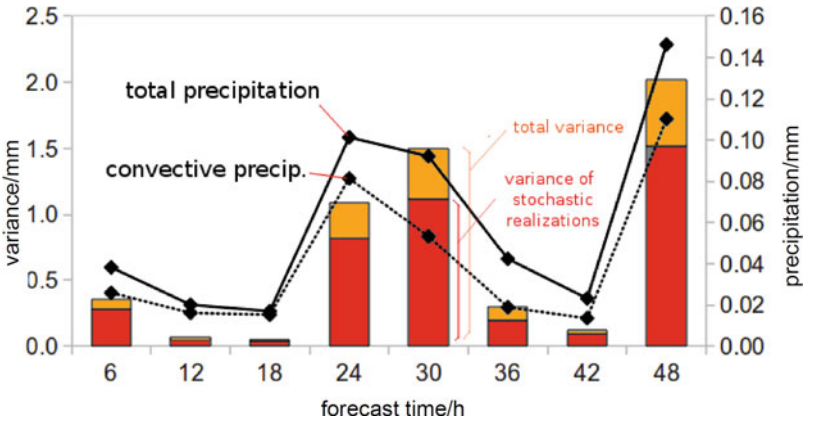

Fig. 40.3 Variance of 1 h accumulated precipitation among ensemble members driven by the same initial and boundary conditions (*dark shaded bars*), the variance among all simulations (*complete bars*), and the spatially averaged 1 h accumulated convective (*dotted line*) and total (*continuous line*) precipitation at eight different forecast times (Groenemeijer and Craig 2011)

40.2.3 Probabilistic Forecasts

A small ensemble of forecasts provides some scenarios of how the atmosphere may evolve, but it remains necessary to turn this information into a quantitative estimate of the probability that an interesting or dangerous weather event will occur. Common examples of high impact weather include precipitation exceeding a *threshold* over a given region, or a shift in *wind* direction at an *airport* requiring change of approach direction for landing *aircraft*. The simplest method to compute a probability is to count the fraction of ensemble members where the event occurs, but this is only correct if the ensemble is a truly random sample, and only accurate if the ensemble is sufficiently large.

Just as the ensemble of simulations is designed based on physical mechanisms that lead to forecast errors, the computation of probabilities can take advantage of physical constraints. This is certainly true for forecasts of convective precipitation, where the large-scale environment controls the probability of precipitation within a region, but not the locations of individual storms. Here we can obtain a probability by considering the frequency of occurrence of convection within a space–time region, centered on the point of interest. This captures the uncertainty in the location of storms without requiring additional simulations. The resulting probabilities can then be averaged over the ensemble to capture uncertainty in the large-scale flow. Figure 40.4 shows examples of forecasts of probability of *radar reflectivity* exceeding a certain threshold, based on the ensemble of forecasts in Fig. 40.2.

The first method computes the probability at each point from the fraction of 20 ensemble members where the threshold was exceeded at that location (Fig. 40.4a). The second method uses the fraction of grid points within a square of side length 75 km centered on the point of interest to give a probability at that location for

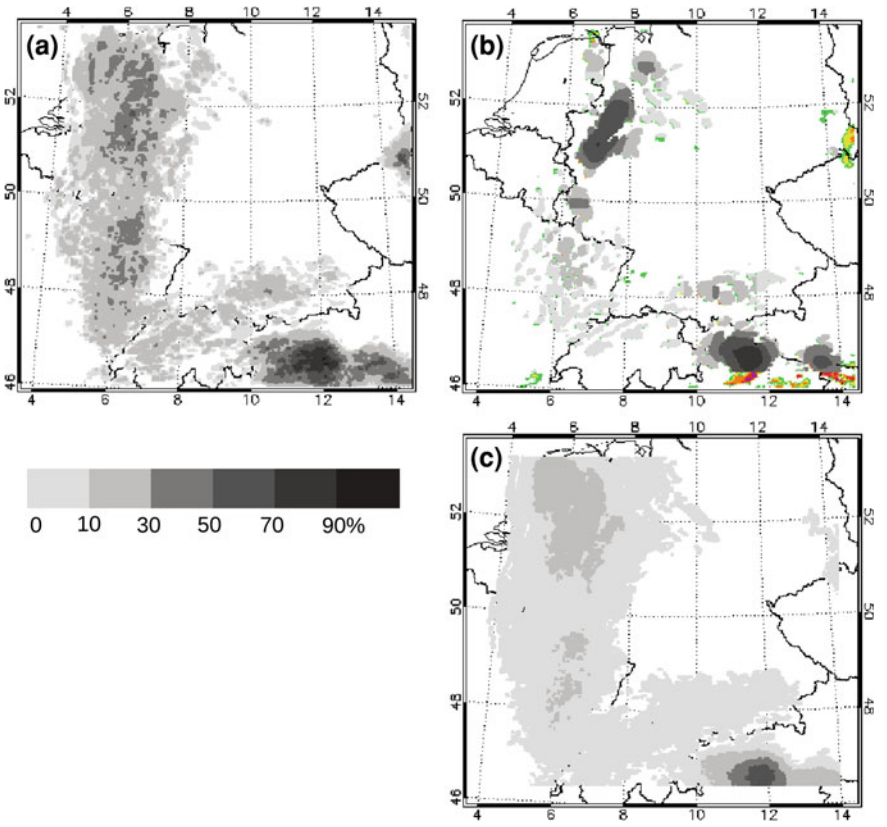

Fig. 40.4 *Grey shading* shows probability of exceeding a radar reflectivity threshold of 19 dBZ for 23:15 UTC 12 August 2007, constructed from the COSMO-DE EPS, using **a** the fraction of ensemble members exceeding the threshold **b** the fraction of neighborhood locations exceeding the threshold in an example ensemble member, and **c** the ensemble mean of the neighborhood probabilities (*grey-shaded*). Colors in **b** show the synthetic radar reflectivity for this example member (Kober et al. 2011)

each ensemble member (a single example is shown in Fig. 40.4b), which is then averaged over the 20 members (Fig. 40.4c).

Regardless of the method used to generate a probabilistic forecast, it is worth remembering that one of the main limitations in practice is the size of the ensemble used to generate the forecast. Events that occur relatively rarely may be missed by all the members of a small ensemble, and the variability shown by the ensemble will substantially underestimate the uncertainty of the forecast. Additionally, the simulations may be biased in a particular direction due to systematic errors in the numerical model. As a result, it is generally necessary to calibrate a probabilistic forecast by using a statistical model to extrapolate the probabilities produced by the ensemble to correct biases and cover the observed range of outcomes.

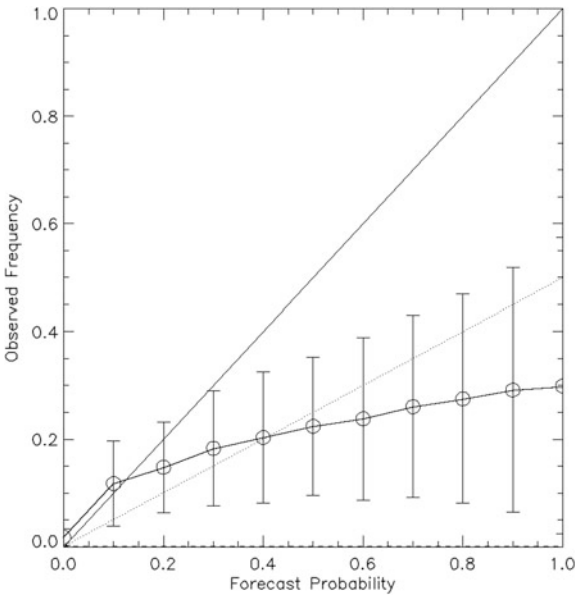

Fig. 40.5 Reliability diagram, showing observed frequency of radar reflectivity exceeding 19 dBZ, as a function of probability for forecasts generated from an ensemble of COSMO-DE forecasts

40.3 Verification and Applications

The final, but crucial, aspect of probabilistic weather forecasting that will be considered here is the question of *verification*: how to evaluate the accuracy and utility of a probabilistic forecast. This is an active area of research, and there are probably as many answers as there are users, so we will confine ourselves to a few examples that illustrate common practices and future challenges.[1]

40.3.1 What Makes a Good Probabilistic Forecast?

A standard method to measure the accuracy of a probabilistic forecast is the Brier score, which is simply the root mean square difference between the probability of the event in observations (always zero or one), and the forecast probability. A perfect forecast with Brier score zero would have predicted probability one every time the event occurred, and zero every time it did not.

One attribute that a good forecast must have, known as reliability, is shown in Fig. 40.5, which plots the frequency of an event occurring whenever a certain probability is forecast. A forecast is said to be reliable if its observed frequencies lie along the diagonal line, indicating that the probability of an event actually

[1] An excellent online resource is provided by the WWRP/WGNE Joint Working Group on Forecast Verification Research at http://www.cawcr.gov.au/projects/verification/.

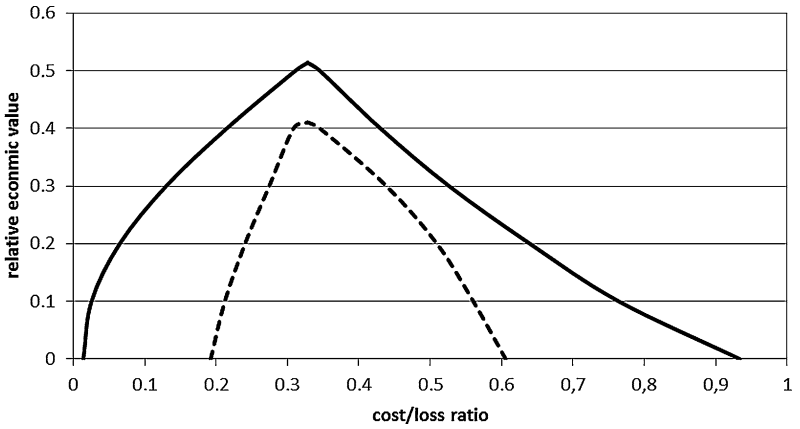

Fig. 40.6 Relative economic value as a function of user cost/loss ratio for 5 d-forecasts of the probability of precipitation exceeding 1 mm in 24 h, generated by the ECMWF ensemble prediction system (*solid line*), and by the deterministic forecast (*dashed line*) (Richardson 2003)

occurring is equal to the forecast probability. For example, Fig. 40.5 shows that on the occasions when a probability of 50 % had been forecast, the event in fact occurred only 22 % of the time. Indeed, the reliability of these forecasts is rather poor, since the event occurs less than 30 % of the time, no matter how high the forecast probability.

Even a perfectly reliable forecast may not be very useful however. It is also desirable to have a confident forecast that accurately predicts probabilities different from the climatological mean, a quality referred to as resolution. It can be shown that the Brier score can be decomposed into three terms, measuring reliability and resolution, plus a correction factor related to the frequency of occurrence of the event in observations. *Calibration* of the probabilities using climatological data can improve the reliability of the forecast by increasing the probabilities of rare events.

40.3.2 What Makes a Useful Probabilistic Forecast?

A forecast is useful if it provides information that a decision maker can act on, whether by evacuating a town before an impending flood or buying electricity in advance of a cloudy period when generation by solar power will be reduced. What probability or degree of certainty the user requires before deciding to act depends on the consequences of the actions. For the energy customer, the cost of being wrong is paying a slightly higher price than necessary, so it is sensible to respond even to forecasts that have a high degree of uncertainty in the expectation that wins will outweigh losses in the long run. In contrast, the high costs of evacuating a town make it unwise to act unless one has a high degree of confidence in the

forecast. A simple measure that captures this difference is the Cost/Loss ratio—the cost of taking action divided by the increase in loss that occurs when the event happens but no action had been taken. A user with a low Cost/Loss ratio will act even if the forecast probability of an event is low, while a high Cost/Loss ratio user will only respond to a very high forecast probability.

Richardson (2000) developed a simple model to compute the economic value of a forecast, that is the reduction in total costs and losses, for users with different Cost/Loss ratios (short: ratio). The relative economic value depends on the skill of the forecast and the climatological frequency of occurrence of the event, and ranges from zero (no forecast information other than the climatological frequency) to one (perfect forecast). An example of such a calculation is shown in Fig. 40.6. As is typically the case, the forecast has most value for users within a certain range of ratios. The predicted probabilities are rarely high enough to convince the high ratio user to act, whereas the low ratio user has little to lose by acting, unless the forecast probability is very low.

The most important result in Fig. 40.6 is that all categories of users benefit from probabilistic forecasts for the situation depicted in this simple model. The deterministic forecast, which predicts yes or no, but is often wrong, is valuable for a limited set of users. The probabilistic forecast is valuable for more users and of more value to all users.

40.4 Concluding Remarks

The problem of making predictions in a chaotic world remains one of the most difficult challenges in atmospheric science. Progress requires a change in the nature of a weather forecast, from a deterministic best guess to a probability distribution, for although exact weather forecasts are only possible for limited times, some statistical properties of the weather are predictable for much longer periods. The benefits of probabilistic weather forecasts are already being realized, as shown in Fig. 40.1. Probabilistic predictions of precipitation from an ensemble of forecasts remain skillful out to 6 days, almost twice as long as a deterministic forecast, and the skill of the probabilistic forecast is increasing rapidly. This rapid progress is likely to continue, A probabilistic forecasting system can make better use of observations of hard to predict fields like *clouds* and precipitation, and take advantage of new observing methods with very high spatial and temporal resolution, including *lidar* and *lightning* detection. Numerical weather prediction models are being adapted for use in ensemble systems by employing stochastic parameterizations of small scale physical processes. Finally, forecast products and verification scores are increasingly being tailored to precise user requirements.

References

Done, J.M., Craig, G. C., Clark, P. A., Gray, S. L., Gray, M. E. B.: Mesoscale simulations of organized convection: Importance of convective equilibrium. Q. J. R. Meteorol. Soc. **132**, 737–756 (2006). doi:10.1256/qj.04.84

Gebhardt, C., Theis, S.E., Paulat, M., Ben Bouallègue, Z.: Uncertainties in COSMO-DE precipitation forecasts introduced by model perturbations and variation of lateral boundaries. Atmos. Res. (2010). doi:10.1016/j.atmosres.2010.12.008

Groenemeijer, P., Craig, G.C.: Ensemble forecasting with a stochastic convective parametrization based on equilibrium statistics. Atmos. Chem. Phys. Disc. **11**, 30457–30485 (2011). doi: 10.5194/acpd-11-30457-2011

Keil, C., Craig, G.C.: Regime-dependent forecast uncertainty of convective precipitation. Meteorol. Z. **20**, 145–151 (2011). doi:10.1127/0941-2948/2011/0219

Kober, K., Craig, G.C., Keil, C., Dörnbrack, A.: Blending a probabilistic nowcasting method with a high resolution ensemble for convective precipitation forecasts. Q. J. R. Meteorol. Soc. (2011). doi:10.1002/qj.939

Leutbecher, M., Palmer, T.N.: Ensemble forecasting. J. Comp. Phys. **227**, 3515–3539 (2008). doi: 10.1016/j.jcp.2007.02.014

Lorenz, E.N.: The essence of chaos. University of Washington Press, Seattle (1996)

Plant, J.M., Craig, G.C.: A stochastic parameterization for deep convection based on equilibrium statistics. J. Atmos. Sci. **65**, 87–105 (2008). doi:10.1175/2007JAS2263.1

Richardson, D.S.: Skill and relative economic value of the ECMWF ensemble prediction system. Q. J. R. Meteorol. Soc. **126**, 649–667 (2000). doi:10.1002/qj.49712656313

Richardson, D.S.: Predictability and economic value. In: Proceedings of the ECMWF Seminar 2002: Predictability of weather and climate, pp. 321–332, 2003

Richardson, D.S., Bidlot, J., Ferranti, L., Ghelli, A., Haiden, T., Hewson, T., Janousek, M., Prates, F., Vitart, F.: Verification statistics and evaluations of ECMWF forecasts in 2010–2011, ECMWF Technical Memorandum, vol. 564, pp. 53 (2011)

Aircraft Emissions at Cruise and Plume Processes

41

Christiane Voigt, Tina Jurkat, Hans Schlager, Dominik Schäuble, Andreas Petzold and Ulrich Schumann

Abstract

The detection of aircraft emissions at cruise altitudes helps to understand and assess the effects of aviation on atmospheric composition and climate. Since the early 1990s, aircraft emissions of carbon dioxide, water vapor, nitrogen and sulfur oxides, aerosol and soot and their processing in the atmosphere as well as contrail formation have been measured in situ with the instrumented DLR research aircraft Falcon. Scientific results from a series of aircraft missions are summarized and explained, uncertainties are discussed and suggestions are made on how to move forward.

C. Voigt (✉) · T. Jurkat · H. Schlager ·
D. Schäuble · A. Petzold · U. Schumann
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Christiane.Voigt@dlr.de

T. Jurkat
e-mail: Tina.Jurkat@dlr.de

H. Schlager
e-mail: Hans.Schlager@dlr.de

D. Schäuble
e-mail: Dominik.Schaeuble@dlr.de

A. Petzold
e-mail: a.petzold@fz-juelich.de

U. Schumann
e-mail: Ulrich.Schumann@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_41,

41.1 Introduction

Aircraft cruising at altitudes between 8 and 13 km are major direct *anthropogenic emission* sources in the *tropopause* region, the transition layer between the *troposphere* and the *stratosphere*. Here, differences in the residence times of stable species between the troposphere (days to weeks) and the stratosphere (months to years) as well as a specific *temperature* minimum cause a characteristic change in the *abundance* of *trace species* with altitude. Low concentrations of *water vapor* (H_2O), nitrogen and sulfur oxides as well as *ozone* (O_3) are typical for the tropopause region. Mean tropopause background mixing ratios expressed in number of molecules per million gas molecules (μmol/mol) are 400 for *carbon dioxide* (CO_2), below 100 for H_2O, less than 0.1 for O_3, and below 10^{-4} for *nitrogen oxides* (NO_x, the sum of NO and NO_2) and *sulfur dioxide* (SO_2); aerosol number densities are below 100 cm^{-3} in this region.

Emissions from *aviation* modify *atmospheric composition*, *air chemistry*, and *climate*. At present, the total climate impact from aviation is highly uncertain (*radiative forcing* 38–139 mW m^{-2}, Lee et al. 2009) and may account for 4.9 % (2–14%, 90 % likelihood range) of total anthropogenic radiative forcing. Projected growth rates of 5 % per year in passenger *air traffic* will even enhance future aviation impact on climate.

Aircraft directly emit trace species in the *troposphere* and *lower stratosphere*. In 2005, the total aircraft *fuel consumption* of ~230 Tg (10^{12} g) *kerosene* (based on data from the International Energy Agency) caused the emission of ~730 Tg of CO_2 into the *atmosphere*. Besides CO_2, other species like H_2O, NO_x and SO_2 are emitted and directly lead to increases in ambient concentrations. Indirect effects such as the formation of secondary products evolving from directly emitted exhaust material also occur. E.g., aircraft NO_x causes ozone production and in turn *methane* (CH_4) reduction in the troposphere. Further, emissions of SO_2, *aerosol* and *soot* may influence particle abundance and cloud properties on global scales. Aircraft emissions are generally quantified with the *emission index* EI_i, a measure of the mass of an emitted species i per mass of burnt fuel (see Table 41.1). The emission indices of CO_2, H_2O and SO_2 are mainly a function of the fuel mass fractions m_i of carbon (C), *hydrogen* (H) and *sulfur* (S) respectively, and of the molar masses M (in g/mol) of the oxidation products formed during *combustion*. Thereby, m_i may vary by about ±1 % depending on the actual fuel composition.

Under realistic, generally incomplete combustion conditions, additional combustion products like NO_x, *carbon monoxide* (CO), *hydrocarbons* (HC), and *soot* are emitted. Typical EI_i for cruise conditions, with the majority derived from DLR measurements, are listed in Table 41.2. (According to accepted standards, EI_{NOx} is defined in mass units of NO_2.)

While emissions from aviation contribute just a small fraction to the global atmospheric budget, the processing of emissions in the atmosphere can enhance their atmospheric impact and their effects on climate. Therefore, a series of scientific programs on the environmental impact of aviation has been launched since

Table 41.1 Mass fraction (m), *molar mass* (M), emission index of species i (EI_i) and typical values of ideal combustion products

	Mass fractions (m)	Molar mass (M)	Emission index (EI)	Typical EI values (range)
C	0.86 (0.85–0.87)	12.01	$EI_{CO_2} = \frac{m_C M_{CO_2}}{M_C}$	3.16 ± 0.06 kg/kg
H	0.138 (0.13–0.145)	1.008	$EI_{H_2O} = \frac{m_H M_{H_2O}}{2M_H}$	1.24 ± 0.02 kg/kg
S	400 (1–3 000) × 10^{-6}	32.07	$EI_{SO_2} = \frac{m_S M_{SO_2}}{M_S}$	0.8 (0.6–1.0) g/kg

Table 41.2 Emission indices of species i (EI_i), range and typical value (from Lee et al. 2010)

Species i	Typical EI_i (range)
NO_x	14 (12–17) g/kg
CO	3 (2–3) g/kg
HC	0.4 (0.1–0.6) g/kg
Soot mass	0.025 (0.01–0.05) g/kg
Soot number	(1–10) × 10^{14}/kg

the early 1970s, stimulated by the debate on stratospheric ozone loss caused by nitrogen oxide emissions from high-flying *supersonic* aircraft (Crutzen 1970; Johnston 1971). Since the 1990s, research addressed various effects of *subsonic* aviation, e.g., DLR performed a series of aircraft measurements (SULFUR 1–7) to determine the impact of fuel sulfur on particle formation. The European Union and NASA funded projects on the atmospheric effects of NO_x and contrails (e.g., The Impact of NO_x Emissions from Aircraft upon the Atmosphere at flight altitudes 8–15 km, AERONOX; Subsonic Aircraft: Contrail and Cloud Effects Special Study, SUCCESS). Results from earlier programs are summarized in the 1999 landmark report on "Aviation and the Global Atmosphere" (IPCC 1999), which has recently been updated (Lee et al. 2009). In addition, the Aircraft Engine Emissions Databank of International Civil Aviation Organization (ICAO) provides a data base of aircraft emission properties at ground. Emissions at cruise, however, are difficult to measure and require the use of instrumented research aircraft.

Hence pioneering in situ measurements of aircraft emissions at cruise altitudes were performed by the DLR Institute of Atmospheric Physics in cooperation with international partners. Here results from the aircraft campaigns Pollution from Aircraft Emissions in the North Atlantic Flight Corridor (POLINAT), SULFUR, Partikel und Zirrus (PAZI) and Contrail and Cirrus Experiment (CONCERT) are summarized. Ground based observations from engine test facilities during PARTEMIS (Measurement and Predictions of the Emission of Aerosols and Gaseous Precursors from Gas Turbine Engines) support the airborne data set. Aircraft engine processes, gaseous and particulate emissions at the engine exit and their physico–chemical processing in the plume are sketched. Further, local and regional effects of aircraft NO_x, SO_2, and aerosol emissions are addressed, and

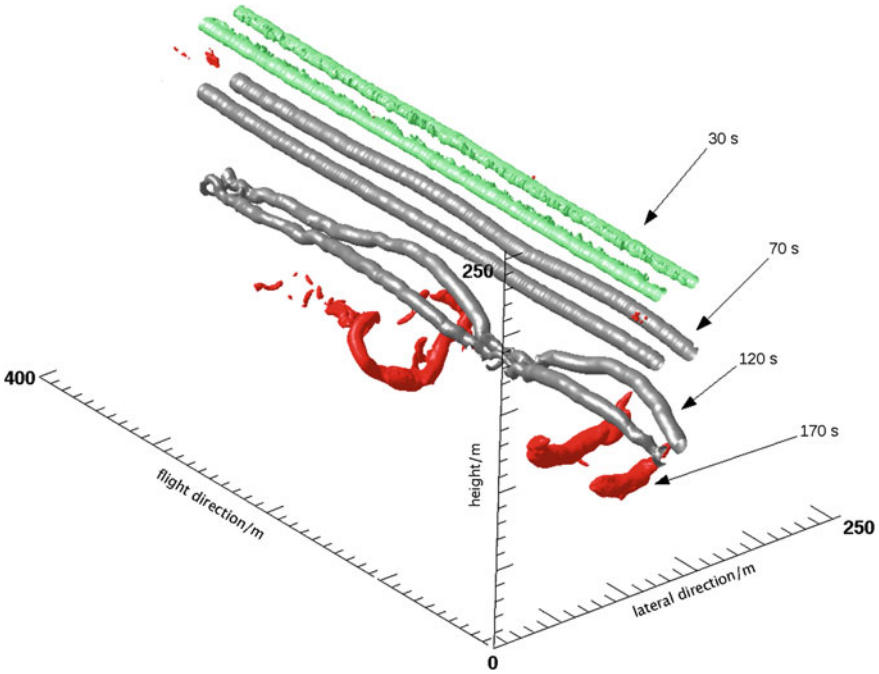

Fig. 41.1 Descent and decay of a segment of the counter-rotating wake vortex pair from an A319 aircraft in a 400 × 256 × 256 m^3 domain for a stably stratified atmosphere with weak turbulence. Shown are vortex tubes, identified as isolines of pressure for plume ages between 30 s and 170 s simulated with the coupled dynamical-microphysical model system EULAG-LCM (Sölch and Kärcher 2010)

the observation of *contrail*s is described. Open gaps in scientific understanding linked to the aviation climate impact are discussed and future directions for aircraft research are presented.

41.2 Aircraft Wake Dynamics and Dilution

In the near field of the aircraft at cruise, pressure gradients between the upper and the lower surface of the aircraft wings induce rotational flow fields and generate wake vortices, which are strongly influenced by the type of aircraft, due to its weight, wing span and speed. Two counter-rotating vortices generally form at cruise at the two wing tips. Within the first few seconds in the *jet regime*, the aircraft engine jet exhaust rapidly mixes with ambient air. At a downstream distance of more than a wingspan behind the engine exit, jet emissions are entrained in the rotational flow field of the wake vortices. Enhanced mixing occurs within the *vortex regime*, caused by stretching and compression of the jet exhaust (Misaka et al. 2012). In the *dispersion regime* starting about 100 s after the release of

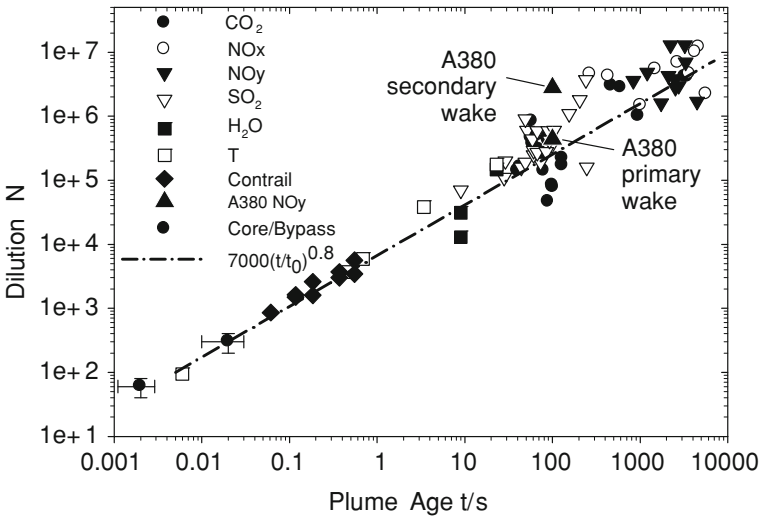

Fig. 41.2 Dilution factor N versus plume age t. Most symbols present previous measurements; adapted from Schumann et al. (1998). *Triangles* (▲): N derived from NO mixing ratios measured in the plume of an A380 aircraft, separated into primary and secondary wake

emissions from the engine exit, the interaction between the vortex and the ambient atmosphere becomes dominant. Then the vortices decay due to wind shear, crosswinds, *turbulence*, and the sinusoidal *Crow instability*, see Fig. 41.1.

Despite the initial inhomogeneous mixing processes, the *dilution* of the exhaust can decently be parameterized by a simple equation (Fig. 41.2). For any emitted passive species i, with molar mass M_i ($M_{air} = 29$ g/mol), the mass-specific concentration increases above ambient levels Δc_i and the corresponding molar mixing ratio $\Delta r_i = (M_{air}/M_i)\ \Delta c_i$ in the exhaust plume can be estimated from EI_i as a function of the dilution ratio N.

$$\Delta c_i = EI_i/N \tag{41.1}$$

Thereby, N is a measure of the mass of air with which the exhaust resulting from a mass of burned fuel per unit flight distance mixes within the bulk of the plume. For $0.006\ s < t < 10^4$ s, the bulk dilution ratio measured in more than 70 plume encounters follows approximately

$$N = 7\,000\,(t/t_0)^{0.8} \tag{41.2}$$

where $t_0 = 1$ s (Schumann et al. 1998). This approximation has been used to estimate the third parameter, when only two of the three parameters (Δc_i, EI_i, N) are known (Anderson et al. 1998; Koike et al. 2000; Brunner et al. 2001; Schäuble et al. 2009; Voigt et al. 2010; Paoli et al. 2011).

Of course this simple relationship does not cover all physical mechanisms within the plume as seen in the data scatter in Fig. 41.2. Parts of this scatter are related to wind shear or turbulence in the wake behind different aircraft. In addition, some scatter results from variation of the local dilution values from the bulk mean dilution. NO measurements in the wake of a large body airliner of type A380 at plume ages of about 100 s (Jurkat et al. 2011) allow separation of the dilution in the *primary wake* (the lower sinking vortices) from that in the *secondary wake* (remaining above the sinking vortices). At 1–2 min plume age, the data show dilution values of $N = 4.4 \times 10^5$ and 2.8×10^6 differing an order of magnitude in the primary and the upper part of the secondary wake, respectively.

As can be derived from a *Gaussian* plume model (Konopka 1995), for even larger plume ages t, the dilution ratio grows as

$$N \sim S\ D_V\ t^2/A_0 \tag{41.3}$$

where S is the vertical shear of the wind perpendicular to the exhaust plume, D_V is the vertical turbulent diffusivity and A_0 the plume cross section at the end of the wake vortex phase.

41.3 Chemical Processing of Aircraft NO_x Emissions

Aircraft engine emissions mainly contain gaseous species like CO_2, H_2O, NO_x, and minor amounts of CO, SO_x, HC, aerosol and soot, where SO_x summarizes SO_2 and SO_3. Thereby, the primary emissions depend on aircraft and engine type, thrust settings, fuel flow and atmospheric parameters. Emissions are generated in the aircraft engine before they are injected into the atmosphere, where they are mixed with ambient air, diluted, and chemically and physically processed in the exhaust plume. Secondary effects can occur when emissions are exposed to ambient conditions (e.g., sunlight, lower temperatures, background aerosol); for example, aircraft NO_x emissions are an important catalyst for the production of the *greenhouse gas* O_3 in the *upper troposphere* (Crutzen 1979). Enhanced ozone production causes a reduction in methane and an ozone *feedback* through formation of the *hydroxyl radical* (OH). Global models predict ozone increases between 3 and 6 % in *mid-latitudes* in the upper troposphere (Brasseur et al. 1996; Grewe et al. 2002; Hoor et al. 2009). These O_3 changes may contribute up to 20 % to the radiative forcing from aviation.

Most NO_x emissions are generated by the dissociation of molecular oxygen (O_2) to atomic oxygen (O) inside the hottest parts of the combustor and turbine at temperatures above 1 200 °C. The oxygen atoms react with molecular (N_2) or atomic nitrogen (N) to form NO. Initially, NO is by far the most abundant ($\sim$80–90 %) reactive nitrogen species (NO_y, the sum of NO, NO_2, HONO, HNO_3, N_2O_5,…) emitted by aircraft. About 5–20 % of the NO_y is emitted as NO_2. Aircraft NO_x mainly reacts with OH in the very young plume up to 30 m behind the engine exit. Thereby, OH is depleted through formation of *nitrous acid*

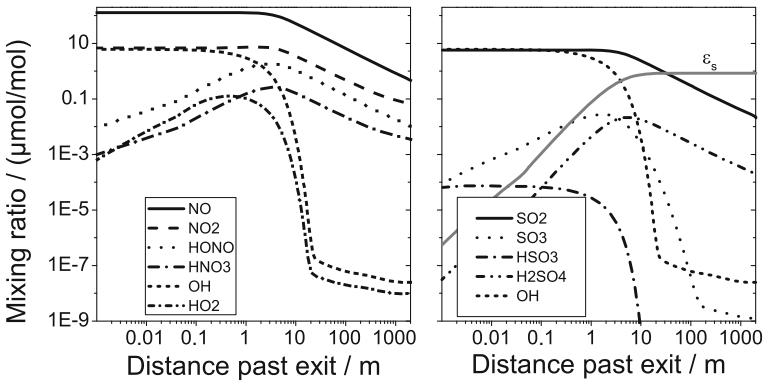

Fig. 41.3 Chemical evolution of hydroxyl radicals, reactive nitrogen and sulfur species, and sulfur conversion efficiency ε_s in the center of the primary vortex exhaust plume from the engine exit to 2 000 m downstream distance, corresponding to approximately 10 s plume age, calculated with a chemistry box model; adapted from Kärcher et al. (1996)

(HONO) and *nitric acid* (HNO_3) by the following three-body reactions, which include a third reaction partner (M).

$$NO + OH + M \rightarrow HNO_2 + M \tag{41.4}$$

$$NO_2 + OH + M \rightarrow HNO_3 + M \tag{41.5}$$

The initial *hydroxyl radical* (OH) mixing ratio is critical for chemical plume evolution as it determines the conversion efficiency of active species (NO_x, SO_x) within the plume. Still, OH mixing ratios at the engine exit are not well known (Tremmel et al. 1998). Figure 41.3 shows the chemical evolution of an aircraft plume from the engine exit to about 10 s plume age, initialized with an OH mixing ratio of 7 µmol/mol (Kärcher et al. 1996).

Further plume processing is mainly caused by dilution. Aircraft plumes a few minutes old have frequently been detected with instrumented aircraft (Schlager et al. 1997; Brunner et al. 2001). Figure 41.4 shows a time series of NO, HONO and SO_2 mixing ratios measured during the plume interception of an A380 aircraft along with particle number density and *extinction*. Short-term variations in NO mixing ratios illustrate the heterogeneity of the plume.

During the CONCERT2008 campaign (Voigt et al. 2010), commercial airliners producing contrails were chased by the DLR *Falcon* at distances between 5 and 86 nautical miles, corresponding to plume ages of 1–10 min. Peak NO mixing ratios up to 58 nmol/mol were observed in the 1 min old exhaust plume from an A380 aircraft and NO mixing ratios were diluted to values < 10 nmol/mol in < 10 min old plumes. NO_x mixing ratios up to 9 nmol/mol were detected in 15 min to 3 h old exhaust plumes in the North Atlantic flight corridor (Schumann et al. 2000). At plume ages of days to weeks the plume further dilutes and extends to regional

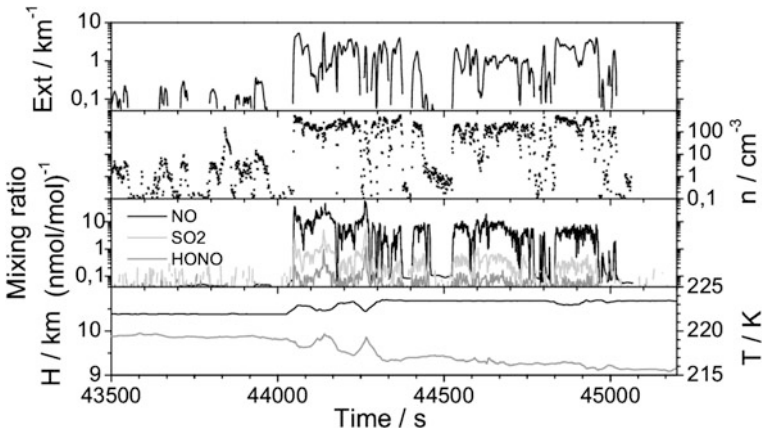

Fig. 41.4 Extinction, particle number density in the size range of 0.3–17 μm, NO, HONO and SO_2 mixing ratios, temperature and flight altitude from a research flight in the contrail from an A380 aircraft as measured on 19 November 2008 above northern Germany; adapted from Voigt et al. (2010)

scales. Slightly enhanced NO mixing ratios were observed in air containing emissions from the North Atlantic flight corridor. Calculations with a transport model suggest that up to 70 % of the observed NO could have originated from aircraft emissions (Schlager et al. 1999).

While most of the nitrogen emissions by aviation are emitted as NO_x, a small fraction is oxidized to short- (HONO) or longer-lived (HNO_3) reservoir gases within the first second of plume evolution (Fig. 41.3). Jurkat et al. (2011) measured HONO/NO_y molar ratios between 1.6 and 3.6 % in 2–5 min old exhaust plumes. Elevated nitric acid (HNO_3) mixing ratios have been detected in gaseous exhaust plumes (Arnold et al. 1992; Tremmel et al. 1998). In addition, some HNO_3 produced in the nascent plume may be incorporated in *volatile* aerosol (Kärcher 1999). Contrail ice crystals nucleating in the volatile aerosol can additionally trap HNO_3 from the gas phase (Kärcher and Voigt 2006; Voigt et al. 2006). Schäuble et al. (2009) observed that the fraction of total HNO_3 incorporated in contrail ice particles (6 %) is higher than in natural cirrus clouds (3 %), probably caused by enhanced aircraft NO_x and HNO_3 emissions.

41.4 Sulfur Emission, Conversion and Aerosol Formation

Sulfur plays a key role in aircraft exhaust chemistry and physics. *Sulfuric acid* (H_2SO_4) is the main precursor for particle formation. Further, it may coat non-volatile hydrophobic soot particles and thus accelerate particle growth in aircraft exhaust plumes. Initially, mainly SO_2 is produced inside the combustion chamber of aircraft engines burning sulfur-containing kerosene. Here temperatures are high enough to produce atomic oxygen that oxidizes fuel sulfur to form SO_2 and also

SO_3. The amount of SO_2 emitted depends exclusively on the *fuel sulfur content* (FSC). Accordingly, the sulfur emission index EI_s depends on the type of kerosene used. Inside the colder turbine segment at temperatures between 1200 °C and 200 °C and in the young exhaust plume, SO_2 is partly converted to H_2SO_4 by oxidation with OH and O_2 (Stockwell and Calvert 1983). The last step (41.8) was confirmed by several studies, e.g., Jayne et al. (1997).

$$SO_2 + OH + M \rightarrow HSO_3 + M \tag{41.6}$$

$$HSO_3 + O_2 \rightarrow SO_3 + HO_2 \tag{41.7}$$

$$SO_3 + 2\,H_2O \rightarrow H_2SO_4 + H_2O \tag{41.8}$$

For a B747 under cruise conditions, an initial SO_2 mixing ratio at the engine exit plane of 7 μmol/mol can be estimated (Kärcher et al. 1996) which decreases during plume evolution mainly due to dilution (Fig. 41.3, right panel). Two orders of magnitude lower amounts of H_2SO_4 form quickly by reactions (41.6–41.8). Gas phase H_2SO_4 is depleted by dilution, and later by nucleation and *condensation* on preexisting particles. Another production mechanism for H_2SO_4 (not included in Fig. 41.3) is the reaction of SO_3 formed directly in the combustor (Starik et al. 2002, 2004). Later, within days after the emission, secondary H_2SO_4 is formed by oxidation of SO_2 with ambient OH.

The *sulfur conversion efficiency* ε_S is a measure for sulfuric acid production in the young plume and is defined as the ratio of oxidized sulfur (S^{VI}) to the total sulfur ($S_T = S^{IV} + S^{VI}$) emitted:

$$\varepsilon_S = S^{VI}/S_T \tag{41.9}$$

While a large part of the sulfur conversion takes place inside the engine, 0.5–2 % of the total sulfur is converted to H_2SO_4 in the expanding exhaust plume shortly behind the engine exit, see Fig. 41.3 and Petzold et al. (2005).

Different methods were applied to measure aircraft sulfur emissions and their conversion (Schumann et al. 2002). Early observations of large volatile aerosol volume densities behind a supersonic aircraft were explained with very high $\varepsilon_S > 12$ % (Fahey et al. 1995). Measurements of the sum of gaseous and aerosol H_2SO_4 behind the *ATTAS* and a B747 suggested far lower conversion efficiencies, and for a B747 an ε_S of (3.3 ± 1.8) % was inferred (Curtius et al. 2002). Based on these measurements, the dependence of ε_S on FSC, reaction rate coefficients, EI_{NO2} and OH mixing ratios was derived from numerical box model simulations (Tremmel and Schumann 1999). Recent observations use an indirect method to derive ε_S (Jurkat et al. 2011). As NO is oxidized by OH to HONO inside the engine and shortly past engine exit, the HONO to NO ratio is a good indicator for the OH abundance in aircraft exhaust. A comparison of the conversion of NO to HONO with that of SO_2 to H_2SO_4 yields the sulfur conversion efficiency. Figure 41.5 shows ε_S of (2.2 ± 0.5) % from measurements in nine different aircraft exhaust

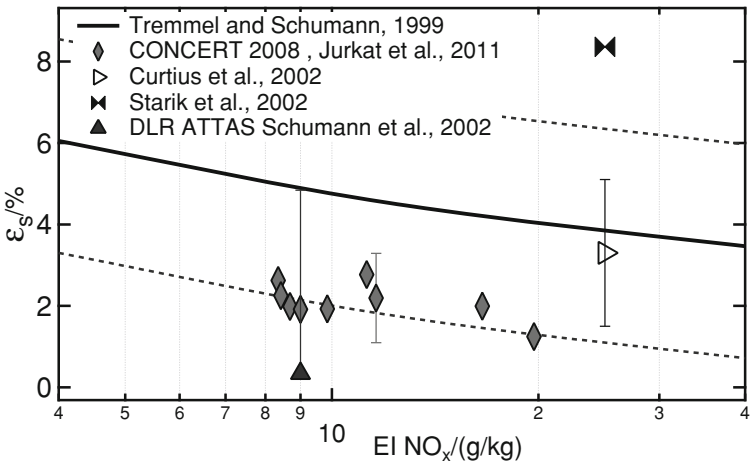

Fig. 41.5 Sulfur conversion efficiencies ε_S in fresh exhaust plumes measured during SULFUR and CONCERT as a function of EI_{NOx}. The lines represent model simulations by Tremmel and Schumann (1999) with upper and lower limits; adapted from Jurkat et al. (2011)

plumes. ε_S shows a downward trend with increasing NO_x emission index caused by OH depletion at higher EI_{NOx}. The CONCERT data represent a lower limit of ε_S and thus are generally in line with former aircraft (Lee et al. 2010) or ground observations at engine test rigs (Petzold et al. 2005).

The aerosol formation properties of aircraft exhaust were investigated in a series of SULFUR experiments. The exhaust plumes from commercial airliners with a wide range of engines and FSCs between 2 and 5 500 μg/g were arbitrarily probed at cruise altitudes under different flight conditions (Schumann et al. 2002). For a systematic approach, each of the two engines of the ATTAS research aircraft was operated with kerosene of different FSC. Additionally, wing-by-wing flights of an older B707 with a PW JT3D-3B engine and an A340 with a CFM56-5C2 engine were performed to investigate the impact of different engines on ε_S. The FSC had a visible effect on exhaust color and contrail onset but no effect on the emitted number of chemi-ions. These positively and negatively charged molecules or molecule clusters form in the engine, enhance coagulation and trigger formation of ultrafine particles. The presence of sulfuric acid allows these small particles to grow faster and finally act as condensation nuclei.

41.5 Soot

Soot particles form in the flame zone of the combustor as a result of incomplete combustion. These particles are composed primarily of carbonaceous material, which is the sum of graphite-like carbon and *organic matter*. The exhaust-soot particle diameter typically ranges within 20–200 nm. In fresh exhaust, modal

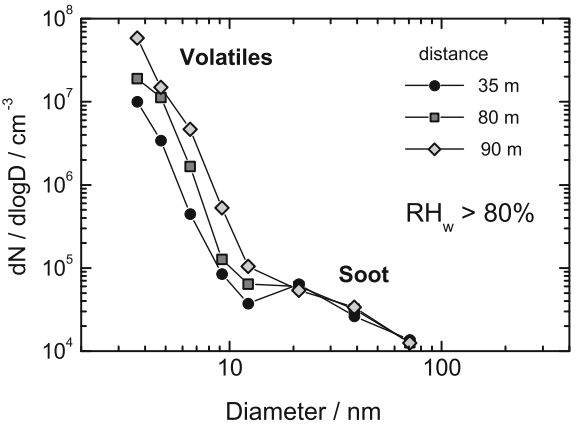

Fig. 41.6 Exhaust particle number size distribution measured in the dry plume 35–90 m behind a Lufthansa B737-300 burning low-sulfur fuel (2.6 μg/g); adapted from Schröder et al. (2000a)

diameters of soot particles are ~20–60 nm (Petzold et al. 2003), hence significantly larger than volatile aerosol with diameters < 10 nm. The engine has an impact on soot properties. Older technology engines emit larger *black carbon* (BC) particles with a modal diameter of 35 nm and a mode of coagulated BC particles with a peak near ~150 nm. In contrast, modern technology engines emit far smaller BC particles. In the expanding and cooling plume these soot particles may gain a coating of condensable gases, become activated, and hence grow in size mainly by uptake of water.

Figure 41.6 shows *particle size distributions* measured 35–90 m behind the leading aircraft. This measurement was the closest approach ever realized for this kind of study. As is clearly visible, the measurable fraction of volatile particles grows with increasing distance (i.e., increasing age) by coagulation of the ultrafine particles and uptake of water and other condensable gases from the gas phase, while the soot particle *mode* remains almost unchanged. Predominantly soot particles can be activated to form contrails at low temperatures in the tropopause region.

41.6 Contrail Formation, Evolution and Detection

Contrails are visual manifestations of aircraft emissions which may have a significant influence on climate by reflecting *solar radiation* back into space (cooling) and by absorbing and re-emitting infra-red radiation from Earth's surface (warming). The net effect is positive, but the exact extent of warming by contrails is highly uncertain. Contrails form when local *humidity* in the exhaust plume exceeds *saturation* with respect to liquid water. In this case, water may condense on the emitted soot and *sulfate* aerosol. These tiny submicrometer sized particles get wetted, activated by water, and grow. At temperatures below −40 °C, ice nucleation in the aerosol occurs. Initially small ice crystals grow by uptake of water vapour, until saturation with respect to ice is reached within the plume.

Fig. 41.7 Contrail from an A380 aircraft at 10.6 km altitude; photo taken from the cabin of the Falcon on 19 November 2008 above northern Germany. Later, this contrail was probed

Ice nucleation and *contrail* formation generally start at distances less than one wingspan behind the aircraft engine. The exhaust jet is caught by the wing tip vortices, and the wake vortex descends several 100 m depending on the aircraft type. During this downward motion, a fraction of the exhaust products is detrained from the plume and thus remains above the sinking vortex. Due to *adiabatic* compression in the descending vortex, the exhaust plume warms up, which may cause partial *evaporation* or *sublimation* of the ice crystals within the primary wake (Unterstrasser et al. 2008).

In situ observations of young contrails started early when Knollenberg (1972) first probed the particle size spectrum from a massive aged contrail. Later, in-situ observations in young contrails were performed during the SULFUR and the PAZI campaigns (Schumann et al. 1996; Petzold et al. 1997; Schröder et al. 2000a; Febvre et al. 2009). More recently, about 20 contrails were probed during the CONCERT campaigns in 2008 and 2011 (Voigt et al. 2010; Voigt et al. 2011). A photo of the contrail from an A380 aircraft chased during CONCERT2008 is shown in Fig. 41.7. Observations within this contrail are also referred to in Figs. 41.2, 41.4, and 41.5.

Large concentrations ($> 1\,000$ cm^{-3}) of micrometer-size ice crystals have been detected in less than 1 min old contrails from a Boeing aircraft (Petzold et al. 1997; Brock et al. 2000; Schröder et al. 2000a). In less than 5 min old contrails (Fig. 41.8), the ice particle concentrations decrease to a few 100 cm^{-3} and ice crystals grow to radii up to 3 μm (Voigt et al. 2011). During further contrail evolution, the ice particle number density becomes more diluted, and in old contrails ice particle concentrations <10 cm^{-3} and radii significantly >5 μm have been observed (Schröder et al. 2000b; Schäuble et al. 2009). Still, *persistent* contrails or *contrail cirrus* with ages of several hours are difficult to distinguish from natural cirrus clouds, and therefore a combination of in-situ observations, *satellite* retrievals and *model* simulations is required to identify contrail cirrus in airborne data sets.

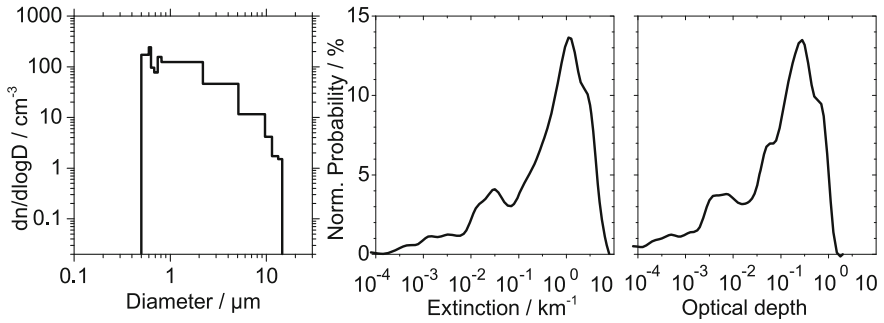

Fig. 41.8 Particle size distribution of 2–5 min old contrails from 14 commercial aircraft detected over Germany in November 2008 (*left*), extinction (*middle*) and optical depth distribution (*right*) derived from the particle size distribution; adapted from Voigt et al. (2011)

The attenuation of light by contrail ice crystals significantly affects the radiative transfer in the atmosphere and causes a radiative forcing which may be of similar magnitude as the effects from CO_2. The extinction of light has been derived from the observed particle size distribution (Fig. 41.8). Multiplying the extinction with the vertical contrail dimension leads to the contrail *optical depth*, which is an important parameter for estimating the climate impact from contrails (Voigt et al. 2011).

41.7 Moving Forward

Scientific progress has been made in the last two decades in detecting aircraft emissions, understanding their processing in the atmosphere, and estimating their climate impact. As also discussed in several other chapters of this book, large uncertainties still remain: How many ice particles form from the aerosol emissions and survive the wake vortex phase for a given meteorological condition? Do ice particle properties in aged contrails approach those of natural cirrus? Does aircraft-induced aerosol cause more or fewer cirrus clouds? Can it be experimentally verified that aviation NO_x emissions enhance ozone formation? Such and further questions as well as future research strategies are discussed below.

Contrail ice crystal nucleation may occur in *solution droplets* forming in the jet phase by condensation on the emitted soot particles and on mixed sulfate aerosol. The ice nucleating properties of different soot samples have been investigated in the laboratory. However, the tested materials may not be representative for aircraft soot emissions at cruise: the type of soot, its processing, ageing, and mixing state may differ from soot emissions under dry and cold atmospheric conditions near 10 km altitude. Kärcher and Yu (2009) suggest that the number of ice particles formed is predominantly determined by the amount of soot emitted. However, at low soot emission indices and at very low temperatures, ice nucleation in liquid or *mixed-phase* aerosol may become important. Recent model simulations are consistent with

observations, assuming that the number of ice particles formed corresponds to the number of emitted soot particles (Schumann 2011). For more precise testing, in situ observations of the number, composition and mixing state of aircraft particle emissions prior to contrail formation would be required. These observations should be compared to in-contrail observations under similar conditions in terms of aircraft operation and meteorology. Such measurements could be performed by chasing an operational aircraft flying into an *ice supersaturated region*. A caveat for this type of observation is that optical particle measurement techniques assume a specific particle composition, density and refractive index. These parameters can be different for individual particles. In particular, particles smaller than 2 μm might be composed of ice or of liquid aerosol. Optical instrumental methods therefore have to be combined with chemical methods such as single particle aerosol *mass spectrometry* or filter sampling with suitable aerosol *inlets*.

The microphysical properties of young contrails are significantly different from those of natural cirrus clouds. During contrail ageing, these properties may become similar to those of natural cirrus. However, young contrails have far higher ice particle densities suppressing ice supersaturation, and large amounts of aerosol are emitted. These differences may also affect the temporal evolution of the ice particle size and *habit*. Hence, though difficult, it would be important to distinguish aged contrails and natural cirrus clouds. Optically thick contrails can be followed over hours from satellite, still, the detection threshold and the fraction of contrails which are detected by satellite may not be well determined in some cases. Hence a combination of satellite observations, contrail prognoses and in situ measurements may help to better quantify microphysical and optical properties from aged contrails and contrail cirrus. These observations could support future contrail *mitigation* strategies.

In addition, aircraft aerosol may not only lead to contrail formation directly in the aircraft wake, but may also influence and modify cloud formation on long time scales (Hendricks et al. 2011). Secondary cloud formation or cloud modification by aircraft emissions is very difficult to detect from aircraft. Observation of cloud formation in regions where trajectories were significantly influenced by aviation compared to regions without aircraft impact could give hints for this difficult research issue.

Ozone increases of 3–6 % caused by aircraft NO_x emissions have been calculated in the upper troposphere (Hoor et al. 2009). However, such calculations have not been validated yet by direct aircraft measurements. Potentially, ozone effects could be enhanced in a meteorological situation caused by a stagnant high where aircraft NO_x emissions could continuously be entrained and accumulate in the same air mass (Schumann et al. 2000). In situ observation of NO_y species and O_3 under these conditions would facilitate the detection of this ozone effect, and combinations with photochemical box models could complement this *scenario*. Complete data sets of the partitioning of NO_y in exhaust plumes are still scarce, especially for the stratosphere where a significant fraction of aircraft emissions are deposited. Simultaneous measurements of NO, NO_2, NO_y, HONO, HNO_3, *halogen* nitrates, and hydroxyl radicals in aging aircraft plumes and contrails could

constrain the range of initial OH mixing ratios and help to validate chemical plume models. Further, experimental investigation of nighttime chemical processes in aircraft exhaust has been neglected in the past. Heterogeneous effects on particles may modify plume chemistry and affect the composition of the tropopause region. Hence, extending the aforementioned observations by measuring particle composition could clarify the effects of heterogeneous processes on aerosol and contrail ice particles.

Alternative fuels such as sulfur-poor *biogenic fuels* present another future research direction, as it has been suggested that bio-fuels might reduce the carbon footprint from aviation. Emission characteristics of biogenic fuels are currently being investigated in engine test facilities, while the effects of biogenic fuels on atmospheric composition and climate still remain unresolved. Plume processing and particle chemistry as well as the ice nucleating ability of biogenic soot could be probed by chasing an aircraft, where one engine is operated with a biogenic fuel, a biogenic fuel/kerosene mixture, or blended fuel (Lufthansa 2012). Thereby the effects of sulfur have to be taken into account and a sulfur-poor biogenic fuel might be used for these tests to enable intercomparability with other kerosene-type fuels.

References

Anderson, B.E., Cofer, W.R., Crawford, J., Gregory, G.L., Vay, S.A., Brunke, K.E., Kondo, Y., Koike, M., Schlager, H., Baughcum, S.L., et al.: An assessment of aircraft as a source of particles to the upper troposphere. Geophys. Res. Lett. **26**, 3069–3072 (1998). doi: 10.1029/1999GL900276

Arnold, F., Scheid, J., Stilp, T., Schlager, H., Reinhardt, M.E.: Measurements of jet aircraft emissions at cruise altitude. 1. The odd-nitrogen gases NO, NO_2, HNO_2, and HNO_3. Geophys. Res. Lett. **19**, 2421–2424 (1992)

Brasseur, G.P., Müller, J.-F., Granier, C.: Atmospheric impact of NO_x emissions by subsonic aircraft: A three-dimensional model study. J. Geophys. Res. **101**, 1423–1428 (1996). doi: 10.1029/95JD02363

Brock, C.A., Schröder, F., Kärcher, B., Petzold, A., Busen, R., Fiebig, M.: Ultrafine particle size distributions measured in aircraft exhaust plumes. J. Geophys. Res. **105**, 26555–26567 (2000)

Brunner, D., Staehelin, J., Jeker, D., Wernli, H., Schumann, U.: Nitrogen oxides and ozone in the tropopause region of the northern hemisphere: measurements from commercial aircraft in 1995/1996 and 1997. J. Geophys. Res. **106**, 27673–27700 (2001). doi:10.1029/2001JD900239

Crutzen, P.J.: The influence of nitrogen oxides on the atmospheric ozone content. Q. J. R. Meteorol. Soc. **96**, 320–325 (1970)

Crutzen, P.J.: The role of NO and NO_2 in the chemistry of the troposphere and stratosphere. Ann. Rev. Earth and Planetary Sci. **7**, 443–472 (1979)

Curtius, J.,Arnold, F., Schulte, P.: Sulfuric acid measurements in the exhaust plume of a jet aircraft in flight: implications for the sulfuric acid formation efficiency. Geophys. Res. Lett. **29**, 17.11–17.14, (2002) doi:10.1029/2001GL013813

Fahey, D.W., Keim, E.R., Boering, K.A., Brock, C.A., Wilson, J.C., Jonsson, H.H., Anthony, S., Hanisco, T.F., Wennberg, P.O., Miake-Lye, R.C., et al.: Emission measurements of the Concorde supersonic aircraft in the lower stratosphere. Science **270**, 70–74 (1995)

Febvre, G., Gayet, J.F., Minikin, A., Schlager, H., Shcherbakov, V., Jourdan, O., Busen, R., Fiebig, M., Kärcher, B., Schumann, U.: On optical and microphysical characteristics of contrails and cirrus. J. Geophys. Res. **114**, D02204 (2009). doi:10.1029/2008JD010184

Grewe, V., Dameris, M., Fichter, C., Sausen, R.: Impact of aircraft NO_x emissions. Part 1: interactively coupled climate-chemistry simulations and sensitivities to climate-chemistry feedback, lightning and model resolution. Meteorol. Z. **11**, 177–186 (2002)

Hendricks, J., Kärcher, B., Lohmann, U.: Effects of ice nuclei on cirrus clouds in a global climate model. J. Geophys. Res. **116**, D18206 (2011). doi:10.1029/2010JD015302

Hoor, P., Borken-Kleefeld, J., Caro, D., Dessens, O., Endresen, O., Gauss, M., Grewe, V., Hauglustaine, D., Isaksen, I.S.A., Jöckel, P., et al.: The impact of traffic emissions on atmospheric ozone and OH: results from QUANTIFY. Atmos. Chem. Phys. **9**, 3113–3136 (2009). doi:10.5194/acp-9-3113-2009

IPCC: Aviation and the Global Atmosphere, p. 373. Cambridge Univ Press, Cambridge (1999)

Jayne, J.T., Pöschl, U., Chen, Y., Dai, D., Molina, L.T., Worsnop, D.R., Kolb, C.E., Molina, M.J.: Pressure and temperature dependence of the gas-phase reaction of SO_3 with H_2O and the heterogeneous reaction of SO_3 with H_2O/H_2SO_4 surfaces. J. Phys. Chem. A **101**, 10000–10011 (1997). doi:10.1021/jp972549z

Johnston, H.S.: Reduction of stratospheric ozone by nitrogen oxide catalysts from supersonic transport exhaust. Science **173**, 517–522 (1971)

Jurkat, T., Voigt, C., Arnold, F., Schlager, H., Kleffmann, J., Aufmhoff, H., Schäuble, D., Schaefer, M., Schumann, U.: Measurements of HONO, NO, NO_y and SO_2 in aircraft exhaust plumes at cruise. Geophys. Res. Lett. **38**, L10807 (2011). doi:10.1029/2011GL046884

Kärcher, B., Hirschberg, M.M., Fabian, P.: Small-scale chemical evolution of aircraft exhaust species at cruising altitudes. J. Geophys. Res. **101**, 15169–15190 (1996). doi:10.1029/96JD01059

Kärcher, B.: Aviation-produced aerosols and contrails. Surv. Geophys. **20**, 113–167 (1999)

Kärcher, B., Voigt, C.: Formation of nitric acid/water ice particles in cirrus clouds. Geophys. Res. Lett. **33**, L08806 (2006). doi:10.1029/2006GL025927

Kärcher, B., Yu, F.: Role of aircraft soot emissions in contrail formation. Geophys. Res. Lett. **36**, L01804 (2009). doi:10.1029/2008GL036649

Knollenberg, R.G.: Measurements of the growth of the ice budget in a persisting contrail. J. Atmos. Sci. **29**, 1367–1374 (1972)

Koike, M., Kondo, Y., Ikeda, H., Gregory, G.L., Anderson, B.E., Sachse, G.W., Blake, D.R., Liu, S.C., Singh, H.B., Thompson, A.M., et al.: Impact of aircraft emissions on reactive nitrogen over the North Atlantic flight corridor region. J. Geophys. Res. **105**, 3665–3677 (2000)

Konopka, P.: Analytical gaussian solutions for anisotropic diffusion in a linear shear flow. J. Non-Equilib. Thermodyn. **20**, 78–91 (1995)

Lee, D.S., Fahey, D.W., Forster, P.M., Newton, P.J., Wit, R.C.N., Lim, L.L., Owen, B., Sausen, R.: Aviation and global climate change in the 21st century. Atmos. Environ. **43**, 3520–3537 (2009). doi:10.1016/j.atmosenv.2009.04.024

Lee, D.S., Pitari, G., Grewe, V., Gierens, K., Penner, J.E., Petzold, A., Prather, M.J., Schumann, U., Bais, A., Berntsen, T., et al.: Transport impacts on atmosphere and climate: aviation. Atmos. Env. **44**, 4678–4734 (2010). doi:10.1016/j.atmosenv.2009.06.005

Lufthansa,: Biofuel flyer. http://verantwortung.lufthansa.com/fileadmin/downloads/de/LH-Biofuel-Flyer.pdf (2012)

Misaka, T., Holzäpfel, F., Hennemann, I., Gerz, T., Manhart, M., Schwertfirm, F.: Vortex bursting and tracer transport of a counter-rotating vortex pair. Phys. Fluids **24**, 025104 (2012). doi:10.1063/1.3684990

Paoli, R., Cariolle, D., Sausen, R.: Review of effective emissions modeling and computation. Geosci. Model Dev. **4**, 643–667 (2011). doi:10.5194/gmd-4-643-2011

Petzold, A., Busen, R., Schröder, F.P., Baumann, R., Kuhn, M., Ström, J., Hagen, D.E., Whitefield, P.D., Baumgardner, D., Arnold, F., et al.: Near-field measurements on contrail properties from fuels with different sulfur content. J. Geophys. Res. **102**, 29867–29880 (1997). doi:10.1029/97JD02209

Petzold, A., Stein, C., Nyeki, S., Gysel, M., Weingartner, E., Baltensperger, U., Giebl, H., Hitzenberger, R., Döpelheuer, A., Vrchoticky, S., et al.: Properties of jet engine combustion

particles during the part Emis experiment: microphysics and chemistry. Geophys. Res. Lett. **30**(13), 1719 (2003). doi:10.1029/2003GL017283

Petzold, A., Fiebig, M., Fritzsche, L., Stein, C., Schumann, U., Wilson, C.W., Hurley, C.D., Arnold, F., Ktragkou, E., Baltensperger, U., et al.: Particle emissions from aircraft engines—a survey of the European project partEmis. Meteorol. Z. **14**, 465–476 (2005). doi:10.1127/0941-2948/2005/0054

Schäuble, D., Voigt, C., Kärcher, B., Stock, P., Schlager, H., Krämer, M., Schiller, C., Bauer, R., Spelten, N., de Reus, M., et al.: Airborne measurements of the nitric acid partitioning in persistent contrails. Atmos. Chem. Phys. **9**, 8189–8197 (2009). doi:10.5194/acp-9-8189-2009

Schlager, H., Konopka, P., Schulte, P., Schumann, U., Ziereis, H., Arnold, F., Klemm, M., Hagen, D.E., Whitefield, P.D., Ovarlez, J.: In situ observations of air traffic emission signatures in the North Atlantic flight corridor. J. Geophys. Res. **102**, 10739–10750 (1997)

Schlager, H., Schulte, P., Flatøy, F., Slemr, F., Velthoven, P.V., Ziereis, H., Schumann, U.: Regional nitric oxide enhancements in the North Atlantic flight corridor observed and modeled during POLINAT 2, a case study. Geophys. Res. Lett. **26**, 3061–3064 (1999). doi:10.1029/1999GL900349

Schröder, F., Brock, C.A., Baumann, R., Petzold, A., Busen, R., Schulte, P., Fiebig, M.: In situ studies on volatile jet exhaust particle emissions: impacts of fuel sulfur content and environmental conditions on nuclei mode aerosols. J. Geophys. Res. **105**, 19941–19954 (2000a)

Schröder, F.P., Kärcher, B., Duroure, C., Ström, J., Petzold, A., Gayet, J.-F., Strauss, B., Wendling, P., Borrmann, S.: The transition of contrails into cirrus clouds. J. Atmos. Sci. **57**, 464–480 (2000b)

Schumann, U., Ström, J., Busen, R., Baumann, R., Gierens, K., Krautstrunk, M., Schröder, F.P., Stingl, J.: In situ observations of particles in jet aircraft exhausts and contrails for different sulfur-containing fuels. J. Geophys. Res. **101**, 6853–6870 (1996). doi:10.1029/95JD03405

Schumann, U., Schlager, H., Arnold, F., Baumann, R., Haschberger, P., Klemm, O.: Dilution of aircraft exhaust plumes at cruise altitudes. Atmos. Environ. **32**, 3097–3103 (1998)

Schumann, U., Schlager, H., Arnold, F., Ovarlez, J., Kelder, H., Hov, Ø., Hayman, G., Isaksen, I.S.A., Staehelin, J., Whitefield, P.D.: Pollution from aircraft emissions in the North Atlantic flight corridor: overview on the POLINAT projects. J. Geophys. Res. **105**, 3605–3631 (2000). doi:10.1029/1999JD900941

Schumann, U., Arnold, F., Busen, R., Curtius, J., Kärcher, B., Petzold, A., Schlager, H., Schröder, F., Wohlfrom, K.H.: Influence of fuel sulfur on the composition of aircraft exhaust plumes: the experiments SULFUR 1-7. J. Geophys. Res. **107**, 4247 (2002). doi:10.1029/2001JD000813

Schumann, U.: A contrail cirrus prediction model. Geosci. Model Dev. Discuss. **4**, 3185–3293 (2011). doi:10.5194/gmdd-4-3185-2011

Sölch, I., Kärcher, B.: A large-eddy model for cirrus clouds with explicit aerosol and ice microphysics and lagrangian ice particle tracking. Q. J. R. Meteorol. Soc. **136B**, 2074–2093 (2010). doi:10.1002/qj.689

Starik, A.M., Savel'ev, A.M., Titova, N.S., Schumann, U.: Modeling of sulfur gases and chemiions in aircraft engines. Aerosp. Sci. Technol. **6**, 63–81 (2002)

Starik, A.M., Savel'ev, A.M., Titiva, N.S., Loukhovitskaya, E.E., Schumann, U.: Effect of aerosol precursors from gas turbine engines on the volatile sulfate aerosols and ion clusters formation in aircraft plumes. Phys. Chem. Chem. Phys. **6**, 3426–3436 (2004)

Stockwell, W.R., Calvert, J.G.: The mechanism of the HO-SO_2 reaction. Atmos. Environ. **17**, 2231–2235 (1983)

Tremmel, H.G., Schlager, H., Konopka, P., Schulte, P., Arnold, F., Klemm, M., Droste-Franke, B.: Observations and model calculations of jet aircraft exhaust products at cruise altitude and inferred initial OH emissions. J. Geophys. Res. **103**, 10803–10816 (1998)

Tremmel, H.G., Schumann, U.: Model simulations of fuel sulfur conversion efficiencies in an aircraft engine: dependence on reaction rate constants and initial species mixing ratios. Aerosp. Sci. Technol. **3**, 417–430 (1999)

Unterstrasser, S., Gierens, K., Spichtinger, P.: The evolution of contrail microphysics in the vortex phase. Meteorol. Z. **17**, 145–156 (2008). doi:10.1127/0941-2948/2008/0273

Voigt, C., Schlager, H., Ziereis, H., Kärcher, B., Luo, B.P., Schiller, C., Krämer, M., Popp, P.J., Irie, H., Kondo, Y.: Nitric acid in cirrus clouds. Geophys. Res. Lett. **33**, L05803 (2006). doi:10.1029/2005GL025159

Voigt, C., Schumann, U., Jurkat, T., Schäuble, D., Schlager, H., Petzold, A., Gayet, J.-F., Krämer, M., Schneider, J., Borrmann, S., et al.: In situ observations of young contrails—overview and selected results from the CONCERT campaign. Atmos. Chem. Phys. **10**, 9039–9056 (2010). doi:10.5194/acp-10-9039-2010

Voigt, C., Schumann, U., Jessberger, P., Jurkat, T., Petzold, A., Gayet, J.-F., Krämer, M., Thornberry, T., Fahey, D.W.: Extinction and optical depth of contrails. Geophys. Res. Lett. **38**, L11806 (2011). doi:10.1029/2011GL047189

Cirrus Clouds and Their Representation in Models

42

Ulrike Burkhardt and Ingo Sölch

Abstract

This article gives a short summary of the physical processes relevant to cirrus and their representation in cloud-resolving models and in global general circulation models. Cloud-resolving models are used to study the evolution of single clouds or cloud systems. With global models the role of clouds in the atmosphere and their interaction with large scale dynamics can be studied. Applications of such models to study cirrus processes and the global contrail cirrus climate impact are discussed. Future research towards a prognostic cloud scheme to include nonequilibrium cirrus cloud physics is laid out.

42.1 Physics of Cirrus Clouds

Cirrus clouds are pure ice clouds found in the *upper troposphere* at temperatures below approximately −40 °C. Above this temperature cirrus clouds may exist, but *mixed-phase* clouds may also be found. The latter are clouds in which liquid droplets and ice crystals coexist. The climatic impact of ice clouds can be estimated from their coverage and *optical depth*. The optical properties of cirrus

U. Burkhardt (✉) · I. Sölch
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Ulrike.Burkhardt@dlr.de

I. Sölch
e-mail: Ingo.Soelch@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_42,

clouds vary widely. Some cirrus clouds are optically thick while many others are optically very thin, so that they can neither be seen by the eye nor detected using ground based or satellite instruments. The latter are therefore called subvisible cirrus. Estimates of cirrus coverage depend on the observational method and the associated optical depth detection threshold. For example, in the tropical Western Pacific about 60 % of cirrus clouds detected with a *lidar* have an optical depth of below 0.2 (Thorsen et al. 2011) and are therefore optically thin. Most of these are not detectable using passive *remote sensing* instruments. Cirrus interacts with radiation, reflecting shortwave solar and trapping longwave *thermal radiation* (Liou 1986). Due to the fact that cirrus persist at high altitude at low temperatures, their effect on the thermal radiation is large. On average, cirrus warms Earth's atmosphere. Cirrus have a significant impact on the upper tropospheric water budget and heating rates and on the water vapor input into the *stratosphere*. Air enters the stratosphere mainly through the tropical tropopause area which connects the convective overturning circulation of the *Hadley cell* and the slow wave-driven upwelling of the lower stratospheric *Brewer–Dobson circulation*. Therefore, tropical cirrus may provide an important control on stratospheric composition by dehydrating the air within subvisible cirrus in the Western Pacific *cold trap* area (Holton and Gettelman 2001). Finally, cirrus clouds can affect low altitude clouds since sedimenting ice crystals from cirrus clouds can lead to the glaciation of liquid clouds due to collisions and subsequent freezing of liquid droplets with sedimenting ice crystals (Herzegh and Hobbs 1980).

Cirrus can form via a number of different pathways. High reaching *convection* transports water rapidly into the upper troposphere where small ice particles (below about 100 μm) that are detrained from the main convective core form an *anvil* cirrus cloud. Within the convective systems water droplets first form and subsequently freeze at temperatures ranging between about −36 °C and about −40 °C, or earlier if they come into contact with ice crystals or interact with heterogeneous *ice nuclei* (IN). These anvils are particularly common in the tropics where they persist in the upper troposphere for around a day (Luo and Rossow 2004).

Cirrus clouds that do not result from the outflow of high reaching convection form in situ, predominantly via *homogeneous nucleation* of supercooled liquid *aerosol* particles (DeMott et al. 2003). High *ice supersaturation* (>0.5) is required to nucleate ice homogeneously in liquid aerosol particles (Koop (2004); Fig. 42.1), as demonstrated in numerous laboratory and field studies. Ice nuclei which constitute a small subset of the atmospheric aerosol may nucleate ice at significantly lower supersaturations than liquid particles. But they have a significant impact on ice cloud properties only when cloud particles form at low cooling rates (DeMott et al. 2003). The fact that large supersaturations relative to ice (in the following, 'ice supersaturation') are needed in order to form cirrus clouds, but only *saturation* relative to ice (ice saturation) is needed to maintain them, means that newly formed ice crystals are not in equilibrium with the surrounding air that is ice supersaturated. This sets cirrus clouds apart from liquid water clouds. The latter form and evaporate slightly above or below saturation (relative to water), respectively. The exact *relative humidity* and temperature at which ice nucleates

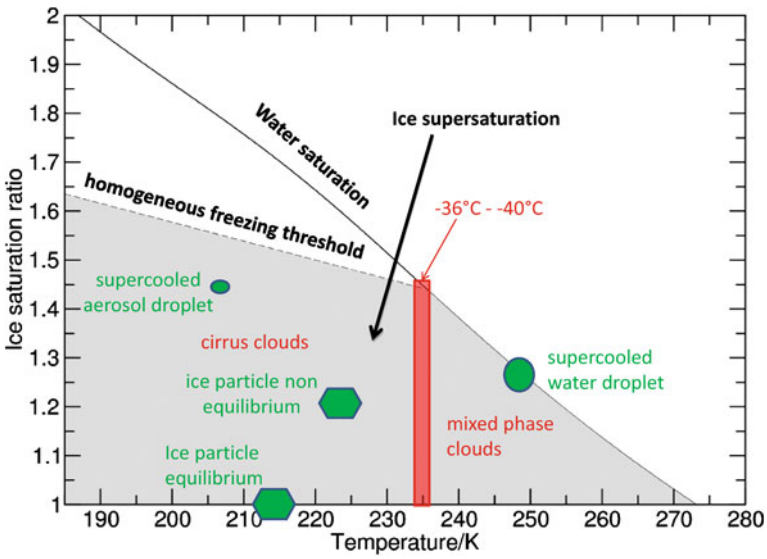

Fig. 42.1 Ice nucleation thresholds and states of water droplets and ice particles. The *grey* area is the area in which air is supersaturated relative to ice. Between −36 °C and −40 °C pure water droplets nucleate ice spontaneously. At colder temperatures, where pure water droplets no longer exist, supercooled aerosol droplets freeze at the homogeneous freezing threshold

depends on the nucleation pathway. *Supercooled water* droplets nucleate at −36 °C to about −40 °C, depending on the radius of the water droplets, the cooling rates and on the degree of purity of the water substance (Koop 2004); supercooled aerosol droplets nucleate at the *homogeneous freezing* threshold (dashed line), and some efficient IN may nucleate at a relative humidity below the homogeneous freezing threshold and above ice saturation (Fig. 42.1). Once formed, ice crystals grow by deposition of water vapor molecules, driving the relative humidity of the air towards ice saturation. *Ice nucleation* and therefore cirrus clouds are often triggered by atmospheric motion such as large-scale ascent within *Rossby waves* or frontal systems, or by vertical motion connected with *gravity waves* acting over a range of scales. Flow over orography often leads to stationary cirrus.

Within ice supersaturated areas in the upper troposphere, which have been shown to exist frequently by in situ and satellite observations (Gierens et al. 1999; Jensen et al. 2001; Lamquin et al. 2012), aircraft-induced *contrails* persist. Due to the mixing of hot and moist exhaust air and the cold surrounding air, contrail ice particles can form and persist already at ice saturation, whereas natural cirrus formation usually occurs at high ice supersaturation. Therefore, *contrail cirrus* can lead to an increase in cirrus cloudiness. The ice water content of contrails initially consists of the moisture emitted by the aircraft engines and increases rapidly due to vapor deposition from the surrounding air. Contrails often develop into extended

contrail cirrus cloud decks that significantly modify the natural cirrus coverage (Burkhardt and Kärcher 2009; 2011).

Ice crystals in cirrus clouds are very diverse as they have many different *habits* (Miloshevich and Heymsfield 1997). The habit of an ice crystal depends on the ice formation process and on the temperature, pressure and humidity experienced by the ice crystal during its lifetime. Within large-scale cirrus pristine ice crystals may have hexagonal, plate-like, or rosette-like shapes, or they may form aggregates which can develop into snowflakes. Within convective clouds liquid water droplets often freeze rapidly to form hail. The ice crystal habit affects the radiative impact of ice clouds and the sedimentation speed of ice crystals.

The presence of IN which might nucleate ice at lower ice supersaturation than the liquid aerosol particles may modify the properties of cirrus clouds (IPCC 2007) and may therefore have an effect on their lifetimes. Solid heterogeneous ice nuclei are by far not as common in the atmosphere as liquid aerosol particles (that nucleate homogeneously) so that a heterogeneous ice nucleation event may only produce few large ice crystals, as opposed to homogeneous nucleation events that may nucleate many small crystals. Heterogeneous ice nucleation may therefore lead to an earlier cirrus onset, to cirrus of lower optical depth, and, because larger ice crystals settle faster, to a decrease in their lifetime. The extent to which these effects are important in the atmosphere could not yet been determined.

42.2 Modeling of Cirrus: Bridging the Scale Gap

One difficulty in simulating clouds in general, or ice clouds in particular, is caused by the large range of scales associated with the processes controlling their formation and evolution. The evolution of ice clouds is determined by interactions between microphysical cloud processes, atmospheric motions, and radiation, all of which can act on very different time and space scales. In Fig. 42.2 a scale diagram is shown, arranging atmospheric motions, cloud structures and microphysical processes according to their spatial and temporal scales. The scales associated with the microphysical and dynamical processes range over several orders of magnitude, starting from the submicrometer spatial and millisecond temporal scale to the planetary scale, which is associated with a time scale of days or more. Tiny ($\sim$100 nm) aerosol particles act as precursors of ice crystals which nucleate and quickly grow into micrometer sized ice particles, eventually reaching sizes of some hundred micrometers due to the deposition of water vapor molecules. Ice crystals sediment in the presence of Earth's gravity field, and when they collide they may stick together in a process called *aggregation* and may modify lower clouds through which they fall. The ice particles form clouds which are affected by atmospheric motions. *Turbulence* on small scales, and larger, organized turbulence such as convection, gravity waves and *synoptic* scale motion, control the occurrence and properties of cirrus.

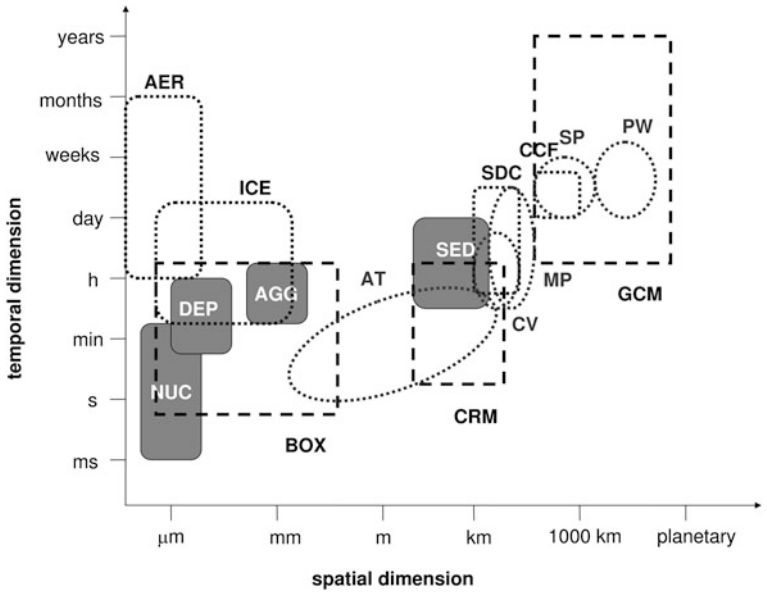

Fig. 42.2 Typical scales of atmospheric motion (*light grey ellipses*), clouds and ice crystals or aerosols (*squares*), and microphysical process (*filled boxes*) relevant for the formation and development of ice clouds. *AT* atmospheric turbulence, *CV* convection, *MP* mesoscale processes, *SP* synoptic-scale processes, *PW* planetary waves, *AER* aerosol particles, *ICE* ice crystals, *SDC* single detached clouds, *CCF* cirrus cloud fields, *NUC* ice nucleation, *DEP* depositional growth, *AGG* aggregation, *SED* sedimentation. *Dashed boxes* represent typical resolution and domain sizes of different modeling approaches: *BOX* box or parcel models, *CRM* cloud-resolving models, *GCM* general circulation models. Adapted from (Sölch and Kärcher 2010)

Several model approaches (Fig. 42.2) are used in atmospheric sciences to study cirrus. They each cover different scales and cirrus-related processes. Using box or parcel models the smallest spatial scales relevant for cirrus clouds are studied. The nucleation and initial evolution of a population of ice crystals in an individual volume of air is treated explicitly. The aerosols acting as precursors for ice crystals are described in detail, which allows study of the relevance of nucleation pathways and microphysical processes in cirrus clouds. *Cloud-resolving models* simulate cirrus formation and evolution on the scale of single clouds with varying degrees of microphysical complexity. Some cloud-resolving models use bulk microphysics and others spectral microphysics, the latter retaining information about the distribution of ice crystal number concentration versus size. The philosophy behind such a model is to retain as much of the microphysical complexity of box models as possible while additionally incorporating the interaction with the radiation and the dynamics, such as cooling of cloud tops and radiative warming of the lower cloud layers, which can lead to turbulence and self-sustaining rising motion.

The first step toward studying cirrus in a cloud-resolving model was taken by Starr and Cox (1985), who developed a two-dimensional cirrus model. Although the microphysical properties of the cirrus were only parameterized on the basis of measurable quantities, e.g., ice water content, the model revealed the important interaction between radiative heating of cloud bases or cooling of cloud tops and the internal dynamics driven by the resulting temperature differences. Over the years more cirrus models were established, refining the treatment of the ice phase and microphysical processes (e.g., Jensen et al. 1994; Spichtinger and Gierens 2009). The importance of different aerosol-related ice formation pathways gained scientific interest and was added to the modeling environments with increasing levels of complexity. In recent years cirrus simulations are also conducted as *large-eddy simulations* (LES) using enhanced resolution (below 100 m). Thus, the shape and structure of cirrus can be simulated increasingly realistically.

In *General Circulation Models* (GCMs), the interaction between cloud fields with dynamics and radiation is studied on long temporal and large spatial scales. GCMs do not explicitly capture the formation and evolution of ice owing to their coarse spatial and low temporal resolution; neither can individual clouds be resolved. The effect of the unresolved processes needs to be captured in terms of the resolved variables using so called *parameterizations*. Especially the representation of cirrus in GCMs is very crude, with cirrus often being modeled using the physics of warm clouds. For a long time, cirrus have not been a focus of research due to a lack of observations and because cloud parameterizations focused to a large degree on liquid clouds since research was driven by the desire to improve precipitation forecasts. Passive remote sensing instruments are not very sensitive to thin cirrus clouds and lidar measurements from the ground, and in situ measurements have not been able to supply the necessary data coverage. This situation has very recently improved with the start of combined space borne lidar and passive remote sensing observations on the A-Train satellites (Stephens et al. 2002). Nevertheless, ice clouds have been a focus of research for a long time at the DLR Institute of Atmospheric Physics partly due to the need to estimate the impact of *air traffic* on ice clouds.

42.3 LES Modeling of Cirrus Clouds at the Institute of Atmospheric Physics

At the DLR Institute of Atmospheric Physics a cloud-resolving LES model, the Lagrangian Cirrus Module used in the LES model EULAG (EULAG-LCM) with size resolved (spectral) representation of ice crystals, was recently developed (Sölch and Kärcher 2010). The model treats the ice phase, including the large variability in the ice particle habit, by tracking single ice crystals in the turbulent flow field. For typical upper-tropospheric temperatures (220 K), pressure (300 hPa), and ice crystal number concentrations (<1 cm^{-3}), at most one ice crystal resides in an air volume of 1 cm^3, compared to 10^{19} air molecules and 10^{14}–10^{15} water molecules. Therefore, treating the air as a continuum and the ice

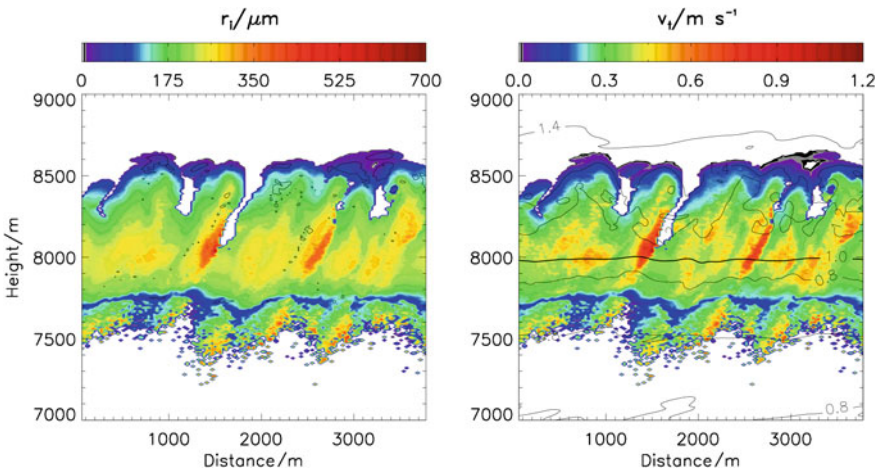

Fig. 42.3 Two-dimensional distribution of (*left*) mean ice crystal maximum dimension (*color shading*) and contour of ice crystal number concentration (>0.05 cm^{-3}) and (*right*) terminal fall speeds (*color shading*) averaged over all ice particles in a grid box, and contours of the ice saturation ratio simulated by the EULAG-LCM for the present case-study at about 66 min after cloud formation. Results highlight the small-scale variability in the cirrus cloud. Adapted from Sölch and Kärcher (2011)

phase as dispersed particles appears to be well suited for cirrus cloud simulations. While gas and aerosol is treated on a fixed Eulerian grid similar to the dynamic and thermodynamic variables, a large number of ice crystals can be tracked in a Lagrangian approach. The suite of explicitly resolved physical processes enables the detailed simulation and analysis of the dynamical-microphysical-radiative *feedbacks* characteristic of cirrus. In a typical simulation the evolution and life cycle of up to tens of millions of ice particles are calculated and from this sample the properties of the cirrus cloud are calculated. Such simulations have become possible only with the advent of supercomputers and increases in memory and processor performance.

Models that simulate an increasing number of processes require increasingly detailed and comprehensive data sets for model validation. The EULAG-LCM was evaluated using such a comprehensive data set, the U.S. Atmospheric Radiation Measurement Program Intensive Operational Period (IOP) in March 2000. In a three-dimensional domain with 5 km horizontal extent the meteorological conditions were reconstructed from the available measurement data, enabling simulation of the cirrus cloud formation on the local morning of 9 March 2000 in a strong south-westerly flow over the Rocky Mountains with the EULAG-LCM (Sölch and Kärcher 2011). The simulation followed the evolution of the cloud from the location where it was first detected by satellite up to the observation site where it was later detected. Comparison of simulations to measurement data allows validation of the numerical cloud laboratory and offers a unique possibility to gain insights into the microphysical structure of the cloud, which ultimately determines

the radiative properties of cirrus. In Fig. 42.3 the water vapor field, the mean terminal fall velocity, and the mean diameter of the ice crystals are shown together with the ice crystal number concentration. At this early stage of cloud evolution, the vertical extent of the cloud increases steadily. In the upper cloud region, where high supersaturation prevails, numerous new ice particles are formed and the ice crystals stay relatively small (<50 μm). At the cloud base several fall streaks are evident. Obviously, the cirrus properties alter significantly on scales down to 100 m or below.

A detailed comparison showed that the ice crystal size distributions (the relative abundance of ice crystal sizes in a cloud volume) measured as it advected over the main observation site, from an airplane descending through the cloud deck, could be reproduced by the model in most parts of the cirrus. Validation of ice crystal size distributions is crucial since it is the most detailed information available on the ice phase properties in cirrus and determines, together with knowledge of the crystal habits and their spatial distribution, almost all of the derived quantities such as radiative properties, number concentration, and ice water content. The validated model was then used to perform sensitivity studies examining the influence of several atmospheric and microphysical parameters on the evolution of the cirrus. Tracking individual ice crystals also permits studies of the life cycle of individual cloud particles. Particle growth, their position in the cloud, their frequency of collision with other crystals to form aggregates, and their sedimentation speed could be analyzed. Studying a large number of simulation particles allows the assembly of statistics on cirrus properties, which are difficult to access with current measurement methods. Sensitivity studies revealed that domain-integrated optical depth of the simulated cirrus cloud is most sensitive to the updraft speed and ice supersaturation. Cloud-averaged ice *particle size distributions* are bimodal, separating two distinct growth regimes in the developed cloud. The small ice mode (ice particle sizes below a few 100 μm) in this cloud was formed by homogeneous freezing of supercooled aerosol droplets and grew by deposition of water molecules from the gas phase. The large mode (sizes up to several 1 000 μm) grew by aggregation. The formation of the largest crystals by aggregation in this deep cirrus was controlled in part by the nucleation of new ice crystals in dynamically active, highly supersaturated upper cloud regions, which settle through the entire underlying cirrus region, triggering a large number of aggregation events.

42.4 Cirrus Modeling in Climate Models

In GCMs clouds are simulated on an equidistant latitude–longitude grid with a spatial and temporal resolution that is far too coarse for resolving single clouds or the processes leading to droplet or ice formation. Instead, the clouds and their properties have to be inferred from the large-scale model variables using so called parameterizations. Convective clouds and large-scale (stratiform) clouds are handled by different parameterizations since convection usually occurs on much smaller spatial and shorter temporal scales than are resolved. In the following, the

parameterization of large-scale clouds and particularly those used in the German community climate model (ECHAM) will be described. In the GCM, large-scale clouds are described by their fractional coverage and, in the case of a model with a one-moment microphysical scheme, by their water content and, in the case of a two-moment microphysical scheme, additionally by the droplet/ice crystal number concentration. The fractional coverage describes the fraction of the grid box that is cloud covered. From the grid-mean variables the cloud fraction can be calculated assuming a distribution of subgrid scale variability of cloud-controlling variables, e.g., total water (water vapor plus condensate). In many cases the parameterization of fractional cloud coverage is diagnostic and often inferred from relative humidity only (relative humidity scheme). In the limit of a very small grid box, that is very high resolution, the *subgrid scale* variability would be negligibly small and, e.g., liquid clouds would form once water saturation is reached on the grid scale. The lower the resolution of a climate model, the larger the subgrid scale variability within a grid box and the lower the grid-mean relative humidity threshold, above which fractional *cloud coverage* is non-zero within the grid box of the GCM. Since the grid boxes in climate models are commonly hundreds of kilometers in size, cloud coverage is non-zero at a grid-mean relative humidity considerably below saturation and increases with relative humidity, reaching full cloud coverage at saturation (in the case of warm clouds). The latter is often assumed for cirrus as well, even though it is only appropriate for liquid phase clouds. This means that ice clouds are usually parameterized using the physics of warm clouds within a GCM.

The increase in cloud coverage is dependent on the assumptions made for the subgrid scale variability of cloud-controlling variables. The likelihood of the cloud-controlling variables to take a given value is usually described in form of a *probability density function* (PDF). The probability density function of subgrid scale variability is in many parameterizations prescribed to be a very simple function such as a top hat (constant probability between a lower and an upper threshold) or a triangle, but it can also take a more complicated form. The shape of the PDF determines the diagnosed cloud fraction insofar as cloud fraction is the integral over the PDF starting from some lower threshold (e.g., saturation) up to infinity. Often the variance (a measure of the width) of the PDF is fixed within a model so that the cloud coverage can be diagnosed from the grid-mean relative humidity alone. Assuming a top hat PDF within a relative humidity scheme and keeping its variance constant (given a specific temperature) leads to an increase in cloud coverage with relative humidity as shown in Fig. 42.4 (dashed curve).

The (ice) water content of clouds is commonly a prognostic variable (requiring simulation of the time evolution). Besides the dynamical processes affecting the prognostic cloud variables, microphysical sources and sinks of (ice) water content are condensation and evaporation (deposition and sublimation in case of ice), sedimentation and precipitation. The latter two processes are parameterized depending on the local (ice) water content (defined as the grid-mean ice water content divided by cloud coverage) when using a one-moment microphysical scheme. Using a two-moment microphysical scheme [additionally using droplet (ice particle) number concentration as a prognostic variable] allows more realistic

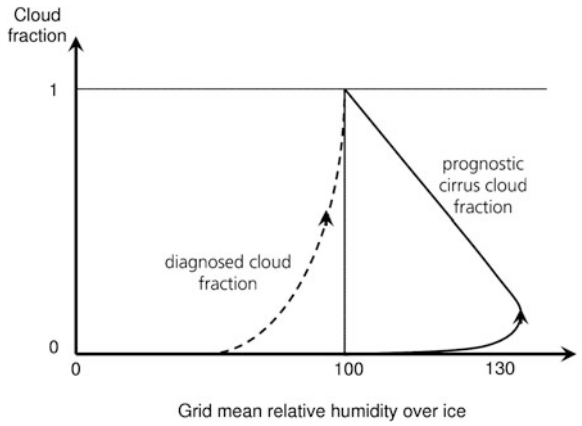

Fig. 42.4 Diagnostic cloud fraction (*dashed curve*) depending on grid-mean relative humidity assuming a top hat probability density function of subgrid scale variability. Within the cloudy area relative humidity is prescribed to be 100 %. Evolution of cirrus cloud fraction in a continuously cooling air parcel (*solid curve*). Cirrus form locally at high ice supersaturation and drive humidity towards saturation. This necessitates prediction of the cloud fraction (requiring a memory). Traditional cloud schemes diagnosing cloud fraction cannot be reconciled with nonequilibrium ice-physics

representation of these two processes and microphysics and radiation in general because the mean radius of the water droplets (ice particles) is known. Microphysical sources and sinks of the droplet/ice particle number concentration include activation/ice nucleation processes, evaporation/sublimation, aggregation, ice splintering, riming, sedimentation and precipitation. Two-moment microphysical schemes enable the simulation of indirect aerosol effects on clouds.

Fixed variance schemes cannot respond to microphysical or dynamical sources or sinks of variance in the cloud related parameters. For example, convection reaching a grid box is not connected with an increase in the moisture variance and does not necessarily result in associated cloud coverage. Instead, convection injects water in the form of water droplets or ice crystals and humidity into a grid box, but in many parameterizations coverage depends only on whether the convection can moisten the grid box enough to reach the threshold relative humidity above which cloud coverage is diagnosed. A clear connection between the PDF moments and the forcing due to dynamical and microphysical processes is made in the cloud parameterization of Tompkins (2002), which is based on a prognostic PDF of subgrid scale variability of total water. A disadvantage using this approach is that it is difficult to determine the tendencies of the moments of the PDF, particularly those due to microphysical processes. In another PDF approach the large-scale cloud scheme is coupled to the convection scheme (Bony and Emanuel 2001) resulting in improved cloud simulations in the tropics.

Most cloud schemes assume that water droplets and ice particles form as soon as saturated conditions are surpassed, as is appropriate for warm clouds. This is despite the fact that upper tropospheric cirrus form and develop in nonequilibrium ice supersaturated conditions and persist at ice saturation, implying a hysteresis in the evolution of cloud coverage (Fig. 42.4 solid curve). Therefore, ice supersaturated air may or may not be cloud covered depending on the history of the cirrus cloud. Likewise, in cloud schemes ice crystals are not allowed to exist in subsaturated conditions despite observational evidence for the transient existence of ice crystals in subsaturated air (Ström et al. 2003). In nearly all cloud schemes it is assumed that within clouds air is saturated. This prohibits the simulation of ice supersaturation within, e.g., thin tropical cirrus clouds. In the case of a warm cloud, cloud coverage and liquid water mass can be calculated as integrals of the PDF of total water mixing ratio, or of the product of the PDF and total water *mixing ratio*, respectively, from the saturation mixing ratio to infinity (or the upper integration bound). However, this is not possible for cirrus. In the case of cirrus, the history of an air parcel needs to be known. Only if the cirrus cloud coverage is known is it possible to infer the ice water content from the PDF of the total water mixing ratio.

Treating cirrus in the same way as liquid water clouds has implications for the water and radiation budget of the upper troposphere, e.g., in the form of dry biases (Ovarlez et al. 2000). Recently, cloud parameterizations have been modified to accommodate ice cloud physics. A physically based parameterization of homogeneous ice formation led to improvements in the upper tropospheric relative humidity, ice water content and ice crystal number concentration (Lohmann and Kärcher 2002). Subsequently, the parameterization was improved to additionally account for effects of IN in order to study their effect on cirrus (Hendricks et al. 2011). Using such a scheme allows the simulation of ice supersaturation on the grid scale, but converts fractional cloud coverage to an unrealistic scheme with either full or zero cloud coverage. Other approaches that were used are to increase the relative humidity at which clouds form and dissolve to ice supersaturated values (Gettelman and Kinnison 2007), or to modify a prognostic cloud scheme allowing for ice supersaturation by forming ice clouds at a high ice supersaturation and by letting them persist at ice saturation while prescribing ice saturation within the ice clouds (Tompkins et al. 2007). This allows representation of cloud free ice supersaturated conditions but not ice supersaturation within ice clouds, as is common in clouds with low ice crystal number concentrations. In order to enable a process-based treatment of ice supersaturation, cirrus microphysics and coverage, it is necessary to track the cirrus cloud fraction and ice water content of these often long-lived large-scale cirrus (Kärcher and Burkhardt 2008).

42.5 Global Modeling of Contrail Cirrus at the Institute of Atmospheric Physics

Air traffic and its effect on climate, in particular through contrails and contrail cirrus, is a major research topic at the Institute of Atmospheric Physics. Climate models are being used in order to study the effect of contrail cirrus on climate, whereas high resolution *numerical weather prediction* models are particularly useful for supporting real time contrail *mitigation* efforts. Here, only the simulation of contrail cirrus in the German community climate model ECHAM is discussed.

One of the least understood components of aviation-induced *climate change* is the climate impact due to contrail cirrus, comprising young *linear contrails* and older irregularly shaped contrails that arise from them. Coverage and *radiative forcing* by linear contrails have been simulated in ECHAM in the past using a diagnostic approach estimating the frequency of contrail formation from the large-scale atmospheric model fields and scaling those with local observations of linear contrails (Ponater et al. 2002). Assuming the scaling to be temporally and spatially constant, the global radiative forcing of linear contrails could be estimated. However, the simulation of longer lived contrail cirrus requires a prognostic treatment of contrail cirrus (including a memory of its history) in the GCM. This allows capturing processes controlling contrail cirrus and its properties, such as advection and spreading.

A prognostic contrail cirrus scheme was implemented in the ECHAM climate model (Burkhardt and Kärcher 2009), introducing contrail cirrus as a prognostic new cloud type. In the model, the formation of contrails in ice supersaturated areas and the processes controlling contrail cirrus coverage (and properties) advection, spreading, deposition, sublimation and precipitation are parameterized. The ice supersaturated areas in which contrail cirrus form and persist are parameterized consistent with the assumptions of subgrid scale total water variability inherent in the model's natural cloud scheme (Burkhardt et al. 2008). Contrail cirrus coverage is calculated assuming that contrail cirrus (or clouds) in adjacent levels of the same grid box overlap maximally and otherwise randomly (maximum random overlap). This assumption is commonly made for natural clouds, but remains a topic of research. The introduction of contrail cirrus in the GCM requires the closure of the water budget including the new cloud class, and leads to changes in the moisture budget affecting natural clouds and their properties. Contrail cirrus and natural cirrus are therefore competing for available water, so that the introduction of contrail cirrus in the model may have a discernible impact on natural cirrus.

According to the model, contrail cirrus coverage amounts to 0.61 % globally (of which approximately a third is optically thicker than 0.02) and is concentrated in the area of the major flight routes in the northern hemisphere (Fig. 42.5a). Contrail cirrus coverage over central Europe reaches 10 %, exceeding contrail cirrus coverage over the USA, which amounts to 6 %. Over the eastern coast of Southeast Asia contrail cirrus coverage reaches 1 %. The optical depth of

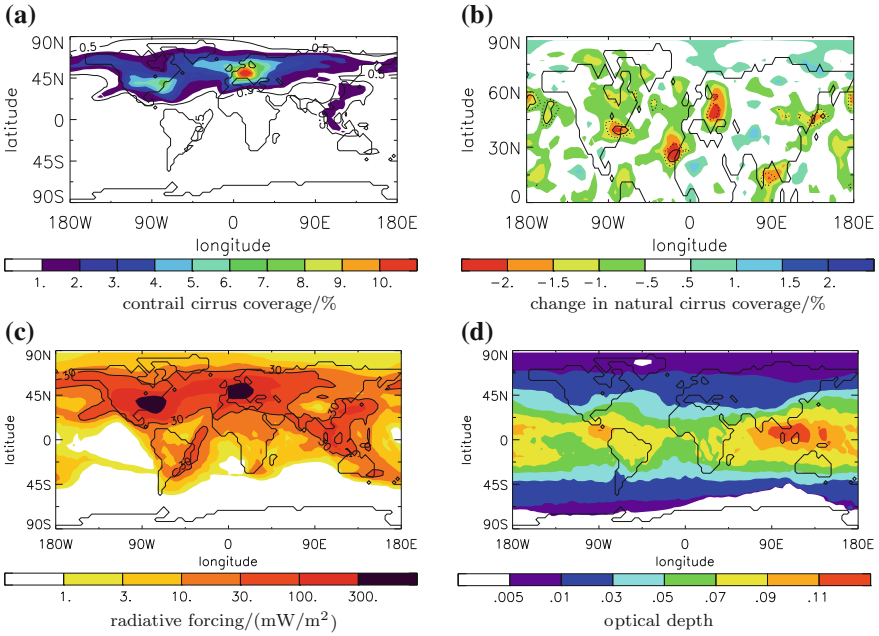

Fig. 42.5 Contrail cirrus coverage in % (a), change in natural cirrus coverage (in % coverage) resulting from the presence of contrail cirrus (b), radiative forcing in mW/m^2 (note the logarithmic scale) (c), and optical depth (d). Statistically significant changes in cirrus coverage are indicated by *solid lines* (95 % significance) and *dotted lines* (90 % significance). For the calculation of contrail cirrus and natural cirrus coverage, coverages in different vertical levels within one grid column have been overlapped using maximum random overlap. Adapted from Burkhardt and Kärcher (2011)

simulated contrail cirrus varies regionally (Fig. 42.5d) and temporally. Close to the equator mean contrail cirrus optical depth is much larger than in the extratropics due to the larger water vapor mixing ratios. This zonal gradient of optical depth has a large influence on the radiative effect of contrails since for optically thin clouds the radiative effect increases linearly with the optical depth. This means that contrail cirrus over the USA have a larger impact on the radiation budget than contrail cirrus over Europe. The change in the radiation budget at the tropopause (Fig. 42.5c), calculated as the difference of the radiative fluxes when including and excluding contrail cirrus, shows that contrail cirrus warm the troposphere and that the radiative forcing exceeds 300 mW/m^2 over the eastern USA and central Europe. Over Southeast Asia radiative forcing exceeds 100 mW/m^2. Over the USA, Europe, the north Atlantic flight corridor and Southeast Asia, the radiative forcing is an order of magnitude larger than over most other areas (note the logarithmic scale used in the figure). Globally, contrail cirrus radiative forcing is calculated to amount to about 38 mW/m^2. The uncertainty of this estimate due to uncertainty in optical depth, ice particle shape and radiative transfer was estimated

to lie around 20 % each. The uncertainty due to the representation of ice supersaturation in the model cannot be meaningfully estimated since existing observations of ice supersaturation differ significantly. An estimate of the combined uncertainty of radiative forcing would need to take into account the interdependence among the different uncertainties.

The presence of contrail cirrus in the upper troposphere changes the humidity field due to the uptake of water vapor in the contrail cirrus and creates local warming. Both effects act to reducing the upper tropospheric relative humidity, causing a decrease in natural cirrus coverage and changes in optical properties. The reduction of natural cloud coverage as estimated by the climate model is shown in Fig. 42.5b. Natural cirrus coverage is reduced mainly in the areas slightly downstream of the maxima in contrail cirrus coverage. The change in natural cirrus coverage and properties cause again a change in the radiative budget at the tropopause. The cooling exerted by the reduction of natural cirrus cloudiness partly compensates the warming due to contrail cirrus so that the overall contrail cirrus radiative forcing (including the cooling due to the change in natural cloudiness) amounts to approximately 31 mW/m^2.

This means that today contrail cirrus radiative forcing exceeds the radiative forcing due to the accumulated CO_2 emitted by aircraft since the start of commercial aviation. Due to the very long lifetime of CO_2 as compared to the lifetime of contrails (hundreds of years as compared to hours or days), CO_2 emitted today will still affect climate in the coming decades, whereas contrail cirrus formed today will have no lasting effect on climate. This means that, as long as air traffic remains at its present level or even increases, contrail cirrus will remain for some time the most important climate forcing associated with aviation. Should air traffic on the other hand decline, then the effect of contrail cirrus on climate will be reduced and may eventually be smaller than the effect of CO_2 emitted by air traffic. This means that mitigating the effect of contrail cirrus is a very effective and fast way of reducing the climate impact of air traffic.

42.6 Conclusions and Future Directions

Clouds constitute a major uncertainty in climate models (IPCC 2007). Differences in the sensitivity of climate models to different forcings are largely due to differences in cloud feedbacks (Cess 1990; Stephens 2005). Hence improvements in the representation of clouds, cloud fraction and optical properties, their vertical overlap and interaction with radiation are important. This contribution gives a short overview over cloud, in particular cirrus, physics, cirrus modeling using cloud-resolving models and the parameterization of cirrus in global models with an emphasis on methods used within the German community climate model ECHAM and puts the research performed at DLR into the context of current research.

Despite great computational advances, global models will for a long time to come be run at resolutions for which cloud processes in general remain unresolved and therefore need to be parameterized. An important step forward in the parameterization of clouds in a GCM would be to base the cloud parameterization on an explicit representation of the joint variability of cloud-controlling variables such as temperature, moisture and vertical velocity on the not resolved (subgrid) scales. This PDF of subgrid scale variability should not only be used in the parameterization of cloud coverage but also for microphysical processes and radiative fluxes (which are strongly nonlinear), thus increasing the consistency in the representation of cloud related processes (Pincus and Klein 2000). Ideally this PDF should also be related to other processes such as turbulence and gravity waves, and should be considered in the calculation of nonlinear chemical reaction rates in CCMs, enhancing the overall consistency of the model. The PDF of subgrid variability should also reflect the impact of large-scale dynamics and convection and therefore needs to be prognostic.

Using a scheme based on a prognostic PDF, prognostic equations for the PDF moments need to be established. This is difficult since it is, e.g., not clear how microphysical processes within clouds change the variance or skewness of the PDF of subgrid scale variability. In order to keep the parameterization tractable, the PDF needs to be described by as few variables as possible. This on the other hand reduces the flexibility of the representation of subgrid scale variability. In the case of convection, the moments of the PDF react to the convective outflow at a particular level, but bimodality in the distribution of the subgrid scale variability, which can result from convection, can be only captured if the form of the PDF is flexible enough. The complexity and sensitivity of a scheme increases strongly with each additional moment that needs to be simulated. This means that it may be difficult to trace sources of instability within a scheme and that a comprehensive data set is needed for an in-depth validation of such a scheme. Since a comprehensive observational data set is not available; validation needs to be based at least in part on model data.

In many GCMs, cirrus are parameterized in the same way as warm clouds, neglecting the fact that cirrus forms at high ice supersaturations. In the last few years, parameterizations of ice nucleation have been implemented in GCMs, leading in many cases to the representation of ice supersaturation on the model's grid scale reverting the fractional cloud scheme to an all-or-nothing scheme. This may be appropriate for studying microphysical processes, but for a realistic simulation of climate, fractional cloud schemes are indispensable in low resolution models. A promising next step toward simulating clouds in low resolution models would certainly be to develop a prognostic PDF cloud scheme to include non-equilibrium cirrus cloud physics. In order to account for this, the assumption that clouds form at saturation needs to be relaxed for the ice phase in order to make simulated cirrus clouds microphysically consistent with the high ice supersaturation observed in the upper troposphere. Cirrus coverage would then describe a hysteresis with increasing relative humidity. The improved representation of ice clouds and the humidity field is likely to reduce model biases in weather

predictions and particularly in climate simulations. This improved PDF based scheme could also be further extended to implement contrail cirrus as a new prognostic cloud class. This scheme would have the advantage of having both a prognostic natural cirrus and prognostic contrail cirrus cloud scheme, which would lead to more realistic representation of the competition between natural cirrus and contrail cirrus. Additionally, the scheme when coupled with a two-moment microphysical scheme would allow physically consistent simulation of the indirect aerosol effects on ice clouds, including the indirect effects of aircraft soot emissions on cirrus.

Improvements in the representation of microphysical processes require at least a two-moment microphysical scheme. This is important since the optical properties and the lifetime of a cloud are dependent on the realistic simulation of the microphysical state of a cloud and since this is a prerequisite for simulating aerosol cloud interactions. A better understanding of cirrus microphysics may be obtained using high resolution LES or cloud-resolving models which can lead to more realistic parameterizations for lower resolution models. This is likely to lead to increasingly complex cloud parameterizations in low resolution models. This poses a challenge since even the most comprehensive measurements available to date are insufficient to constrain the atmospheric and microphysical parameters within cloud-resolving models or cloud parameterizations. Therefore, progress is likely to come from high resolution modeling serving as virtual reality for parameterization development.

References

Bony, S., Emanuel, K.A.: A parameterization of the cloudiness associated with cumulus convection; evaluation using TOGA COARE data. J. Atmos. Sci. **58**, 3158–3183 (2001)

Burkhardt, U., Kärcher, B., Ponater, M., Gierens, K., Gettelman, A.: Contrail cirrus supporting areas in model and observations. Geophys. Res. Lett. **35**, L16808 (2008). doi: 10.1029/2008GL034056

Burkhardt, U., Kärcher, B.: Process-based parameterization of contrail cirrus in a global climate model. J. Geophys. Res. **114** (2009). doi:10.1029/2008JD011491

Burkhardt, U., Kärcher, B.: Global radiative forcing from contrail cirrus. Nature Clim. Change **1**, 54–58 (2011)

Cess, R.D., et al.: Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. J. Geophys. Res. **95**, 16601–16615 (1990)

DeMott, P.J., Cziczo, D.J., Prenni, A.J., Murphy, D.M., Kreidenweis, S.M., Thomson, D.S., Borys, R., Rogers, D.C.: Measurements of the concentration and composition of nuclei for cirrus formation. Proc. Nat. Acad. Sci. U.S.A. **100**, 14655–14660 (2003)

Gettelman, A., Kinnison, D.E.: The global impact of supersaturation in a coupled chemistry climate model. Atmos. Chem. Phys. **7**, 1629–1643 (2007)

Gierens, K., Schumann, U., Helten, M., Smit, H., Marenco, A.: A distribution law for relative humidity in the upper troposphere and lower stratosphere derived from three years of MOZAIC measurements. Ann. Geophys. **17**, 1218–1226 (1999)

Hendricks, J., Kärcher, B., Lohmann,U.: Effects of ice nuclei on cirrus clouds in a global climate model. J. Geophys. Res. **116**(D18206), 1–24 (2011). doi:10.1029/2010JD015302

Herzegh, P.H., Hobbs, P.V.: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. II: Warm frontal clouds. J. Atmos. Sci. **37**, 597–611 (1980)

Holton, J.R., Gettelman, A.: Horizontal transport and the dehydration of the stratosphere. Geophys. Res. Lett. **28**, 2799–2802 (2001)

IPCC: Climate Change 2007: The physical basis, Contribution of working group I to the 4th assessment report of the Intergovernmental Panel on Climate Change. In: Solomon, S., et al. (eds.) Cambridge University Press, Cambridge (2007)

Jensen, E.J., Toon, O.B., Westphal, D.L., Kinne, S., Heymsfield, A.J.: Microphysical modeling of cirrus. 1. Comparison with 1986 FIRE IFO measurements. J. Geophys. Res. **99**, 10421–10442 (1994)

Jensen, E.J., Toon, O.B., Vay, S.A., Ovarlez, J., May, R., Bui, P., Twohy, C.H., Gandrud, B., Pueschel, R.F., Schumann, U.: Prevalence of ice-supersaturated regions in the upper troposphere: implications for optically thin ice cloud formation. J. Geophys. Res. **106**, 17253–17266 (2001)

Kärcher, B., Burkhardt, U.: A cirrus cloud scheme for general circulation models. Quart. J. R. Meteorol. Soc. **134**, 1439–1461 (2008). doi:10.1002/qj.301

Koop, T.: Homogeneous ice nucleation in water and aqueous solutions. Z. Phys. Chem. **218**, 1231–1258 (2004)

Lamquin, N., Stubenrauch, C., Cros, S., Smit, H., Gierens, K., Burkhardt, U.: A 6-year global climatology of occurrence of upper tropospheric ice supersaturation inferred from the Atmospheric Infrared Sounder and its synergy with MOZAIC. Atmos. Chem. Phys. **12**, 381–405 (2012)

Liou, K.-N.: Influence of cirrus clouds on weather and climate processes: a global perspective. Mon. Weather Rev. **114**, 1167–1199 (1986)

Lohmann, U., Kärcher, B.: First interactive simulations of cirrus clouds formed by homogeneous freezing in the ECHAM GCM. J. Geophys. Res. **107**, 4105 (2002). doi:10.1029/2001JD000767

Luo, Z., Rossow, W.B.: Characterizing tropical cirrus life cycle, evolution and interaction with upper-tropospheric water vapor using lagrangian trajectory analysis of satellite observations. J. Clim. **17**, 4541–4563 (2004)

Miloshevich, L.M., Heymsfield, A.J.: A balloon-borne continuous cloud particle replicator for measuring vertical profiles of cloud microphysical properties: instrument design, performance, and collection efficiency analysis. J. Atmos. Oceanic Technol. **14**, 753–768 (1997)

Ovarlez, J., van Velthoven, P., Sachse, G., Vay, S., Schlager, H., Ovarlez, H.: Comparison of water vapor measurements from POLINAT 2 with ECMWF analyses in high humidity conditions. J. Geophys. Res. **105**, 3737–3744 (2000)

Pincus, R., Klein, S.A.: Unresolved spatial variability and microphysical process rates in large-scale models. J. Geophys. Res. **105**, 27059–27065 (2000)

Ponater, M., Marquart, S., Sausen, R.: Contrails in a comprehensive global climate model: parameterization and radiative forcing results. J. Geophys. Res. **107**(D13), 4164 (2002). doi: 10.1029/2001JD000429

Sölch, I., Kärcher, B.: A large eddy model for cirrus clouds with explicit aerosol and ice microphysics and Lagrangian ice particle tracking. Quart. J. R. Meteorol. Soc. **136**, 2074–2093 (2010)

Sölch, I., Kärcher, B.: Process-oriented large-eddy simulations of a midlatitude cirrus cloud system based on observations. Quart. J. R. Meteorol. Soc. **137**(655), 374–393 (2011)

Spichtinger, P., Gierens, K.: Modelling of cirrus clouds—Part 1a: Model description and validation. Atmos. Chem. Phys. **9**, 685–706 (2009)

Starr, D.O.C., Cox, S.K.: Cirrus clouds. Part I: A cirrus cloud model. J. Atmos. Sci. **42**, 2663–2681 (1985)

Stephens, G.L., Vane, D.G., Boain, R.J., Mace, G.G., Sassen, K., Wang, Z., Illingworth, A.J., O'Connor, E.J., Rossow, W.B., Durden, S.L., et al.: The CloudSat mission and the A-train: a

new dimension of space-based observations of clouds and precipitation. Bull. Am. Meteorol. Soc. **83**, 1771–1790 (2002). doi:10.1175/BAMS-83-12-1771

Stephens, G.L.: Cloud feedbacks in the climate system: a critical review. J. Clim. **18**, 237–273 (2005)

Ström, J., Seifert, M., Kärcher, B., Ovarlez, J., Minikin, A., Gayet, J.-F., Krejci, R., Petzold, A., Auriol, F., Haag, W., et al.: Cirrus cloud occurrence as function of ambient relative humidity: a comparison of observations obtained during the INCA experiment. Atmos. Chem. Phys. **3**, 1807–1816 (2003)

Thorsen, T.J., Fu, Q., Comstock, J.: Comparison of the CALIPSO satellite and ground-based observations of cirrus clouds at the ARM TWP sites. J. Geophys. Res. **116**, D21203 (2011). doi:10.1029/2011JD015970

Tompkins, A.: A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover. J. Atmos. Sci. **59**, 1917–1942 (2002)

Tompkins, A., Gierens, K., Rädel, G.: Ice supersaturation in the ECMWF integrated forecast system. Quart. J. R. Meteorol. Soc. **133**, 53–63 (2007)

Climate Impact of Transport

43

Robert Sausen, Klaus Gierens, Veronika Eyring,
Johannes Hendricks and Mattia Righi

Abstract

Transport impacts the atmospheric composition and the climate by CO_2 and non-CO_2 emissions. The atmospheric lifetime of most non-CO_2 emissions is much shorter than the CO_2 lifetime. Nevertheless, the non-CO_2 climate effects are large in comparison to the CO_2 effect, in particular for aviation and shipping. This is mainly due to triggering new clouds and modifying existing clouds, and to the impact of nitrogen oxides emissions on the abundances of ozone and methane.

R. Sausen (✉) · K. Gierens · V. Eyring · J. Hendricks · M. Righi
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Robert.Sausen@dlr.de

K. Gierens
e-mail: Klaus.Gierens@dlr.de

V. Eyring
e-mail: Veronika.Eyring@dlr.de

J. Hendricks
e-mail: Johannes.Hendricks@dlr.de

M. Righi
e-mail: Mattia.Righi@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_43,

43.1 Introduction

The *transport* sector contributes about one-fifth to total global *anthropogenic carbon dioxide* (CO_2) *emissions*. The annual growth rate of transport-related *greenhouse gas* emissions is larger than that from other mature industrial sectors. In light of United Nations Framework Convention on Climate Change (UNFCCC) recommendations, the *Kyoto Protocol*, and possible follow-up protocols, this rate of increase creates a severe problem when trying to achieve emission reduction targets. Additionally, the impact of the transport sector on climate is complex and is not only caused by the gases listed in the Kyoto Protocol.

Nonetheless, the global economic system is largely dependent on an efficient transport network. This dependency has grown during recent decades. Lifestyle and the availability of cheaper transport in developed countries also enhance demand. In addition, strong growth of the transport sector is expected in developing countries. In the long term, a sustainable transport system is needed that satisfies in an optimal way the demands of economy and population whilst adhering to the constraints of climate change control. In order to meet these constraints clear information on the climatic impact of different transport emissions is required.

Transport can impact climate in several ways: (1) by direct emission of greenhouse gases, mainly CO_2, but also nitrous oxide (N_2O) and other gases; (2) by emission of indirect greenhouse gases, e.g., ozone *precursors*, such as *nitrogen oxides* (NO_x), *carbon monoxide* (CO) or *volatile organic compounds* (VOCs); (3) by emission of *aerosols particles*, mainly *soot* containing *black carbon* (BC) and *organic carbon* (OC), and particle precursor gases such as *sulfur dioxide* (SO_2), which are directly and indirectly (via cloud formation and cloud modification) radiatively and chemically active; (4) by directly triggering additional clouds (e.g., *contrails*, *contrail cirrus*).

The climate effects and further effects of transport emissions are visualized in Fig. 43.1. The chain of impacts starts with the emission of substances that become mixed with the ambient air, whereby they undergo chemical reactions (processing) and spatial *dispersion*, resulting in a changed chemical *composition* of the *atmosphere*. *Particulate matter* (PM) emissions may lead to additional clouds and may modify natural cloud properties. All these effects lead to changes in *longwave* and *shortwave radiation*, resulting in *radiative forcing*. Ultimately, the Earth and its atmosphere have to react in order to restore the energy balance. This manifests itself as *climate change*, which can appear in many forms, e.g., higher temperatures, rising sea level, changed *precipitation* patterns, etc., so that a multitude of climate impacts on *health* and economy are possible. While it is relatively straightforward to compute transport emissions, the results become more and more uncertain when going down the cause-effect chain, because the complexity of the problem grows from stage to stage and the uncertainties from one stage are transferred down the chain.

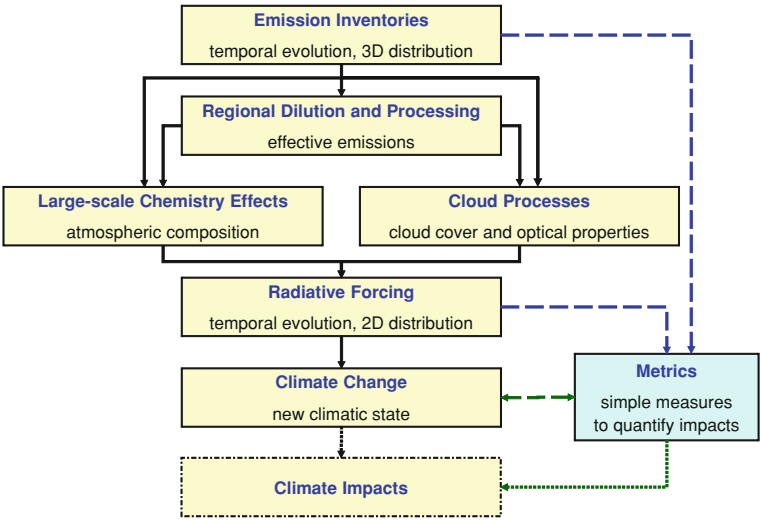

Fig. 43.1 Schematic illustration of the chain of impact for transport emissions. Metrics provide simple methods to bypass the complex chain for comparing and regulating the impacts of emissions

More than two decades ago the DLR Institute of Atmospheric Physics (IPA) began research on the global atmospheric impact of aviation (Schumann 1990) via internal, national and international research projects, many of them initiated and co-ordinated by IPA, e.g., "*Schadstoffe in der Luftfahrt*" (BMBF)[1], AERONOX (EC), and POLINAT (EU). IPA significantly contributed to several *assessment* reports, including the IPCC (1999) Special Report "Aviation and the Global Atmosphere". Later the institute extended its research to other modes of transport, and initiated and co-ordinated the Helmholtz-University Young Investigators Group SeaKLIM (Impact of Ship Emissions on Atmosphere and Climate) and the European Integrated Project QUANTIFY (Quantifying the Climate Impact of Global and European Transport Systems). This led to the European Assessment of Transport Impacts on Climate Change and Ozone Depletion (ATTICA; Sausen 2010), which assessed global climate impacts of all modes of transport in a consistent way. Such an assessment enables a comparison of individual contributions and therefore a fair judgment of the impact of changes in the mode of transport or the potential of *mitigation* strategies. In parallel, aviation research projects at DLR continued investigating aviation particle effects (*Partikel und Zirrus*, PAZI) and prepared for a climate compatible air transport system (CATS).

[1] German national research program "Pollutants from Aviation".

43.2 Transport Emissions

In 2000, global *road transport* emitted more than 4 000 Tg CO_2, corresponding to 14 % of all anthropogenic CO_2 emissions (Fig. 43.2, Eyring et al. (2005)). CO_2 emissions from *shipping* and *aviation* contributed around 2.7 and 2.2 %, respectively (Fig. 43.2). In the same year, *transport* contributed nearly 30 % to the anthropogenic emissions of CO_2-equivalent[2] in EU-15 (htttp://www.unfccc.int/). The fractions of land transport, aviation and shipping were 21.6, 3.6, and 3.4 %, respectively, according to UNFCCC data. If the EU15 fulfill their 2020 target of 20 % reduction relative to 1990, and if transport CO_2 emissions increase at rates similar to those during the last two decades, the fractions will reach 30.9, 7.4, and 6.9 %, respectively, adding up to a total of about 45 %.

Transport also significantly contributes to *non-CO_2 emissions* not regulated in the Kyoto Protocol: e.g., in 2000 road transport, aviation and shipping contributed about 19, 1.6 and 15 %, respectively, to anthropogenic NO_x emissions (Fig. 43.2). The corresponding fractions for SO_2 are 2.9, 0.1 and 8 %. While CO_2 emissions from aviation and shipping were of similar magnitude and each only about one-fifth of the road transport emissions, the NO_x emissions of shipping were approximately as large as the emissions from road transport and one order of magnitude larger than the aviation emissions. Concerning SO_2, shipping dominates as a consequence of the high sulfur content of marine fuel oil. The particle emissions (PM_{10} in Fig 43.2) of road transport and shipping are of similar magnitude, while those of aviation are several orders of magnitude lower.

Transport emissions are geographically inhomogeneous. As an example, Fig. 43.3 shows the *source* strengths of NO_x from three transport modes in comparison to the nontraffic sector. Globally, road NO_x emissions are dominated by the eastern USA and Western Europe, as well as by India and eastern China. Ship traffic over the northern central Atlantic, along the east coast of Asia and along routes from the Mediterranean Sea to Southeast Asia is the most significant source of NO_x over large areas. Outside the continental eastern USA and Western Europe, large emissions from *air traffic* occur over the North Atlantic, though further north than the shipping maximum. In the southern hemisphere, NO_x emissions are largely dominated by nontraffic sources, in particular by *biomass burning* (Hoor et al. 2009).

While most emissions occur close to the surface, aviation is the only significant source of pollutants emitted at high altitude. As shown below, aviation emissions of *short-lived species* have a much higher relative *climate impact* due to the altitude of the emissions.

[2] Here CO_2-equivalent is the sum of all emissions of gases regulated in the Kyoto Protocol (in particular CO_2, CH_4) weighted with their 100-year GWP (Global Warming Potential).

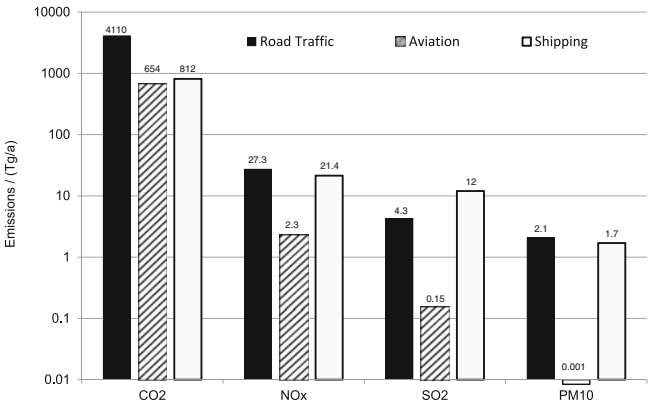

Fig. 43.2 Transport-related annual mass emission rates of CO_2, NO_x (as NO_2), SO_2 and PM_{10} estimated for the year 2000 (modified from Eyring et al. 2005). Note the logarithmic scale of the y-axis

43.3 Impact on Atmospheric Ozone

Pollutant species emitted from vehicles are initially spatially confined. When the species become diluted in the atmosphere they undergo chemical reactions and physical transformations. Emissions of NO_x can lead to the formation of ozone. This is mainly triggered by the oxidation cycles of carbon monoxide (CO) and volatile organic compounds, which are either released with NO_x or are present in the background. The ozone production rate due to these mechanisms shows a strong nonlinear dependence on the NO_x concentration. At very low NO_x levels, ozone is even destroyed. Above a critical NO_x concentration, ozone is produced and its production rate increases with the increasing availability of NO_x. Above another critical NO_x level, ozone production decreases again. Both NO_x *thresholds* as well as the ozone production rate itself strongly depend on the concentrations of several other chemical compounds and additional parameters such as temperature and the intensity of *solar radiation*. Outside of highly polluted areas, ozone production mostly increases with the available amount of NO_x and, consequently, NO_x emissions lead to the formation of additional ozone.

Simulations performed in the QUANTIFY project with an *ensemble* of different global atmospheric chemistry models (Hoor et al. 2009) show that global transport emissions, on average, lead to an increase in ozone. The simulations further reveal strong seasonal variations in ozone changes, with the largest ozone increase during summer. Another important result is that the sensitivity of ozone formation per NO_x molecule emitted is highest for aircraft emissions because the *lifetime* of NO_x is longer in the *upper troposphere* than near the surface: In the case of road transport 0.33 ± 0.05 ozone molecules are produced per emitted NO_x molecule, while the corresponding numbers for shipping and aviation are 0.54 ± 0.07 and 1.63 ± 0.58, respectively.

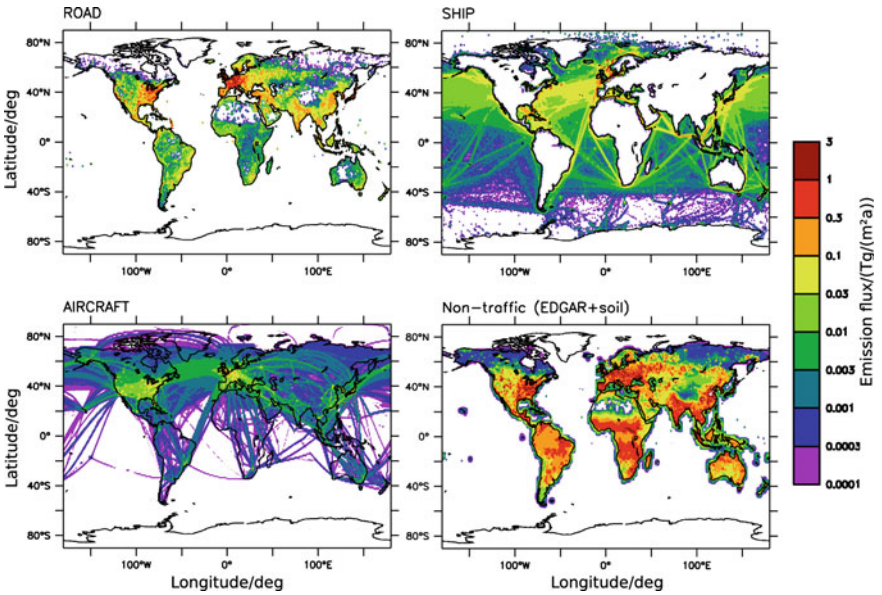

Fig. 43.3 Annual mean NO_x (as N) emission fluxes in Tg/(m^2 a) from road transport, aviation, shipping, and non-transport sectors such as industry, agriculture, biomass burning, etc. (from Hoor et al. 2009)

As an example of model results obtained in QUANTIFY, Fig. 43.4 shows the impacts of transport on the ozone concentration in July as *zonal* means. The ozone perturbation was calculated from the difference between a reference simulation considering all relevant emissions and sensitivity simulations where the emissions from road transport, aviation or shipping were reduced by 5 %. As expected, the impact of aviation is largest around the northern extra-tropical *tropopause*. The strongest (absolute) impact of road transport is not close to the surface, but at higher altitudes. The ozone precursors are transported to these regions with lower background NO_x concentrations mainly from low latitude regions. Shipping causes the strongest signal close to the surface between 30 and 50°N. As ships mostly emit in relatively clean air, their impact on ozone is larger than that of road transport, which emits in many cases into polluted air.

In addition to the formation of ozone, NO_x emissions lead to many further chemical impacts, in particular to a reduced lifetime of atmospheric methane. The latter effect also results in a reduced ozone production from the methane oxidation cycle (primary mode ozone) as well as a diminished production of stratospheric *water vapor*. While initially induced ozone changes have a lifetime in the order of few months (corresponding to the atmospheric lifetime of ozone), the primary mode ozone inherits the response time from the lifetime of methane. The effects via methane are differently effective for the different *transport modes*. Here shipping is most efficient (Hoor et al. 2009).

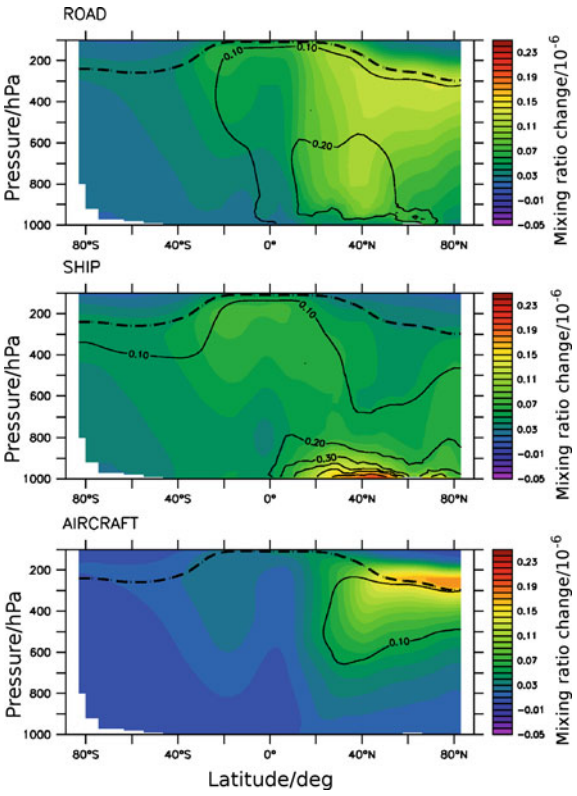

Fig. 43.4 Zonal mean perturbations of ozone volume mixing ratios (10^{-6}) in July for the different modes of transport, applying a 5 % emission reduction (difference between a reference case and a simulation with reduced emissions). *Solid contours* show the changes relative to the reference simulation. The *dashed lines* indicate the tropopause (from Fig. 6 of Hoor et al. 2009)

43.4 Impact on Aerosol

As explained in Sect. 43.2, the transport sector emits aerosol particles as well as particle precursor gases, such as SO_2. The directly released particles mainly contain black and organic carbon as well as small amounts of sulfate (SO_4). Aerosol particles can also be formed from gaseous emissions. Chemical reactions in the gas phase lead to the conversion of SO_2 into gaseous *sulfuric acid* (H_2SO_4), which either nucleates to very small new *sulfate* aerosol particles or condenses on pre-existing aerosol. Aerosol sulfate can also be generated from SO_2 by liquid phase chemistry within cloud droplets. Atmospheric aerosol particles can also take up significant amounts of *nitrate* (NO_3) generated by the chemical conversion of NO_x. Hence NO_x emissions from transport also affect *particulate matter*.

Model results reveal how transport emissions can perturb atmospheric aerosol. Global simulations performed within the DLR project VEU (*Verkehrsentwicklung und Umwelt*[3]) suggest that road traffic emissions cause a strong increase of the global

[3] Transport Development and Environment.

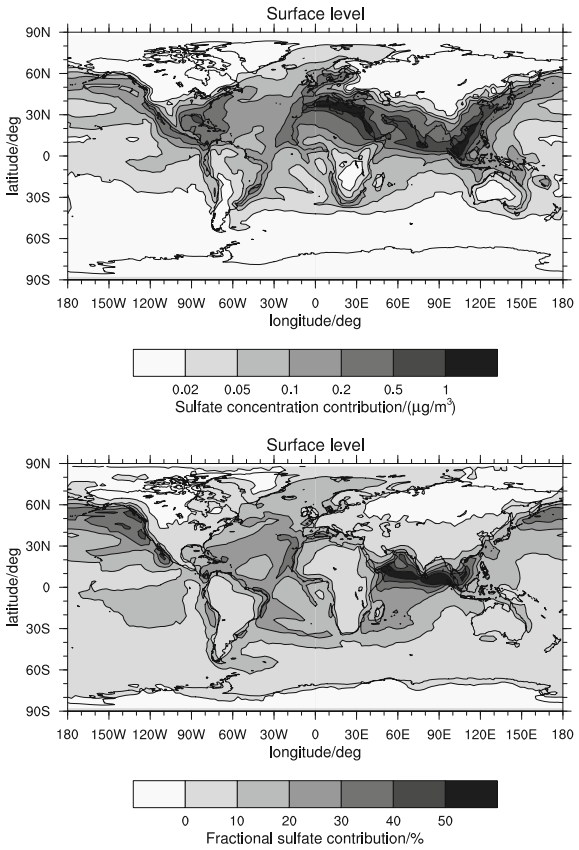

Fig. 43.5 Contribution of shipping (multiyear average) to sulfate aerosol concentration at the surface, as simulated by EMAC-MADE; *top* absolute contribution, *bottom* fraction of total sulfate aerosol concentration

black carbon burden, particularly over Europe and the eastern USA where black carbon from road traffic contributes about 50 % to the mean black carbon concentration at Earth's surface. The simulations also reveal that emissions from aviation can markedly increase the black carbon concentration in the upper troposphere. In the case of shipping, the SO_2 emissions are particularly relevant. Results from the SeaKLIM project indicate that SO_2 emissions from shipping cause important increases in the sulfate aerosol burden over the oceans and some coastal areas (Lauer et al. 2007). In areas of high ship traffic density, the sulfate aerosol (mass) concentration is increased by up to 1 $\mu g/m^3$. Over the oceans shipping contributes more than 50 % of the total near-surface aerosol sulfate (Righi et al. (2011), Fig. 43.5).

Particulate matter resulting from transport sector emissions has a multitude of potential impacts on climate. Particles lead to negative radiative forcing since solar radiation is scattered back into space. Soot absorbs radiation, which causes heating and thus leads to positive forcing. Besides these direct effects on radiation, indirect effects also occur. The latter are initiated by the involvement of aerosol particles in cloud microphysical processes, that is, the formation and further development of cloud particle populations.

43.5 Impact on Clouds

Clouds are of utmost importance in the *climate system* because of their interaction with the *hydrological cycle* and the radiative energy transfer. Since cloud formation requires the presence of aerosol particles that can act as *condensation* and *freezing nuclei*, cloud formation conditions can be altered by changes in aerosol composition, size and number concentration. Hence, transport-induced aerosol changes can affect cloud formation, cloud evolution, and cloud properties. Such processes are called *indirect effects*, because clouds are affected via the aerosol pathway.

But transport can also directly produce clouds in the form of (linear) *contrails* (Schumann 1996; Gierens 2010). There are two types of contrails: exhaust contrails, which form from water vapor and (mainly) soot emissions in the exhaust plume of an aircraft, and *aerodynamic contrails*, which form on ambient aerosol particles that allow condensation and freezing when the airflow over the wings becomes strongly cooled for a fraction of a second. The relative importance of the second type is not yet clear. Both types of contrails can exist for only a few minutes in *ice subsaturated* air. However, in sufficiently cold and humid air, *linear contrails* can continue to grow and eventually form *contrail cirrus* (e.g., Burkhardt and Kärcher (2011)) with *lifetimes* in the order of hours. Contrail cirrus is difficult to distinguish from natural cirrus clouds unless the history of evolution is known from observations (Vazquez-Navarro et al. 2010).

There are a number of indirect aerosol effects on clouds, but for cloud formation the most important one usually is the so-called *Twomey effect*: the addition of anthropogenic aerosol in regions where water clouds form leads to more numerous but smaller droplets. Since *scattering* of *radiation* becomes stronger with decreasing droplet size (for a constant water mass) the Twomey effect makes clouds more reflective for *solar radiation*. *Ship tracks* in *maritime stratus clouds* are a good example of this effect (Schreier et al. 2006). The smaller droplet size has further consequences, e.g., for *drizzle* formation and cloud lifetime.

The situation is more intricate in the case of *cirrus* clouds, which might be affected by particles from aviation (Hendricks et al. 2011). The addition of anthropogenic soot aerosol to the background air can lead to smaller *crystal number concentrations* (negative Twomey effect) in regions where cirrus clouds form predominantly via *homogeneous freezing* (typically in relatively clean air), while the number of ice crystals can be increased in regions of *heterogeneous freezing* (where industrial and traffic emissions enrich the background aerosol with ice-forming nuclei).

While direct cloud effects of traffic are easily observed, it is rather difficult to observe and monitor the more widespread indirect effects (an exception are ship tracks), particularly when emission of the particles and their effect on cloud formation are spatially and temporally separated. Therefore, current knowledge of contrails and ship tracks is to a large extent drawn from observations and measurements (Mannstein et al. 1999; Voigt et al. 2011), while global scale values

such as contrail *coverage* and associated radiative forcing are based on numerical simulations (Burkhardt and Kärcher 2011; Frömming et al. 2011; Schumann 2011). Indirect effects have mainly been studied by numerical simulation (Kärcher and Yu 2009; Hendricks et al. 2011).

43.6 Radiative Forcing and Other Metrics of Climate Change

As discussed above, transport emissions change the chemical composition of the atmosphere, both in the gas phase and the aerosol component, as well as cloud coverage and optical properties. These changes in turn affect the flow of radiative energy in the atmosphere and thus perturb the energy balance of the planet. The climate system responds to this imbalance by either cooling or warming, in an attempt to re-establish the planetary energy balance. In order to better understand and compare the climate effects of the various CO_2 and non-CO_2 contributions from different modes of transport, the contributions must be placed on a common scale. The *metric* (or weighting function) for evaluating what transport has so far contributed to climate change will be different, for example, from that for regulating future emissions in order to minimize climate impact (Fuglestvedt et al. 2010).

A rather frequently applied metric for comparing different contributions to climate change is *radiative forcing* (RF), which is a measure of the magnitude of the radiative imbalance resulting from a perturbation of the atmosphere. There is an approximately linear relationship between the magnitude of the RF and the expected equilibrium global-mean *surface temperature* change and this is what has motivated the use of RF as an indicator of the strength of different climate change mechanisms (IPCC 2007).

Within the QUANTIFY and ATTICA projects, and further studies, the RF caused by emissions from all modes of transport was estimated using a variety of models. Table 43.1 shows the mean RF values and associated uncertainties for the most relevant transport-related atmospheric perturbations.

The ratio of RF values from CO_2 from different transport sectors differs from the ratio of the respective CO_2 emissions (cf. Fig 43.1). This is the result of the longer emission periods for shipping and railways, which started many decades before road transport and, in particular, aviation. As CO_2 typically stays in the atmosphere for more than a century, CO_2 from shipping and railways has been accumulated in the atmosphere over a much longer period than has been the case for road transport and aviation.

The RF due to emissions of NO_x (because of the increase of short-lived ozone and the destruction of ambient methane) shows broad agreement among different models. Land transport and aviation lead to net positive forcing, while the shipping sector leads to negative forcing because here the effect of the loss of methane is greater than the increase in ozone. The ozone primary mode effect (with negative RF) is rather uncertain and therefore not included in Table 43.1.

Table 43.1 Radiative forcings and associated uncertainty measures (numbers in parentheses) in mW/m^2 for aviation, shipping and land transport for the year 2005 (2000 in the case of land transport) based on the ATTICA assessment (Eyring et al. 2010; Lee et al. 2010; Uherek et al. 2010) and on Burkhardt and Kärcher (2011). In the case of aviation, the 90 %-confidence intervals are provided as measures of uncertainty; in the case of shipping, the range of published data is shown, while the mean ± one standard deviation is used for road transport and railways. As a proxy for uncertainty of the RF from contrail cirrus, the range provided by Sausen et al. (2005) for aircraft-induced cloudiness is displayed

	Aviation	Shipping	Road transport	Railways
Carbon dioxide (CO_2)	28 (15, 41)	37 (28, 47)	150 (143, 167)	21 (15, 28)
Ozone (O_3)	26 (9, 82)	26 (10, 50)	54 (19, 91)	2 (*)
Methane (CH_4)	−13 (−76, −2)	−33 (−69, −14)	−12 (−29, −5)	−1 (*)
Water vapor	3 (0, 20)	*	*	*
Sulfate aerosol (direct)	−5 (−29, −1)	−31 (−58, −15)	−12 (−17, −7)	−7 (−10, −4)
Soot aerosol (direct)	3 (1, 21)	2 (1, 4)	23 (14, 32)	1 (*)
Linear contrails	12 (6, 26)	n.a.	n.a.	n.a.
Contrail cirrus, including linear contrails	38 (10, 80)	n.a.	n.a.	n.a.
Aerosol indirect effect	*	−409 (−737, −47)	*	*
CFC-12 (direct)	*	*	30 (21, 40)	*
CFC-12 (ozone depletion)	*	*	−10 (−12, −5)	*
Total	80	−408	223	16

n.a. not applicable
* No reliable estimate available

The RF from the direct effect of aerosols by *scattering* and *absorption* of *radiation* is particularly large for sulfate aerosol from shipping as a result of the high sulfur content of marine fuel oil. The black carbon effect is largest for road transport, mainly resulting from diesel engines on trucks and cars. The indirect aerosol effect of shipping is negative with about a factor two uncertainty. This negative RF from shipping is as large as 20–25 % of the RF from all anthropogenic CO_2. Current estimates of indirect aerosol effects from aviation and land transport are highly uncertain and not listed here. (In the case of aviation, even the sign of RF is unclear.)

While reliable estimates of linear contrail RF were already included in the assessments by Sausen et al. (2005) and Lee et al. (2009), the RF from contrail cirrus has recently been estimated by means of an atmosphere *general circulation model* (Burkhardt and Kärcher 2011). Their result (~ 38 W/m^2) is within the range of previous estimates for aviation-induced cirrus obtained from correlation analyses of aviation activity and cloudiness observed from *satellites*. Schumann et al. (2011) estimated 37 W/m^2 as a lower bound (Schumann personal communication)

Table 43.2 CO_2-equivalent emissions in Tg on the basis of GWP_{100} for aviation, shipping and land transport in the year 2005 (2000 in the case of land transport) based on the ATTICA assessment (Lee et al. 2010) and on Burkhardt and Kärcher (2011)

	Aviation	Shipping	Road transport	Railways
Carbon dioxide (CO_2)	641	660	4174	120
Nitrogen oxides (NO_x)	63[a]	−150	−98	−4
Water vapor	35	*	*	*
Sulfate aerosol (direct)	−7	−300	−75	−24
Soot aerosol (direct)	3	*	297	47
Linear contrails	135	n.a.	n.a.	n.a.
Contrail cirrus (including linear contrails)	465[b]	n.a.	n.a.	n.a.
Sulfate total (including sulfate aerosol direct)	*	−2 900	*	*
CFC-12 (direct)	*	*	652	*
Total (NO_x high)	1 200	−2390	4950	139

[a] Based on the NO_x high estimate of Lee et al. (2009) (giving an upper bound)
[b] Scaled from Table 43.1
n.a. not applicable
* No reliable estimate available

from a combination of modeling and data analysis. Contrail cirrus and natural cirrus compete for the same atmospheric water. If this *feedback* process is also considered, the *net forcing* from contrail cirrus is reduced to about 31 W/m^2 in the simulation by Burkhardt and Kärcher (2011).

In the case of road transport, RF from the emission of *chlorofluorocarbon* CFC-12 (CCl_2F_2) is a significant contribution. CFC-12 is a gas used as a refrigerant in the air conditioning systems of cars. It is a *greenhouse gas* and also an *ozone-depleting substance*. Hence, indirect cooling results from the depletion of *stratospheric ozone*.

As a backward looking *metric*, *radiative forcing* is suitable for determining what transport has so far contributed to *climate change*. However, RF is not an appropriate metric for estimating the relative importance of potential future impacts from the various transport effects, and hence not suitable for regulating transport in order to reduce its climate impact. Some other metrics can serve for this purpose. As an example, CO_2-equivalent emissions on the basis of the 100-year *global warming potential* (GWP_{100}), which is the metric of the *Kyoto Protocol* (for details see Fuglestvedt et al. (2010)), are considered. This is a forward looking metric. Table 43.2 shows the values for different contributions from transport.

Note that the definition of *CO_2-equivalent emissions* on the basis of GWP_{100} has a shortcoming that can lead to misinterpretations, such as in the case of NO_x, where pulse emissions first cause a warming due to rapid ozone production and then a cooling due to long-lived methane and associated ozone destruction (the *ozone primary mode* effect). A GWP-based climate metric may result in negative

CO_2-equivalent emissions for *pulses*, while in the long term the temperature would rise in the case of continuous (*sustained*) emissions or in the case of an emission *scenario* (Fuglestvedt et al. 2009). A potentially better metric might be the *average temperature response* (Dahlmann 2011).

43.7 Conclusions and a Perspective for the Future

The results of QUANTIFY and other projects on the impact of emissions from the transport sector on global atmosphere and climate formed the core of the ATTICA assessment reports, as published in a special issue of Atmospheric Environment (Sausen 2010). The main findings are:

- The impact on climate of transport, in particular aviation and shipping, presently grows faster than the impact from other sectors of human activity.
- The non-CO_2 effects of aviation and shipping (e.g., NO_x effects or induced clouds) are particularly large in comparison to other modes of transport.
- Aircraft-induced clouds warm the atmosphere in the global mean. The associated RF is of a magnitude similar to RF from aviation CO_2.
- Ship-induced cloud changes cool the atmosphere. The associated RF may be as large as one-third of the total anthropogenic RF.
- The total RF from road transport and aviation is positive, while that from shipping is negative.

Most of the quantitative results have substantial uncertainty; for example, the climate impact of indirect aerosol effects on clouds. It is not yet known whether aviation-induced soot has a significant effect on natural cirrus clouds. The net effect of aircraft NO_x emissions via *ozone* and *methane* also deserves further investigation. Furthermore, many of the emitted species responsible for the non-CO_2 effects also impact air quality, e.g., soot or sulfur emissions.

By applying a hierarchy of models from *process models* (Kärcher et al. 2007) and *cloud-resolving models* (Unterstrasser and Gierens 2010) to comprehensive *climate-chemistry models* (Dahlmann et al. 2011) and by analyzing measurements from *satellites* (Mannstein and Schumann 2005) and dedicated *field campaigns* (Jurkat et al. 2011), the main *uncertainties* with respect to *short-lived non-CO_2 effects* on the atmospheric *composition* and on *climate* could be reduced:

- The importance and magnitude of *trade-offs* in the NO_x effect (short-term increase in ozone versus longer-term decrease in methane and the associated ozone decrease) should be explicitly simulated by comprehensive models covering the long response time of methane.
- The climate impact of direct effects on clouds (in particular contrail cirrus) should be simulated by climate models including *parameterizations* of contrail cirrus and evaluated, e.g., by *tracking algorithms* identifying these clouds in satellite data.

- The climate impact of aerosols on clouds (e.g., soot effects on background cirrus, aerosol-induced effects on low level clouds) should be simulated by now available climate-chemistry models including aerosol-cloud interactions; the necessary parameters should be obtained from dedicated airborne measurement campaigns and laboratory studies.
- The regional response pattern in atmospheric composition and climate resulting from transport emissions needs to be determined consistently in a multimodel approach.
- The effects from *biofuels* replacing *fossil fuel* should be considered.
- Finally, trade-offs between climate and the air quality effects of emissions from different sectors, as well as traffic *noise*, should be examined as a precondition for robust policy formulation.

References

Burkhardt, U., Kärcher, B.: Global radiative forcing from contrail cirrus. Natl. Clim. Change. **1**, 54–58 (2011). doi:10.1038/NCLIMATE1068

Dahlmann, K.: Bewertung neuer Flugzeugtechnologien hinsichtlich ihrer Klimawirkung. Dissertation, Ludwig-Maximilians-Universität München (2011)

Dahlmann, K., Grewe, V., Ponater, M., Matthes, S.: Quantifying the contributions of individual NO_x sources to the trend in ozone radiative forcing. Atmos. Environ. **45**, 2860–2868 (2011b). doi:10.1016/j.atmosenv.2011.02.071

Eyring, V., Köhler, H.W., van Aardenne, J., Lauer, A.: Emissions from international shipping: 1. The last 50 years. J. Geophys. Res. **110**, D17305 (2005). doi:10.1029/2004JD005619

Eyring, V., Isaksen, I.S.A., Berntsen, T., Collins, W.J., Corbett, J.J., Endresen, O., Grainger, R.G., Moldanova, J., Schlager, H., Stevenson, D.S.: Transport impacts on atmosphere and climate: Shipping. Atmos. Environ. **44**, 4735–4771 (2010). doi:10.1016/j.atmosenv.2009.04.059

Frömming, C., Ponater, M., Burkhardt, U., Stenke, A., Pechtl, S., Sausen, R.: Sensitivity of contrail coverage and contrail radiative forcing to selected key parameters. Atmos. Environ. **45**, 1483–1490 (2011). doi:10.1016/j.atmosenv.2010.11.033

Fuglestvedt, J., Berntsen, T., Eyring, V., Isaksen, I., Lee, D.S., Sausen, R.: Shipping emissions: From cooling to warming of climate. Environ. Sci. Technol. **43**, 9057–9062 (2009)

Fuglestvedt, J.S., Shine, K.P., Berntsen, T., Cook, J., Lee, D.S., Stenke, A., Skeie, R.B., Velders, G.J.M., Waitz, I.A.: Transport impacts on atmosphere and climate: Metrics. Atmos. Environ. **44**, 4648–4677 (2010). doi:10.1016/j.atmosenv.2009.04.044

Gierens, K.: Contrails and Contrail Cirrus. Encyclopedia of Aerospace Engineering, pp. 3683–3694. Wiley, Chichester (2010)

Hendricks, J., Kärcher, B., Lohmann, U.: Effects of ice nuclei on cirrus clouds in a global climate model. J. Geophys. Res. **116**, D18206 (2011). doi:10.1029/2010JD015302

Hoor, P., Borken-Kleefeld, J., Caro, D., Dessens, O., Endresen, O., Gauss, M., Grewe, V., Hauglustaine, D., Isaksen, I.S.A., Jöckel, P., et al.: The impact of traffic emissions on atmospheric ozone and OH: Results from QUANTIFY. Atmos. Chem. Phys. **9**, 3113–3136 (2009). doi:10.5194/acp-9-3113-2009

IPCC: Aviation and the Global Atmosphere, p. 373. Cambridge University Press, Cambridge (1999)

IPCC: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, p. 996. Cambridge University Press, Cambridge (2007)

Jurkat, T., Voigt, C., Arnold, F., Schlager, H., Kleffmann, J., Aufmhoff, H., Schäuble, D., Schaefer, M., Schumann, U.: Measurements of HONO, NO, NO_y and SO_2 in aircraft exhaust plumes at cruise. Geophys. Res. Lett. **38**, L10807 (2011). doi:10.1029/2011GL046884

Kärcher, B., Möhler, O., DeMott, P.J., Pechtl, S., Yu, F.: Insights into the role of soot aerosols in cirrus cloud formation. Atmos. Chem. Phys. **7**, 4203–4227 (2007)

Kärcher, B., Yu, F.: Role of aircraft soot emissions in contrail formation. Geophys. Res. Lett. **36**, L01804 (2009). doi:10.1029/2008GL036649

Lauer, A., Eyring, V., Hendricks, J., Jöckel, P., Lohmann, U.: Global model simulations of the impact of ocean-going ships on aerosols, clouds, and the radiation budget. Atmos. Chem. Phys. **7**, 5061–5079 (2007). doi:10.5194/acp-7-5061-2007

Lee, D.S., Fahey, D.W., Forster, P.M., Newton, P.J., Wit, R.C.N., Lim, L.L., Owen, B., Sausen, R.: Aviation and global climate change in the 21st century. Atmos. Environ. **43**, 3520–3537 (2009). doi:10.1016/j.atmosenv.2009.04.024

Lee, D.S., Pitari, G., Grewe, V., Gierens, K., Penner, J.E., Petzold, A., Prather, M.J., Schumann, U., Bais, A., Berntsen, T., et al.: Transport impacts on atmosphere and climate: Aviation. Atmos. Environ. **44**, 4678–4734 (2010). doi:10.1016/j.atmosenv.2009.06.005

Mannstein, H., Meyer, R., Wendling, P.: Operational detection of contrails from NOAA-AVHRR data. Int. J. Remote Sens. **20**, 1641–1660 (1999)

Mannstein, H., Schumann, U.: Aircraft induced contrail cirrus over Europe. Meteorol. Z. **14**, 549–554 (2005). doi:10.1127/0941-2948/2005/0058

Righi, M., Klinger, C., Eyring, V., Hendricks, J., Lauer, A., Petzold, A.: Climate impact of biofuels in shipping: Global model studies of the aerosol indirect effect. Environ. Sci. Technol. **45**, 3519–3525 (2011). doi:dx.doi.org/10.1021/es1036157

Sausen, R. (ed.): Transport impacts on atmosphere and climate: The ATTICA assessment report. Special Issue Atmos. Environ. 44, 4645–4816 (2010)

Sausen, R., Isaksen, I., Hauglustaine, D., Grewe, V., Lee, D.S., Myhre, G., Köhler, M.O., Pitari, G., Schumann, U., Stordal, F., et al.: Aviation radiative forcing in 2000: An update on IPCC (1999). Meteorol. Z. **14**, 555–561 (2005). doi:10.1127/0941-2948/2005/0049

Schreier, M., Kokhanovsky, A.A., Eyring, V., Bugliaro, L., Mannstein, H., Mayer, B., Bovensmann, H., Burrows, J.P.: Impact of ship emissions on the microphysical, optical and radiative properties of marine stratus: A case study. Atmos. Chem. Phys. **6**, 4925–4942 (2006). doi:10.5194/acp-6-4925-2006

Schumann, U. (ed.): Air Traffic and the Environment—Background, Tendencies and Potential Global Atmospheric. Lecture Notes in Engineering, p. 170. Springer, Berlin (1990)

Schumann, U.: On conditions for contrail formation from aircraft exhausts. Meteorologische Zeitschrift **5**, 4–23 (1996)

Schumann, U.: A contrail cirrus prediction model. Geosci. Model Dev. Discuss. **4**, 3185–3293 (2011). doi:10.5194/gmdd-4-3185-2011

Schumann, U., Graf, K., Mannstein, H.: Potential to reduce the climate impact of aviation by flight level changes. In: 3rd AIAA Atmospheric and Space Environments Conference, AIAA paper 2011–3376, 1–22 (2011)

Uherek, E., Halenka, T., Borken-Kleefeld, J., Balkanski, Y., Berntsen, T., Borrego, C., Gauss, M., Hoor, P., Juda-Rezler, K., Lelieveld, J., et al.: Transport impacts on atmosphere and climate: Land transport. Atmos. Environ. **44**, 4772–4816 (2010). doi:10.1016/j.atmosenv.2010.01.002

Unterstrasser, S., Gierens, K.: Numerical simulations of contrail-to-cirrus transition—Part 1: An extensive parametric study. Atmos. Chem. Phys. **10**, 2017–2036 (2010). doi:10.5194/acp-10-2017-2010

Vazquez-Navarro, M., Mannstein, H., Mayer, B.: An automatic contrail tracking algorithm. Atmos. Meas. Tech. **3**, 1089–1101 (2010). doi:10.5194/amt-3-1089-2010

Voigt, C., Schumann, U., Jessberger, P., Jurkat, T., Petzold, A., Gayet, J.-F., Krämer, M., Thornberry, T., Fahey, D.W.: Extinction and optical depth of contrails. Geophys. Res. Lett. **38**, L11806 (2011). doi:10.1029/2011GL047189

Climate Optimized Air Transport

44

Sigrun Matthes, Ulrich Schumann, Volker Grewe, Christine Frömming, Katrin Dahlmann, Alexander Koch and Hermann Mannstein

Abstract

Aviation climate impact is caused by CO_2 and non-CO_2 emissions where the climate effect of non-CO_2 emissions depends on weather and aircraft route. An aviation system with minimum climate impact differs from a system with minimum emissions. Considerable potential exists to reduce the climate impact of aviation by weather- and cost-dependent climate-optimized air traffic management ("smart routing") and aircraft design ("green aircraft"). Current research provides a unique opportunity to systematically investigate the

S. Matthes (✉) · U. Schumann · V. Grewe ·
C. Frömming · K. Dahlmann · H. Mannstein
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Sigrun.Matthes@dlr.de

U. Schumann
e-mail: Ulrich.Schumann@dlr.de

V. Grewe
e-mail: Volker.Grewe@dlr.de

C. Frömming
e-mail: Christine.Froemming@dlr.de

K. Dahlmann
e-mail: Katrin.Dahlmann@dlr.de

H. Mannstein
e-mail: Hermann.Mannstein@dlr.de

A. Koch
DLR, Air Transportation Systems, Blohmstr 18,
21079 Hamburg, Germany
e-mail: Alexander.Koch@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_44,

trade-offs between various mitigation concepts and cost functions. Here various approaches are presented to minimize the climate impact on a climatological and weather basis, some being applicable to aircraft designs for reduced climate impact and others offering alternative operational concepts.

44.1 Introduction

Aviation is one of the *transport* sectors with currently moderate *climate* impact. However, *air traffic* has been expanding faster than the global economy for many decades and this *trend* is likely to continue for a long period in the future. Hence, the *climate impact* of aviation is increasing both in absolute and relative terms. Rising concern about potential damage in light of *climate change* has led to joint efforts to satisfy the growing demand for *mobility* in the most efficient and climate-*sustainable* manner (Green 2005). The development of *mitigation* methods for this purpose is in line with aviation visions and research programs, such as the ACARE "Vision 2020" of 2001, the EU Vision 'Flight Path 2050' of 2011 (HLGAR 2011), the European aeronautics projects Clean Sky and *SESAR*, and the U.S. Next Generation strategy (Waitz et al. 2004).

The climate impact of aviation results from CO_2 and *non-CO_2 emissions* (IPCC 1999). CO_2 is of concern mainly because it is an important *anthropogenic greenhouse gas* with long *lifetimes* in the *atmosphere* and because of its considerable contribution to *radiative forcing* (*RF*). The non-CO_2 emissions (*nitrogen oxides*, *water vapor*, *aerosols*, etc.) have shorter lifetimes but contribute a large share to aviation climate impact. They have a higher climate impact when emitted at cruise than at ground levels (Lee et al. 2009, 2010). The relative importance of CO_2 and non-CO_2 depends strongly on the time horizon for evaluation of climate impacts and *scenarios*, e.g., future air traffic development. The non-CO_2 effects are more important for short time horizons than for long horizons (Sausen and Schumann 2000; Grewe and Stenke 2008; Fuglestvedt et al. 2010).

Strategies for minimizing the climate impact of air traffic include identifying the most efficient options for airframe and propulsion technology, *air traffic management*, and alternative route network concepts. Economic measures and market-based incentives may also contribute, but these are out of the scope of this chapter. Depending on the objectives and mitigation options, various strategies are followed by aviation system stakeholders, such as *aircraft* manufacturers (e.g., Airbus), airline operators (e.g., Lufthansa), and air traffic management (ATM) system operators (e.g., EUROCONTROL). DLR uses its system expertise to investigate mitigation concepts. The role of the DLR Institute of Atmospheric Physics is to quantify climate impact, evaluate mitigation options and provide an interface between technology and climate impact.

Various approaches have been investigated in the past: generally flying higher or lower (Sausen et al. 1998; Grewe et al. 2002; Williams et al. 2002; Fichter et al. 2005; Gauss et al. 2006; Köhler et al. 2008; Rädel and Shine 2008), flying around *contrail* forming regions vertically or laterally (Mannstein et al. 2005; Gierens et al. 2008; Campbell et al. 2009; Sridhar et al. 2011), climate optimized aircraft design (Koch et al. 2011; Schwartz Dallara et al. 2011), and *weather*-dependent *route optimization* for minimum contrail climate impact (Schumann et al. 2011) and overall climate impact (Matthes 2011). These studies indicate significant *mitigation potential* for reducing climate impact by modifying the *air transport system*. This chapter describes the principles of route optimization, mitigation options, and specific mitigation concepts under development.

44.2 Climate Optimized Routing: Principles

Optimization often addresses a single objective that can be represented by a scalar *cost function* which is to be minimized. A single-objective approach has been discussed earlier (Sausen et al. 1994). When only considering the climate impact I_C, one can use a physical climate *metric* such as the *average temperature response* (*ATR*), representing the mean global *surface temperature* change for a given emission scenario. Alternatives such as the *global warming potential GWP* may be useful as well. However, the costs of climate change are difficult to assess. Monetization of climate change is uncertain and requires socioeconomic and political assumptions which are out of the scope of *atmospheric* research. Moreover, *trade-offs* in complex problems are difficult to understand with a single-objective approach. Alternatively, one can use a multiobjective approach considering a vector $\boldsymbol{I}$ of objective functions with separate components, e.g., for climate impact (I_C), for direct operating costs *DOC* (I_{DOC}) or *fuel consumption* (I_{fuel}), without specifying monetizing factors. Optimal solutions with identified trade-offs in complex problems may be represented in the multiobjective approach by the associated climate impact (e.g. in K) and costs for priced goods (in €) on a so-called *Pareto front* (see Sect. 44.4.2).

Formally, climate impact may be measured by

$$I_C = \int_{flight} \int_{time} R(\vec{s}(t), t, e(t))\, dt, \tag{44.1}$$

where a time-dependent function R quantifies, e.g., the *ATR* of an emission rate $e(t)$ at time t during flight along an aircraft trajectory $\vec{s}(t) = (s_x(t), s_y(t), s_z(t))^T$. In principle, climate impact may be computed using a full *climate model*. However, for practical optimization purposes one uses climate response functions describing an approximated functional relationship between the emission and climate change (Sausen and Schumann 2000; Grewe and Stenke 2008; Fuglestvedt et al. 2010;

Schwartz Dallara et al. 2011). The function R may be approximated by a Taylor series to the first order:

$$R(x, y, z, t, e) \approx \frac{\partial}{\partial e} R(x, y, z, t, 0) \cdot e(t) = r(x, y, z, t) \cdot e(t) \tag{44.2}$$

When using ATR, R has the unit of K, and the specific impact function r has the unit K/kg. Different components of total climate impact (e.g., CO_2, ozone, contrails) are represented as vector $\vec{r}$. The respective emissions are represented as vector $\vec{e}$.

$$I_C = \int_{flight\ \ time} \vec{r}(s_x, s_y, s_z, t) \cdot \vec{e}(t)\, dt \tag{44.3}$$

For given weights w and given monetization factors (m, in units of €/K providing $f_C = m_C \cdot I_C$), a single objective (f) can be defined, e.g., for simultaneous optimization of climate and DOC,

$$f = w_C m_C I_C + w_{DOC} I_{DOC}, \tag{44.4}$$

with positive normalized weighting factors w.

In this paper various mitigation concepts are described which differ in using either single- or multi-objective approaches, different specific impact functions r, and different emission functions $e(t)$, as described below.

44.3 Mitigation Options

Minimizing the climate impact of aviation would require addressing all climate impact components. In the following, emissions of CO_2, the climate impact of NOx and water vapor emissions, and contrail climate impact are considered.

44.3.1 Minimizing CO_2 Emissions

Minimum fuel consumption is of primary interest for the aviation industry because it reduces costs. However, fuel is not the only cost driver and various constraints cause fuel penalties. Although fuel reduction below the current state is challenging, further reduction of fuel consumption and hence of fossil CO_2 climate impact is nevertheless feasible.

Fossil CO_2 emissions can be reduced by increasing the fuel efficiency of aircraft. For a given aircraft and mission, fuel consumption can be minimized by optimal routing, altitudes and flight speeds. The fuel burn of an aircraft is inversely proportional to the so-called specific air range (SAR). The SAR is the distance

flown per unit of fuel mass. The SAR is directly proportional to the aircraft lift-to-drag ratio, and speed, and inversely proportional to the thrust specific fuel consumption and weight. There is an optimal altitude for each speed and weight combination at which the *SAR* is maximized. During cruise flight with given speed, the fuel burn can be minimized by increasing flight altitude according to weight reduction by fuel burn. Such a cruise flight procedure is denoted as a continuous climb cruise. In practice, continuous climb cruise is seldom possible with present ATM. Instead, regulatory *safety* and *capacity* constraints require step climbs performed at discrete *flight levels*. A future ATM might enable increased fuel efficiency by a more flexible system, retaining safety and efficiency. Concepts like two-step operations on long-haul flights might be considered in the future to further reduce fuel consumption at the expense of longer flight times (Green 2005; Linke et al. 2011). Aircraft design, optimized for lower flight altitudes and speeds, could also reduce fuel burn. New aircraft concepts such as laminar wings and blended wing bodies are being investigated to reduce fuel consumption. *Kerosene* may be replaced by new *nonfossil renewable* fuels to make aviation sustainable despite climate change and reduced petrol oil resources. The potential climate impact of *hydrogen*-powered aircraft has been investigated in an EU project CRYOPLANE (1996–2003) (Klug et al. 1996). The drawbacks of more frequent contrail formation and larger contrail particles are balanced by radiatively less effective and quickly sedimenting particles (Schumann 1996; Marquart et al. 2001; Ström and Gierens 2002; Ponater et al. 2006). Aircraft fuelled with *liquid hydrogen* need to carry a larger fuel volume, and liquid hydrogen would have to be produced without fossil carbon emissions. It is expected to take a long time before such concepts become effective in practice. Nevertheless, in view of the urgency of reducing fossil fuel consumption, such options may be considered.

Fuel can also be saved by using optimized routes. Besides non-optimal routing in terms of fuel burn caused by ATM constraints due to the air space structure or congestions, improved use of *weather forecasts* could reduce fuel consumption. Routes may be optimized by varying the selected track, flight altitude, speed, and departure time. The shortest connection between departure and destination *airports* follows a great circle. Because of head or tailwinds the great circle route is not necessarily the one with the shortest flight time or lowest fuel consumption. Weather-dependent route optimization with minimum head *wind* and maximum tail wind is common practice in aeronautics. However, present routing is not optimal. For example, the ocean track system for air navigation over the North Atlantic is calculated daily, and thus depends on daily mean weather forecast, and fixed vertical and horizontal flight separation distances because of traditional safety and capacity constraints. Improved accuracy of wind, *temperature* and weather prediction in general may contribute to further improvements and mitigation, in particular for long-distance flights. The fuel consumption of today's fleet can be reduced by several percent by overcoming such constraints. In addition, fuel consumption can be reduced by further improving weather prediction and weather dependent flight routing (Klima 2005).

44.3.2 Minimizing Aviation NO_x Climate Impact

The non-CO_2 aviation emissions have a large potential for quick climate mitigation because of their shorter lifetime and their large immediate radiative impact (Sausen and Schumann 2000; Grewe and Stenke 2008). Nitrogen oxides (NOx, i.e., NO and NO_2) are one of the major non-CO_2 emissions (Brasseur et al. 1998). NOx emissions in the *troposphere* and *lower stratosphere* contribute to *ozone* (O_3) formation and *methane* (CH_4) reduction. Both are important greenhouse gases. On average, the O_3 impact of aviation NOx is expected to be stronger than the impact on CH_4, which increases the *greenhouse effect*, though the precise amounts are uncertain (Lee et al. 2010; Holmes et al. 2011).

The amount of NOx emissions depends on fuel consumption and the *engine*'s type-specific *emission index*. The emission index for NOx depends on the engine and combustor architecture, power setting, flight speed, ambient *pressure*, temperature and *humidity* (Döpelheuer and Lecht 1999). This dependence has to be taken into account when considering changes in aircraft design and operations. A reduction of NOx mass emissions per seat kilometer has been achieved in the past and further reductions are among the major technological goals of the aviation industry. However, over the last decades, the total mass of NOx emissions of the global fleet increased more strongly than fuel consumption (Lee et al. 2010).

NOx enhances O_3 formation by photochemical processes involving *hydroxyl radicals* (OH) and the photolysis of nitrogen dioxide (NO_2) (Brasseur et al. 1998; Grooß et al. 1998). Enhanced OH reduces the lifetime of CH_4. Ozone *chemistry* depends strongly on the background concentration of many trace gases, aerosols, *clouds*, temperatures and *radiation*. The latter depends on altitude, solar zenith angle, and the *Earth*-atmosphere *albedo*. The amount of O_3 formed depends strongly on the lifetime of NOx. Typical lifetimes are days to weeks in the *upper troposphere*. The mean NOx lifetime increases with altitude and is larger in the lower stratosphere than in the troposphere. The amount of CH_4 and O_3 mass changes per unit NOx mass emission is dependent on altitude and latitude (Grewe and Stenke 2008; Fichter 2009). On average, higher cruise altitudes tend to yield increased O_3 perturbations, whereas lower cruise altitudes result in reduced O_3, according to model studies (Grewe et al. 2002; Gauss et al. 2006). The ratio between ozone *RF* and negative methane *RF* increases with altitude up to roughly 10–12 km and decreases at higher altitudes (Grewe and Stenke 2008). In addition, NOx induced changes are weather dependent and differ regionally from the climatological annual mean. Thus, routes with minimal NOx-induced impact depend on the daily weather situation.

44.3.3 Minimizing Aviation Water Vapor Climate Impact

The climate impact of water vapor emissions without contrail formation is relatively small for *subsonic* aviation. The relative impact increases with altitude because of longer lifetimes and lower background concentrations at higher

altitudes in the *stratosphere*, and would be more important for *supersonic aircraft*; water vapor would also become more important when using hydrogen-powered aircraft (Schumann 1994; Gauss et al. 2003; Grewe and Stenke 2008). The total route time in the stratosphere can be used as an indicator for water vapor climate impact (Sausen et al. 1994; Irvine et al. 2012).

44.3.4 Minimizing Contrail Climate Impact

Water vapor emissions contribute to contrail formation. Contrails are special *cirrus* clouds (Schumann 2002) which contribute a large share to the non-CO_2 aviation climate impact. Thin cirrus clouds in general warm the Earth atmosphere system mainly by reducing *longwave radiation* to space. Contrails form when aircraft fly in sufficiently cold air, as determined by the *Schmidt-Appleman criterion* (SAC). Contrails are *persistent* and may form wide and thick cirrus in *ice supersaturated regions*, mainly in the upper troposphere (Burkhardt and Kärcher 2011). The lifetime of *contrail cirrus* is estimated to range from minutes to possibly a day (Haywood et al. 2009; Vazquez-Navarro 2009).

Contrails can be avoided by flying lower, in warmer atmospheric layers, at temperature above the *SAC threshold* temperature. However, the *SAC* threshold temperature also decreases with altitude. Therefore, large altitude changes are required to reach strong reduction in contrail formation by lowering flight levels (Sausen et al. 1998; Fichter et al. 2005). Alternatively, contrails can be avoided by flying higher, in particular at *mid-latitudes*, in the lower stratosphere, which often is far drier than the upper troposphere (Fichter et al. 2005; Schumann 2005). However, flying higher does not reduce contrail formation in the *tropics*. Here the *tropopause* is typically above 16 km altitude, far higher than the maximum cruise levels of subsonic aircraft. Therefore, generally flying higher or lower globally is certainly not the optimal mitigation strategy.

Instead, one may consider weather-dependent changes in cruise altitudes and lateral routing. Contrails can be avoided by avoiding flights in ice supersaturated regions (ISSR) in the atmosphere. As ISSR regions are often rather thin, in the order of 500 m (Spichtinger et al. 2003) and have horizontal dimensions which vary strongly but are often below 100 km (Gierens and Spichtinger 2000), changes of one flight level (1000 ft) up or down (Mannstein et al. 2005) or small lateral deviations (Sridhar et al. 2011) may be sufficient to avoid contrail formation.

Though contrails are expected to warm the atmosphere in the global mean, contrails may have a cooling effect during the day, in particular over cool and dark surfaces (such as the *ocean*), and at high solar zenith angles (such as during morning and evening) (Meerkötter et al. 1999; Schumann et al. 2012). Hence, a minimum climate impact of aviation may also be achieved by enhancing the formation of such cooling contrails (Schumann et al. 2011). For an implementation of this strategy one has to know the evolution and radiative impact of the contrail over its full lifetime, e.g., whether a cooling contrail will eventually become a warming contrail.

The contrail climate impact is a complex function of aircraft, atmosphere, ground and solar parameters. It does not only depend on the presence or nonpresence of an ISSR. Therefore, avoiding ISSR regions does not necessarily provide minimum climate impact.

Aviation induced cloudiness also depend on the amount of *soot* particles, and possibly on *volatile* aerosol emitted or formed from aircraft emissions (Kärcher and Yu 2009). Hence, the climate impact of aviation may be mitigated at least partly by reducing soot emissions. This opens a technological option for mitigating aviation induced cloud changes. This aspect is uncertain however, because soot and other aviation induced aerosol may impact cirrus properties also without contrail formation and in various directions (Kärcher et al. 2007; Penner et al. 2009; Hendricks et al. 2011).

44.4 Mitigation Approaches

Different approaches for mitigation strategies can be distinguished in climate- or weather-based routing concepts. Further mitigation potential can be achieved by developing aircraft optimized for these alternative aircraft trajectories. The climate change mitigation potential will strongly depend on the flexibility of future ATM. The largest climate impact reduction is likely to be achieved when ATM allows individual aircraft to fly trajectories avoiding climate sensitive regions for all components: fossil fuel consumption, contrails, chemical impacts, water vapor, etc. This requires more frequent cruise altitude changes, putting more work load on ATM and hence potentially increased costs. Moreover, in cases of congested air space as in central Europe, such a strategy is limited. On the other hand, a general change in cruise altitude (and related fuel optimal speeds) for specific routes can be handled easier in comparison by ATM, but leads for actual aircraft to an increase in *DOC* and assumingly, because the comparison of mititgation potentials is not yet done (part of WECARE) has lower mitigation potential. It is an open question to what extent future ATM will be able to handle future air traffic which is optimized with respect to its climate impact on a weather basis. In any case, evaluating the best strategy requires reliable estimates of aviation climate impact.

Both weather-dependent and climatologically mean optimization of aircraft trajectories may result in the need for an optimized aircraft. Generally, aircraft design is based on one typical flight trajectory and not for a specific weather situation. Today, design trajectories are set for certain cruise altitudes with a few step climbs depending on the range of the aircraft. For weather-dependent optimization, this trajectory may include more frequent altitude changes to reflect re-routing caused by the avoidance of climate-sensitive regions, whereas for a climatologically mean optimization this trajectory will be characterized by a more general change in cruise altitude and speed.

44.4.1 Climatological Approach for Air Transport System Climate Optimization

The DLR research project CATS (Climate compatible Air Transport System) has the goal to identify the combined operational and technological potential to reduce the climate impact of the actual and future air transport system. Two research questions are addressed in CATS: first, how much can the climate impact be reduced with a world fleet of current *long-range* aircraft by flying at different altitudes and speeds and at which cost can this reduction be achieved? The answer to this question provides for every route in the analyzed network a new cruise altitude and speed that corresponds to a chosen trade-off between climate impact reduction and increased operating costs, e.g., X % climate impact reduction for Y % cost increase. Obviously the highest *mitigation potential* X/Y is obtained for highest reduction in climate impact and lowest cost increases. As actual aircraft are specifically designed for current cruise levels and speeds, they would be operated in nonoptimal conditions at lower altitudes and slower cruise speeds. In order to prevent the related performance losses under off-design conditions, actual aircraft must be redesigned for the new cruise conditions. This leads to the second research question in CATS: How do aircraft need to be designed for cruise conditions with reduced climate impact? The answer to this question shows the potential of novel aircraft configurations and technologies specifically designed to reduce climate impact.

In a collaborative effort, experts from the fields of air transport system analysis, aircraft and engine design, emissions, route network, air traffic management and atmospheric scientists developed an integrated *simulation* framework which allows *assessment* and optimization of novel operational and technological concepts with respect to overall mean climate impact and costs on a per route basis for a global route network. The CATS simulation workflow considers flight plans with city pairs, aircraft fleet *composition*, resulting aircraft movements, and trajectories to compute the changes in *ATR* and *DOC*. Within this model chain, the climate impact of aircraft emissions is calculated with the climate-chemistry *response model* AirClim (Grewe and Stenke 2008). AirClim quantifies the atmospheric response in terms of climate change of emissions dependent on altitude and latitude of emissions, while considering climatologically mean impact response functions $r(y,z)$ as a function of latitude y and altitude z, determined from detailed simulations with a three-dimensional state-of-the-art *chemistry-climate model*. The CATS simulation workflow was applied to systematically quantify the potential for climate impact reduction and related cost changes through reduced cruise flight altitudes and speeds. The world fleet of twin-*jet* long-range aircraft is assessed on a global route network. The simulated global route network comprises all routes served by the selected aircraft type in the year 2006, resulting in 1 178 city pairs worldwide. Both an aircraft of present design and a novel aircraft specifically designed for cruise conditions with reduced climate impact are being investigated.

Figure 44.1 shows the climate impact for a modelled aircraft similar to an Airbus A330, expressed as *ATR* per km, for flight conditions with minimum *DOC*

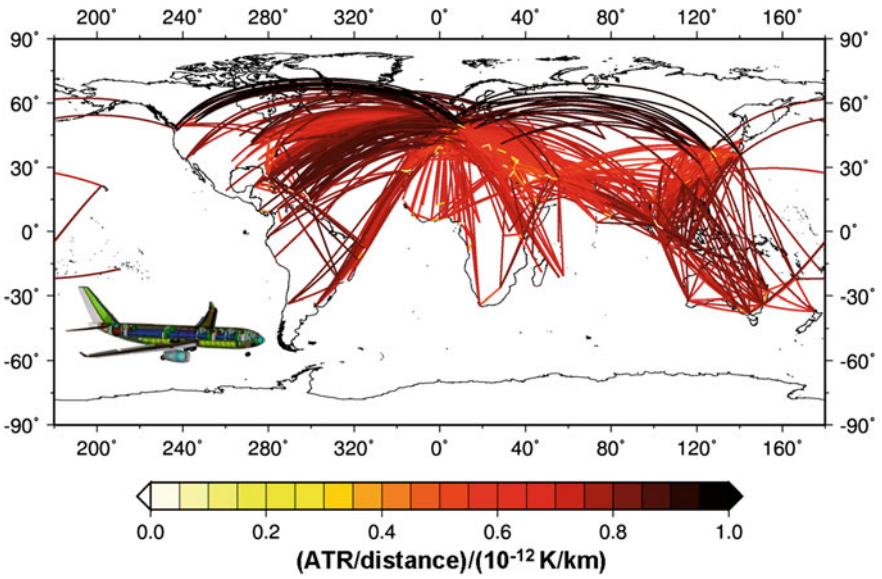

Fig. 44.1 Average temperature response (*ATR*, the global mean surface temperature response integrated over 100 years) per flown kilometer for selected routes for minimum direct operating costs (*DOC*), taking into account effects from contrail cirrus, NOx (on O_3 and CH_4), CO_2 and H_2O. Inlay: Aircraft model of an A330-200 as analyzed in CATS

on each network route. In this study uniform changes of cruise altitudes and aircraft speeds were considered for each route and for all weather situations throughout a year. Based on the AirClim calculations, short-range flights generally have a smaller climate impact per flown distance (yellow), since the aircraft fly at a lower altitude, resulting in lower impact of H_2O, O_3, and contrail cirrus. Individual long-range flights at mid-latitudes usually have a smaller impact from contrail cirrus than at higher latitudes, since the high air traffic density already leads to *saturation* effects.

Figure 44.2 shows the trade-off (the Pareto front) between *DOC* and *ATR* changes exemplarily for the route Detroit-Frankfurt. The changes in *ATR* and *DOC* for different cruise altitudes and speeds are evaluated relative to flight conditions with minimum *DOC*. The study shows a considerable potential to reduce the climate impact of today's aircraft with relatively small cost increases by reduced cruise altitudes and speeds (Koch et al. 2011). Roughly, 30 % reduction in *ATR* can be achieved with a 10 % increase in *DOC*. This analysis is presently repeated and reassessed including novel aircraft configurations (in progress).

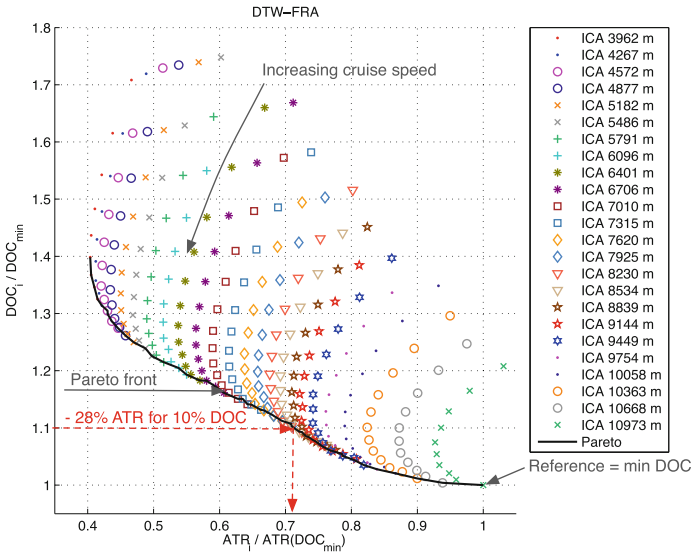

Fig. 44.2 Trade-off between climate impact reduction in terms of average temperature response *ATR* (for 100 years) and increased direct operating cost *DOC*, exemplarily shown for the route Detroit-Frankfurt flown with an A330-200 aircraft for different cruise altitudes (*symbols*) and different speeds, where increasing speed at the same altitude gives lower *DOC*. The Pareto front (*black line*) identifies the best compromises between reduced *ATR* and increased *DOC* for different cruise altitudes and speeds

44.4.2 Flight Route Optimization for Minimum Contrail and Fuel Climate Impact (UFO and CoCiP)

The feasibility of measures to avoid contrail cirrus using *smart routing* (weather dependent minimum contrail climate impact) (Mannstein et al. 2012) has been studied within the IPA-coordinated BMBF climate research project "Environmentally Compatible Flight Route Optimization" (UFO—Umweltgerechte Flugroutenoptimierung) (Mannstein 2011). The basic UFO strategy is to avoid warming contrail cirrus (day and night; Fig. 44.3, left) and to enhance the formation of cooling contrails (during daytime for suitable conditions; Fig. 44.3, right). The method includes climate impact *r(x,y,z,t)* for fossil CO_2 and contrails, which were considered as the most important contributions. The methods were implemented as part of the operational flight planning tool of Lufthansa Systems (LIDO-OC). For CO_2, the operation-dependent fuel consumption and the related climate costs are parameters available in LIDO. For contrails, LIDO was extended to include the potential *energy forcing* (i.e., the contrail forcing in case of flights) for a route segment by contrail-induced radiation flux changes at the top of the atmosphere as an integral over estimated contrail width and contrail lifetime per unit flight distance. Within UFO the German Weather Service (DWD) provided *numerical weather prediction* (NWP) data to estimate the potential energy forcing

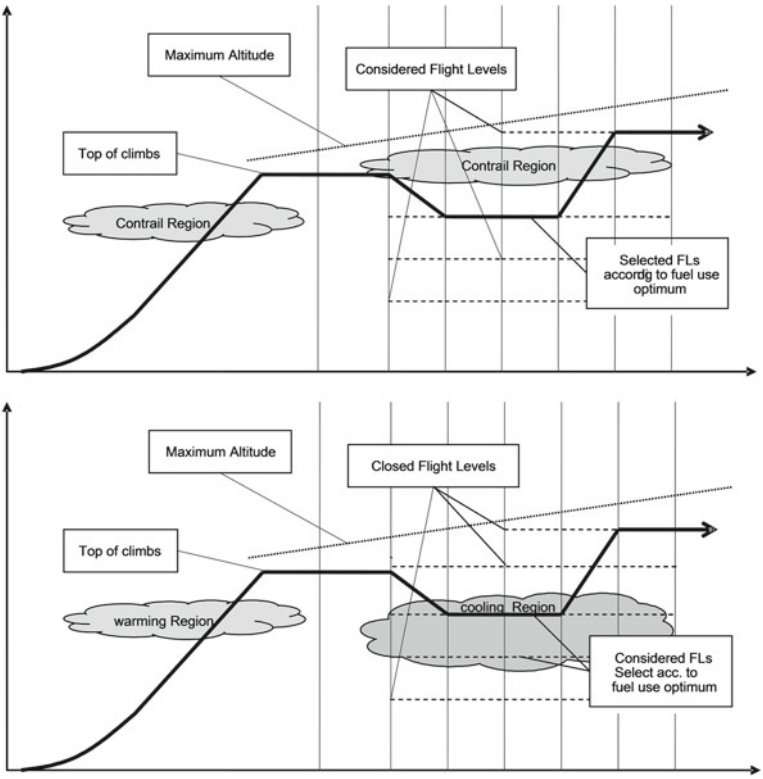

Fig. 44.3 Schematic view of trajectory-avoiding or -seeking contrail regions during cruise for minimum climate impact (figure by Stefanie Meilinger, Lufthansa)

in addition to the global *forecast* of wind and temperature fields. Contrail energy forcing was approximated as a parametric function of humidity, temperature and vertical motion at the emission location in the weather prediction data. An explicit simulation of the contrail life cycle was intended but not realized within the project, as it would have required major changes in the operational schedules of the DWD global model, the data *transmission*, and the LIDO-OC. The parametric energy forcing provides estimates with considerable *uncertainty*. The derived contrail energy forcing is highly variable at small scales in space and time, and may change sign during daytime. The UFO results indicate that the largest amount of warming contrail cirrus is produced by a small fraction of route segments. Hence the contrail avoidance strategy is already efficient even when applied to a limited number of route segments only.

The UFO idea can be improved by more accurate computation of the warming or cooling effects of contrails. The energy forcing by contrails depends not only on the local meteorological state at the time of flight but also on the weather thereafter affecting contrail evolution. It further depends on the aircraft, emissions, flight

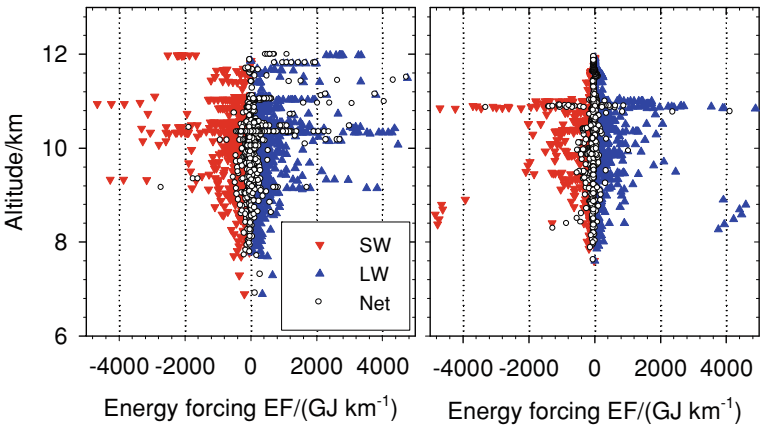

Fig. 44.4 Example results of contrail energy forcing (*EF*) of individual flight segments for global air traffic at a certain time instance (07:00 UTC 12 April 2005) versus flight altitude (*SW*, shortwave; *LW,* longwave, net = SW + LW). *Left*: reference; *right*: flight levels changed to minimum EF (Schumann et al. 2011)

direction relative to wind shear, vertical and horizontal motions, wind shear, *turbulence*, and changes in terrestrial and *solar radiation*. These influences are included in the IPA contrail prediction tool CoCiP (Schumann 2012). The method follows the evolution of the contrail in a Lagrangian manner from aircraft dependent formation until final decay. The model computes the drift of contrail segments with the wind. Contrail properties (width, depth, age, ice mass content, number of *ice particles* etc.) are computed with a bulk ice model using ambient meteorology taken from a NWP model. For efficiency and according to present understanding, the method still contains considerable simplifications. Inclusion of *feedback* of contrail formation onto the background meteorology would require interactive integration of CoCiP as part of the NWP model. CoCiP may be combined with models, as used in REACT4C, in the future.

The CoCiP model is efficient in computing contrail induced radiative forcing and from this the energy forcing (*EF*) by integrating over *contrail cover* and time. This is computed for each flight segment along the given route or any alternative route. Figure 44.4 (left) shows example results demonstrating the high variability of *EF* as caused by shortwave and longwave radiative forcing. Contrail net *EF* may be positive or negative depending on the shortwave and longwave contributions. The net *EF* of single route segments reaches several thousand GJ/km warming or cooling not only because of weather dependent *RF* values, but also because of different lifetimes. For a given time horizon, *EF* can be converted to *ATR* following a standard linear response model. Also the fuel consumption and its change with route changes are computed to estimate the *ATR* from CO_2 changes. This method requires accurate NWP data not only for the flight time period but also for some time thereafter until the last piece of contrail decays, which may be several hours after flight completion.

Table 44.1 Changes in radiative forcing and fuel consumption, illustrating the potential to mitigate the climate impact of aviation by flight level changes. Results from a) Schumann et al. (2011) and b) Fichter et al. (2005)

Case	References (%)	Flying 2 000 ft higher (%)	Flying 2 000 ft lower (%)	2 000 ft higher or lower depending on relative humidity (%)	2 000 ft higher or lower depending on local *RF* (%)
(a) Net contrail *RF*	100	95	92	46	−12
Fuel consumption	100	96.69	108.21	99.55	100.11
(b) Net contrail *RF*	100	107	86	n.a.	n.a.
Fuel consumption	100	95.5	103.2	n.a.	n.a.

The CoCiP model has been applied for considering vertical flight level changes of ±2 000 ft (±600 m). The avoidance strategies considered are: (1) flying always higher or (2) always lower, or (3) flying at the nearest level with lowest *relative humidity* when flying in an ISSR, or (4) flying at the nearest level with the lowest (possibly negative) *EF*. Figure 44.4 (right) shows example results when selecting flight levels with minimal *EF*, demonstrating the reduction of longwave *EF* and partly enhanced shortwave *EF*. First results, Table 44.1, show that the route changes may not only reduce the climate impact of contrails but may change the sign of the overall contrail forcing from warming to cooling. This reduction in climate impact can be achieved with a moderate amount of additional fuel consumption (far less than 1 % of block fuel) and related costs. This shows the high mitigation potential of changing contrail climate impact by individual flight level changes (Schumann et al. 2011). The same method can be combined with lateral flight route changes. Table 44.1 also compares results from CoCiP and from ECHAM4 for flying always higher or always lower. There are significant differences between the different meteorology, traffic, and performance models, which deserve further investigations.

44.4.3 Flight Route and Aircraft Design Optimization for Minimum Weather-Dependent Climate Impact (REACT4C)

In the EU-funded European collaborative project REACT4C (Reducing Emissions from Aviation by Changing Trajectories for the benefit of Climate) coordinated by IPA, a concept to identify alternative routes with minimal climate impact based on weather-dependent route optimization is explored (Matthes 2011). REACT4C determines climate cost functions for a particular *synoptic* situation as four-dimensional

functions *r(x,y,z,t)*. These functions quantify the climate impact of a unit emission for given synoptic weather conditions as a function of emission location, altitude and time of emission. This allows identification of the mitigation potential with respect to both fuel consumption and overall climate impact. The climate cost functions are calculated with the comprehensive global chemistry-climate model EMAC/ATTILA (Jöckel et al. 2010) using a Lagrangian approach developed at IPA (AIRTRAC). Here, CO_2, NOx (via O_3 and CH_4), H_2O, and contrails are considered. Separate cost functions are calculated for each disturbance. The climate cost function, e.g., of NOx-induced O_3, is attributed to the local emission from a flight per unit aviation NOx emission mass by a *tagging* attribution technology (Grewe et al. 2010). The fate of these emissions is studied following a set of Lagrangian trajectories for several weeks after the flight in order to identify a corresponding O_3 concentration change (Δc) pattern (increase). The NO_x emission further causes a CH_4 concentration change (decrease). From these concentration change fields the overall radiative imbalance is calculated. The space and time integral of local radiative imbalance is used to derive standard climate metrics, i.e., *RF* and *ATR*. This impact is attributed to the initial unit emission.

The climate cost functions represent the interface between climate-chemistry modeling and flight planning of aircraft trajectories. This procedure is presently used for a finite set of typical weather patterns over the North Atlantic in order to estimate overall mitigation potential. These patterns are defined based on meteorological data and track system data for typical weather patterns on synoptic scales (Irvine et al. 2012). For the North Atlantic, the temporal evolution of the NOx concentration change resulting from a unit NOx emission, the contribution to ozone, and the induced radiative imbalance are exemplarily shown for two distinct emission locations (Fig. 44.5) within one weather pattern. Large differences of O_3 radiative forcing, of one order of magnitude, can be found between these two emission locations A and B. These differences are mainly caused by different prevailing local transport characteristics at the respective emission location. Either the emitted species are transported to lower altitudes and towards the tropics (in the case of emission location B), or the emitted species remain at high altitudes and latitudes (in the case of emission location A), with different lifetimes of NO_x, different local ozone production efficiency, and different specific radiative impact. The approach is employed for each weather pattern and multiple emission locations, altitudes and times, providing individual cost functions of climate impact simultaneously for all emission species considered.

The concept is tested as part of the flight planning tools of EUROCONTROL. Implementing climate optimized flight planning in current ATM systems requires a re-design of specific system components. Computation of cost functions requires information on the evolution of the emissions and the induced climate relevant species over their full lifetime (weeks to months). After some time, with increasing *dispersion*, the decay of induced species becomes less sensitive to the specific weather pattern selected. The cost functions are presently computed for typical weather patterns in a climate model. For operational daily route planning, weather forecast data are to be used. At present, NWP forecasts cover time periods of a few days. This period is not sufficient to assess the full climate impact. One may

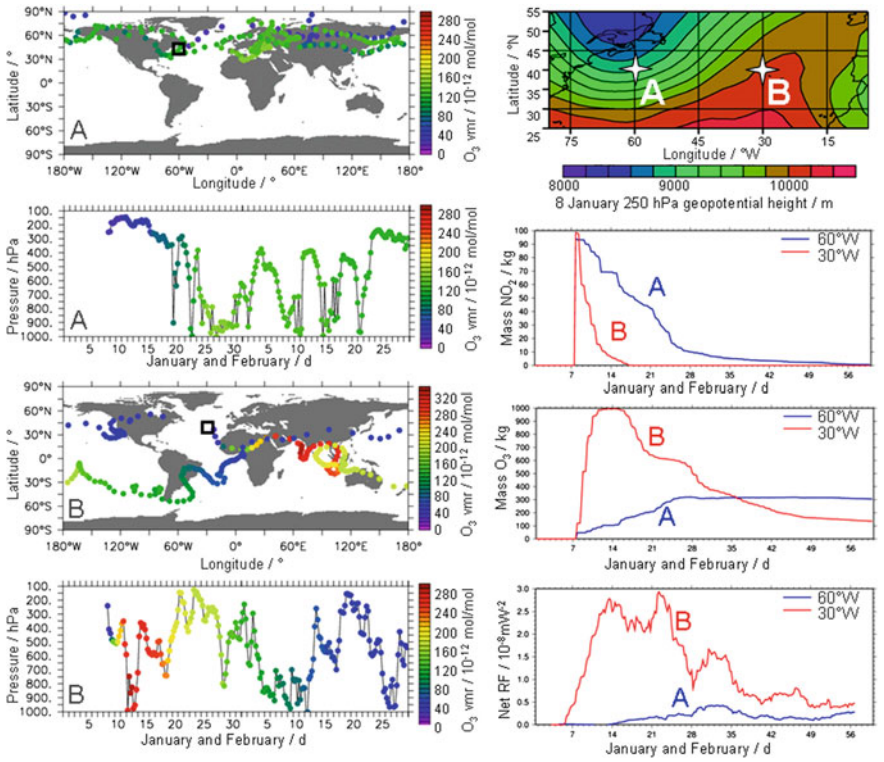

Fig. 44.5 Temporal development of perturbations following aviation NOx emissions at two different emission locations A and B (*stars*) for a particular weather pattern over the North Atlantic as indicated by the geopotential height (*upper right*). The trajectory pathways and altitudes for both emission locations A and B are shown on the *left*, where the color indicates the ozone mixing ratio. The time series of the NO_x perturbation, ozone perturbation and induced net O_3 radiative forcing (10^{-8} mW/m^2) for the two emission locations are shown on the *right*

approximate the climate impact using results for that weather pattern which is closest to the one actually predicted. Apart from the exploitation of the concept in future ATM, it is planned to feed the results into the aircraft predesign process (for Airbus, etc.) to optimize aircraft design ("*green aircraft*"). Potential changes in aircraft design that allow for weather and climate dependent route changes with a smaller fuel penalty than is the case for present aircraft are explored.

44.4.4 Comparison of Optimization Objectives

Three different mitigation approaches have been described with different applications and optimization principles. CATS follows a climatological strategy while UFO and REACT4C use weather-dependent routing concepts. CATS and

REACT4C use a multiobjective approach in which the climate impact in terms of ATR is considered simultaneously with the costs of priced goods. UFO considers a single cost function including costs of operations and costs for climate change with a monetization factor related to the costs of emission trading. The formulations differ also for the specific impact functions *r*. For optimization under climate mean conditions, CATS so far uses climate cost functions depending on latitude and altitude *r(y,z)* without utilizing for weather-induced variability. UFO aims at weather-dependent contrail avoidance and fuel optimization and, hence, uses a time- and space-dependent response function $r(x, y, z, t)$. For overall weather and climate impact optimization, REACT4C uses vectors $\bar{r}(x, y, z, t)$ and $\bar{e}(t)$ to represent individual emission effects (CO_2, NOx, H_2O and contrails). For weather-dependent optimization, aircraft design should allow for more frequent altitude changes to avoid climate-sensitive regions and for weather dependent variability in the response functions.

44.5 Conclusions

Climate optimized flight planning and aircraft design offer the potential for a reduction of the climate impact of aviation for a given transport volume. The mitigation potential, i.e., the potential to reduce climate impact for given costs, is dependent on scenarios and time horizons. Longer flights offer more possibilities for lateral route changes and hence the mitigation effects by cruise optimization may be larger for long range than for short range flights. Because of the larger variability of climate cost functions in a weather-based approach, the mitigation potential increases when weather adapted operation changes are included in the climate impact calculation. If aviation puts the effort on selective avoidance of warming contrails and cirrus, net contrail cirrus forcing can be reduced substantially. Similarly, ozone formation can be reduced by avoiding sensitive regions of the atmosphere.

Climate optimization introduces climate costs in common flight planning and design optimization procedures. Implementation of mitigation options requires accurate and efficient methods to compute the climate impact of emissions and to compare them with the additional cost for more complex operations. Weather dependent cost functions require additional meteorological and air-chemistry information. For climate based aircraft design, the mean variability of the cost functions from cost uncertainties or weather dependent operations should be included (e.g., by *Monte Carlo* methods). The optimization approaches, once established, have to be validated and certified before they can be brought into practical operation.

The research described in this article aims to identify the mitigation potential of and the emerging methods for establishing climate optimal air traffic. Current research provides the opportunity to systematically investigate trade-offs between various mitigation concepts and cost functions. The practical implementation of various mitigation schemes would also require regulatory incentives to industry, including a suitable accounting system for CO_2 and non-CO_2 effects.

References

Brasseur, G.P., Cox, R.A., Hauglustaine, D., Isaksen, I., Lelieveld, J., Lister, D.H., Sausen, R., Schumann, U., Wahner, A., Wiesen, P.: European scientific assessment of the atmospheric effects of aircraft emissions. Atmos. Environ. **32**, 2329–2418 (1998)

Burkhardt, U., Kärcher, B.: Global radiative forcing from contrail cirrus. Nat. Clim. Change. **1**, 54–58 (2011). doi:10.1038/NCLIMATE1068

Campbell, S.E., Neogi, N.A., Bragg, N.B.: An operational strategy for persistent contrail mitigation. 9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO), pp. 1–14 (2009)

Döpelheuer, A., Lecht, M.: Influence of engine performance on emission characteristics. Gas Turbine Engine Combustion, Emissions and Alternative Fuels, RTO MP-14, ISBN 92-837-0009-0, p. 20, 11 (1999)

Fichter, C., Marquart, S., Sausen, R., Lee, D.S.: The impact of cruise altitude on contrails and related radiative forcing. Meteorol. Z. **14**, 563–572 (2005). doi:10.1127/0941-2948/2005/0048

Fichter, C.: Climate impact of air traffic emissions in dependency of the emission location and altitude. DLR Forschungsbericht 2009-22, ISSN 1434-84543, Oberpfaffenhofen, 152 pp (2009)

Fuglestvedt, J.S., Shine, K.P., Berntsen, T., Cook, J., Lee, D.S., Stenke, A., Skeie, R.B., Velders, G.J.M., Waitz, I.A.: Transport impacts on atmosphere and climate: Metrics. Atmos. Environ. **44**, 4648–4677 (2010). doi:10.1016/j.atmosenv.2009.04.044

Gauss, M., Isaksen, I.S.A., Wong, S., Wang, W.-C.: Impact of H_2O emissions from cryoplanes and kerosene aircraft on the atmosphere. J. Geophys. Res. **108**, 4304 (2003). doi:10.1029/2002JD002623

Gauss, M., Isaksen, I.S.A., Lee, D.S., Søvde, O.A.: Impact of aircraft NO_x emissions on the atmosphere—tradeoffs to reduce the impact. Atmos. Chem. Phys. **6**, 1529–1548 (2006)

Gierens, K., Spichtinger, P.: On the size distribution of ice-supersaturated regions in the upper troposphere and lowermost stratosphere. Ann. Geophys. **18**, 499–504 (2000). doi:10.1007/s00585-000-0499-7

Gierens, K., Lim, L., Eleftheratos, K.: A review of various strategies for contrail avoidance. Open Atmos. Sci. J. **2**, 1–7 (2008)

Green, J.E.: Future aircraft—greener by design? Meteorol. Z. **14**, 583–590 (2005). doi:10.1127/0941-2948/2005/0052

Grewe, V., Dameris, M., Fichter, C., Lee, D.S.: Impact of aircraft NO_x emissions. Part 2: Effects of lowering the flight altitude. Meteorol. Z. **11**, 197–205 (2002)

Grewe, V., Stenke, A.: AirClim: An efficient tool for climate evaluation of aircraft technology. Atmos. Chem. Phys. **8**, 4621–4639 (2008). doi:10.5194/acp-8-4621-2008

Grewe, V., Tsati, E., Hoor, P.: On the attribution of contributions of atmospheric trace gases to emissions in atmospheric model applications. Geosci. Model Dev. **3**, 487–499 (2010). doi:10.5194/gmd-3-487-2010

Grewe, V., Tsati, E., Hoor, P.: On the attribution of contributions of atmospheric trace gases to emissions in atmospheric model applications. Geosci. Model Dev. **3**, 487–499 (2010). doi:10.5194/gmd-3-487-2010

Grooß, J.-U., Brühl, C., Peter, T.: Impact of aircraft emissions on tropospheric and stratospheric ozone. Part I: Chemistry and 2-D model results. Atmos. Environ. **32**, 3173–3184 (1998)

Haywood, J.M., Allan, R.P., Bornemann, J., Forster, P.M., Francis, P.N., Milton, S., Rädel, G., Rap, A., Shine, K.P., Thorpe, R.: A case study of the radiative forcing of persistent contrails evolving into contrail-induced cirrus. J. Geophys. Res. **114**, D24201 (2009). doi:10.1029/2009JD012650

Hendricks, J., Kärcher, B., Lohmann, U.: Effects of ice nuclei on cirrus clouds in a global climate model. J. Geophys. Res. **116**, D18206 (2011). doi:10.1029/2010JD015302

HLGAR.: Flightpath 2050—Europe's Vision for Aviation 2011: Report of the High Level Group on Aviation Research (HLGAR), Luxembourg: Publications Office of the European Union,

ISBN 978-92-79-19724-6, http://ec.europa.eu/transport/air/doc/flightpath2050.pdf (2011). doi: 10.2777/50266

Holmes, C.D., Tang, Q., Prather, M.J.: Uncertainties in climate assessment for the case of aviation NO. PNAS, 6, (2011). doi:10.1073/pnas.1101458108

IPCC: Aviation and the Global Atmosphere. Cambridge University Press, Cambridge (1999). 373 pp

Irvine, E.A., Hoskins, B.J., Shine, K.P., Lunnon, R.W., Frömming, C.: Characterizing North Atlantic weather patterns for climate-optimal aircraft routing. Meteorol. Appl. (2012). doi: 10.1002/met.1291

Jöckel, P., Kerkweg, A., Pozzer, A., Sander, R., Tost, H., Riede, H., Baumgaertner, A., Gromov, S., Kern, B.: Development cycle 2 of the modular earth submodel system (MESSy2). Geosci. Model Dev. **3**, 717–752 (2010). doi:10.5194/gmd-3-717-2010

Kärcher, B., Möhler, O., DeMott, P.J., Pechtl, S., Yu, F.: Insights into the role of soot aerosols in cirrus cloud formation. Atmos. Chem. Phys. **7**, 4203–4227 (2007)

Kärcher, B., Yu, F.: Role of aircraft soot emissions in contrail formation. Geophys. Res. Lett. **36**, L01804 (2009). doi:10.1029/2008GL036649

Klima, K.: Assessment of a global contrail modeling method and operational strategies for contrail mitigation. Thesis for a Master of Science, Aeronautics and Astronautics at the Massachusetts Institute of Technology (2005)

Klug, H.G., Bakan, S., Gayler, V.: Cryoplane—Quantitative Comparison of Contribution to Anthropogenic Greenhouse Effect of Liquid Hydrogen Aircraft Versus Conventional Aircraft. European Geophysical Society, XXI, General Assembly, The Hague, The Netherlands (1996). 22

Koch, A., Lührs, B., Dahlmann, K., Linke, F., Grewe, V., Litz, M., Plohr, M., Nagel, B., Gollnick, V., Schumann, U.: Climate impact assessment of varying cruise flight altitudes applying the CATS simulation approach. CEAS 2011 The International Conference of the European Aerospace Societies, p. 12 (2011)

Köhler, M.O., Rädel, G., Dessens, O., Shine, K.P., Rogers, H., Wild, O., Pyle, J.A.: Impact of perturbations to nitrogen oxide emissions from global aviation. J. Geophys. Res. **113**, D11305 (2008). doi:10.1029/2007JD009140

Linke, F., Langhans, S., Gollnick, V.: Global fuel analysis of intermediate stop operations on long-haul routes. 11th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, AIAA 2011-6884. Virginia Beach, USA (2011)

Lee, D.S., Fahey, D.W., Forster, P.M., Newton, P.J., Wit, R.C.N., Lim, L.L., Owen, B., Sausen, R.: Aviation and global climate change in the 21st century. Atmos. Environ. **43**, 3520–3537 (2009). doi:10.1016/j.atmosenv.2009.04.024

Lee, D.S., Pitari, G., Grewe, V., Gierens, K., Penner, J.E., Petzold, A., Prather, M.J., Schumann, U., Bais, A., Berntsen, T., et al.: Transport impacts on atmosphere and climate: Aviation. Atmos. Environ. **44**, 4678–4734 (2010). doi:10.1016/j.atmosenv.2009.06.005

Mannstein, H., Spichtinger, P., Gierens, K.: How to avoid contrail cirrus. Transp. Res. **D 10**, 421–426 (2005)

Mannstein, H.: Umweltgerechte Flugrouten-Optimierung (UFO)—Endbericht, DLR and Lufthansa, 61 pp. (2011)

Mannstein, H., Schumann, U.: Gierens, K., Meilinger, S., Waibel, A.: Verfahren und Vorrichtung zur klimaoptimierten Flugplanung. Submitted Patent, 2012

Marquart, S., Sausen, R., Ponater, M., Grewe, V.: Estimate of the climate impact of Cryoplanes. Aerosp. Sci. Technol. **5**, 73–84 (2001)

Matthes, S.: Climate-optimised flight planning—REACT4C in Innovation for a Sustainable Avation in a Global Environment, Proceedings of the Sixth European Aeronautics Days 2011, IOS Press & European Union (2012) ISBN 978-92–79-22968-8

Meerkötter, R., Schumann, U., Minnis, P., Doelling, D.R., Nakajima, T., Tsushima, Y.: Radiative forcing by contrails. Ann. Geophys. **17**, 1080–1094 (1999). doi:10.1007/s00585-999-1080-7

Penner, J.E., Chen, Y., Wang, M., Liu, X.: Possible influence of anthropogenic aerosols on cirrus clouds and anthropogenic forcing. Atmos. Chem. Phys. **9**, 879–896 (2009). doi: 10.5194/acp-9-879-2009

Ponater, M., Pechtl, S., Sausen, R., Schumann, U., Hüttig, G.: Potential of the cryoplane technology to reduce aircraft climate impact: A state-of-the-art assessment. Atmos. Environ. **40**, 6928–6944 (2006). doi:10.1016/j.atmosenv.2006.06.036

Rädel, G., Shine, K.P.: Radiative forcing by persistent contrails and its dependence on cruise altitudes. J. Geophys. Res. **113**, D07105 (2008). doi:10.1029/2007JD009117

Sausen, R., Nodorp, D., Land, C.: Towards an optimal flight routing with respect to minimal environmental impact. In: Schumann, U., Wurzel, D. (eds.) Impact of Emissions from Aircraft and Spacecraft upon the Atmosphere. Procedings of an International Science Colloquium, Köln (Cologne), Germany, April 18–20, pp. 473–478 (1994). ISSN 0939-298X

Sausen, R., Gierens, K., Ponater, M., Schumann, U.: A diagnostic study of the global distribution of contrails. Part I: Present day climate. Theor. Appl. Clim. **61**, 127–141 (1998)

Sausen, R., Schumann, U.: Estimates of the climate response to aircraft CO_2 and NO_x-emission scenarios. Clim. Change **44**, 27–58 (2000)

Schumann, U.: On the effect of emissions from aircraft engines on the state of the atmosphere. Ann. Geophys. **12**, 365–384 (1994)

Schumann, U.: On conditions for contrail formation from aircraft exhausts. Meteorol. Z. **5**, 4–23 (1996)

Schumann, U.: Contrail Cirrus. In: Lynch, D.K., Sassen, K., Starr, D.O'C., Stephens, G. (eds.) Cirrus. Oxford University Press, Oxford, pp. 231–255 (2002)

Schumann, U.: Formation, properties and climate effects of contrails. Compt. Rend. Phys. **6**, 549–565 (2005)

Schumann, U.: A contrail cirrus prediction model. Geosci. Model Dev. **5**, 543–580 (2011). doi: 10.5194/gmd-5-543-2012

Schumann, U., Graf, K., Mannstein, H.: Potential to reduce the climate impact of aviation by flight level changes. 3rd AIAA Atmospheric and Space Environments Conference AIAA paper 2011–3376, 1–22 (2011)

Schumann, U., Mayer, B., Graf, K., Mannstein, H.: A parametric radiative forcing model for contrail cirrus. J. Appl. Meteorol. Clim. 51 (2012) 10.1175/JAMC-D-11-0242.1

Schwartz Dallara, E., Kroo, I.M., Waitz, I.: Metric for comparing lifetime averaged climate impact of aircraft. AIAA J. **49**, 1600–1613 (2011)

Spichtinger, P., Gierens, K., Leiterer, U., Dier, H.: Ice supersaturation in the tropopause region over Lindenberg. Ger. Meteorol. Z. **12**, 143–156 (2003). doi:10.1127/0941-2948/2003/0012-0143

Sridhar, B., Chen, N.Y., Ng, H.K., Linke, F.: Design of aircraft trajectories based on trade-offs between emission sources. 9th USA/Europe Air Traffic Management Research and Development Seminar (ATM2011). http://www.atmseminar.org/. (2011)

Ström, L., Gierens, K.: First simulations of cryoplane contrails. J. Geophys. Res. **107**, 4346 (2002). doi:10.1029/2001JD000838

Vazquez-Navarro, M.R.: Life cycle of contrails from a time series of geostationary satellite images, DLR-FB 2010-19, 146 pp (2009)

Waitz, I., Townsend, J., Cutcher-Gershenfeld, J., Greitzer, E., Kerebrock, J. (eds.) Aviation and the Environment—Report to the United States Congress, A National Vision Statement, Framework for Goals and Recommended Actions. Massachusetts Institute of Technology, under FAA Cooperative Agreement No. 03-C-NE-MIT, 52 pp (2004)

Williams, V., Noland, R.B., Toumi, R.: Reducing the climate change impacts of aviation by restricting cruise altitudes. Transp. Res. **D7**, 451–464 (2002)

Recent and Future Evolution of the Stratospheric Ozone Layer

45

Martin Dameris and Diego Loyola

Abstract

Since the early 1980s significant depletion of the ozone layer in the stratosphere, in other words the ozone hole, has been observed every year over the South Pole area in Antarctic spring. In the meantime destruction of stratospheric ozone has been detected globally. Emissions of man-made halogenated chemicals play a dominant role in ozone loss. Combined analyses of observations and numerical modeling help to understand the complex interplay of the dynamic and chemical processes involved. Evaluated models provide a base for predicting the future recovery of the ozone layer expected for the middle of this century.

45.1 Introduction

Nearly 100 years ago the *ozone* layer in the stratosphere was discovered by (Fabry and Buisson 1913). The largest amount of ozone molecules (O_3) is found in the *stratosphere* with about 90 % of it at altitudes between 15 and 30 km. Here, the

M. Dameris (✉)
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Martin.Dameris@dlr.de

D. Loyola
DLR, Remote Sensing Technology Institute (IMF), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Diego.Loyola@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_45,

molar mixing ratio maximizes with about ten O_3 molecules per 1 million other air molecules, i.e., 10 μmol/mol. Stratospheric O_3 filters out a large part of the *ultraviolet (UV)* radiation emitted by the sun, protecting life on Earth. The almost complete absorption of the energy-intensive solar UV-B radiation (wavelength range from 280 to 320 nm) is of particular importance. Enhanced UV-B radiation can have a negative impact on photosynthesis, cause skin cancer and weaken the immune system. On the other hand, absorption of solar UV radiation by stratospheric O_3 causes the temperature of the stratosphere to increase with height, creating a stable layer that limits strong vertical air movement. This plays a key role for Earth's *climate system*.

A first report about massive *depletion* of the *ozone layer* during spring over *Antarctica* (Fig. 45.1) was published by (Farman et al. 1985). Shortly after the discovery of the *ozone hole*, this extreme thinning of the ozone layer was explained as a combination of special meteorological conditions and changed chemical composition of the atmosphere, based on industrially manufactured *chlorofluorocarbons* (CFCs) (Solomon et al. 1986). In the following years it became clear that the ozone layer was getting thinner over many other regions, too, although to a lesser extent. In 1987 this led to a first-time regulation of the production and use of ozone-depleting substances in the form of the *Montreal Protocol*. In a number of subsequent agreements, the signature states committed to reducing and eventually almost fully abandoning the production and use of *chlorine* and *bromine*-containing chemicals that are responsible for O_3 depletion in the stratosphere. Moreover, in 1988 the German government reacted to the significant thinning of the ozone layer with the creation of a national ozone research program (*Ozonforschungsprogramm*, OFP). Between 1989 and 1999 three large scientific measurement campaigns (EASOE, SESAME, and THESEO) were organized by research groups in Europe. In addition, laboratory experiments and investigations with numerical models of the atmosphere started. The DLR Institute of Atmospheric Physics was involved in many OFP projects and in the coordination of OFP activities.

The ozone layer's thickness is not solely controlled by chemical processes in the stratosphere. As an atmospheric trace gas, O_3 is also transported over large distances by stratospheric winds, which significantly affects global ozone distribution. Dynamic and chemical processes in the atmosphere interact in highly complex ways. Radiative processes also play a crucial role since O_3 is one of the most important radiatively active gases in the atmosphere. It absorbs both *shortwave* and *longwave radiation* and thus influences the vertical *temperature* distribution of the stratosphere. Understanding all these processes and their interactions is the key to characterizing both short- and long-term fluctuations in the ozone layer's thickness and explaining trends. This knowledge is necessary for reliably estimating the future evolution of the ozone layer. In this context alterations of atmospheric processes due to *climate change* and respective *feedbacks* must be considered.

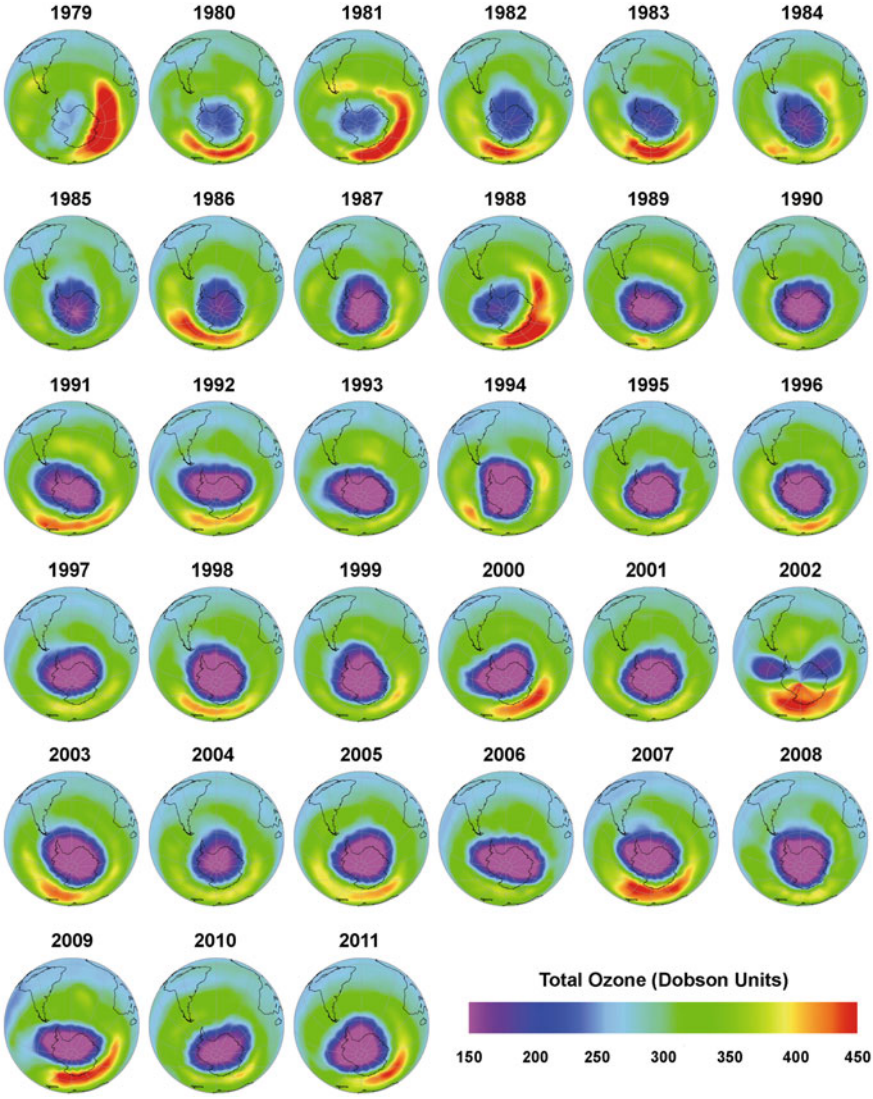

Fig. 45.1 3-day mean ozone column densities (in Dobson units) in Antartic spring from 1979 to 2010 showing the evolution of the ozone hole. The *purple* area indicates the area of the ozone hole (Fig. 4 in Dameris and Loyola 2011)

45.2 The Chemistry of Stratospheric Ozone

In the atmosphere, O_3 is formed exclusively by *photochemical* processes. Direct O_3 sources, e.g., at the Earth's surface, do not exist. Stratospheric formation of O_3 is initiated by *photolysis* of molecular *oxygen* (O_2) at *wavelengths* of less than

240 nm. Two oxygen atoms (O) are produced, recombining with O_2 to form O_3 (with the aid of another molecule, generally nitrogen or oxygen). Since O_3 is photochemically formed, it is mainly produced at low latitudes, where sunshine is most intensive throughout the year. At the same time, O_3 is destroyed by the photolysis of O_3 at wavelengths of less than 300 nm and by reaction with O. These reactions describe the basis of stratospheric ozone chemistry and are known as the *Chapman mechanism*.

If stratospheric O_3 amounts are solely determined via the Chapman mechanism with the known rate constants and photolysis rates, the results obtained are about twice as high as measured values. Since the early 1950s it has been known that fast, so-called *catalytic* cycles reduce the determined O_3 amounts to the observed values (Bates and Nicolet 1950), whereby the *catalyst* X is not wasted in the *reaction cycle*:

$$\begin{aligned} &X + O_3 \rightarrow XO + O_2 \\ &O + XO \rightarrow X + O_2 \\ &\text{net}: \; O + O_3 \rightarrow 2\,O_2 \end{aligned}$$

X is oxidized to XO and reformed in the subsequent reaction with O. In this way, a catalyst can take part in the reaction cycle several thousand times. This reaction sequence significantly enhances ozone depletion. By the early 1970s, the catalysts had been identified as the radical pairs (X/XO) OH/HO_2 and NO/NO_2, which are formed from *water vapor* (H_2O) and *nitrous oxide* (N_2O), respectively (Crutzen 1971). In the mid-1970s, the radical pairs Cl/ClO (from CFCs) and Br/BrO (from *halons*) were identified as further significant contributors (Molina and Rowland 1974). In 1995, Crutzen, Molina and Rowland received the Nobel Prize for Chemistry for their work concerning the formation and decomposition of ozone.

The natural stratospheric *chlorine* mixing ratio is around 0.6 nmol/mol, the main *source* of chlorine being *methyl chloride* (CH_3Cl), which is generated in large amounts in oceans and soil. Due to the large quantities of CFCs emitted into the atmosphere over recent decades, stratospheric chlorine mixing ratios of around 3.5 nmol/mol arose at the turn of the millennium.

The increased occurrence of CFCs and halons due to *anthropogenic emissions* has significantly accelerated the stratospheric O_3 depletion cycle over recent decades, triggering a negative stratospheric O_3 *trend*. In the troposphere, CFCs and halons are mostly inert. Due to air mass motion they are transported from the *troposphere* into the *lower stratosphere*, e.g., by *deep convection*, where they are moved further upward. Only there are they photolysed and converted into active *chlorine* or *bromine* compounds. This effect is solely slowed down by the *radicals* reacting with each other, e.g., via

$$\begin{aligned} &ClO + NO_2 + M \rightarrow ClONO_2 + M \\ &ClO + HO_2 \rightarrow HOCl + O_2 \\ &Cl + HO_2 \rightarrow HCl + O_2 \end{aligned}$$

This creates reservoir gases such as *chlorine nitrate* ($ClONO_2$) and *hydrogen chloride* (HCl), which reduce the ozone-depleting potential. However, they can also be converted back to their active forms. A similar schematic can be derived for the bromine oxides (Br, BrO).

Most source gases for O_3-depleting catalysts (H_2O, N_2O, CFCs, and halons) enter the stratosphere via the tropical troposphere. The catalysts (HO_x, NO_x, ClO_x, BrO_x; abbreviations for radical pairs) are mainly released by photochemical reactions. Due to the varying degrees of *absorption* of solar UV radiation and the availability of *solar radiation*, not all catalysts are equally important at the same stratospheric altitudes. ClO_x is dominant in the middle and lower stratosphere (ca. 15–30 km), where most O_3 is found. At heights below around 15 km, HO_x determines O_3 depletion, and above 30 km, it is NO_x that dominates. That is why the negative ozone trend of the past decades is primarily attributed to anthropogenic ClO_x.

Just like at lower and middle latitudes, O_3 can also be depleted via HO_x-, NO_x-, ClO_x-, and BrO_x- catalytic cycles in polar winter. However, the kinetics of these processes are very slow because the amount of UV radiation is limited due to the prevailing twilight conditions. In the polar stratosphere, mainly *heterogeneous* chemical reactions on the surface of stratospheric ice and nitric-acid-containing particles are responsible for activating chlorine (or bromine) and then driving it to deplete O_3, as observed, immediately after the end of *polar night* (Solomon 1999).

In the very cold lower polar stratosphere, *polar stratospheric clouds* (PSCs) form during polar night. PSCs develop at *temperatures* below around −78 °C and *nitric acid trihydate* (NAT, i.e., HNO_3 $3H_2O$) crystals can form (Voigt et al. 2000). Under given standard conditions in the lower stratosphere, NAT has an equilibrium temperature that is around 7–8 °C higher than that of ice, developing at temperatures below ca. −85 °C. Thus, it is generated more often during the cooling process. Due to different land-sea distributions on the northern and southern *hemispheres*, the lower stratosphere in winter cools significantly more over the south than over the north polar region. The climatological mean polar winter temperature of the lower *Arctic* stratosphere is around 10 °C higher than that of the lower *Antarctic* stratosphere. While the Antarctic stratosphere reaches temperatures below PSC-forming temperatures for several weeks every year, there is a high year-to-year variability observed in the north polar stratosphere: relatively warm winters, where hardly any PSCs develop are observed, as well as very cold winters, with conditions similar to those of Antarctica. This means that expansive synoptic PSC fields develop in the Antarctic stratosphere every year; similar kinds are seldom seen over the Arctic. PSCs in the Arctic region can also be formed close to mountain regions. The meteorology in the Arctic is characterized by uneven land–ocean distribution and *orography*, and air flowing over mountains can be set into vertical oscillation, for example over the Scandinavian Alps (Volkert and Intes 1992). These *gravity waves* cause pressure gradients and therefore by courtesy of *adiabatic* expansion and compression one can get

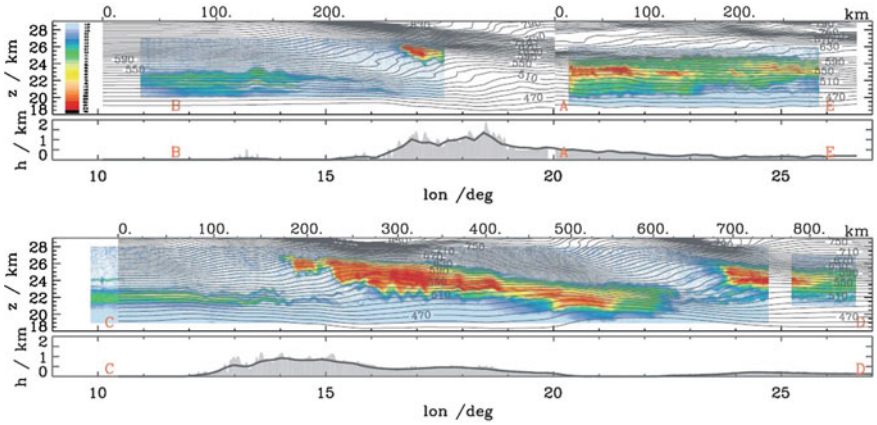

Fig. 45.2 *Backscatter ratio* at 1064 nm along cross-mountain flight legs of the DLR Falcon aircraft on 26 January 2000 (*color* shading). Measurements were carried out with the IPA's lidar OLEX system. Colored areas indicate PSCs in the lee of the mountains; the orography is indicated in the bottom parts of both cross sections. Superimposed is the *potential temperature* (*grey lines*) which was calculated with the *mesoscale model* MM5. *White isolines* show the temperature difference relative to the frost-point temperature, i.e., regions where ice particles may persist (similar to upper part of Fig. 8 in Dörnbrack et al. 2002)

temperature variations of 10 °C or more (Dörnbrack et al. 2002). In the wave region, cold spots appear with a horizontal size of about 10–100 km, which allows the formation of PSCs. Examples of "mountain PSCs" are shown in Fig. 45.2; complementary results of streamlines calculated by a *mesoscale model* are presented in the figure, indicating large temperature fluctuations in the lee of the mountain ridge, allowing PSC formation.

One of the first measurements of Arctic PSCs was performed in the early 1990s by the DLR airborne Ozone Lidar Experiment (OLEX) on board the German military aircraft *Transall* C-160; later on the DLR *Falcon* aircraft, the Russian high altitude research aircraft *Geophysica*, as well as *balloons* were used as carriers for respective PSC observations. Fundamental information about the specific condition of PSC formation (e.g., size and phase, i.e., liquid or solid) were provided by the first OLEX observations, laying the basis for further studies which enabled more detailed investigation of *microphysical* and *chemical* processes relevant for ozone destroying chemistry (Wirth and Renger 1996; Carslaw et al. 1997; Wirth et al. 1999). PSC particles allow heterogeneous reactions to take place on their surfaces, which enables halogen compounds to be released from reservoir compounds and then converted to active forms. Due to their size and weight, PSC particles can sediment out and thus transport nitrogen compounds to lower stratospheric layers (*denitrification*). This reduces the chance of the destructive chlorine compounds being deactivated again shortly after the PSC particles melt. At certain heights, both the release and photolysis of gaseous HNO_3 and the formation of new NO_2 are reduced, contributing to the deactivation of ClO radicals. Denitrification in the Arctic vortex was observed in

the EUPLEX and RECONCILE projects in winter 2002/2003 and 2009/2010, respectively. Measurements of *total reactive nitrogen* (NO_y) contained in the gas-phase and in ice particles were performed aboard the *Geophysica*. So-called '*NAT rocks*' were detected which did not form on ice but possibly on *meteoritic condensation nuclei* (Voigt et al. 2005).

The NO_x cycle thus loses importance and the ClO_x cycle accelerates due to the released chlorine. In the lower polar stratosphere, the efficiency of the ClO_x cycle increases mainly due to the low number of O present under *twilight* conditions. The cycle does not run via

$$ClO + O \rightarrow Cl + O_2$$

but via the formation of ClOOCl:

$$ClO + ClO + M \rightarrow ClOOCl + M$$
$$ClOOCl + h\nu \rightarrow Cl + ClOO$$
$$ClOO + M \rightarrow Cl + O_2 + M$$

ClOOCl is photolyzed slightly, even when the sun is low and releases Cl atoms that destroy O_3. As a result, the photolysis of ClOOCl plays a key role in polar O_3 depletion.

Attention must be also drawn to the fact that violent *volcanic eruptions* inject large amounts of *aerosols* and chemical substances directly into the stratosphere. When Mount *Pinatubo* erupted on the Philippines in June 1991, almost 20 million tons of *sulfur dioxide* (SO_2) were injected into the lower stratosphere. The SO_2 oxidized in the stratosphere and formed *sulfuric acid* droplets. In the atmosphere these have an effect very similar to that of PSCs, but with the difference that the heterogeneous chemical reactions already start at higher temperatures. However, volcanic effects on the ozone layer and climate are limited to a few years, mainly due to the removal of particles and respective chemical substances by *sedimentation*.

Finally, it is important to note that many chemical reactions in the atmosphere depend on the prevailing temperature, connecting the consequences of climate change to atmospheric chemistry, including reactions affecting O_3 (Dameris and Baldwin 2011). The main O_3-depletion cycles in the middle and upper stratosphere (brought about by the catalysts HO_x, NO_x, ClO_x and BrO_x) slow down with decreasing temperature, resulting in higher O_3 concentrations. However, low temperatures in the lower polar stratosphere lead to increased PSC formation and thus potentially to increased depletion of O_3 via heterogeneous chemistry.

45.3 Stratospheric Circulation and Ozone Transport

Atmospheric *trace gas* concentrations are not solely influenced by chemical processes, but also by air mass transports, which are affected by *wind speed* and direction, and *mixing* of reacting trace gases, which is of particular importance for quickly reacting

substances (Schumann 1989). The extent to which transport of trace gases takes place depends on the *lifetime* of the chemical species in question. In the upper and middle stratosphere, O_3 is rather *short-lived* (an order of minutes to hours), since its depletion reactions via photolysis and with oxygen are relatively fast. Therefore, O_3 transport at these altitudes only plays a minor role. In the lower stratosphere the chemical lifetime of O_3 is significantly longer (an order of days to weeks), which is why transport processes play a key role in geographical distribution there. At these heights, O_3 can be transported to latitudes where, photochemically, it is only produced to an insignificant extent. In this way, O_3 generated at lower and middle latitudes is transported effectively in the direction of the winter pole due to a large-scale *meridional* circulation in the stratosphere, the *Brewer-Dobson circulation* (BDC). There, O_3 is mixed in with the local air. In the end this leads to a globally asymmetric distribution of total ozone with maximum values at higher latitudes during the corresponding spring months; the total ozone amounts found in tropical regions are relatively low (Figs. 45.1 and 45.3). Regarding the vertical distribution of stratospheric ozone in the *tropics*, the highest O_3 number concentrations occur at heights around 25 km; at higher latitudes the O_3 maximum is located at lower altitudes. The dynamics of the stratosphere and related O_3 transport processes are responsible for this distribution. With respect to atmospheric dynamics, it is expected that climate change will modify processes that drive the circulation systems. This would change both the intensity of air mass transports and the transportation routes, with possible long-term consequences for the atmospheric distribution of climate-influencing trace gases, including O_3. This, in turn, can affect Earth's climate.

45.4 Recent Changes in the Amount of Stratospheric Ozone

At the end of the 1920s, Gordon Dobson began to establish a worldwide network of stations to monitor the distribution of O_3 in the stratosphere. So-called Dobson and Brewer *spectrometers* have been installed at the ground to observe continuously the total amount of ozone in the atmosphere; later balloon-borne *radiosondes* were used to measure the vertical ozone distribution. In 1970, *satellite*-based instruments started to detect O_3 from space, providing a consistent global view of the O_3 distribution and its temporal evolution including short- and long-term fluctuations and trends. Among other purposes, such data are required to evaluate *numerical models*, helping to achieve a better understanding of relevant *atmospheric processes* (Dameris et al. 2005). An outstanding task is still to combine multiyear observations derived from different sensors in a way that provides consistent and homogeneous data products enabling solid scientific investigations of processes causing the basic state of the atmosphere and its variability. Satellite measurements of the O_3 amount are systematically compared with ground-based measurements; differences are typically lower than 1%. Nevertheless, measurements of O_3 data from different space-borne instruments show spatial and temporal differences due to sensor-specific characteristics and drifts. Therefore corrections are needed before data from different satellite instruments can be merged to long-term homogenous data records useful for ozone trend studies (Loyola et al. 2009).

Figure 45.1 indicates the evolution of the *ozone hole* as measured by different satellite sensors in the *Antarctic* from 1979 to 2010. The average total ozone *column density* from October 1–3 is shown for all years with the exception of 1993 and 2002 where data from September 23–25 are used. In 1993 no satellite data were available at the beginning of October, and in 2002 the data from September are plotted to show the atypical split of the ozone hole due to the unusual meteorological conditions in the stratosphere occurring only in that year. An ozone hole is said to exist when the *total ozone column* sinks to values below 220 *Dobson units* (DU), which is around 30 % under the norm. Dobson units are column densities, a measure of the total ozone amount in a column over a specific place. At standard temperature and pressure (STP; 1 000 hPa, 0 °C), a 0.01 mm thick ozone layer corresponds to 1 DU. A 300 DU thick ozone layer at Earth's surface would thus correspond to a pure ozone column of 3 mm.

Corresponding results for the *Arctic* in northern hemisphere spring are presented in Fig. 45.3. There, average total ozone columns from March 25–27 are plotted for all years between 1979 and 2011, except 1995 for which no satellite data are available. Obviously the ozone depletion is not as strong as in the southern hemisphere and the trend towards a lower amount of O_3 is much less visible. *Interannual* variability is high, which can be explained by the internal variability of stratospheric *dynamics* in the northern hemisphere (Dameris et al. 1995). Nevertheless, most clearly seen in years like 1997 and 2011, the dynamic situation of the Arctic stratosphere can be very similar to the Antarctic's, showing a well-pronounced and undisturbed *polar vortex* in winter with temperatures low enough to form PSCs to a large extent. On the other hand, in years like 1998 and 2010 when stratospheric temperatures are higher due to disturbed stratospheric dynamic conditions, total ozone values are enhanced. It is also obvious that total ozone column values at low latitudes are naturally low.

In addition to detailed analyses of measurements, numerical models are used to reproduce as well as possible recent atmospheric conditions and their modulation in space and time. Sensitivity studies help to identify those processes important for describing climatological mean conditions as well as spatial and temporal changes. For example, changes in climate, the temporal evolution of the ozone layer and the connections between them are simulated by atmospheric models which consider all known and relevant dynamic and chemical processes. In such numerical studies it is important to consider natural processes and their variations as well as human activities affecting atmospheric processes. Comprehensive evaluation of data derived from numerical model simulations together with the respective observations is required. This serves to identify the strengths and weaknesses of the applied model systems which to a large part reflect the current state of knowledge about processes acting in Earth's atmosphere. A good understanding of all crucial processes is a prerequisite for reliable estimates of the future development of the ozone layer (Loyola et al. 2009).

Since the early 1990s a coupled *Chemistry-Climate Model* (CCM) has been developed at the DLR Institute of Atmospheric Physics in joint collaboration with the Max-Planck-Institutes for Meteorology in Hamburg and for Chemistry in

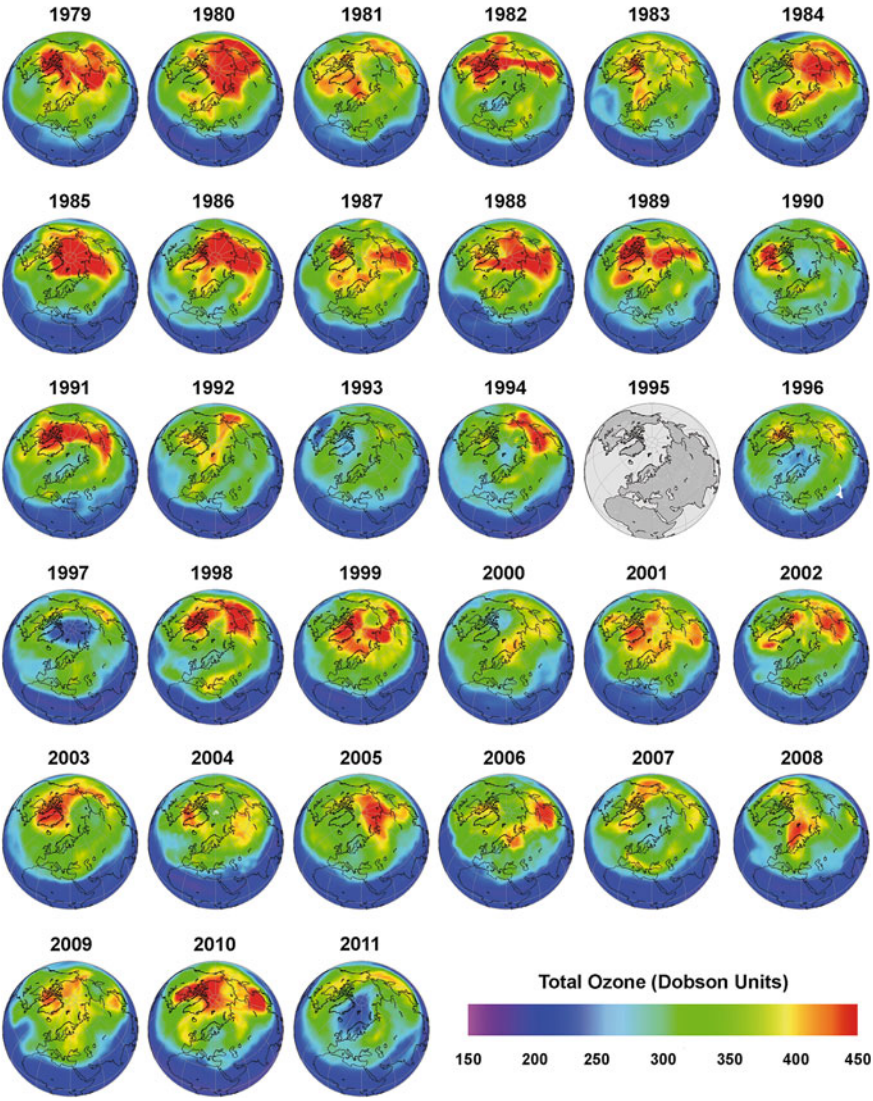

Fig. 45.3 As Fig. 45.1, but for the northern hemisphere and using a different color scale. Evolution of total ozone derived from satellite measurements in spring from 1979 until 2011 (see text; no data available for 1995). (Fig. 5 in Dameris and Loyola 2011)

Mainz (Steil et al. 1998), initially for understanding the impact of *aviation emissions* on ozone. Over the years this model has been used for several process-oriented investigations, providing a better understanding of dynamic and chemical processes and respective feedback mechanisms. At the end of the 1990s, CCMs became mature enough for first tentative estimates of the potential future evolution of the ozone layer (Dameris et al. 1998). Among others, the DLR CCM was one of

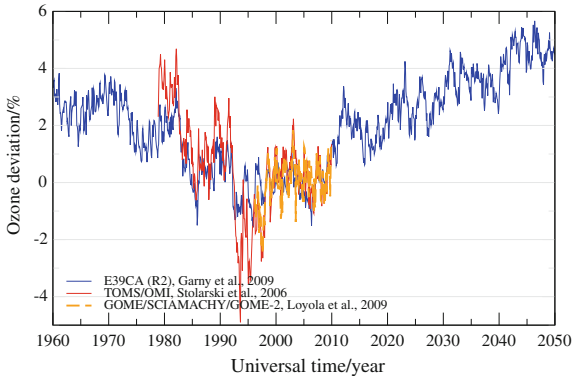

Fig. 45.4 Average anomalies of the total ozone column (in %) for the region between 60°N and 60°S. The mean annual cycle for the period 1995–2004 was subtracted in each case. The *orange* and *red curves* represent data obtained from *satellite instruments*. The *blue curve* shows results from numerical simulations using the CCM of DLR (i.e., E39CA). (Fig. 10 in Dameris and Loyola 2011)

the first models used to predict the possible future evolution of the ozone layer and how the expected onset of *recovery* is affected by the 11-year activity *solar cycle* (Dameris et al. 2006). Figure 45.4 provides a recent example of a comparison of data from multiyear satellite observations and a multidecadal simulation performed with an updated version of the DLR CCM (Stenke et al. 2009). Looking at the model results for past atmospheric conditions, it is obvious that the DLR CCM is able to reproduce seasonal and interannual fluctuations in a sufficient manner. Natural processes, such as the 11-year solar activity cycle and large volcanic eruptions (*Agung* in 1963; *El Chichon* in 1982; *Pinatubo* in 1991) cause a clear change in ozone layer thickness. The anomaly amplitudes are slightly underestimated; for example the minimum values observed in the years 1993–1995, which are caused by the eruption of Pinatubo volcano, are not adequately reproduced. The overall decrease in the amount of stratospheric O_3 until about 2000 is a direct consequence of the decrease of stratospheric chlorine content which is prescribed in the model. The close similarity between the changes in the thickness of the ozone layer derived from simulations and observations proves the good quality of the numerical simulation. The CCM properly takes into account the key factors influencing the stratospheric ozone layer.

45.5 Future Ozone Developments and Consequences

As a result of the Montreal Protocol and its amendments, tropospheric CFC content has declined since the mid-1990s. Consequently, a slight decrease in stratospheric chlorine concentrations has also been detected for several years now. However, due to the long lifetimes of CFCs in the atmosphere, it will take until the

middle of this century for the stratosphere's chlorine content to return to values resembling those observed in the 1960s. Therefore, the ozone layer's thickness should start to build up again and the ozone hole over Antarctica should close up with time. However, the speed at which the ozone layer will rebuild in future also depends on a range of other factors.

Due to further increasing *greenhouse gas* (GHG) concentrations, global atmospheric temperatures will continue to change over the coming decades: the troposphere will continue to warm up (*greenhouse effect*) and it is expected that the stratosphere will cool down further due to radiation effects; Chap. 4 in WMO (2011). Owing to *climate change*, it is highly unlikely that the ozone layer will return to exactly the state it was before the time of increased concentrations of O_3-depleting substances. Exact *predictions* of the extent of the expected *trends* are difficult to make, since future GHG emissions are uncertain. In addition, it must also be taken into account that, due to the expected build-up of the ozone layer, stratospheric ozone *heating rates* (*absorption* of solar UV radiation by O_3) will increase again, to some extent counteracting the increased cooling due to rising GHG concentrations. However, as the O_3 concentration depends largely on the background temperature, there will be some feedback. Since climate change also involves a change in stratospheric dynamics, "dynamic" heating of the stratosphere can also occur, depending on the time of year and place, leading to local heating, rather than cooling. It is obvious that estimates of future stratospheric O_3 concentration developments are not trivial and bring uncertainties with them. Due to the interactions described above, a further cooling of the lower polar stratosphere could lead to a delay in the ozone hole's closing up above Antarctica, while the ozone layer could rebuild faster in other areas of the stratosphere.

Changes in climate and the ozone layer can be estimated by means of climate models and CCMs. In numerical studies, it is important to take into account both natural processes and their variations and human interventions affecting atmospheric processes. Important factors to consider are changes in *solar activity* and *eruptions* of large *volcanoes* on the one hand, and the atmosphere's altered *chemical composition* due to industrial processes on the other.

CCMs consistently show that the ozone layer will build up again in the next decades (SPARC CCMVal 2010). CCM simulations indicate that the regeneration of the ozone layer will be faster in some areas than others, where it is quite possible that the recovery will be delayed. A continued increase in GHG concentrations will cause the stratosphere to cool further, resulting in faster rebuilding of O_3 in the middle and upper stratosphere. In the lower polar stratosphere, the re-creation of the ozone layer will slow down during spring. There, lower temperatures lead to an increased formation of PSCs. These results confirm that the regeneration of the ozone layer will vary from region to region and does not represent a simple reversal of the depletion observed over recent years. Full recovery of the ozone layer is expected by the middle of the century. Due to climate change, it seems possible that the thickness of the ozone layer will exceed that of earlier years from then on (Fig. 45.4). The results of many CCM simulations are uniform in this conclusion, even if the time that this "*super-recovery*" of the ozone layer will take place is forecast differently. The DLR CCM shows that climate-change-related processes will result in an accelerated recovery of the ozone

layer overall. O_3 values of a magnitude similar to those measured in the 1960s are predicted to occur even before the middle of the century, and the ozone layer in the middle of the century will be thicker than it was in the period 1960–1970, a time when the massive stratospheric O_3 depletion due to CFCs had not yet shown a significant impact.

Prognostic studies regarding the future evolution of the ozone layer contain a number of *uncertainties*. Some of them arise from the numerical models themselves, others from the future development of stratospheric H_2O concentrations, which is currently highly uncertain. It is predicted that tropospheric H_2O concentrations will increase with increasing temperatures in the troposphere, which could also increase the amount of H_2O being transported into the lower tropical stratosphere. Estimates based on numerical studies indicate that higher H_2O concentrations would increase the amount of hydroxides (HO_x) in the stratosphere, which, in turn, would influence depletion of O_3. Higher H_2O concentrations would lead to an increased PSC-forming potential in the polar regions during winter and thus influence O_3 chemistry. Since H_2O is a key greenhouse gas, changes in its stratospheric concentration would affect the atmosphere's *radiation balance*. Stratospheric H_2O concentrations would also increase with rising *methane* (CH_4) concentrations (due to the oxidation of CH_4), which would increase O_3 production in the lower stratosphere. On the other hand, rising CH_4 concentrations would bind reactive chlorine in the atmosphere. However, CH_4 is also an important greenhouse gas. So, higher atmospheric CH_4 concentrations would influence both climate and atmospheric chemistry. Changes in emissions of both NO_x and non-methane-*hydrocarbons* influence tropospheric concentrations of the *hydroxyl radical* (OH) and thus affect the lifetimes and concentrations of stratospheric trace gases such as CH_4. Finally, a further rise in atmospheric *laughing gas* or *nitrous oxide* (N_2O) concentrations (e.g., due to the production and use of artificial fertilizer) would increase the amount of *nitrogen oxides* (NO_x), decreasing the O_3 content of the middle and upper stratosphere (Ravishankara et al. 2009); N_2O is also a key greenhouse gas.

45.6 Summary

The concerted activities of the scientific community and political decision makers to protect the ozone layer are a success story. Based on analyses of observations and comprehensive numerical studies, quick political decisions could be made (Montreal Protocol and amendments). In the last 25 years, science programs have been initiated to better understand the ozone layer and the changes it is undergoing, new measurement techniques have been developed on different platforms, and numerical simulations of atmospheric processes have made significant progress. Scientific investigations with satellite and aircraft measurements (in situ and *lidar*) in combination with results from atmospheric *models* (mesoscale and climate-chemistry modeling) have enhanced knowledge in several respects.

Explanations of the stratospheric ozone layer's variability are complex. Natural effects such as the variability of solar radiation and particle emissions due to strong volcanic eruptions affect stratospheric ozone. Fluctuations of stratospheric *circulation* influence the stratosphere's thermal structure and air mass transport. Chemical production and depletion of O_3 are determined by photochemical processes, homogeneous gas-phase reactions and heterogeneous chemistry on the surface of particles. Understanding atmospheric processes and their interconnections is made even more difficult because atmospheric conditions change over the long term due to increased greenhouse gas concentrations. Climate change influences the amount of stratospheric O_3 and its distribution. Therefore, for *assessing* atmospheric changes, dynamic and chemical processes cannot be looked at independent of each other. Changes in climate and in the atmosphere's chemical composition are closely related, sometimes in complex ways. Nevertheless, the reasons for the formation of the ozone hole are well explained and numerical models are able to describe the recent spatial and temporal evolution of the ozone layer in a satisfactory manner. However, surprising developments cannot be ruled out in the future, either. Therefore, continuous global *monitoring* of the ozone layer and related chemical substances (e.g., N_2O and CH_4) is needed. Moreover, the synergistic use of earth observation data and model simulations will help to further close gaps in understanding. Due to many influencing factors and feedbacks, a reliable *forecast* of the future evolution of the stratospheric ozone layer is still a scientific challenge.

References

Bates, D.R., Nicolet, M.: The photochemistry of atmospheric water. J. Geophys. Res. **55**, 301–327 (1950)

Carslaw, K.S., Peter, T., Clegg, S.L.: Modeling the composition of liquid stratospheric aerosols. Rev. Geophys. **35**(2), 125–154 (1997). doi:10.1029/97RG00078

Crutzen, P.J.: Ozone production rates in an oxygen-hydrogen-nitrogen oxide atmosphere. J. Geophys. Res. **76**, 7311–7327 (1971)

Dameris, M., Nodorp, D., Sausen, R.: Correlation analysis of tropopause height and TOMS-data for the EASOE-winter 1991/1992. Beitr. Phys. Atmos. **68**, 227–232 (1995)

Dameris, M., Grewe, V., Hein, R., Schnadt, C., Brühl, C., Steil, B.: Assessment of the future development of the ozone layer. Geophys. Res. Lett. **25**, 3579–3582 (1998)

Dameris, M., Grewe, V., Ponater, M., Deckert, R., Eyring, V., Mager, F., Matthes, S., Schnadt, C., Stenke, A., Steil, B., et al.: Long-term changes and variability in a transient simulation with a chemistry-climate model employing realistic forcing. Atmos. Chem. Phys. **5**, 2121–2145 (2005)

Dameris, M., Matthes, S., Deckert, R., Grewe, V., Ponater, M.: Solar cycle effect delays onset of ozone recovery. Geophys. Res. Lett. **33**, L03806 (2006). doi:10.1029/2005GL024741

Dameris, M., Baldwin, M.P.: Impact of climate change on the stratospheric ozone layer. In: Müller, R. (ed.) Stratospheric Ozone Depletion and Climate Change, pp. 214–252. RSC Publishing, Cambridge CB40WF (2011)

Dörnbrack, A., Birner, T., Fix, A., Flentje, H., Meister, A., Schmid, H., Browell, E., Mahoney, V.: Evidence for inertia gravity waves forming polar stratospheric clouds over Scandinavia. J. Geophys. Res. **107** (D20), 8287 (2002). doi:10.1029/2001JD000452

Fabry, C., Buisson, H.: L'absorption de l'ultraviolet par l'ozone et la limite du spectre solaire. J. Phys. **3**(Série 5), 196–206 (1913)

Farman, J.C., Gardiner, B.G., Shanklin, J.D.: Large losses of total ozone in Antarctica reveal seasonal ClO_x/NO_x interaction. Nature **315**, 207–210 (1985)

Loyola, D.G., Coldewey-Egbers, M., Dameris, M., Garny, H., Stenke, A., van Roozendael, M., Lerots, C., Balis, D., Koukouli, M.: Global long-term monitoring of the ozone layer—a prerequisite for predictions. Int. J. Rem. Sens. **30**, 4295–4318 (2009)

Molina, M.J., Rowland, F.S.: Stratospheric sink for chlorofluoromethanes: chlorine atom-catalyzed destruction of ozone. Nature **249**, 810–812 (1974)

Ravishankara, A.R., Daniel, J.S., Portmann, R.W.: Nitrous oxide (N_2O): the dominant ozone-depleting substance emitted in the 21st century. Science **326**, 123–125 (2009)

Schumann, U.: Large-eddy simulation of turbulent diffusion with chemical reactions in the convective boundary layer. Atmos. Environ. **23**, 1713–1727 (1989)

Solomon, S., Garcia, R.R., Rowland, F.S., Wuebbles, D.J.: On the depletion of Antarctic ozone. Nature **321**, 755–758 (1986)

Solomon, S.: Stratospheric ozone depletion: a review of concepts and history. Rev. Geophys. **37**(3), 275–316 (1999)

SPARC CCMVal: SPARC Report on the Evaluation of Chemistry-Climate Models. In: Eyring, V., Shepherd, T.G., Waugh, D.W. (eds.) SPARC Report No. 5, WCRP-132, WMO/TD-No. 1526 (2010)

Steil, B., Dameris, M., Brühl, C., Crutzen, P.J., Grewe, V., Ponater, M., Sausen, R.: Development of a chemistry module for GCMs: first results of a multiannual integration. Ann. Geophys. **16**, 205–228 (1998)

Stenke, A., Dameris, M., Grewe, V., Garny, H.: Implications of Lagrangian transport for simulations with a coupled chemistry-climate model. Atmos. Chem. Phys. **9**, 5489–5504 (2009)

Voigt, C., Schreiner, J., Kohlmann, A., Zink, P., Mauersberger, K., Larsen, N., Deshler, T., Kröger, C., Rosen, J., Adriani, A., et al.: Nitric acid trihydrate (NAT) in polar stratospheric clouds. Science **290**, 1756–1758 (2000)

Voigt, C., Schlager, H., Luo, B.P., Dörnbrack, A., Roiger, A., Stock, P., Curtius, J., Vössing, H., Borrmann, S., Davies, S., et al.: Nitric acid trihydrate (NAT) formation at low NAT supersaturation in polar stratospheric clouds (PSCs). Atmos. Chem. Phys. **5**, 1371–1380 (2005)

Volkert, H., Intes, D.: Orographically forced stratospheric waves over northern Scandinavia. Geophys. Res. Lett. **19**, 1205–1208 (1992)

Wirth, M., Renger, W.: Evidence of large scale ozone depletion within the Arctic polar vortex 94/95 based on airborne LIDAR. Geophys. Res. Lett. **13**, 813–816 (1996)

Wirth, M., Tsias, A. Dörnbrack, A., Weiß, V., Carslaw, K.S., Leutbecher, M., Renger, W., Volkert, H., Peter, T.: Model-guided Lagrangian observation and simulation of mountain polar stratospheric clouds. J. Geophys. Res. 104 (D19) (1999). doi:10.1029/1998JD100095

WMO (World Meteorological Organization): Scientific Assessment of Ozone Depletion: 2010. Global Ozone Research and Monitoring Project, Report No. 52, Geneva (2011)

How Good are Chemistry-Climate Models?

46

Veronika Eyring, Pierre Friedlingstein, Heidi Huntrieser, Theodore G. Shepherd and Darryn W. Waugh

Abstract

Chemistry-climate models (CCMs) are important tools for simulating stratospheric ozone changes that arise from changes in ozone-depleting substances and greenhouse gases. Interest in providing reliable future projections of stratospheric ozone and its impact on climate has created a need to address how well the related key processes are represented in CCMs. The Chemistry-Climate Model Validation (CCMVal) activity is a response to this need. Here, the key processes and observations identified by CCMVal are

V. Eyring (✉) · H. Huntrieser
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Veronika.Eyring@dlr.de

H. Huntrieser
e-mail: Heidi.Huntrieser@dlr.de

P. Friedlingstein
University of Exeter, Harrison Building, Streatham Campus,
North Park Road, Exeter, EX4 4QF, UK
e-mail: P.Friedlingstein@exeter.ac.uk

T. G. Shepherd
Department of Physics, University of Toronto, 60 St. George Street,
Toronto, ON M5S 1A7, Canada
e-mail: tgs@atmosp.physics.utoronto.ca

D. W. Waugh
Department of Earth and Planetary Sciences, Johns Hopkins University,
320 Olin Hall 3400 N. Charles Street, Baltimore, MD 21218, USA
e-mail: waugh@jhu.edu

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_46,

reviewed. Results are presented to demonstrate the importance of a process-oriented and quantitative global model evaluation. In the future, similar evaluation strategies will be needed for Earth system models.

46.1 Introduction

Stratospheric ozone has been depleted over the last decades of the twentieth century until present. This is of concern since it elevates surface ultraviolet radiation, thereby posing a threat to humans and the Earth's biosphere (WMO 2011). Scientific evidence that *anthropogenic* emissions of halogenated *chlorine* and *bromine* species are responsible for stratospheric ozone depletion has led to the *Montreal Protocol*. Observations show that the protocol and its amendments and adjustments have been successfully implemented. The stratospheric *halogen* loading, with a peak in the late 1990s, is now decreasing. This slow decline is expected to continue through the *twenty-first century* in future emission *scenarios* consistent with current protocol restrictions. In response, the stratospheric ozone layer is expected to recover (WMO 2011). On the other hand, atmospheric concentrations of long-lived *greenhouse* gases such as *carbon dioxide* (CO_2) are expected to increase further in the future (IPCC 2007). CO_2-induced cooling of the upper and middle *stratosphere* may enhance ozone in this region due to slower gas-phase ozone loss cycles. In addition, greenhouse gas induced changes in the meridional circulation and increases in *nitrous oxide* (N_2O) and *methane* (CH_4) may impact stratospheric ozone (WMO 2011). The importance of these factors varies with region and time. Thus the evolution of stratospheric ozone in the *twenty-first century* also varies with region.

Considerable interest in the interaction between *ozone recovery* and *climate change* therefore still exists. To project the future evolution of stratospheric ozone and to attribute its behavior to different forcings, three-dimensional coupled *Chemistry-Climate Models* (CCMs) are needed. In CCMs chemical reactions drive changes in atmospheric *composition* which in turn change the atmospheric radiative balance and hence atmospheric dynamics. They produce a range of results concerning the timing and extent of *ozone layer* recovery. Interest in reducing these *uncertainties* and improving future *projections* has created a need to address how the key *processes* that determine the long-term behavior of ozone are represented in the models (Eyring et al. 2005).

The *Chemistry-Climate Model Validation* (CCMVal) Activity of the WCRP's (*World Climate Research Programme*) SPARC (*Stratospheric Processes And their Role in Climate*) project is a response to the need to quantitatively assess the confidence that can be placed in the models by a comprehensive evaluation of the ability of CCMs to represent key processes of stratospheric ozone and its impact on climate. To eliminate many of the uncertainties in the conclusions of earlier multi-CCM evaluations that resulted from differences in anthropogenic and natural forcings, CCMVal has defined reference *simulations* for the past and for the future.

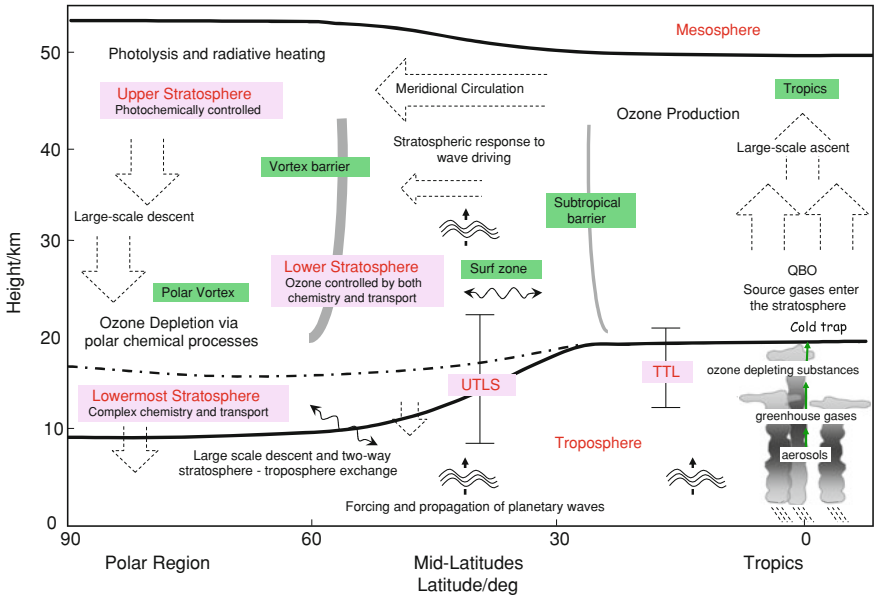

Fig. 46.1 Schematic diagram of the principal regions and processes in the winter hemispheric stratosphere. *Broad arrows* denote diabatic circulation and *wavy arrows* denote transport along *isentropic* surfaces. The average position of the tropopause is shown by the lower *thick black line*, the average position of the *stratopause* by the upper *thick black line*, and the 380 K isentropic surface by the *thick black dashed line*. The *vertical bars* denote the range of the UTLS region and the tropical tropopause layer (TTL). The key processes are described in Sect. 46.2. Modified from Fig. 1 of Eyring et al. (2005)

The first round of CCMVal (CCMVal-1) evaluated only a limited set of key processes in 13 CCMs, focusing mainly on dynamics and transport (Eyring et al. 2006). Recently, the SPARC-CCMVal (2010) report was completed; it describes the second round of CCMVal (CCMVal-2) which represents a more complete effort to assess the performance of 17 CCMs. These results provided a strong foundation for the 2006 and 2010 WMO/UNEP Scientific Assessments of Ozone Depletion (WMO 2007; 2011).

In this chapter, the key processes that determine stratospheric ozone (Sect. 46.2) along with the CCMVal concept for process-oriented model evaluation (Sect. 46.3) are first described. CCMVal was established in 2003 as an international research activity and has since then been led by the DLR Institute of Atmospheric Physics (IPA) in cooperation with other institutions. Section 46.4 presents results from CCMVal, which on the long-term has been proven to lead to significant model improvements. In addition, the extension of this concept to Earth System Models (ESMs) is discussed by showing examples for diagnostics that have been successfully applied to other components of ESMs, in particular the carbon cycle (Sect. 46.5). The role of observations in model evaluation is addressed in Sects. 46.6, and 46.7 closes with a summary and outlook.

46.2 Key Processes in the Stratosphere

A detailed description of the key processes that determine stratospheric ozone and its impact on climate is repeated here as given in Eyring et al. (2005). They are separated into processes important for *transport* (Sect. 46.2.1), *dynamics* (Sect. 46.2.2), *radiation* (Sect. 46.2.3), and *chemistry* and *microphysics* (46.2.4). For illustration, a schematic diagram showing the principal regions and processes in the stratosphere is displayed in Fig. 46.1.

46.2.1 Transport

Transport in the stratosphere involves both meridional overturning (the residual circulation) and mixing, which together represent the Brewer–Dobson circulation. The most important aspects are vertical (diabatic) mean motion and horizontal mixing. Horizontal mixing is highly inhomogeneous, with *transport barriers* in the *subtropics* and at the edge of the wintertime polar vortex; mixing is most intense in the wintertime "*surf zone*", i.e., the region surrounding the polar vortex (see Fig. 46.1), and is comparatively weak in the summertime extratropics. Accurate representation of this structure in CCMs is important for the ozone distribution itself, as well as for the distribution of chemical families and species that affect ozone chemistry (e.g., total reactive nitrogen (NO_y) inorganic chlorine (Cl_y), water vapor (H_2O), and CH_4). Within both the tropics and the polar vortex, the key physical quantities to be represented are the degree of isolation and the diabatic ascent or descent. The impact of diabatic ascent or descent on the actual vertical motion of chemical species depends on the degree of isolation. Transport in the *Upper Troposphere* and *Lower Stratosphere* (UTLS) region is also complex. The extratropical *tropopause* 'break' is a barrier to quasi-horizontal mixing, causing significant contrasts in chemical species between the extratropical *lowermost stratosphere* and the tropical upper troposphere. The *Tropical Tropopause Layer* (TTL) is marked by changes in the vertical *stability* and in chemical species. Processes in this layer are important for setting chemical boundary conditions for the stratosphere. In addition, convective processes and microphysics affect water vapor and the chemistry of ozone and other minor species. These radiatively active gases can have large impacts on the climate of the UTLS.

46.2.2 Dynamics

The basic dynamic state of the stratosphere within which transport takes place is determined by a number of physical processes. These include: the forcing mechanisms and propagation of planetary-scale Rossby and *gravity waves*, wave-mean-flow interaction, and the diabatic circulation. The *Quasi-Biennial Oscillation* (QBO) is an internal mode of variability of the atmosphere that dominates the interannual variability in wind in the tropical stratosphere and contributes to the

variability in extratropical dynamics. The QBO is important for understanding interannual variability in ozone and other constituents of the middle atmosphere, in the tropics and extratropics. Correct reproduction of the climatological mean state of the stratosphere by CCMs, including interhemispheric differences, and interannual and intraseasonal variability, is important but not sufficient: the basic dynamical mechanisms must be well represented in the underlying GCMs on which the CCMs are based if future changes are to be modeled credibly.

46.2.3 Radiation

Radiative calculations are used in CCMs to derive *photolysis* and *heating rates*. Photolysis rates in the stratosphere control the abundance of many chemical constituents that in turn control chemically active constituents, such as ozone. At the same time these trace gases have an impact on *temperature* and, thus, circulation through the radiative heating rates. There are currently not enough high quality measurements that can be used to validate the important radiative processes in global models. Presently, the best radiative models (currently not included in CCMs) provide an important complement to available measurements for CCM validation. Accordingly the approach taken (unlike the other three categories) is to perform model comparisons between the best radiative models and the radiation modules used in CCMs.

46.2.4 Chemistry and Microphysics

One of the ways in which chemistry and dynamics are coupled is through the temperature dependence of many chemical reaction rates. The importance of local control of ozone by chemistry relative to transport varies substantially between various times and places. In the upper stratosphere transport plays a role by controlling the concentrations of long-lived tracers such as inorganic chlorine, but photochemical timescales are so short that transport has a minimal direct impact on ozone. However, in the lower stratosphere, the photochemical timescales are rather longer (typically of the order of months) and interactions with dynamics are complex and more challenging to model accurately. *Aerosols* have an important role on chemistry in the lower stratosphere, since reactions can take place within or on the particles of Polar Stratospheric Clouds (PSCs). Consequently, even though the photochemical *lifetime* of ozone is typically many months in the lower stratosphere, rapid chemical loss of ozone occurs in the Antarctic lower stratosphere, following exposure of air to polar stratospheric clouds and sunlight.

46.3 Concept for Process-Oriented Model Evaluation

The goal of CCMVal is to improve understanding and performance of CCMs through process-oriented evaluation and to provide reliable projections of stratospheric ozone and its impact on climate. The CCMVal concept takes as a starting point the premise

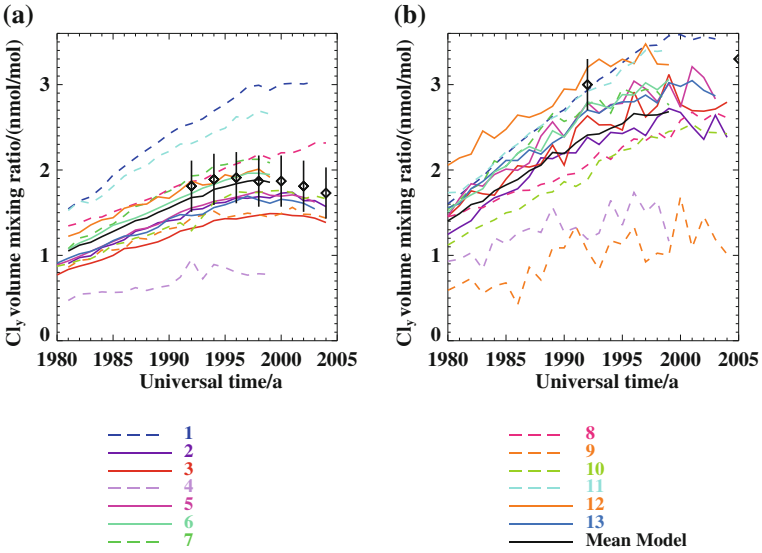

Fig. 46.2 Times series of **a** annual-mean, 35°–60°N, and **b** october-mean, 80°S Cl_y at 50 hPa from CCM simulations (*colors*) and observations (*symbols* plus *vertical bars*). Observations of Cl_y are derived from satellite data of hydrogen chloride (HCl) measured by the Halogen Occultation Experiment (HALOE) and the Microwave Limb Sounder (MLS). The numbers 1–13 refer to individual models participating in CCMVal-1 and the *black solid line* shows the multi-model mean. Modified from Fig. 1 of Waugh and Eyring (2008)

that model performance is most accurately assessed by examining the representation of key processes, rather than just the model's ability to reproduce long-term ozone *trends*, as the latter can be more easily tuned and can include compensating errors. Thus a priority is placed on high-quality observations that can be used to assess the representation of key processes in the models (SPARC-CCMVal 2010).

Each of the identified key processes described in Sect. 46.2 is assigned with model diagnostics (i.e., a quantity derived from model output) which can be used for the evaluation with relevant observations. As an example, Fig. 46.2 shows a comparison between 13 CCMs and observations for mid-latitude and polar inorganic chlorine (Cl_y) at 50 hPa, with the observational uncertainty displayed as vertical bars. Cl_y is an important process-oriented diagnostic because it is a strong indicator of chlorine induced ozone loss which has been the primary driver of stratospheric ozone depletion in particular over Antarctica. In the mid-latitude lower stratosphere, several models are within the uncertainty range of the observations (Waugh and Eyring 2008). However, substantial differences among the models and observations have been identified in polar regions. Peak Cl_y at 80°S varies from around 1 nmol/mol to over 3.5 nmol/mol in spring, but the limited measurements available to date show that peak values of Cl_y close to or less than 2.5 nmol/mol, as simulated by several CCMs, are unrealistic (Eyring et al. 2006). Eyring et al. (2007) found that models with lower peak Cl_y concentrations tend to

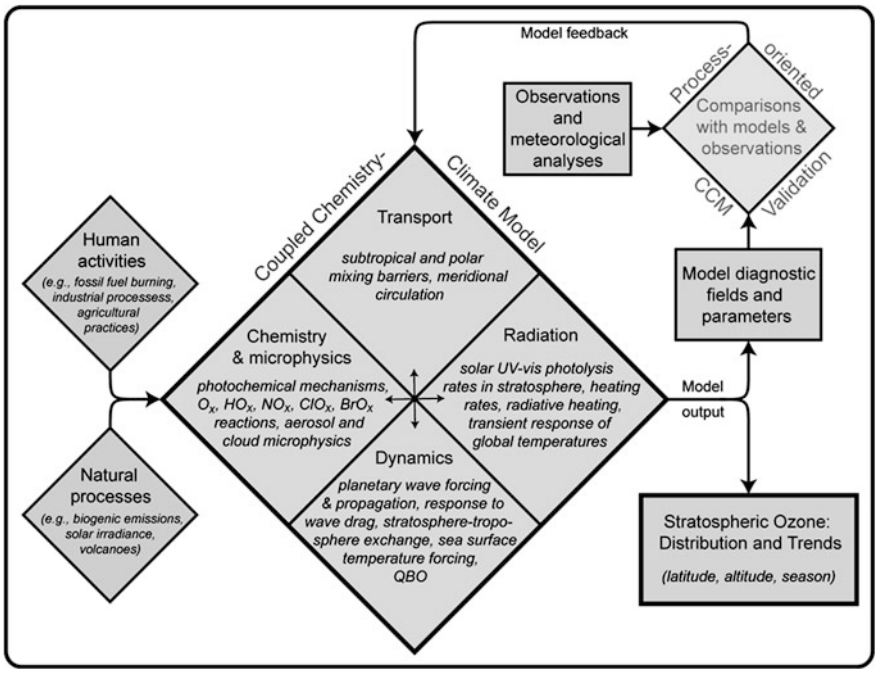

Fig. 46.3 Schematic diagram of the CCMVal evaluation approach. The centerpiece is a CCM comprised of four basic ingredients: transport, dynamics, radiation, and stratospheric chemistry and microphysics. From Eyring et al. (2005), a black-and-white version of their Fig. 3

simulate an earlier return of Cl_y to 1980 values. This result suggests that the recovery of ozone in the CCMVal-1 models might have occurred too early, and helps to better understand model projections.

The overall concept for the model evaluation in CCMVal is displayed in Fig. 46.3. The CCM output includes a wide array of diagnostics associated with the processes defined within the four ingredients (radiation, dynamics, transport, and stratospheric chemistry and microphysics). Comparisons of model diagnostics with observations and meteorological analyses are the key to process-oriented evaluation. The results of the comparisons can be used to provide *feedback* to the representation of processes in CCMs in order to improve the models on the long-term. In this way, the uncertainties in future changes in stratospheric ozone and other key model outputs can be reduced, and errors better quantified (Eyring et al. 2005).

46.4 Assessment of Chemistry-Climate Models

The approach described in Sect. 46.3 has been successfully applied to CCMs over the past few years. Some key results are summarized here to demonstrate the importance of a process-oriented and quantitative model evaluation. For further reading reference is made to the SPARC-CCMVal (2010) report.

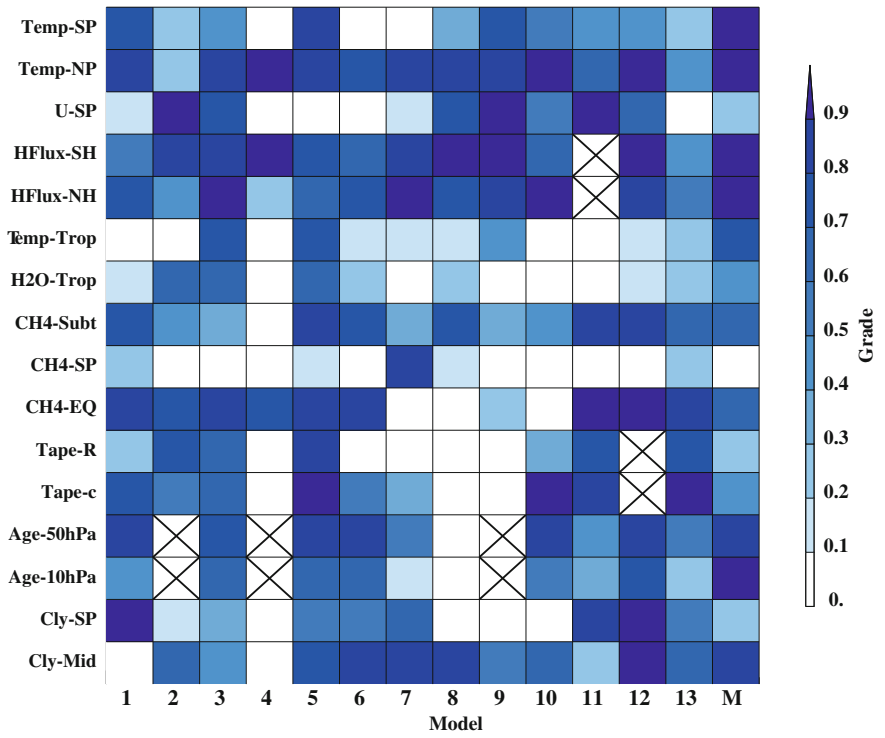

Fig. 46.4 Matrix displaying performance metrics (grades) for different diagnostics to each CCM participating in CCMVal-1. Each row shows a different metric, and each column a different CCM (indicated by numbers 1–13). The *right* most column is the "mean model". A *cross* indicates that this test could not be applied. The metrics are explained in detail by Waugh and Eyring (2008) and abbreviated as follows: south polar temperatures (Temp-SP), north polar temperatures (Temp-NP), transition to easterlies (U-SP), southern hemisphere heat flux (HFlux-SH), northern hemisphere heat flux (HFlux-NH), tropical tropopause temperature (Temp-Trop), entry of stratospheric water vapor (H2O-Trop), subtropical tracer gradients (CH_4-Subt), polar transport (CH_4-SP), tropical transport (CH_4-EQ), amplitude of water vapor tape recorder (Tape-R), phase speed of water vapor tape recorder (Tape-c), middle stratospheric mean age of air (Age-50 hPa), lower stratospheric mean age of air (Age-10 hPa), southern hemisphere polar Cly (Cly-SP), mid-latitude Cly (Cly-Mid). From Fig. 2 of Waugh and Eyring (2008)

A way to summarize model performance is to apply observationally-based *performance metrics*. A performance metric is a statistical measure of agreement between a simulated and observed quantity which can be used to assign a quantitative measure of performance to individual models (Knutti et al. 2010a). Applying quantitative metrics to a range of observationally-based diagnostics provides several benefits, including easy recognition of the models' performance for multiple aspects of the simulations, the identification of missing or incompletely modeled processes, and a quantitative *assessment* of model improvements (e.g., CCMVal-1 vs. CCMVal-2). Figure 46.4 illustrates such a comparison from

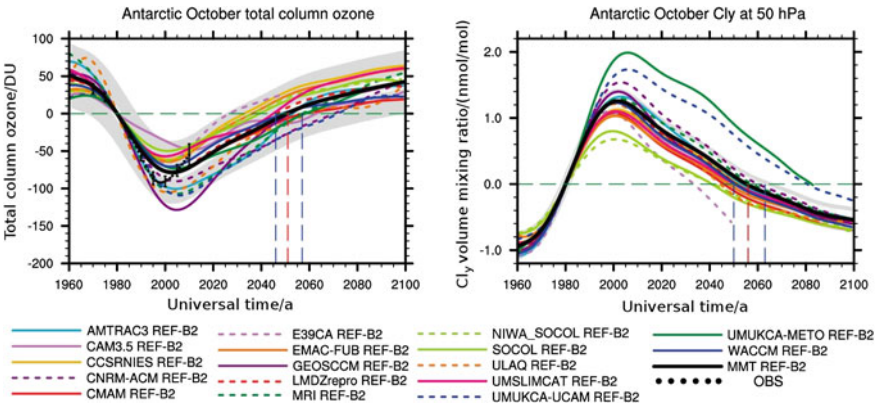

Fig. 46.5 Multi-model mean time series of total column ozone (DU; *left*) and Cl_y at 50 hPa (nmol/mol; *right*) anomalies for the latitude range 60–90°S in October (*thick dark gray line*) with 95 % confidence and 95 % prediction intervals appearing as light- and *dark-gray* shaded regions. The CCMs participating in CCMVal-2 are additionally plotted (*colored lines*). The *red vertical dashed line* indicates the year when the multi-model mean returns to 1980 values and the *blue vertical dashed lines* indicate the uncertainty in these return dates. The *black dotted lines* in the *left* panels show observed total column ozone shifted vertically so that it equals 0 in 1980. The observations include ground-based measurements, merged satellite data, the National Institute of Water and Atmospheric Research combined total column ozone database, and solar backscatter ultraviolet retrievals. Modified from Fig. 3.11 (c,d) of WMO (2011)

CCMVal-1. It summarizes the performance of the Cl_y diagnostic (see Sect. 46.3) and a set of other process-oriented diagnostics. The statistical measure applied here quantifies the difference between the climatological mean simulated by the model and the observed mean by considering the observational uncertainty, see details in Waugh and Eyring (2008). If the performance metric equals 1 (dark blue in Fig. 46.4), the simulated climatological mean is in good agreement with the observation (i.e. within observational uncertainty). Values smaller than 1 correspond to appreciable differences between model and observations. CCMs that are more than three standard deviations from the observed mean have a grade of 0 (white square), indicating large deviations from observations. The wide range of grades for all diagnostic tests shows that there are no tests where all CCMVal-1 models perform well or poorly. The majority of models do well in simulating north polar temperatures and northern and southern hemispheric heat fluxes, and, to a lesser degree, the mid-latitude age of air. However, most models perform poorly for tropical tropopause temperature, entry of water vapor, southern polar methane, and for Cl_y (see also Fig. 46.2). These differences between observations and some of the models affect the capability of the models to simulate realistic ozone distributions (Waugh and Eyring 2008).

The identification of model deficiencies in CCMVal-1 led to quantifiable improvements in some models (e.g. transport, Cl_y abundance, tropical tropopause temperatures). Compared to CCMVal-1, the simulation of Cl_y, on average, has

improved in the CCMVal-2 models due to improvements in individual CCMs. However, differences in Cl_y of up to 1 nmol/mol in the polar lower stratosphere in October still remain, see Fig. 46.5 (right panel), indicating the necessity for further model improvements. These Cl_y differences, together with differences in the dynamical processes determining the strength of the stratospheric polar vortices and temperatures, cause the spread in simulated polar ozone in the past and future (Fig. 46.5, left).

46.5 Earth System Model Evaluation

Until recently, the components of the Earth System (e.g., *ocean* dynamics, cryosphere, marine *biogeochemistry*, tropospheric and stratospheric chemistry, atmospheric dynamics and physics, terrestrial ecosystems, ecology) have been investigated separately by different disciplines. For example, in the CCMs discussed in the previous sections *sea-surface temperatures* are so far mostly prescribed from observations or climate model simulations, which hinder the feedback between chemistry-climate interactions. In addition, most of the CCMs participating in CCMVal-2 did not consider detailed tropospheric chemistry. Consequently the SPARC-CCMVal (2010) report concluded that development should continue towards comprehensive troposphere-stratosphere CCMs, which include an interactive ocean, tropospheric chemistry, a naturally occurring QBO, spectrally resolved solar *irradiance*, and a fully resolved stratosphere. The ECHAM/MESSy Atmospheric Chemistry (EMAC) model developed at DLR IPA in cooperation with other institutes participated in CCMVal-2 and is currently extended to a fully coupled ESM in response to these needs. It includes submodels describing tropospheric and middle atmosphere processes and their interaction with oceans, land and human influences. The chemical mechanism of this model is integrated in the entire model domain, i.e., consistently from the surface to the stratosphere (Jöckel et al. 2010), and the ocean can be coupled interactively (Pozzer et al. 2011).

In the move towards more complex models which include different components of the Earth System, the strategy of setting up benchmarks and criteria for process-oriented model evaluation is also important for other components of an ESM to advance understanding of the various processes and to ensure that employing such complex models in coupled mode would be beneficial. ESMs are commonly understood as extensions of conventional physical climate models with the addition of a range of interactive components adding further feedback loops. So far, models used for climate projections have been predominantly of the physical type only, i.e. so called atmosphere–ocean *general circulation models* (AOGCMs). For example, the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report included physical climate models, and only a few of the models featured in this report, the models participating in the Coupled Climate-Carbon Cycle Model Intercomparison Project (C^4MIP) and some Earth Models of

Intermediate Complexity (EMICs), explored carbon cycle feedbacks by computing CO_2 concentrations interactively (IPCC 2007). Overall, while some aspects in the simulations of current climate by AOGCMs have improved in recent years, considerable uncertainties in their projections remain (IPCC 2007; Knutti et al. 2010b; Curry 2011; Tebaldi et al. 2011). Therefore, the evaluation of processes in the AOGCM (Bony et al. 2006) is a critical aspect of ESM evaluation, since biases in the mean climate and in climate variability could directly lead to biases in ESM performance.

Using more comprehensive ESMs may also lead to increased uncertainty in future climate projections, as components (such as greenhouse gas concentrations) are not prescribed anymore. For centennial climate simulations, the inclusion of the global carbon cycle in ESMs allowed the models to be driven by anthropogenic CO_2 emissions, as opposed to atmospheric CO_2 concentration. Although this might be seen as an improvement, this introduced an additional source of uncertainty in the projections, for example for an SRES A2 emission scenario the simulated atmospheric CO_2 concentration was between 700 and 1 000 μmol/mol by 2 100 (IPCC 2007). For a given emission scenario, the simulated atmospheric CO_2 increase depends on the strength of the land and ocean carbon sinks. These latter are driven by a) the biogeochemical response of land and ocean to the increase in atmospheric CO_2, and b) the sensitivity of these sinks to climate change. None of these processes is currently simulated with a sufficient level of confidence, in particular on the land side (Friedlingstein et al. 2006). This new source of uncertainty initiated several process-oriented evaluation activities with a focus on the land carbon components of ESMs, for example the International Land Model Benchmarking (ILAMB) under the AIMES (Analysis, Integration and Modeling of the Earth System) program of IGBP (International Geosphere Biosphere Programme). ILAMB builds on previous independent evaluations of water, energy and carbon fluxes to evaluate the *Dynamic Global Vegetation Models* (DGVMs) that are now used in ESMs (Randerson et al. 2009; Cadule et al. 2010). Combining *satellite*, atmospheric, and surface datasets spanning from the recent decades to the full *twentieth* century allows model evaluation on seasonal, interannual and centennial *time scales*. The atmospheric CO_2 concentration record from Mauna Loa contains invaluable information on the global *carbon cycle*. An example for the comparison of simulated and observed CO_2 at two different stations is shown in Fig. 46.6. The trend in CO_2 is a direct measure of the strength of the land and ocean sinks (assuming past anthropogenic emissions are known); the interannual variability of the CO_2 growth rate is a direct measure of the land ecosystem's response to climate variability, while the evolution of the atmospheric CO_2 seasonal cycle could provide information on the land carbon sink and its response to increase in atmospheric CO_2 (Cadule et al. 2010). In addition, manipulative ecosystem experiments such as the ones from the Free Atmospheric CO_2 Enrichment (FACE) sites and from the soil warming sites provide direct constrain on ecosystems response to CO_2 increase and *global warming*. Recent advances in the development of DGVMs indicate that nitrogen is critical for properly simulating the ecosystem's response to change in these environmental drivers. On the

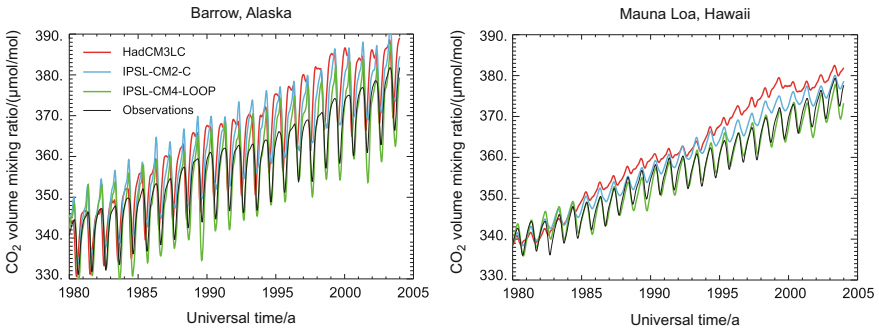

Fig. 46.6 *Left* atmospheric CO_2 concentration at Point Barrow, Alaska as simulated by the atmospheric transport of the surface fluxes from three coupled models: HadCM3LC (*red*), IPSL-CM2-C (*blue*), and IPSL-CM4-LOOP (*green*) over the historical period. Observed global CO_2 concentration, from combined ice core and atmospheric measurements, is shown in *black*. *Right* same but for Mauna Loa, Hawaii. Modified from Fig. 1 (b, c) of Cadule et al. (2010)

one hand, nitrogen limitation severely reduces the strength of the land carbon sink in response to increased atmospheric CO_2, but on the other hand, it also slightly reduces the magnitude of the land carbon feedback. Soil warming induces a remineralization of nitrogen and hence allows for more photosynthesis (Zaehle et al. 2010). However, the majority of ESMs participating in the fifth round of the Coupled Model Intercomparison Project (CMIP5) do not yet account for nutrient limitation on land, suggesting an overestimation of the land carbon sinks in CMIP5 simulations.

46.6 The Role of Observations in Model Evaluation

The role of observations in model evaluation is crucial since both the opportunities and the limitations in the available data need to be well known. A large number of observations from a variety of different platforms and instruments have been used to assess ESMs (Cadule et al. 2010; SPARC-CCMVal 2010). Model evaluation generally requires the availability of consistent, error-characterized global and regional *Earth observations* as well as accurate globally gridded reanalyses. Uncertainty in observations can be accounted for in model evaluation either by using error estimates provided with the observational data set, or by using more than one data set. If the observational uncertainty is large for a particular variable, then it is hard to discriminate models. However, for many observational data sets formal error estimates are lacking.

Recently, the earth observation community has embarked on a large effort to develop consistent data sets of numerous essential climate variables (ECVs) to monitor climate. These initiatives implement the recommendations for specific actions outlined in the Global Climate Observing System (GCOS) Implementation Plan for the UNFCCC. The variables are chosen because of their key importance

for understanding and *monitoring* the global climate system and because their systematic observation is currently technically and economically feasible. This list will develop slowly over time. The data sets are based on existing satellite data and data from new satellites recently placed in orbit. Along with ongoing efforts to improve global reanalysis products, such efforts offer an enormous potential for evaluating the next generation of chemistry-climate and *Earth system models*. In addition to sustained observations that monitor climate on a global scale, process study observations are made, which are usually more localized and cover limited time periods. Regional field experiments provide the basis for much understanding about key processes in the Earth system, for example on surface atmosphere interactions, cloud physics, and atmospheric composition. Examples include field projects such as the African Monsoon Multidisciplinary Analyses (AMMA) experiment, the Tropical Convection, Cirrus and Nitrogen Oxides Experiment (TROCCINOX) aircraft *campaign*, the HIAPER Pole-to-Pole Observations (HIPPO) of the carbon cycle and greenhouse gases, and the FLUXNET initiative on ground-based CO_2 flux measurements.

At IPA a number of field experiments have been carried out with the DLR research aircraft *Falcon* and the Russian high-flying aircraft Geophysica since the beginning of the 1990s (Huntrieser et al. 2011). Airborne in situ measurements from more than 30 projects with different focuses (e.g., aircraft and ship emissions, intercontinental transport of pollution, polar stratospheric chemistry, trace gas transport in convection, *lightning*-produced *nitrogen oxides*) are archived at IPA and have been used for model evaluation of CCMs (Ziereis et al. 1999; Meijer et al. 2001; Kurz 2006). The main target area was the European continent and the eastern North Atlantic. More recently, several field experiments were performed in tropical regions, for example TROCCINOX in Brazil, the Stratospheric-Climate Links with Emphasis on the UTLS (SCOUT-O3) campaign in Australia, and the AMMA experiment in Burkina Faso. The climate-sensitive *Arctic* region was also investigated during a number of field experiments, e.g., the Polar Study using Aircraft, Remote Sensing, Surface Measurements and Models of Climate, Chemistry, Aerosols, and Transport (POLARCAT), the Reconciliation of Essential Process Parameters for an Enhanced Predictability of Arctic Stratospheric Ozone Loss and its Climate Interactions (RECONCILE), and the Polar Stratospheric Aerosol Experiment (POLSTAR). The DLR Falcon measurements cover the altitude range between the boundary layer up to the UTLS region ($\sim$13 km). The high-flying *Geophysica* aircraft reaches heights even up to 20 km. Important trace species such as ozone (O_3), *carbon monoxide* (CO), CO_2, the ratio of nitric oxide and total reactive nitrogen (NO/NO_y) and aerosol particles are measured with a high temporal resolution ($\sim$s). In the near future, a new research aircraft for atmospheric research and Earth observations by the German science community, the High Altitude and Long Range Research Aircraft (*HALO*), will become available. With a maximum cruise altitude up to 15.5 km and a flight duration of more than 10 h, measurements of a new dimension will become available. In 2012, first field experiments such as the Transport and composition in the UTLS (TACTS)/Earth System Model Validation (ESMVal) campaign will be carried out.

The distribution of trace species over a variety of altitude and latitude ranges (subtropics to mid-latitudes), and their seasonal variety, is expected to become a unique data set for comprehensive evaluations of CCMs and ESMs.

A proper comparison between observations and models requires sampling the model output at the times and locations of the measurements and interpolating the model data to the observed vertical levels. Comparisons to satellite data should in addition consider a priori profiles and averaging kernels from the retrievals when sampling model output to, for example, calculate tropospheric columns for trace gas species. During the last few years, several satellite simulators have been developed which either involve on-line calculations or postprocessing to provide model output more directly comparable to *remote sensing* observations from satellites. Comparisons to more local measurements made for example during in situ aircraft campaigns or flux measurements at the ground exhibit the problem of a mismatch of spatial and temporal scales between observations and models. ESMs usually run at horizontal resolutions of a few 100 km, whereas field experiments sample local air masses. In addition, ESM simulations are free running, meaning that the meteorological situation and atmospheric dynamics do not match those observed in a particular year. A comparison to observations is thus only meaningful if longer time records are considered. A possibility to compare with in situ data is to combine different campaigns into one database with a horizontal grid comparable to that used in ESMs (Emmons et al. 2000). However, it has to be kept in mind that as aircraft campaigns are often targeted at specific events they do not necessarily provide a good representation of the mean climate or composition.

46.7 Summary and Outlook

The coupling of stratospheric chemical models with climate models has led to a new generation of models far more complex than those available when the Montreal Protocol was signed around 25 years ago. This increased complexity allows questions about future stratospheric ozone and UV radiation levels to be studied in much more detail than could be done in the past. However, the performance of these CCMs is also much harder to fully understand. Model comparison has allowed some model groups to identify and correct previously unrecognized model errors, a prerequisite for a better representation of these processes in future model configurations.

Conclusions on the performance of state-of-the-art CCMs were summarized in the executive summary of the SPARC-CCMVal (2010) report. This report concluded that the radiation schemes in the set of 17 state-of-the-art CCMs are sufficient for representing the major causes of observed temperature changes in the stratosphere and the main radiative drivers of surface climate. Chemistry schemes have been found to generally agree with benchmark schemes. Model performance in the UTLS was found to be better than might have been expected based on the spatial resolution of the models, but uncertainties in observations limit this conclusion. Common systematic errors in CCM results relate to the simulation of (1) tropical lower

stratospheric temperatures, water vapor, and transport, (2) the response to *volcanic* eruptions, and (3) details of the *Antarctic polar vortex*, the ozone hole, and lower stratospheric Cl_y. There are also some areas, where missing observations or disagreement about existing observations impede progress or *verification*. For example, CCMs consistently predict a strengthening of the *Brewer–Dobson circulation* and a decrease in the mean age of air as a result of climate change. However, there is not yet unambiguous observational evidence that the increase in the Brewer–Dobson circulation, leading to increases in tropical lower stratospheric ozone, actually occurs. This trend might be detectable in a long, stable time series of ozone observations in this region, but currently the length of the observational record is not long enough (WMO 2011). Similarly, UTLS temperature and *water vapor* trends are not fully understood. Long-term vertically resolved data sets of constituent observations in the stratosphere and the UTLS are required to further assess model behavior. This includes ozone, but also other species that can be used to diagnose transport and chemistry. As CCMs evolve towards including tropospheric chemistry, the present lack of observations in the UTLS will become a major limitation on model evaluation (SPARC-CCMVal 2010).

On the other hand, while there is a spread in the projected evolution of ozone, the CCMs agree in several important regards, so that general conclusions can be drawn and some confidence placed in their projections. For example, the SPARC-CCMVal (2010) report concluded that CCMs consistently project (1) a partial recovery of tropical ozone followed by a decrease in the second half of the *twenty-first century* such that tropical column ozone is predicted not even to return to 1980s values within this century, (2) a steady increase in northern mid-latitude and polar ozone such that 1980s values will be exceeded well before halogens return to 1980s values, and (3) a slow recovery of southern mid-latitude and polar ozone, with mid-latitude ozone returning to 1980 values slightly before halogens do, and polar ozone returning roughly in line with halogens. Major contributors to these changes include the recovery of ozone from *ozone-depleting substances*, the strengthened Brewer–Dobson circulation, and the cooling of the upper stratosphere.

Overall, the CCMVal studies have demonstrated the advantage of a multi-model process-oriented evaluation strategy. In the future, closer links with the tropospheric chemistry and climate community will be established, partly in collaboration with IGBP and the Working Group on Numerical Experimentation (WGNE)/Working Group on Coupled Modeling (WGCM) metrics panel. Until recently, detailed tropospheric and stratospheric chemistry have been treated separately. In addition, chemical models are evolving towards ESMs, which include the interaction between the atmosphere, the ocean and the biosphere. The emission of greenhouse gases, aerosols and *ozone precursors* by *vegetation* and *soils* and the response of these emissions to a changing climate are now recognized as key issues in the interaction of chemistry and climate. A detailed ESM evaluation would provide a means to identify possible model weaknesses and increase the level of understanding of uncertainties affecting climate predictions. Improved understanding of the key processes and feedback mechanisms is needed if credible ESM projections of climate change and its implications for human activities are to

be made. Ongoing international model intercomparison projects (see a catalogue at http://www.wcrp-climate.org/wgcm/projects.shtml) and the DLR Earth System Model Validation (ESMVal) project will contribute to improvements in this area. In view of the importance of reducing uncertainty in climate projections, these efforts should be accorded high priority. In addition, efforts to improve global reanalysis and observational products will offer an enormous potential for evaluating the next generation of ESMs and should be further pursued. Examples include the SPARC data initiative, the Satellite Application Facilities (SAFs), the *European Space Agency's Climate Change Initiative*, the NASA observation initiative that provides observations specifically for the evaluation of ESMs participating in CMIP5, and the development of an in situ database for model evaluation as part of ESMVal.

References

Bony, S., Colman, R., Kattsov, V.M., Allan, R.P., Bretherton, C.S., Dufresne, J.L., Hall, A., Hallegatte, S., Holland, M.M., Ingram, W., et al.: How well do we understand and evaluate climate change feedback processes? J. Clim. **19**, 3445–3482 (2006)

Cadule, P., Friedlingstein, P., Bopp, L., Sitch, S., Jones, C.D., Ciais, P., Piao, S.L., Peylin, P.: Benchmarking coupled climate-carbon models against long-term atmospheric CO(2) measurements. Global Biogeochem. Cycles **24**, Gb2016 (2010). doi:10.1029/2009gb003556

Curry, J.A.: Climate science and the uncertainty monster. Abstr, Pap. Amer. Chem. S (2011). 242

Emmons, L.K., Hauglustaine, D.A., Muller, J.F., Carroll, M.A., Brasseur, G.P., Brunner, D., Staehelin, J., Thouret, V., Marenco, A.: Data composites of airborne observations of tropospheric ozone and its precursors. J. Geophys. Res. **105**, 20497–20538 (2000)

Eyring, V., Harris, N.R.P., Rex, M., Shepherd, T.G., Fahey, D.W., Amanatidis, G.T., Austin, J., Chipperfield, M.P., Dameris, M., Forster, P.M.F. et al.: A strategy for process-oriented validation of coupled chemistry-climate models. Bull. Am. Meteorol. Soc. **86**, 1117–1133 (2005). doi:10.1175/Bams-86-8-1117

Eyring, V., Butchart, N., Waugh, D.W., Akiyoshi, H., Austin, J., Bekki, S., Bodeker, G.E., Boville, B.A., Bruhl, C., Chipperfield, M.P., et al.: Assessment of temperature, trace species, and ozone in chemistry-climate model simulations of the recent past. J. Geophys. Res. **111**, D22308 (2006). doi:10.1029/2006jd007327

Eyring, V., Waugh, D.W., Bodeker, G.E., Cordero, E., Akiyoshi, H., Austin, J., Beagley, S.R., Boville, B.A., Braesicke, P., Bruhl, C., et al.: Multimodel projections of stratospheric ozone in the 21st century. J. Geophys. Res. **112**, D16303 (2007). doi:10.1029/2006jd008332

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., et al.: Climate-carbon cycle feedback analysis: results from the C(4)MIP model intercomparison. J. Clim. **19**, 3337–3353 (2006)

Huntrieser, H., Schlager, H., Lichtenstern, M., Stock, P., Hamburger, T., Holler, H., Schmidt, K., Betz, H.D., Ulanovsky, A., Ravegnani, F.: Mesoscale convective systems observed during AMMA and their impact on the NO_x and O_3 budget over West Africa. Atmos. Chem. Phys. **11**, 2503–2536 (2011). doi:10.5194/acp-11-2503-2011

IPCC: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (eds.) C. U. Press, Cambridge (2007)

Jöckel, P., Kerkweg, A., Pozzer, A., Sander, R., Tost, H., Riede, H., Baumgaertner, A., Gromov, S., Kern, B.: Development cycle 2 of the Modular Earth Submodel System (MESSy2). Geosci. Model Dev. **3**, 717–752 (2010). doi:10.5194/gmd-3-717-2010

Knutti, R., Abramowitz, G., Collins, M., Eyring, V., Gleckler, P.J., Hewitson, B. Mearns, L.: Good Practice Guidance Paper on Assessing and Combining Multi Model Climate Projections. Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Assessing and Combining Multi Model Climate Projections. In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Midgley, P.M. (eds.) IPCC Working Group I Technical Support Unit, University of Bern, Bern (2010a)

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., Meehl, G.A.: Challenges in combining projections from multiple climate models. J. Clim. **23**, 2739–2758 (2010b). doi:10.1175/2009jcli3361.1

Kurz, C.: Entwicklung und Anwendung eines gekoppelten Klima-Chemie-Modellsystems, Ludwig-Maximilians-Universität München, pp. 142 (2006)

Meijer, E.W., van Velthoven, P.F.J., Brunner, D.W., Huntrieser, H., Kelder, H.: Improvement and evaluation of the parameterisation of nitrogen oxide production by lightning. Phys. Chem. Earth. C **26**, 577–583 (2001)

Pozzer, A., Jöckel, P., Kern, B., Haak, H.: The atmosphere-ocean general circulation model EMAC-MPIOM. Geosci. Model Dev. **4**, 771–784 (2011). doi:10.5194/gmd-4-771-2011

Randerson, J.T., Hoffman, F.M., Thornton, P.E., Mahowald, N.M., Lindsay, K., Lee, Y.H., Nevison, C.D., Doney, S.C., Bonan, G., Stockli, R., et al.: Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models. Glob. Change Biol. **15**, 2462–2484 (2009). doi:10.1111/j.1365-2486.2009.01912.x

SPARC-CCMVal: SPARC Report on the Evaluation of Chemistry-Climate Models.In: Eyring, V., Shepherd, T.G., Waugh, D.W. (eds.) SPARC Report No. 5, WCRP-132, WMO/TD-No. 1526. http://www.atmosp.physics.utoronto.ca/SPARC (2010)

Tebaldi, C., Arblaster, J.M., Knutti, R.: Mapping model agreement on future climate projections. Geophys. Res. Lett. **38**, L23701 (2011). doi:10.1029/2011gl049863

Waugh, D.W., Eyring, V.: Quantitative performance metrics for stratospheric-resolving chemistry-climate models. Atmos. Chem. Phys. **8**, 5699–5713 (2008)

WMO: Scientific Assessment of Ozone Depletion: 2006, World Meteorological Organization, Global Ozone Research and Monitoring Project, Geneva, 50, pp. 572 (2007)

WMO: Scientific Assessment of Ozone Depletion: 2010, World Meteorological Organization, Global Ozone Research and Monitoring Project Geneva, Switzerland, 52, pp. 516 (2011)

Zaehle, S., Friedlingstein, P., Friend, A.D.: Terrestrial nitrogen feedbacks may accelerate future climate change. Geophys. Res. Lett. **37**, L01401 (2010). doi:10.1029/2009gl041345

Ziereis, H., Schlager, H., Schulte, P., Kohler, I., Marquardt, R., Feigl, C.: In situ measurements of the NO_x distribution and variability over the eastern North Atlantic. J. Geophys. Res. **104**, 16021–16032 (1999)

Methane Modeling: From Process Modeling to Global Climate Models

47

Andrea Stenke, Rudolf Deckert and Klaus-Dirk Gottschaldt

Abstract

Methane is a long-lived greenhouse gas playing an important role in global climate change. Since preindustrial times atmospheric CH_4 concentrations have more than doubled. The global CH_4 budget is determined by the balance between surface emissions and different sink processes. Model projections of future atmospheric CH_4 concentrations are complicated by uncertainties not only in the strength of individual CH_4 sources, but also by uncertainties in other trace species affecting the major CH_4 loss reaction.

47.1 Introduction to Atmospheric Methane

During the last few years, increasing attention in climate research has been given to *methane* (CH_4), the second most important long-lived greenhouse gas (GHG) after carbon dioxide (CO_2). Ice core records show that atmospheric CH_4

A. Stenke (✉)
Institute for Atmospheric and Climate Science, ETH Zürich,
Universitätstrasse 16, 8092 Zürich, Switzerland
e-mail: andrea.stenke@env.ethz.ch

R. Deckert · K.-D. Gottschaldt
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Rudolf.Deckert@dlr.de

K.-D. Gottschaldt
e-mail: Klaus-Dirk.Gottschaldt@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_47,

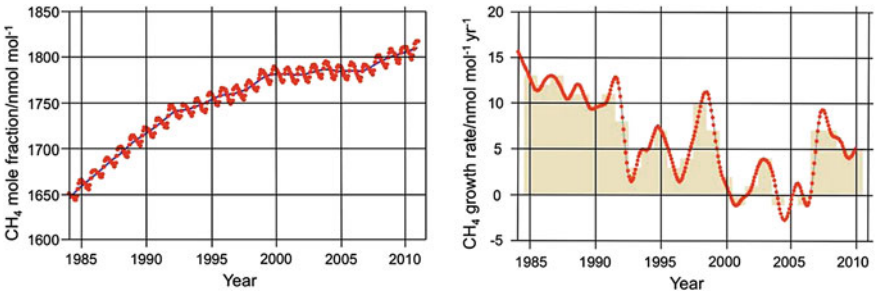

Fig. 47.1 Global average abundances of CH_4 at the Earth's surface from the NOAA global air sampling network (*left*) and its growth rate (*right*) plotted since 1984 (based on WMO 2011). The columns at (*right*) show the annually averaged growth rate. During the late 1980s the accumulation of CH_4 in the atmosphere has declined to nearly steady-state. Since 2007 global abundances of CH_4 have been rising again

concentrations have more than doubled over the past 200 years, from around 700 nmol/mol up to presently about 1 800 nmol/mol (Khalil 2000). Especially in the last 25 years the mixing ratio increased from 1 650 to 1 810 nmol/mol (Fig. 47.1, left). In historical times, atmospheric CH_4 abundances have been closely correlated with falling and rising temperatures during the transition to glacial and interglacial episodes, respectively. However, current CH_4 levels have been unprecedented, at least for the last 650 000 years (Forster et al. 2007). Despite nearly a doubling in CH_4 emissions since preindustrial times, the accumulation of CH_4 in the atmosphere has declined to nearly steady-state during the late 1980s until 2005 (Fig. 47.1, right). Since 2007 atmospheric CH_4 concentrations have been rising again (WMO 2011).

The global mean budget of atmospheric CH_4 is determined by the balance between CH_4 emissions at the surface and different removal processes. The main formation pathway of methane is the bacterial decomposition of *organic matter* under *anaerobic* conditions, e.g., in *wetlands* or in the digestive tract of animals. Methane-producing bacteria, so-called methanogens, are a special group of microbes that produce a hydrocarbon as a major catabolic product. The biogenic production of CH_4 begins with the primary fermentation of complex organic matter by non-methanogenic bacteria into a limited number of substrates that can be catabolized by methanogens to CH_4:

$$CH_3COOH \rightarrow CH_4 + CO_2 \tag{47.1}$$

$$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O \tag{47.2}$$

Furthermore, CH_4 is produced by the *combustion* of biomass material or fossil fuels. The nature and amount of gaseous *fire* emissions depend on the characteristics of both the fire and the material burned. The flaming phase of a fire is characterized by complete combustion under hot and dry conditions with sufficient

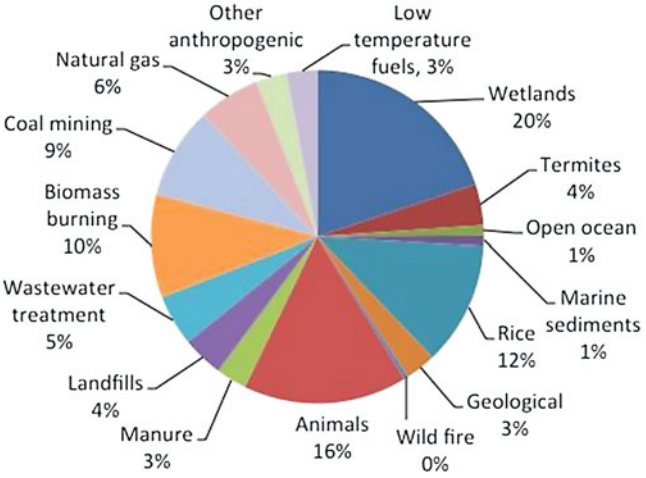

Fig. 47.2 Relative contributions of various methane sources to overall methane release. Figure based on Khalil (2000)

supply of oxygen, producing primarily CO_2, while the smoldering phase of a fire is characterized by incomplete combustion and a greater production of carbon monoxide (CO), CH_4 and non-methane hydrocarbons (NMHCs) (Khalil 2000).

The total amount of global CH_4 emissions is presently estimated as 500–600 Tg/a (Khalil 2000; Denman et al. 2007). Methane has various man-made and natural *sources* (Fig. 47.2). The major sources of CH_4 with an individual emission rate of more than 50 Tg/a are natural wetlands (25–40 % of total CH_4 emissions), *ruminants* such as cattle and sheep (16 %), and *rice agriculture* (12 %). Minor sources with an emission rate of 10–50 Tg/a include *biomass burning*, the anaerobic decomposition of waste in landfills and wastewater treatment lagoons, as well as the microbial degradation of manure, the decomposition of biomass in the digestive tract of termites, geological sources like volcanic, geothermal and hydrothermal activity, and the extraction, distribution and combustion of fossil fuels (coal and gas). Though there are still large uncertainties associated with the strength of individual CH_4 sources, it is very unlikely that there are any unknown major sources.

The major atmospheric sink of methane is the reaction with the hydroxyl *radical* (OH) in the *troposphere* ($\sim$90 %), followed by soils and the destruction of CH_4 in the stratosphere, both accounting for roughly 5 % of the total CH_4 sink (Khalil 2000). The OH radical is the most important tropospheric oxidant and is therefore also called the "detergent of the troposphere." The OH radical is formed by the *photolysis* of ozone (O_3) and the subsequent reaction of excited oxygen ($O(^1D)$) with water vapor (H_2O):

$$O(^1D) + H_2O \rightarrow 2OH \tag{47.3}$$

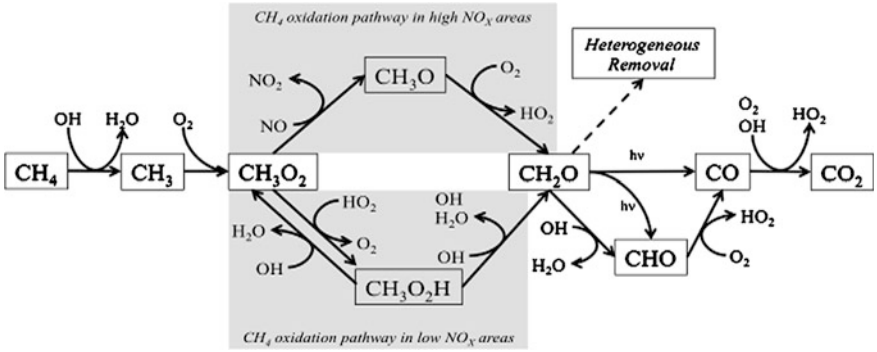

Fig. 47.3 Pathways of methane oxidation by OH in areas with high (*upper branch*) and low (*lower branch*) concentrations of nitrogen oxides (NO_x). Figure based on Khalil (2000)

The CH_4 oxidation pathway with the OH radical is initialized by the following reactions (Fig. 47.3):

$$CH_4 + OH \rightarrow CH_3 + H_2O \tag{47.4}$$

$$CH_3 + O_2 + M \rightarrow CH_3O_2 + M \tag{47.5}$$

Depending on the ambient concentrations of nitrogen oxides (NO_x, i.e., $NO + NO_2$) the subsequent reaction pathway splits into two different channels, either via the formation of the methoxy radical (CH_3O) in areas with high NO_x concentrations or via the formation of methyl peroxide (CH_3O_2H) in areas with low NO_x concentrations, both leading to the formation of formaldehyde (CH_2O). Formaldehyde is then oxidized to CO, which is finally oxidized to CO_2.

Methane chemistry has an important impact on the tropospheric oxidizing capacity. Depending on the NO_x concentrations, CH_4 oxidation leads either to a net production (high NO_x areas) or to a net destruction (low NO_x areas) of odd-hydrogen (HO_x, i.e., $OH + HO_2$). Compared to other organic compounds, the reaction of CH_4 with OH is about 100–1 000 times slower. Because of this, atmospheric concentrations of CH_4 are significantly higher than those of other organic compounds.

In the stratosphere and above, CH_4 is not only removed by the reaction with OH, but also by reactions with chlorine atoms (Cl) and $O(^1D)$ as well as by photochemical processes (Fig. 47.4). Within soils, CH_4 is oxidized by methanotrophic bacteria, resulting in the production of CO_2 and water:

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O \tag{47.6}$$

The strength of the soil sink depends strongly on the soil gas diffusivity and the activity of the methanotrophs, which are affected by soil bulk density and soil moisture.

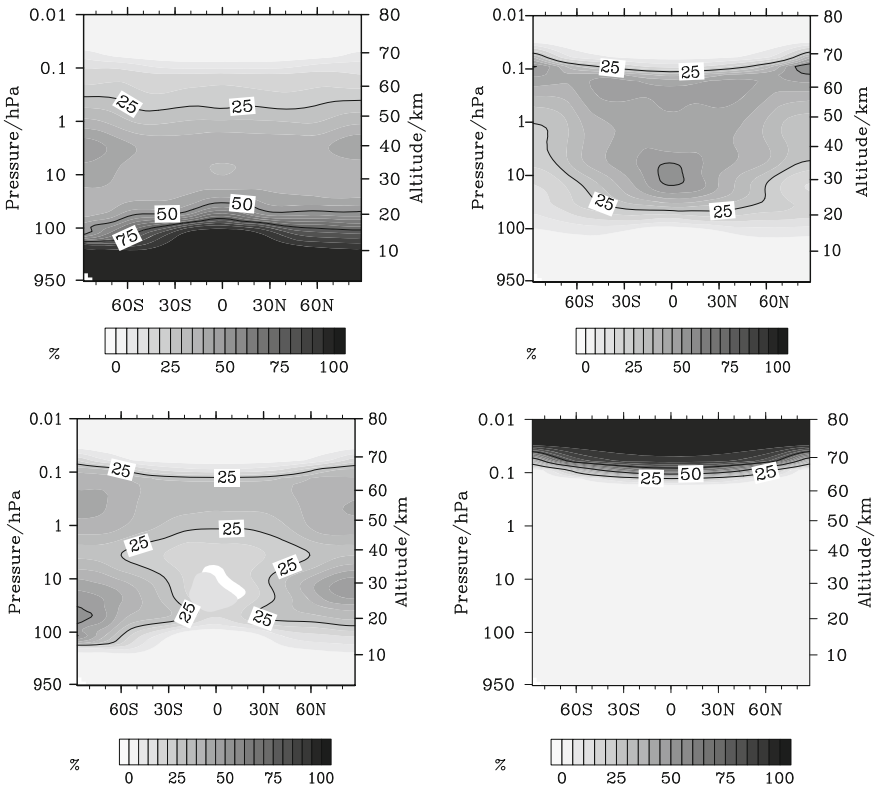

Fig. 47.4 Relative importance of different chemical loss reactions for methane, from *top left* to *bottom right*: $CH_4 + OH$, $CH_4 + O(^1D)$, $CH_4 + Cl$, photolysis of CH_4. Annual zonal mean loss (%) in a global climate chemistry model simulation with the EMAC model

The tropospheric reaction with OH leads to an atmospheric *lifetime* of CH_4 of about 9.6 years. The CH_4 lifetime is calculated by dividing the total atmospheric burden by the mean global loss rate. While CH_4 levels are rather well known, the short-lived OH-radical, the most important CH_4 sink, is hard to assess on a global scale (Jöckel et al. 2003). Recent discoveries of new reaction pathways (Butkovskaya et al. 2009; Taraborrelli et al. 2012) might shift estimates for global OH—and therefore CH_4 lifetime—by at least 10 %. Furthermore, uncertainties in the rate constant of reaction (47.4) translate to about 10–30 % uncertainty in CH_4 lifetime (IUPAC 2007; Sander et al. 2011). The two minor CH_4 *sinks* give a lifetime of 160 years (soils) and 120 years (stratosphere), resulting in an overall atmospheric CH_4 lifetime of about 8.7 years (Denman et al. 2007).

The geographic distribution of CH_4 sources leads to a strong interhemispheric gradient with higher values in the northern hemisphere. Methane concentrations vary seasonally with minimum values during summer, following the photochemical production of the OH radical, the major CH_4 sink.

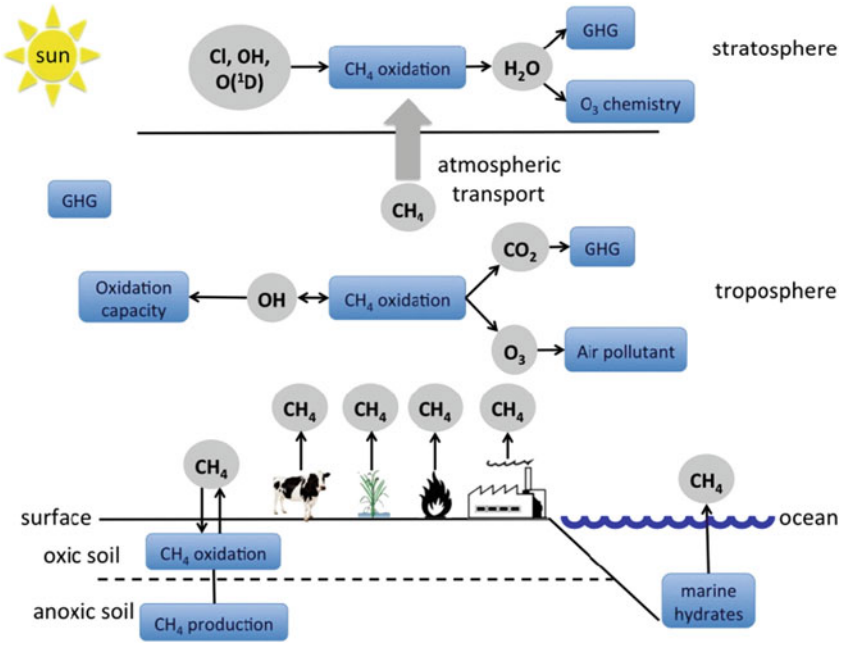

Fig. 47.5 Schematic overview of the CH_4 cycle in Earth's climate system. Methane is emitted by various sources like soils, cattle, rice paddies, wildfires, anthropogenic emissions and marine hydrates. Methane acts as a greenhouse gas in the atmosphere. In the troposphere as well as in the stratosphere CH_4 is oxidized by OH, Cl and $O(^1D)$, resulting in the formation of the greenhouse gases H_2O and CO_2 as well as the tropospheric air pollutant O_3

Changing CH_4 concentrations have the potential to affect global atmospheric chemistry and climate in several ways. Like CO_2, methane is a *greenhouse gas*: it absorbs *longwave radiation* emitted from the warm surface of the earth and emits radiation at lower atmospheric temperatures to space. The net effect is a warming of the troposphere and a cooling of the stratosphere. Methane absorbs mainly in the 7.66 μm spectral band. The strong increase in atmospheric CH_4 concentrations since pre-industrial times results in an additional *radiative forcing* of 0.48 W m^{-2}, which is the second largest radiative forcing after CO_2 (+1.66 W m^{-2}) (Forster et al. 2007). On a per unit basis CH_4 is about 25 times more effective in warming the earth globally than CO_2, i.e., CH_4 has a *global warming potential* (GWP) of 25 (on a time horizon of 100 years). The GWP describes the time-integrated global mean radiative forcing of a pulse emission of 1 kg of some gas relative to that of 1 kg of CO_2. Given this high GWP, methane's importance as a driver of climate change would grow even further if its concentrations continued to rise. On the other hand, compared with CO_2 methane has a rather short atmospheric lifetime of about 9 years, making it a promising candidate for mitigating global warming in the near future. Several options to reduce anthropogenic CH_4 emissions and their

mitigation potential have been widely discussed (Bogner et al. 2007; Smith et al. 2007; Shindell et al. 2012). However, human efforts to mitigate global warming by reducing CH_4 emissions could become inefficient if natural CH_4 emissions were to substantially increase in a warmer future climate.

Besides its direct role as greenhouse gas, CH_4 indirectly influences global climate through various chemical interactions (Fig. 47.5): CH_4 oxidation by the OH radical leads to the formation of CO_2. The amount of CO_2 produced from CH_4 oxidation is estimated to be about 5 % of the CO_2 emissions from anthropogenic fossil fuel use and cement production (Khalil 2000). In polluted areas (high NO_x) CH_4 oxidation leads to the formation of tropospheric O_3, which is a strong oxidant itself and affects human health. Additionally, in what is called the "lifetime effect," an increase in atmospheric CH_4 reduces the abundance of the hydroxyl radical (OH), the most important chemical CH_4 sink, thereby increasing the atmospheric lifetime of CH_4 and of a range of other non-CO_2 greenhouse gases. This close interaction between atmospheric OH and CH_4 constitutes a strong mutual relationship between the global atmospheric oxidation capacity and air quality on the one hand, and global climate on the other hand.

In the stratosphere, oxidation of CH_4 is the most important in situ source of H_2O, with approximately two H_2O molecules being produced for each CH_4 molecule oxidized. Water vapor plays an important role in the stratosphere: Firstly, it acts as a greenhouse gas. Secondly, water vapor is a source of HO_x, which is involved in the *catalytic* destruction of *stratospheric* O_3, and thirdly, increases in stratospheric H_2O concentrations due to CH_4 oxidation or other processes are expected to enhance the formation of *polar stratospheric clouds* and therefore heterogeneous O_3 depletion. Lower stratospheric temperatures caused by the radiative effect of increasing GHG concentrations could further enhance the formation of polar stratospheric clouds. Furthermore, the reaction of CH_4 with the Cl radical affects the abundance of reactive chlorine and, therefore, stratospheric O_3 destruction.

Although the key processes driving atmospheric CH_4 concentrations seem to be identified, substantial uncertainties in the relative importance of single processes as well as in the interplay of the different drivers remain. The main open questions are:

- What are the reasons for the decrease in CH_4 growth rates during the 1980s and 1990s? Various studies attribute the observed slowdown (Fig. 47.1, right) to a reduction in source emissions, mainly in fossil fuels due to the economic breakdown in the former Soviet Union and Eastern Europe, plus a smaller concurrent decrease in wetland emissions during the late 1990s. Conversely, several other studies suggested a decrease in CH_4 lifetime. However, different measurements of proxy tracers for global OH do not provide indications of significant long-term changes in OH and, therefore, of changes in CH_4 lifetime (Forster et al. 2007; Montzka et al. 2011).
- What are the drivers of the large year-to-year variability of CH_4 growth rates? Two outstanding years are 1991/1992 and 1998 (Fig. 47.1, right). The increase in 1991 is most likely related to the volcanic eruption of Mount Pinatubo: The enhanced stratospheric *aerosol* loading after the eruption led to reduced UV radiation reaching the lower atmosphere, resulting in reduced photochemical

OH formation. This increase was followed by a strong slowdown of the CH_4 growth rate in 1992, which might be related to reduced CH_4 emissions from natural wetlands due to lower *surface temperatures* observed after the eruption of Mount Pinatubo. In contrast, 1998 was one of the warmest years in the instrumental surface air temperature record. Several studies argued that this peak in CH_4 growth rates might have been primarily related to thermally driven wetland emissions. Furthermore, large wildfires led to enhanced CO emissions in 1998, which also increases CH_4 lifetime (Forster et al. 2007).

- Why have atmospheric CH_4 concentrations been rising again since 2007? This renewed increase could be related to anomalously high Arctic temperatures and large amounts of precipitation in the *tropics*, resulting in increased CH_4 emissions from natural wetlands (O'Connor et al. 2010).

The future evolution of atmospheric CH_4 concentrations is difficult to predict. Obviously, future CH_4 concentrations will strongly depend on anthropogenic emissions. Without adopting suitable mitigation measures, CH_4 emissions from agriculture are expected to increase due to expected increases in food demand and diet changes as the world's population continues to grow. In the framework of the upcoming fifth assessment report of the IPCC, several emission scenarios have been developed. Depending on the particular scenario, CH_4 concentrations for the year 2100 have been estimated to range between 1250 nmol/mol and 3750 nmol/mol. However, CH_4 emissions will also respond to climate change. Wetland emissions are expected to increase by 10–35 % per °C warming. Thawing *permafrost* soils, longer ice-free periods for northern wetlands and degassing of *marine hydrates* have the potential to increase CH_4 emissions in the future. Near-surface permafrost sediments on land, which contain a significant portion of the carbon content, have already begun to thaw. Yet, the sensitivity to the future rise in temperature is uncertain. The current consensus is that the increase in CH_4 emission rates is probably going to be moderate and elevated emissions will rather occur chronically over a long time. However, soil-climate interactions are considerably more complex than anticipated and much research remains to be done. Recently, CH_4 venting into the atmosphere from the subsea permafrost of the shallow East Siberian Arctic shelf has been detected. This release is small compared to the global CH_4 budget, but on a par with estimates for the entire world ocean. This subsea permafrost layer developed on land during several glacial periods when East Siberia was not ice covered, and it displays a disturbed thermal regime already from post-glacial flooding. As a result, this huge potential source of CH_4 might be particularly sensitive to the pronounced warming of Arctic coastal waters. For a review of possible CH_4 emissions from wetlands, permafrost and hydrate in a future climate the reader is referred to (O'Connor et al. 2010).

Besides changing CH_4 sources, the main CH_4 sink via the reaction with OH might also change in a changing climate. The future development of atmospheric OH abundances depends on several factors, which include future emissions of air pollutants that react with OH and changes in the amount of UV radiation reaching the lower atmosphere. The latter is caused by changes in cloud cover and stratospheric ozone.

47.2 Modeling: From Process-Oriented Models to Global Climate Models

The spatial distribution of atmospheric *trace gases* like CH_4 is determined by the interaction of temporally and spatially varying sources and sinks, and by atmospheric transport. Different modeling approaches are available to study processes related to atmospheric CH_4.

Global three-dimensional *chemistry-transport models* as well as coupled *chemistry-climate models* (CCMs) are valuable tools in the research of atmospheric CH_4. Modeling global distributions of trace gases and comparing model results with available observations can help to identify and quantify key drivers of atmospheric CH_4. Chemistry-transport models focus on the *advection* and chemical production/loss of one or more chemical species. They are driven by prescribed meteorological fields and, therefore, do not consider *feedback* processes between atmospheric dynamics and chemistry, while CCMs take those important feedback effects into account. CCMs are, for instance, capable of dealing with the "lifetime effect" in a consistent manner. Such interactions are of particular interest in the context of global-climate projections where the aim is to quantify the climate impact from a given perturbation in atmospheric composition. For example, a perturbation in the flux of CH_4 from the ground into the atmosphere will not only affect Earth's climate via its immediate GWP. It would also act via secondary chemically-induced repercussions, altering the atmospheric concentrations of greenhouse gases such as O_3 and CH_4 itself.

The DLR Institute of Atmospheric Physics (IPA) utilizes the CCM system ECHAM/MESSy Atmospheric Chemistry (EMAC) to tackle questions about Earth's climate system. EMAC is designed to handle the feedbacks between atmospheric dynamics and chemistry comprehensively (Jöckel et al. 2010). For the global-climate projections mentioned above, a horizontal resolution of about 2.8° by 2.8° in latitude and longitude is practicable because the required computing resources still permit multidecadal simulations. In order to capture the impact of changes in CH_4 concentrations on stratospheric chemistry, this set-up is combined with a vertical resolution of 90 model levels from the surface up to the *middle atmosphere* (approximately 80 km). EMAC is applied with a mixed-layer ocean scheme, which keeps the integration time feasibly short whilst allowing for interactively changing sea surface temperatures.

Until recently, the ground source of atmospheric CH_4 in global CCM simulations was associated with the relaxation towards given concentration fields in the lowermost model layer, a so-called mixing ratio boundary condition. For the past, the prescribed concentration fields are based on atmospheric CH_4 measurements or on ice core records. For future climate projections, however, representative concentration pathways of atmospheric greenhouse gases have to be developed assuming different social, political and economic scenarios (Moss et al. 2010). Those pathways are achieved by so-called integrated assessment models, which consider different key aspects of human systems, such as population density,

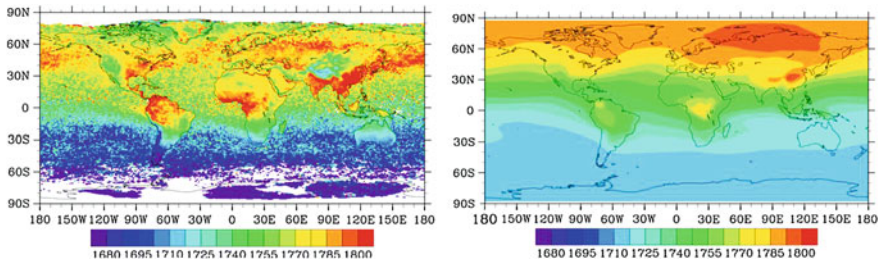

Fig. 47.6 CH_4 column-averaged mole fraction (nmol/mol) as observed with the SCIAMACHY instrument on board ENVISAT (*left*, average over the years 2003–2005) and as simulated with EMAC (*right*). The SCIAMACHY/ENVISAT data are provided by the Institut für Umweltphysik, University of Bremen. The EMAC results are only preliminary as they rely on a heavily simplified scheme of methane chemistry and do not take into account the satellite averaging kernel

energy use, economics and land use, including a simplified representation of the climate system. The mixing ratio boundary condition, however, is incapable of fully determining the interactions between CH_4 concentrations and climate. A flux boundary condition for CH_4, i.e., applying surface emission fluxes, is more realistic. In the recent past, large efforts have been made toward consistent representation of atmospheric CH_4 in CCMs. First EMAC simulations with simplified chemistry and a flux boundary condition have shown promising results (Fig. 47.6): Several features of the observed CH_4 distribution like the strong interhemispheric gradient or the high CH_4 abundances over China are well reproduced by the model. However, the shown intercomparison of the model results with CH_4 observations of the SCIAMACHY instrument on board the ENVISAT satellite also reveals several differences, which calls for a profound analysis of uncertainties in the modeling approach and also in the measurements. As a next step, a reference simulation with a comprehensive chemistry scheme including CH_4 emission fluxes is planned. This simulation will be done in "nudged mode," i.e., the model dynamics are forced into an analyzed meteorological state. In general, the nudging technique permits direct comparison of the simulated CH_4 concentrations against observations. Once the *validation* is accomplished, the road is free towards meaningful global-climate projections taking into account the interaction between atmospheric CH_4 concentrations and climate.

CCMs like EMAC are designed to handle the feedbacks between atmospheric physical and chemical processes. Yet, they cannot handle the feedbacks between global warming and changing natural sources of CH_4 and other relevant trace gases. This is an aspect of earth system modeling. To overcome this gap, CCMs are beginning to be coupled with models of the terrestrial biosphere. In order to provide EMAC with such capability, it will be coupled to a *Dynamic Global Vegetation Model* (DGVM). DGVMs are designed to represent large-scale terrestrial vegetation dynamics as well as associated land–atmosphere carbon and water fluxes. They are driven by observed or simulated meteorological input

parameters like temperature, precipitation, cloud cover and atmospheric CO_2 concentrations. The vegetation coverage is described in terms of different plant functional types which take into account the variety of structure and function among plants. More sophisticated models not only focus on carbon, but also consider the nitrogen cycle, simulate emissions of biogenic *volatile* organic compounds, and include wetland hydrology (e.g., Wania et al. (2010)).

A prerequisite for reliable model studies of interactions between atmospheric chemical processes and the climate is profound qualitative and quantitative understanding of individual emission sources and sink processes in the present climate. Estimates of individual CH_4 emission source strengths can be obtained by so-called "bottom-up" and "top-down" methods. Bottom-up approaches include a variety of techniques from integrating or up-scaling single flux measurements to process-oriented modeling. For example, CH_4 emissions by domestic ruminants are often estimated by integrating or scaling-up single laboratory measurements with available statistics of the population density and distribution of ruminants. This approach is limited by uncertainties in feeding and livestock management practices, which have a large influence on CH_4 emissions from livestock. In general, emission estimates based on up-scaling methods strongly depend on the representativeness of the available field data.

For sources with a large natural variability such as rice agriculture, natural wetlands or biomass burning, up-scaling techniques are not appropriate. In such cases, process-oriented modeling approaches ranging from local to global scale are used to estimate related CH_4 emissions. The amount of CH_4 emitted from natural wetlands depends on the water table height and the amount of available biomass that can be degraded. Process-oriented models of different complexity are available, starting from DGVMs including an empirical functional relationship between available biomass, soil hydrology and CH_4 emissions, up to DGVMs including mechanistic wetland hydrology modules. The latter even consider detailed transport processes like ebullition (degassing of CH_4 from lakes), diffusion, or plant-mediated transport of CH_4. Nevertheless, there are still large uncertainties concerning global wetland CH_4 emissions, which is partially due to uncertainties in the location of wetlands as well as their seasonal and interannual extent and inundation. Fire process modules implemented in DGVMs are becoming more capable of estimating dynamic surface emissions, but uncertainties about where fires occur also limit the accuracy of estimated trace gas emissions from fire. Some biogeochemical models have the capacity to model multiple land sources of CH_4, e.g., from wetlands, rice cultivation, and biomass burning, on a regional scale, but not yet on the global scale.

In addition to bottom-up techniques, so-called top-down techniques can be used to estimate sources. Top-down methods start from measurements of atmospheric greenhouse gases like CH_4 and use information about atmospheric transport, mixing and sink processes to determine the emissions of these gases. This information is usually provided by *numerical models*. A good overview of inverse problems in atmospheric research is given in the book of (Enting 2002). Mathematically, the inverse problem can be described as follows:

$$c \approx y = Mx \tag{47.7}$$

where c describes a vector of CH_4 measurements at different locations corresponding to the model results y. The matrix M describes the causal relationship (source-receptor relationship) between CH_4 emission fluxes (x) and simulated CH_4 concentrations (y), defined by atmospheric transport and sink processes.

Generally, the number of available measurements is a limiting factor in this approach. In the case of CH_4, the Global Atmosphere Watch Programme of the *World Meteorological Organization* and the National Oceanic and Atmospheric Administration (NOAA)/Global Monitoring Division (GMD) provide a global network of in situ sample sites. However, most of these sample sites are located in the northern hemisphere. Observations in tropical regions, the southern hemisphere and over the oceans are rather sparse. For such an inverse problem to have a single solution, either the number of emission parameters must be adjusted according to the available measurements or different constraints must be introduced. Often it is required to introduce a first guess or a priori information about the sources.

A general problem associated with the modeling of atmospheric CH_4 is limited knowledge of the global OH distribution and its temporal development, and, therefore, of the strength of the major CH_4 sink. The OH radical has a very short lifetime of about a second. OH formation strongly depends on ambient conditions such as cloudiness, humidity, solar *irradiance*, and the presence of other reactive *trace species*, making OH concentrations highly variable in space and time. Though there are a limited number of local OH measurements available, no global OH observations exist. Information about the tropospheric CH_4 sink from the reaction with OH and its temporal evolution is usually obtained by indirect methods, e.g., *inverse modeling* involving atmospheric chemicals whose sources are well established and for which OH-initiated oxidation is the dominant removal mechanism. One of these species is methyl chloroform ($C_2H_3Cl_3$), which is entirely man-made and provides good concentration data. Nevertheless, the strength of the chemical CH_4 sinks is still highly uncertain.

There exist various *inversion* techniques with different representations of the chemical CH_4 sink. All of them are somehow based on chemical-transport models, but the complexity of the chemical scheme and, therefore, the description of the chemical CH_4 sink differs, ranging from a simplified CH_4 loss reaction using prescribed OH concentrations and lumped reaction rates up to comprehensive schemes of tropospheric chemistry including hundreds of species and even more reactions.

A further approach to refine estimates of individual emission sources is the isotopic ratio (Khalil 2000). Two stable carbon isotopes and two hydrogen isotopes are relevant for CH_4. These are ^{12}C and ^{13}C, with the carbon atom's core containing either 12 or 13 neutrons, as well as ^{1}H and ^{2}H, i.e., common hydrogen and deuterium (D). The different masses of the isotopes determine their chemical kinetic behavior such that each emission source, but also each sink, is

characterized by a specific isotopic signature. The isotopic ratio is expressed as the fractional deviation of the isotopic ratio of the sample from a reference standard:

$$\delta R_X(‰) = \left[\left(\frac{R_a}{R_b}\right)_{sample} \Big/ \left(\frac{R_a}{R_b}\right)_{reference} - 1 \right] \times 10^3 \qquad (47.8)$$

where R_a/R_b are the *isotope ratios*, i.e., $^{13}C/^{12}C$ or $D/^1H$, of the sample and of the reference standard. Biogenic processes emit $^{13}CH_4$ at a very low rate compared to $^{12}CH_4$, while burning processes present the most equilibrated ratio between these two carbon isotopes. The hydrogen isotopes of methane can help to distinguish high latitude sources, such as wet tundra, versus low latitude sources, such as swamps and rice farms. The unstable carbon isotope $^{14}CH_4$ mainly has a biogenic origin. In fossil fuels or geological methane sources, ^{14}C has usually decayed below the detection limit. The implementation of such isotopic information in CCMs and subsequent comparisons with isotope measurements can be used for model validation.

47.3 Global Modeling and Interpretation of Satellite Data

Long records of temporally and spatially highly resolved measurements of atmospheric trace species are a common request of climate modelers. In the case of methane, the NOAA/GMD provides weekly flask measurements from a globally distributed network of sample sites, a service which started in the 1980s. Furthermore, information about the vertically integrated CH_4 amount in the atmosphere is available from the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) instrument aboard the ENVISAT satellite (2003–2009) (Schneising et al. 2011) and from the TANSO instrument aboard GOSAT (since 2009) (Morino et al. 2011). Both satellites are sun-synchronous orbiters, i.e., any given point of Earth's surface is passed by the satellite at the same local mean solar time. Observations from space-borne CH_4 instruments are not only desirable for model evaluation, but are probably the only way to quantify CH_4 emissions from the individual emission sectors on a global basis via inverse modeling.

The Methane Remote Sensing Lidar Mission MERLIN, a joint project of DLR and the French space agency, will yield CH_4 data of unprecedented quality due to an innovative laser-based measuring principle (Amediek et al. 2009). The mission is, however, scheduled for a period of less than 6 years and its data product will not contain information on the vertical distribution of CH_4 within the atmosphere. In this context, EMAC simulations may permit a greater spatiotemporal perspective and also contribute to tuning the height-dependent sensitivity of the space-borne measuring equipment towards an optimum quantification of CH_4 emissions.

EMAC is particularly useful for these aims. There is a mode for sampling simulated concentrations along sun-synchronous orbits (Jöckel et al. 2010), indispensable for a direct comparison of CH_4 concentrations from simulations and MERLIN data. EMAC provides a quasi-chemistry-transport mode, where the chemistry does not feed back into meteorology (Deckert et al. 2011). This mode is important when quantifying the impact of small CH_4 emission anomalies on atmospheric composition. Finally, using a horizontal resolution of $1.875° \times 1.875°$ may increase the comparability of simulations and satellite data further, but due to computational costs it is feasible only for episodic investigations.

The ground measuring network is not dense enough and time series from satellite missions—including MERLIN—are too short to properly resolve CH_4 interannual variability and trends. Therefore, multidecadal simulations with EMAC will contribute to putting the MERLIN data in a broader temporal context. Simulations will be done both in nudged mode to reproduce observed atmospheric transport, and in chemistry-climate mode to fully account for feedbacks between chemical and meteorological processes.

The MERLIN CH_4 data product comes with a significant contribution from the middle and *upper troposphere*. Additionally, contributions from the *lowermost stratosphere* might be noticeable during winter at high latitudes, when the *tropopause* descends to heights as low as 400 hPa. Furthermore, there are indications that seasonal cycles in CH_4 are different at different altitudes. At high latitudes, for instance, the near-surface seasonal cycle is mainly controlled by the balance between surface emissions and losses through the reaction with OH. In contrast, the seasonal variation of CH_4 in the upper troposphere and lower stratosphere is mainly determined by the height of the tropopause, as well as by air exchange between stratosphere and troposphere. Finally, due to the lack of observations, not much is known about the altitude dependence of interannual CH_4 variability. When interpreting height-averaged satellite data it is thus important to consider modeled CH_4 profiles. EMAC simulations will give an indication of height resolved seasonal cycles and variability of CH_4 on a global basis over a multi-decadal period.

For measurements from space-borne instruments, sensitivity to atmospheric CH_4 concentrations as a function of height is to a certain degree adjustable in orbit. However, the optimum choice as to the inference of methane emission rates by inverse modeling is not known. A range of EMAC simulations will yield the responses of height-resolved CH_4 concentrations to small CH_4 emission anomalies. Then, a range of different MERLIN sensitivity profiles will be employed with the responses to determine which profiles are associated with the strongest signal.

47.4 Outlook

To overcome existing uncertainties in qualitative and quantitative understanding of atmospheric CH_4 concentrations and to provide reliable model projections of the future climate, further scientific efforts will be necessary, both in terms of modeling and measurements.

Inverse models ("top-down methods") are a useful tool to determine global CH_4 emissions. However, insufficient observational data limit these approaches. DLR-IPA contributions to the MERLIN mission will have a share in improved global-scale observational data of CH_4. Improved global data sets on OH-related quantities, global wetlands and land surface hydrology are also needed to decrease the uncertainties of top-down methods. Measurements of CH_4 isotopes can be used to further diminish uncertainties in the strength of different emission processes.

Additional detailed field measurements for a large number of different ecosystems, especially in the tropics and the southern hemisphere, are required for improving process-based models ("bottom-up methods"). This would help to determine the respective methane emission estimates, e.g., from wetlands or fires.

To understand the complex interaction between climate and the terrestrial ecosystems, earth system models that include all relevant processes, starting from wetland hydrology, permafrost dynamics and vegetation, up to atmospheric chemistry and transport are required (Fig. 47.5). Especially the implementation of permafrost soils and feedbacks with climate is becoming more and more important in view of global warming and thawing permafrost soils. Peat-land and permafrost models are starting to be implemented in land surface models and DGVMs. Different temporal and spatial scales as well as unknown feedback mechanisms, however, make the implementation of such models in chemistry-climate models challenging. Marine CH_4 hydrates pose a particular challenge to modeling. Uncertainties in the occurrence and amount of hydrates are still very large. Further field measurements of CH_4 hydrates and the development of process-based models for hydrate destabilization and the CH_4 ocean cycle will be necessary before marine hydrates can be included in large-scale earth system models. Nevertheless, reproducing the present-day CH_4 distribution with global chemistry-climate models is a prerequisite for reliable future projections.

References

Amediek, A., Fix, A., Ehret, G., Caron, J., Durand, Y.: Airborne lidar reflectance measurements at 1.57 μm in support of the A-SCOPE mission for atmospheric CO_2. Atmos. Meas. Tech. **2**, 755–772 (2009). doi:10.5194/amt-2-755-2009

Bogner, J., Abdelrafie Ahmed, M., Diaz, C., Faaij, A., Gao, Q., Hashimoto, S., Mareckova, K., Pipatti, R., Zhang, T.: Waste management. In: Metz, B., Davidson, O.R., Bosch, P.R., Dave, R., Meyer, L.A. (eds.) Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge (2007)

Butkovskaya, N., Rayez, M.-T., Kukui, A., Le Bras, G.: Water vapor effect on the HNO_3 yield in the HO_2+ NO reaction: experimental and theoretical evidence. J. Phys. Chem. A **113**, 11327–11342 (2009). doi:10.1021/jp811428p

Deckert, R., Jöckel, P., Grewe, V., Gottschaldt, K.-D., Hoor, P.: A quasi chemistry-transport model mode for EMAC. Geosci. Model Dev. **4**, (2011). doi:10.5194/gmd-4-195-2011

Denman, K. L., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P.M., Dickinson, R. E., Hauglustaine, D., Heinze, C., Holland, E., Jacob, D. et al.: Couplings between changes in the climate system and biogeochemistry. In: Solomon, S., and Coauthors (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 (2007)

Enting, I.G.: Inverse Problems in Atmospheric Constituent Transport. Cambridge University Press, Cambridge, 408 pp. (2002)

Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T. Betts, R., Fahey, D.W., Haywood, J. Lean, J. Lowe, D.C., Myhre, G., et al.: Changes in Atmospheric Constituents and in Radiative Forcing. In: Solomon, S., and Coauthors (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 (2007)

IUPAC, cited 2007: IUPAC Subcommittee on Gas Kinetic Data Evaluation - Data Sheet HOx_VOC1, updated 12th December 2007. Available at http://www.iupac-kinetic.ch.cam.ac.uk/index.html

Jöckel, P., Brenninkmeijer, C.A.M., Crutzen, P.J.: A discussion on the determination of atmospheric OH and its trends. Atmos. Chem. Phys. **3**, 107–118 (2003). doi:10.5194/acp-3-107-2003

Jöckel, P., Kerkweg, A., Pozzer, A., Sander, R., Tost, H., Riede, H., Baumgaertner, A., Gromov, S., Kern, B.: Development cycle 2 of the modular earth submodel system (MESSy2). Geosci. Model Dev. **3**, 717–752 (2010). doi:10.5194/gmd-3-717-2010

Khalil, M.A.K. (ed.): Atmospheric Methane: Its Role in the Global Environment. Springer, Berlin, 351 pp. (2000)

Montzka, S.A., Krol, M., Dlugokencky, E., Hall, B., Jöckel, P., Lelieveld, J.: Small interannual variability of global atmospheric hydroxyl. Science **331**, 67–69 (2011). doi:10.1126/science.1197640

Morino, I., Uchino, O., Inoue, M., Yoshida, Y., Yokota, T., Wennberg, P.O., Toon, G.C., Wunch, D., Roehl, C.M., Notholt, J., et al.: Preliminary validation of column-averaged volume mixing ratios of carbon dioxide and methane retrieved from GOSAT short-wavelength infrared spectra. Atmos. Meas. Tech. **4**, 1061–1076 (2011). doi:10.5194/amt-4-1061-2011

Moss, R.H., Edmonds, J.A., Hibbard, K.A., Manning, M.R., Rose, S.K., van Vuuren, D.P., Carter, T.R., Emori, S., Kainuma, M., Kram, T., et al.: The next generation of scenarios for climate change research and assessment. Nature **463**, 747–756 (2010). doi:10.1038/nature08823

O'Connor, F.M., Boucher, O., Gedney, N., Jones, C.D., Folberth, G.A., Coppell, R., Friedlingstein, P., Collins, W. J., Chappellaz, J., Ridley, J. et al.: Possible role of wetlands, permafrost, and methane hydrates in the methane cycle under future climate change: a review. Rev. Geophys. **48**, RG4005, (2010). doi:10.1029/2010RG000326

Sander, S.P., Barker, J.R., Golden, D.M., Kurylo, M.J., Wine, P.H.: Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies, Evaluation Number 17, Pasadena, CA, Jet Propulsion Laboratory, JPL Publication 10-6 (2011)

Schneising, O., Buchwitz, M., Reuter, M., Heymann, J., Bovensmann, H., Burrows, J.P.: Long-term analysis of carbon dioxide and methane column-averaged mole fractions retrieved from SCIAMACHY. Atmos. Chem. Phys. **11**, 2863–2880 (2011). doi:10.5194/acp-11-2863-2011

Shindell, D., Kuylenstierna, J.C.I., Vignati, E., van Dingenen, R., Amann, M., Klimont, Z., Anenberg, S.C., Muller, N., Janssens-Maenhout, G., Raes, F., et al.: Simultaneously mitigating near-term climate change and improving human health and food security. Science **335**, 183–189 (2012). doi:10.1126/science.1210026

Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H.H., Kumar, P., McCarl, B., Ogle, S., O'Mara, F., Rice, C., et al.: Greenhouse gas mitigation in agriculture. Philos. Trans. Roy. Soc. B **363**, (2007). doi:10.1098/rstb.2007.2184

Taraborrelli, D., Lawrence, M.G., Crowley, J.N., Dillon, T.J., Gromov, S., Groß, C.B.M., Vereecken, L., Lelieveld, J.: Hydroxyl radical buffered by isoprene oxidation over tropical forests. Nature Geosci. **5**, 190–193 (2012). doi:10.1038/ngeo1405

Wania, R., Ross, I., Prentice, I.C.: Implementation and evaluation of a new methane model within a dynamic global vegetation model: LPJ-WHyMe v1.3.1. Geosci. Model Dev. **3**, 565–584 (2010). doi:10.5194/gmd-3-565-2010

WMO: WMO Greenhouse Gas Bulletin, The State of Greenhouse Gases in the Atmosphere Based on Global Observations through 2010, No. 7 (2011)

Towards a Greenhouse Gas Lidar in Space

48

Gerhard Ehret, Axel Amediek and Mathieu Quatrevalet

Abstract

Highly accurate measurements of atmospheric carbon dioxide (CO_2) and methane (CH_4) by a space-borne lidar will help to substantially improve knowledge of greenhouse gas fluxes. The method of integrated-path differential-absorption lidar for total column measurements has proven to be a suitable means for CH_4 detection in natural gas leak surveillance and active remote sensing of CO_2. This pioneering work facilitated the instrument development of an advanced greenhouse gas lidar on HALO and set the stage for the development of a CH_4-lidar in space instrument foreseen in the Franco-German climate mission MERLIN.

G. Ehret (✉) · A. Amediek · M. Quatrevalet
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Gerhard.Ehret@dlr.de

A. Amediek
e-mail: Axel.Amediek@dlr.de

M. Quatrevalet
e-mail: Mathieu.Quatrevalet@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_48,

48.1 Introduction

Long-life atmospheric species such as *carbon dioxide* and *methane* are the most important *greenhouse gases* in the atmosphere which are regulated by the *Kyoto protocol*. The atmospheric abundance of these gases has increased rapidly due to human activity since the industrial revolution. Any prediction of future climate and any policy advice on greenhouse gas avoidance require reliable estimates of the natural and *anthropogenic sources* and *sinks* of these gases in terms of location, magnitude, and variability on a global basis. The latter involves at least satellite instruments for the measurement of the column-integrated dry-air mixing ratio of corresponding greenhouse gas (GHG) abundance, commonly referred to as XGHG. Together with appropriate modelling activities, these data can be used to infer GHG sources and sinks by means of inverse models that describe atmospheric *transport* and mixing (Houweling et al. 2004).

The major anthropogenic sources of CO_2 are the burning of fossil fuels, producing carbon fluxes of about 7 Pg/a (1 Pg $= 10^{15}$ g $= 10^{12}$ kg), and changes in land use by deforestation to establish agricultural areas ($\sim$2 Pg/a). Atmospheric measurements indicate that only about half of the emitted CO_2 remains in the atmosphere, the remainder being taken up by both the oceans and various processes on land. The problem of the exact quantification of the surface sinks of anthropogenic CO_2 is complicated by the fact that the human perturbation is superimposed on the global natural *carbon cycle*. The natural background carbon fluxes of about 60 Pg/a between the atmosphere and the land (primarily photosynthesis and respiration) or those between the atmosphere and the ocean are significant larger than anthropogenic ones, thus detection of the anthropogenic sink signals is very difficult. On the other hand, better quantification of the natural spatio-temporal pattern of CO_2 sources and sinks is needed for better validation of the coupled carbon cycle—*climate models*.

For CH_4 the situation is somewhat different because the anthropogenic emissions associated with fossil fuels (gas leaks, incomplete combustion) and agricultural practice (rice cultivation and cattle grazing) are significant larger than the natural background emissions from *wetlands*, which amount to about 30 % of total CH_4 emissions. However, wetland emissions vary substantially in different parts of the globe and *global warming*, in particular in arctic regions, might foster the melting of *permafrost* soils, which could give rise to a significant increase of CH_4. There exist also very large deposits of CH_4 as hydrates on the ocean shelves that are vulnerable to ocean warming. Paleorecords indicate that both processes have provided important *feedback* to the *climate system*. The development and implementation of an observational system for the detection of CH_4 emissions in these vulnerable areas has therefore high scientific priority.

On the other hand, the well-mixed GHGs in the atmosphere (caused by their long *lifetimes*) pose great challenges for the measurement system. A significant improvement in current knowledge of regional fluxes would require a relative measurement precision as high as 0.3 % for CO_2 and 1 % in the case of CH_4

(Ehret and Kiemle 2005) in relation to the background mixing ratios of 380 μmol/mol and 1.8 μmol/mol for CO_2 and CH_4, respectively. Most important, the residual biases in the measurement that cannot be corrected should even be an order of magnitude lower than the random errors. Significant progress has been made to retrieve GHG data from SCIAMACHY (Buchwitz et al. 2005) and GOSAT instruments that measure the reflected solar radiation from the Earth surface in the near IR. However, their results may be questionable as they involve complex bias correction schemes to deal with aerosol/cloud scattering in the light path, which can have strong regional variability. The main drawback of passive sounders in the infrared spectral region is related to their atmospheric weighting functions, which favor the middle and the upper troposphere rather than the lower troposphere where the sources and sinks reside.

In contrast, high measurement sensitivity is expected from making use of the *Integral-Path Differential-Absorption* (IPDA) technique where the strong *lidar* echoes from cloud tops or the Earth surface can be used to infer XGHG data (Menzies and Tratt 2003; Ehret et al. 2008). In contrast to passive instruments, a lidar will have its own light source emitting pulsed narrow-line *laser* radiation, not relying on sunlight. The XGHG values can be provided by applying the conventional DIAL technique to the signals from a range-gated receiver for detection of the echoes from the Earth surface. Such a technique can distinguish surface from cloud or aerosol *backscatter*, permitting high-precision retrievals of XGHG in the presence of particle layers with small/moderate *optical depth*, such as thin *cirrus* or *aerosols* (e.g., from megacity plumes, deserts, biomass burning). Also, measurements are possible under cloudy conditions. The lidar beam can reach the surface when gaps between clouds occur due to the near-nadir view and the small lidar footprint. In addition, XGHG measurements above dense stratiform clouds used as reflective targets can provide partial columns for profile information.

48.2 The Integral-Path Differential Absorption Lidar Principle

Airborne or space-borne *DIfferential Absorption Lidar* (DIAL) measurements of trace gases in the atmosphere rely on the highly wavelength-selective absorption of laser light by molecules. The column of air underneath the instrument is illuminated with two monochromatic laser beams that only differ in their wavelengths, one near the peak of a molecular absorption line of the trace gas of interest, the other further away in the wing of the line. The former is referred to as the on-line wavelength λ_{on}, the latter as the off-line wavelength λ_{off}.

In the *Integral-Path Differential-Absorption* (IPDA) implementation of the DIAL principle, as illustrated in Fig. 48.1, use is made of the small fraction of emitted photons which is scattered back by a so-called hard target (Earth's surface or a cloud top) towards the instrument's receiver, and subsequently converted into electrical signals by a photodetector. Using a short-pulse laser and a small

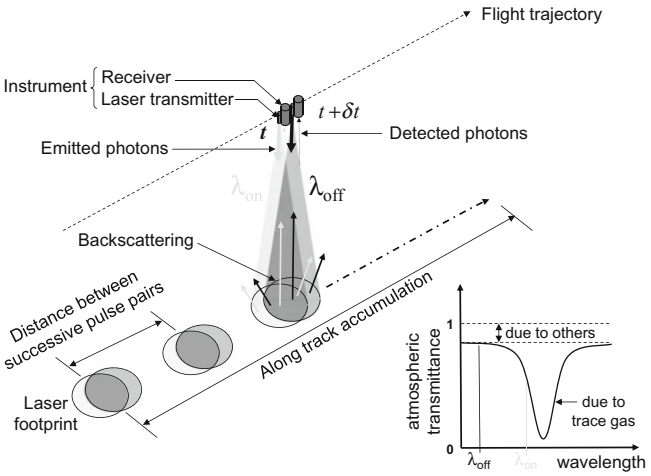

Fig. 48.1 Principle of double-pulse Integral-Path Differential Absorption lidar for trace gas measurements

temporal separation δt between on-line and off-line pulse, the two signals can be distinguished. The pulses are digitized and corrected for the number of emitted photons per pulse which are monitored within the instrument. Many such pulse pairs are accumulated along the instrument's ground track to increase the signal-to-noise ratio. The exact fractions of generated photoelectrons per emitted photons ε_{on} and ε_{off} are products of many instrumental factors and highly variable geophysical factors that depend on local aerosol loading, cloudiness and ground reflectivity, but if the pulses of each pair are very closely co-located in space and time and for a carefully designed instrument, these factors are the same for both. The ground reflectivity and beam attenuation from contributors other than the trace gas are typically constant at the very narrow spectral scale of a molecular absorption line, as shown using dashed lines in the right-hand corner of Fig. 48.1. As a result, when taking the ratio of ε_{off} to ε_{on} these factors cancel out and only the ratio of off-line to on-line atmospheric transmittance due to the trace gas of interest remains, squared because of the return trip. Taking half the logarithm of this ratio turns it into a difference and converts the atmospheric transmittances into atmospheric one-way optical depths Δ_{off} and Δ_{on}:

$$d\Delta = \Delta_{on} - \Delta_{off} = \frac{1}{2} \cdot \ln\left(\frac{\varepsilon_{off}}{\varepsilon_{on}}\right) \tag{48.1}$$

This quantity, $d\Delta$, commonly referred to as the Differential Atmospheric Optical Depth (DAOD), can be shown by direct application of the Beer–Lambert law to be proportional to a weighted average of the trace gas' mixing ratio vertical profile $q(z)$ between the instrument at altitude z_{inst} and the hard target at altitude z_{ht}:

$$d\Delta = \int_{z=z_{ht}}^{z_{inst}} q(z) \cdot WF(z) \cdot dz$$

$$= \left(\int_{z=z_{ht}}^{z_{inst}} WF(z) \cdot dz \right) \cdot \underbrace{\left(\int_{z=z_{ht}}^{z_{inst}} q(z) \cdot \frac{WF(z)}{\left(\int_{z=z_{ht}}^{z_{inst}} WF(z) \cdot dz \right)} \cdot dz \right)}_{=XGHG \text{ by definition}} \tag{48.2}$$

with

$$WF(z) = n(z) \cdot \left(\sigma(\lambda_{on}, p(z), T(z)) - \sigma\left(\lambda_{off}, p(z), T(z)\right) \right) \tag{48.3}$$

In Eq. (48.2), the weighted average of $q(z)$, usually referred to as *XGHG* (with $GHG = CO_2$ or CH_4), appears as the second factor on the right-hand side. The quantity $WF(z)$ singled out in Eq. (48.3) is the vertical weighting function, with $n(z), p(z)$ and $T(z)$ the vertical profiles of air molecule number density, pressure and temperature, respectively, and $\sigma(\lambda, p, T)$ the wavelength-, pressure- and temperature-dependent absorption cross-section of the trace gas. The key to IPDA measurements is that all these parameters can be obtained independently from the measurement of $d\Delta$ itself; $n(z)$, $p(z)$ and $T(z)$ from *numerical weather prediction* models, $\sigma(\lambda, p, T)$ from laboratory characterizations, and z_{ht} from the inherent ranging capability of a pulsed lidar. This enables the conversion using Eq. (48.2) of $d\Delta$ to *XGHG*.

The choice of an absorption line and the fine-tuning of the wavelength pair around the line are governed by many constraints. The absorption line itself must be sufficiently strong for the contrast between Δ_{off} and Δ_{on} to emerge over the measurement noise, but sufficiently weak for a sufficient number of on-line photons to reach the photodetector. It is of utmost importance to ensure that no strong absorption lines of other trace gases overlap spectrally with the useful absorption line, especially water vapor lines because of the high variability of *water vapor* in the atmosphere. For carbon dioxide, two spectral windows with several appropriate absorption lines and minimal interference from water vapor can be found in the near-infrared region near 1.6 and 2 μm. Figure 48.2a is an overview of the 1.6-micron region, showing that water vapor lines are typically between a hundred to a thousand times weaker than the strongest CO_2 lines.

Figure 48.2b is a closer look at one of the suitable CO_2 lines, and illustrates how optimization of the weighting function is possible by means of the only degree of freedom in Eq. (48.3), i.e., the on-line wavelength. It is a discrete graphical representation of the leftmost integral in Eq. (48.2), assuming an off-line wavenumber far enough in the wing for Δ_{off} to be negligible (i.e., for $d\Delta \approx \Delta_{on}$). The U.S. standard atmosphere is split into 20 layers each containing the same number of CO_2 molecules, and the average molecular absorption cross-section in

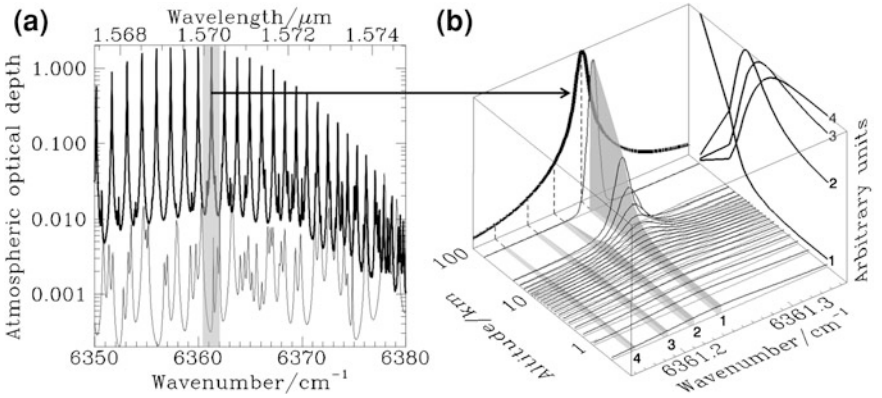

Fig. 48.2 **a** Atmospheric optical depth of CO_2 (*thick*) and water vapor (*thin*) as a function of wavelength/wavenumber near 1.6 μm, for a U.S. *standard atmosphere*. A suitable CO_2 absorption line for Integral-Path Differential Absorption measurements, analyzed in more details in **b**, is indicated with a grey area. **b** Absorption cross-section of atmospheric CO_2 as a function of altitude and wavenumber (*thin*) and CO_2 atmospheric optical depth as a function of wavenumber (*thick*) near the CO_2 absorption line at 6361.25 cm^{-1}, for a U.S. standard atmosphere. Numbered gray volumes and curves: illustration of the weighting functions obtained with various on-line wavenumbers

each layer is plotted in the plane corresponding to the upper altitude of the layer. Thanks to this representation, a transversal section of the resulting 3D surface qualitatively shows the relative contribution of each atmospheric layer to the total atmospheric optical depth—also plotted on the back face of the cube—for the corresponding on-line position. Four such profiles represented in Fig. 48.2b with numbered gray volumes are projected onto the right-hand face of the cube after normalization for ease of comparison. Dashed lines point to the corresponding Δ_{on}. Because of pressure-dependent line broadening, the absorption line progressively widens from the top of the atmosphere to the lower layers, while conservation of the spectrally integrated absorption cross-section means that the peak value decreases accordingly. As a result, for an on-line wavelength at the peak of the line (profile number 1), the contribution of the upper layers of the atmosphere to Δ_{on} is larger than the contribution of the lower layers, while the opposite is true for an on-line wavelength far away in the wing (profile number 4), with intermediate cases in between (profiles number 2 and 3). For the detection of sources and sinks of CO_2, a weighting function of the second type is desirable so that variations in CO_2 concentration near the ground produce the largest possible variations of XCO_2. However, the DAOD is insufficient at on-line positions such as those corresponding to profiles number 3 or 4, so that in practice a trade-off must be found between an acceptable weighting function and sufficient DAOD, such as profile number 2.

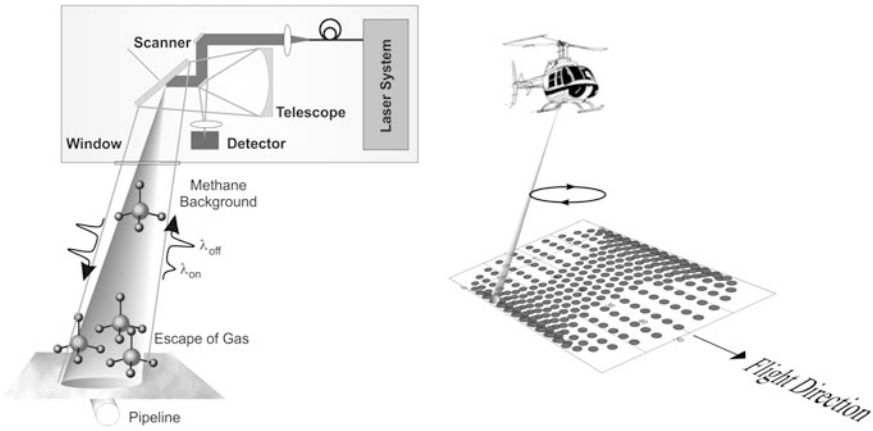

Fig. 48.3 Schematic of the CHARM® instrument subsystems (*left*) and the scanning principles in helicopter-based gas pipeline surveillance (*right*)

48.3 The Helicopter-Based Methane Leak Detection System CHARM®

A prominent example of the application of the IPDA lidar principle is related to the detection of CH_4 escaping from gas pipelines, using a helicopter as the system carrier. Such an application is of considerable relevance in energy mining or natural gas transmission industries. In a joint effort by the Institute of Technical Physics of DLR Stuttgart and the Institute of Atmospheric Physics (IPA) of DLR Oberpfaffenhofen on one hand, and the companies Adlares and Open Grid Europe (former gas transportation branch of E.ON Ruhrgas) on the other hand, the CHARM® (CH_4 Airborne Remote Monitoring) system for pipeline surveillance has been developed.

The basic technical features of this system are illustrated in Fig. 48.3. In order to achieve high detection probability even in cases of tiny gas leaks, a scanning configuration has been implemented. This enabled an 8-m wide observational corridor along the flight direction of the helicopter where the ground spots are rather uniformly distributed across the width of the pipeline corridor. For the development of the laser transmitter, an *optical parametric oscillator* (OPO) was used, a technology that is well established at DLR for lidar applications. In order to provide high measurement sensitivity, the OPO is operated at a sounding wavelength of 3.3 μm. In this spectral region the strong CH_4 absorption lines led to optimum column measurements over short distances of about 100 m, given by the helicopter flight height above ground. Two frequency-controlled, spectrally narrow-band OPO pulses alternatively emitted at slightly different wavelengths serve for the IPDA lidar measurements. The ground echoes are detected by means of a fast IR photodiode in combination with a suitable amplifier and digitizer and

Fig. 48.4 CHARM® measurement in densely built-up area indicated by the point cloud (*blue* and *red dots*) along the helicopter flight track (*blue solid line*). The *blue dots* illustrate the background CH_4 level of 1.8 μmol/mol, the *red* show enhanced CH_4 concentrations from a gas leak. The fine red lines mark the cross hairs of the camera. Credit: Adlares GmbH

stored together with beam steering and navigation data on the computer. The CHARM® system fits into a container about 2.3 × 1.1 × 0.5 m^3 in size with a weight less than 430 kg. The electrical power consumption amounts to ∼1.5 kW.

Figure 48.4 illustrates an example from the first test phase with significant signal levels above the natural background of 1.8 μmol/mol. It is found that small gas flows of 0.15 m^3/h can easily be detected from a flight height of 100 m above natural and man-made surfaces. In particular, the automatic control of the scanner head using a global position sensor and an inertial navigation sensor in combination with geographic information on the pipeline route enabled the pipeline position to be determined with high accuracy which is better than 0.5 m. The CHARM® system has been in routine operation for gas pipeline surveillance since 2008. It is currently the only *remote sensing* system in the world that monitors high pressure pipelines in accordance with the technical code of the Deutscher Verein des Gas- und Wasserfaches.

48.4 Ground-Based IPDA Measurement of Atmospheric Carbon Dioxide

Since the beginning of the new millennium, the development of lidar instruments for the measurement of atmospheric carbon dioxide has attracted increasing attention. Several experimental setups have been tested by different lidar groups around the world, each using slightly different approaches (Gibert et al. 2008; Abshire et al. 2010; Ishii et al. 2010; Spiers et al. 2011). In 2003, as part of a PhD thesis project a prototype instrument was built for the first ground-based CO_2 measurements at IPA. The goal was to investigate a configuration that was not used before in any other CO_2 measurement setup and that took advantage of IPA knowledge of OPO technology for the lidar transmitter. However, numerical simulations indicated in an early stage of the study that the required measurement

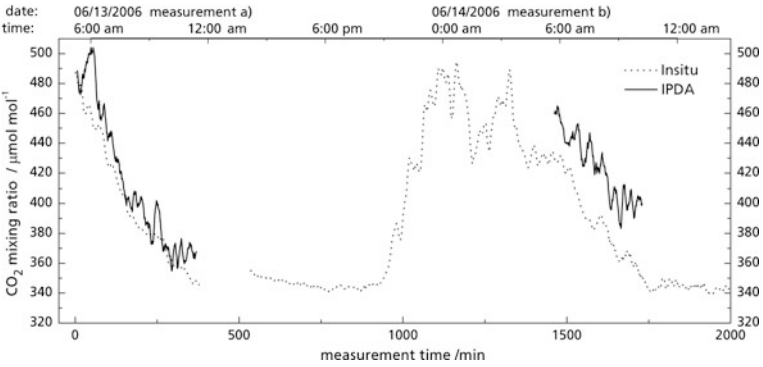

Fig. 48.5 Results of IPA's first IPDA lidar measurements of atmospheric CO_2. Diurnal variations. With kind permission from Springer Science + Business Media: Applied Physics B, "Development of an OPO system at 1.57 µm for integrated path DIAL measurement of atmospheric carbon dioxide", 92, 2008, p. 295-302, Axel Amediek, Andreas, Fix, Martin Wirth, Gerhard Ehret

precision for a meaningful retrieval of the CO_2 mixing ratio which allows the identification of sources, sinks and fluxes places the highest technical demands on a lidar system. This includes in particular stringent requirements on wavelength stability, spectral purity, bandwidth, output power, and beam profile of the envisaged lidar transmitter, which could not be met at that time at IPA. The proposed IPDA lidar approach was known from its application in the CHARM® system, but it was not technically mature enough to meet the high measurement standards required for CO_2 soundings.

The experimental setup used an injection seeded OPO to generate the transmitted light pulses, which allows tuning to the selected trace gas absorption lines. At that time, this was the first application of OPOs for CO_2 lidar measurements worldwide. Similar to the CHARM® system, laser wavelength stabilization was performed by means of a CO_2-filled absorption cell serving as the reference. A particular challenge was to simultaneously achieve all the above-mentioned stringent requirements on the laser radiation with an OPO-based system. So, one purpose of the PhD thesis was to identify the limitations which accompany implementation of the envisaged technical concept. Furthermore, many effects that influence the measurement precision were investigated in numerical simulations.

The first test measurements of atmospheric CO_2 with the developed IPDA lidar prototype took place in late spring 2006. The CO_2 mixing ratio was measured along a 2-km horizontal path using trees as the reflecting target. A commercially available *in-situ* device served to evaluate the results. In contrast to other research groups that also used the IPDA technique, i.e., hard target reflexes, here a natural (i.e., nonideal and thus more realistic) target was used instead of an artificial, defined target. Figure 48.5 shows the results of the test measurements. The IPDA lidar soundings match very well the in situ observations. The statistical deviations of about 2 % correspond to the value that was expected for this setup. However,

the varying biases of 2–9 % for different measurement runs could not be explained by the simulated error sources.

With the developed transmitter setup, many of the defined requirements for the laser radiation could be well met, such as wavelength stability, bandwidth, spectral purity and output power. Only some lack of spatial laser beam quality was noticed, a possible reason for the observed biases. This was noted as an important point for further development and investigation. Another lesson learned was attributed to the power monitoring of emitted pulses, which turned out to be more complex in the IPDA lidar measurement approach for CO_2. In summary, the experience gained from this first prototype setup is beneficial for the next step, which is to develop an airborne *demonstrator* for simultaneous lidar soundings of CO_2 and CH_4 using DLR's new research aircraft *HALO* as the system carrier.

48.5 DLR's Airborne Lidar System for Carbon Dioxide and Methane Monitoring

The demonstration on airborne platforms of new remote sensing instruments intended for use in space is a common and efficient method employed during the development period to reach the needed level of maturity for space applications. Especially in the case of CO_2-sounding lidar systems, the use of aircraft systems is a helpful intermediate step for optimizing the technology and retrieval algorithm. Within the CHARM-F project an airborne demonstrator for the simultaneous sounding of CO_2 and CH_4 applying the IPDA lidar technique is being developed in cooperation with the Fraunhofer Institut für Laser Technik in Aachen. It is designed for use on board DLR's research aircraft HALO. CHARM-F was regarded as supporting the A-SCOPE (Advanced Space Carbon and Climate Observation of Planet Earth) mission proposal on space-borne active remote sensing of CO_2 which was selected for evaluation at the Pre-Phase A level by ESA. The optical features of the transmitting and receiving system are laid out in a way that the geometry of the laser spot on ground as well as the receiver's field of view on ground are very similar to the proposed space-borne system. Thus, this experimental system allows investigation of the IPDA lidar method under similar conditions. Furthermore, it employs different types of semiconductor detectors (with and without internal amplification) which are operated simultaneously, each using specially adapted receiving telescopes. In this way, the advantages and disadvantages of these two detector types can be directly compared to find the optimal configuration. The envisaged instrument consists of two independent, almost identical lasers and receiving systems, each for CO_2 and methane. The sounding of CO_2 uses wavelengths at 1.57 μm; methane requires 1.64 μm wavelengths. The generation of the specific laser wavelengths is achieved using injection seeded OPO systems pumped by Nd:YAG laser sources, which are improved versions of the OPO used for the laboratory prototype, presented above. The wavelength stabilization scheme is also a further development of the

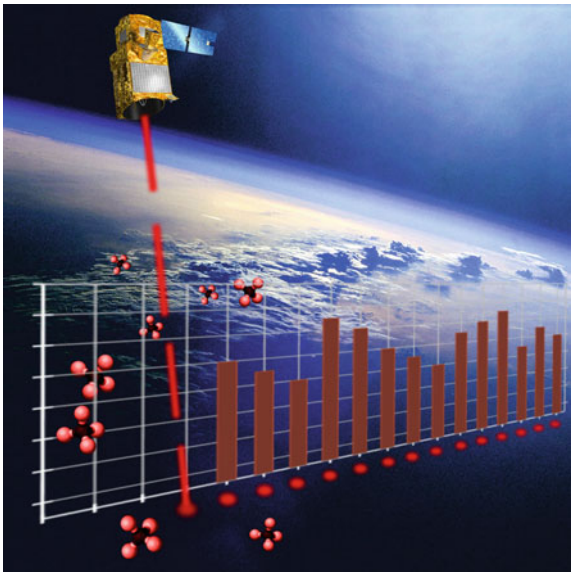

Fig. 48.6 Observational principle of the MERLIN instrument. *Vertical bars* symbolize the amount of XCH_4. Credit: CNES for satellite artist view, NASA for background photo

prototype setup's corresponding subsystem. Currently, the design of the airborne demonstrator is finished and the hardware is being set up. Ground-based test measurements of atmospheric CO_2 and methane are expected for midyear 2012. A first airborne campaign to validate the system by overflying the Total Carbon Column Observing Network ground sites is planned for 2013/2014. This instrument on HALO is intended to serve as an airborne demonstrator for the French-German climate mission MERLIN.

48.6 Space-Borne Observation of Atmospheric Methane by the MERLIN Mission

The acronym MERLIN (Methane Remote Lidar Mission) stands for the Franco-German climate monitoring initiative that will make use of the IPDA lidar principle for global observations of atmospheric methane from space. A further data product of this mission will be canopy height, which is of key importance in biomass detection. In addition, information on cloud boundaries and elevated aerosol layers will be provided by MERLIN observations. The main scientific objective of this mission is related to the improvement of knowledge of natural and anthropogenic methane emissions around the globe where the uncertainties often exceed the 100 % level. This goal will be achieved in concert with observations from surface network and other satellite measurements using passive sensors.

As illustrated in Fig. 48.6, MERLIN will measure the gradient of XCH_4 along the satellite footprint on Earth. This data will serve as a reference in flux inversion models on a monthly basis to capture the seasonal cycle. A sun-synchronous polar

orbit with the satellite crossing the equator at 6:00 a.m. will provide the best conditions for the measurements. It is important to note that such a "polar orbit" will allow for many observations at high latitudes during all seasons. New insights into *climate feedback* from changing wetlands and thawing permafrost are expected from the MERLIN observations, since current observations of methane emissions in those regions are sparse.

The mission, currently in the evaluation period of the Phase A level, is scheduled for launch in the timeframe 2016. It is based on a joint instrument/mission concept where CNES delivers the satellite bus from their MYRIADE series and DLR provides the IPDA lidar instrument. In analogy to former instrument concepts that have been successfully applied in airborne and ground-based measurements, a Nd:YAG-laser pumped OPO will serve as a suitable MERLIN transmitter. This technical concept will also strongly benefit from ESA's laser development programs in connection with the Earth Explorer core missions ADM-Aeolus and EarthCARE. Further heritage on the instrument readiness level is provided by NASA's CALIPSO mission that uses high average power Nd:YAG lasers as lidar transmitters in space over a period of more than three years. This suggests that a mission lifetime of 3 years is feasible if a Nd:YAG laser is used for the OPO as the pump source. Such a period of continuous operation would allow capture of the inter-annual trend in methane emissions where uncertainties are large and the observed variability is not well understood.

In general, MERLIN measurement performance in terms of random error strongly depends on the surface reflectivity and *cloud coverage* along the satellite flight track (Amediek et al. 2009; Kiemle et al. 2011), whereby both parameters can vary significantly.

A suitable means to achieve more or less similar measurement precision everywhere on the globe is the method of signal averaging along the satellite flight track. In Fig. 48.7 the required horizontal integration length is plotted for a measurement precision of 1 %, as indicated by the color bar. The results show that in the case of the envisaged 1 % precision level, the spatial resolution of statistically independent MERLIN observations can be kept reasonably small (< 50 km) for most regions over land. Over the ocean, the situation is more difficult because of the strong surface winds and clouds, both of which impact unfavorably on MERLIN performance. This is also true over the tropics where cirrus clouds may require a larger averaging cell for the MERLIN observations. The figure was obtained by simulations of expected MERLIN performance using realistic instrument and environmental parameters.

In Fig. 48.8, the improvement in knowledge of CH_4 surface fluxes by MERLIN observations is shown. This result was obtained by "Inverse Modelling" using the simulated precision field from hypothetical MERLIN observations indicated in Fig. 48.7 as the reference. It is obvious that MERLIN observations will lead to a substantial (more than 50 %) reduction in the prior CH_4 flux errors in most regions of the globe on a monthly time scale, particularly over source regions of key importance to the global carbon cycle: arctic permafrost, boreal forests, tropical wetlands.

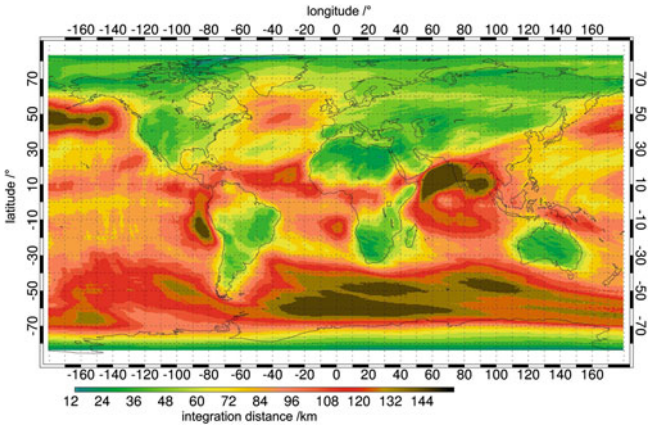

Fig. 48.7 Required along track integration distance in km (*color bar*) required to achieve equal measurement precision of 1 % around the globe by MERLIN observations

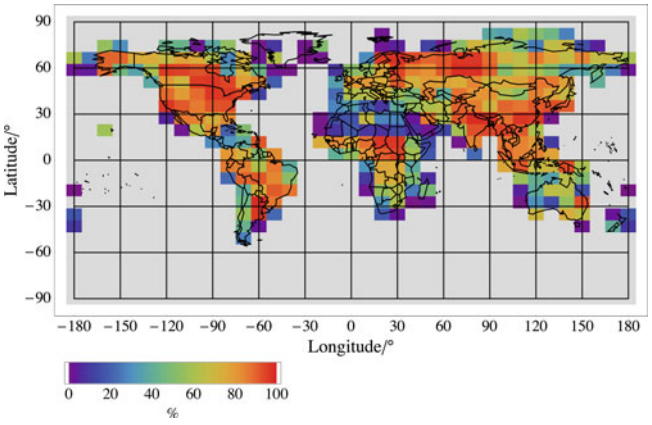

Fig. 48.8 Relative error reduction (*color bar*) of monthly methane surface fluxes for July expressed by $1 - {}^{\sigma_{post}}/_{\sigma_{pri}}$ where $\sigma_{pri}, \sigma_{post}$ correspond to the prior and posterior flux uncertainties, respectively. Credit: Martin Heimann, Max Planck Institute for Biogeochemistry, Jena

48.7 Conclusion

The method of integrated-path differential-absorption lidar is regarded as a suitable means for the measurement of the greenhouse gases carbon dioxide and methane in airborne and space-borne applications. Selection of the sounding wavelength in the wing of an appropriate GHG absorption line will provide high measurement

sensitivity on ground where greenhouse gas sources reside. Using a range-gated receiver and a pulsed laser transmitter any biases from aerosols and clouds in the light path can be avoided, which is an indispensable prerequisite for inferring sources and sinks by the method of "Inverse Modelling." The method can be applied in pipeline surveillance by measuring spurious methane leaks with high precision using a helicopter that flies along the pipeline. The IPDA lidar measurement of CO_2 is more challenging as observational requirements are more stringent than in the case of methane. The next step towards a space instrument will be the CHARM-F project, which aims to develop an airborne demonstrator for CO_2 and CH_4 using the HALO aircraft. These observations strongly support the French-German climate mission MERLIN, which is on the way to becoming the first lidar in space experiment for atmospheric trace gases. This mission is expected to be a significant step forward toward improvement of knowledge of CH_4 surface fluxes around the globe. On longer time scales, an active IPDA lidar mission for CO_2 may become feasible, whereby MERLIN is regarded as a valuable space-based demonstrator. For accurate, responsive and comprehensive monitoring of the Kyoto protocol regulations, however, several of these high-accuracy instrumented satellites will be required.

References

Abshire, J.B., Riris, H., Allan, G.R., Weaver, C.J., Mao, J., Sun, Z., Hasselbrack, W.E., Kawa, S.R., Biraud, S.: Pulsed airborne lidar measurements of atmospheric CO_2 column absorption. Tellus B **62**, 770–783 (2010)

Amediek, A., Fix, A., Ehret, G., Caron, J., Durand, Y.: Airborne lidar reflectance measurements at 1.57 μm in support of the A-SCOPE mission for atmospheric CO_2. Atmos. Meas. Tech. **2**, 755–772 (2009). doi:10.5194/amt-2-755-2009

Buchwitz, M., de Beek, R., Noel, S., Burrows, J.P., Bovensmann, H., Bremer, H., Bergamaschi, P., Korner, S., Heimann, M.: Carbon monoxide, methane and carbon dioxide columns retrieved from SCIAMACHY by WFM-DOAS: year 2003 initial data set. Atmos. Chem. Phys. **5**, 3313–3329 (2005). doi:10.5194/acp-5-3313-2005

Ehret, G., Kiemle, C.: Requirements Definition for Future DIAL Instruments, Final Report, 10880/03/NL/FF (2005)

Ehret, G., Amediek, A., Kiemle, C., Wirth, M., Fix, A., Houweling, S.: Space-borne remote sensing of the greenhouse gases CO_2, CH_4, N_2O by integrated path differential absorption lidar: a sensitivity analysis. Appl. Phys. B **90**, 593–608 (2008). doi:10.1007/s00340-007-2892-3

Gibert, F., Flamant, P.H., Cuesta, J., Bruneau, D.: Vertical 2 μm heterodyne differential absorption lidar measurements of mean CO_2 mixing ratio in the troposphere. J. Atmos. Oceanic Technol. **25**, 1477–1497 (2008)

Houweling, S., Bréon, F.M., Aben, I., Rödenbeck, C., Gloor, M., Heimann, M., Ciais, P.: Inverse modeling of CO_2 sources and sinks using satellite data: a synthetic inter-comparison of measurement techniques and their performance as a function of space and time. Atmos. Chem. Phys. **4**, 523–538 (2004)

Ishii, S., Mizutani, K., Fukuoka, H., Ishikawa, T., Philippe, B., Iwai, H., Aoki, T., Itabe, O., Sato, A., Asai, K.: Coherent 2 μm differential absorption and wind lidar with conductively cooled laser and two-axis scanning device. Appl. Opt. **49**(10), 1809–1817 (2010)

Kiemle, C., Quatrevalet, M., Ehret, G., Amediek, A., Fix, A., Wirth, M.: Sensitivity studies for a space-based methane lidar mission. Atmos. Meas. Tech. **4**, 2195–2211 (2011). doi: 10.5194/amt-4-2195-2011

Menzies, R.T., Tratt, D.M.: Differential laser absorption spectrometry for global profiling of tropospheric carbon dioxide: selection of optimum sounding frequencies for high-precision measurements. Appl. Opt. **42**, 6569–6577 (2003)

Spiers, G., Menzies, R.T., Jacob, J., Christensen, L.E., Phillips, M.W., Choi, Y., Browell, E.V.: Atmospheric CO_2 measurements with a 2 μm airborne laser absorption spectrometer employing coherent detection. Appl. Opt. **50**, 2098–2111 (2011)

The Spaceborne Wind Lidar Mission ADM-Aeolus

49

Oliver Reitebuch

Abstract

Measurements of winds throughout the atmosphere are crucial for both numerical weather prediction and for climate studies. Nevertheless, observations of height profiles of the global wind field are lacking. Thus the European Space Agency decided to establish a Doppler wind lidar mission named Atmospheric Dynamics Mission ADM-Aeolus. The mission will provide vertical profiles from ground up to the lower stratosphere of one component of the horizontal wind vector, which is perpendicular to the flight track. Pre-launch validation of the instrument with an airborne prototype was performed by DLR.

49.1 Importance of Wind Observations and Aeolus Mission Requirements

Global observations of *wind* profiles at all levels are recognized as the number one priority for global *numerical weather prediction* (NWP, WMO *World Meteorological Organization* (2011)). The wind profile is still poorly observed globally because most satellite sounding instruments provide temperature rather than wind information. Global satellite wind observations are obtained from scatterometers and microwave radiometers on *polar-orbiting satellites* for the sea surface, and as atmospheric motion vectors and cloud-drift winds from imaging instruments on geostationary satellites.

O. Reitebuch (✉)
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Oliver.Reitebuch@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_49,

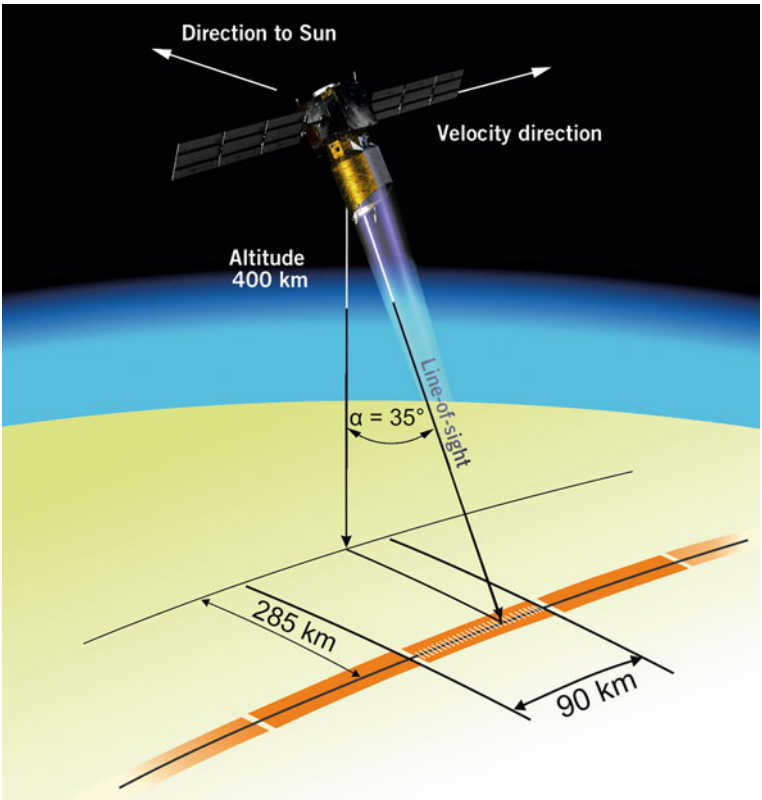

Fig. 49.1 Schematic view of Aeolus measurement geometry with a polar-orbiting satellite at 400 km altitude and a telescope pointing direction $\alpha = 35°$ off-nadir and perpendicular to the satellite velocity direction, which separates satellite and measurement track by 285 km; line-of-sight wind profiles are obtained after on-ground averaging over 90 km (adapted from European Space Agency ESA (2008), Credits: ESA/Aeolus Project)

A spaceborne *Doppler wind lidar* (DWL) is considered to be the only candidate to provide the global coverage of wind profile observations throughout the troposphere and lower stratosphere considered essential for improved NWP. Typically, advancements over the past 10 years in medium-range weather forecast skills are in the order of 1 day and result from improvements in the numerical model, observing system, and computing power. ADM studies show that satellite DWLs can contribute a significant share to this advancement in the coming decades.

In 1999 the *European Space Agency* (ESA) decided to establish a DWL mission named *Atmospheric Dynamics Mission* ADM-Aeolus (nowadays called only "Aeolus") to fill the gap in the global observing system (Stoffelen et al. 2005; European Space Agency ESA 2008). The mission will provide profiles of one component of the horizontal wind vector in the *line-of-sight* (LOS) direction of the telescope, which is tilted by 35° from nadir (Fig. 49.1). As the mission is realized

Table 49.1 Summary of Aeolus key parameters

	Parameters	Values, descriptions
Mission requirements	Mission lifetime	3 a
	Availability of data products	< 3 h after observation
	Random error	Horizontal LOS wind: < 1 m/s up to 2 km, < 2 m/s up to 16 km
	Systematic error	< 0.4 m/s in horizontal LOS; < 0.7 % wind-speed dependent error
	Vertical resolution	250 m – 2000 m
	Along-track averaging length	Onboard: 3 km to 5 km; on-ground: 90 km
	Altitude range	Up to 30 km
Orbit and satellite	Mean altitude	408 km
	Orbit	Polar, dawn dusk sun-synchronous; orbital period of 93 min; orbit repeat cycle of 7 d
	Mass	1100 kg plus about 250 kg fuel
	Volume	4.3 m × 1.9 m × 2.0 m
	Power	1.4 kW average
	Data downlink	Once per orbit with X-band to Svalbard
ALADIN instrument	Principle	Direct-detection lidar with two spectrometers for molecular and aerosol/cloud backscatter
Laser transmitter	Wavelength	354.9 nm
	Pulse	120 mJ pulse energy; 50 Hz repetition rate
Telescope	Diameter	1.5 m
	LOS pointing direction	35° off-nadir, perpendicular to satellite track
Receiver	Molecular backscatter	Sequential *Fabry-Perot interferometer* for double-edge technique
	Aerosol and cloud particle backscatter	*Fizeau interferometer* for fringe imaging technique
	Detectors	Two accumulation charge-coupled devices CCDs
Data products	Level 1b	Profile of horizontal LOS wind speed (uncorrected for actual atmospheric temperature)
	Level 2a	Cloud and aerosol optical parameters, e.g., backscatter, extinction profiles, optical depth, cloud boundary
	Level 2b	Profile of horizontal LOS wind speed with atmospheric temperature/pressure correction
	Level 2c	Profile of horizontal wind vector (two components) after assimilation into ECMWF model

with a polar-orbiting satellite and the LOS pointing direction is perpendicular to the satellite ground track, mainly the zonal component of the wind vector (west–east direction) will be measured, except for high latitudes. The main benefit of Aeolus is expected for NWP applications, where the single LOS component is assimilated into a NWP model and combined with other observations and the model fields. Aeolus will be the first European lidar and the first wind lidar worldwide in space, with a launch currently planned for 2014.

The mission requirements for Aeolus are mainly derived to be beneficial for global NWP with an altitude range for the wind profiles up to the lower stratosphere (20–30 km) and a vertical resolution of 0.5 to 2 km depending on altitude (Table 49.1). Additional wind observations from a space-borne DWL are capable of improving the weather forecast skill, if the random errors are smaller than for the other existing satellite sensors and the related model output, and if the observations have very low systematic errors (biases). This sets very demanding requirements for the random and systematic error for space-borne DWL missions and the main focus of Aeolus is put on accuracy and precision, rather than on a high quantity of observations. Both random and systematic errors are expressed for an observation with an along-track horizontal integration length of 90–100 km, and after projection to the horizontal plane (called HLOS: horizontal LOS direction). The random error should be below 1 m/s up to 2 km and below 2 m/s up to 16 km altitude. A systematic error which cannot be corrected by a-priori knowledge (e.g., calibration, instrumental corrections) is called an "unknown bias" and is detrimental for NWP applications. Thus the requirement for the "unknown bias" is very stringent and challenging and is separated into a wind-independent contribution of below 0.4 m/s and a wind-dependent contribution of below 0.7 %. The observations should be available at the NWP centers within 3 h after observations were performed in-orbit to be useful as input for the forecast model. This requires downlinking all the data from the satellite to the ground for each orbit and processing the data within 30 min.

The impact of space-based wind lidar observations on analysis and forecast skills can be determined with an *Observing System Simulation Experiment* (OSSE). An OSSE determines the influence of certain new observation types, such as wind data from a DWL, by performing numerical forecast experiments with and without the simulated new observations. DLR supported such OSSE studies with end-to-end simulations, which included atmospheric, orbit, instrument and retrieval modules (Leike et al. 2001). Results with a data assimilation ensemble technique at the *European Centre for Medium-Range Weather Forecasts* (ECMWF) show that ADM-Aeolus will provide benefits comparable to the radiosonde and wind-profiler network with impact on the analysis results particularly over oceans and the tropics up to 5 days (Tan et al. 2007). An OSSE for ADM-Aeolus by Stoffelen et al. (2006) shows an average improvement of the 5-day wind forecast for 500 hPa by 0.5 day for Europe and 0.25 day for the Northern Hemisphere; this means that the 5.5-day forecast will be as accurate as the 5-day forecast for Europe, when ADM-Aeolus observations are available. Major improvements from Aeolus are expected for the tropics (Zagar 2004) where additional wind information is even more important due to the weak or missing

geostrophic balance of the wind field which dominates the mid-latitude regions. The ADM-Aeolus related studies for assessing the impact on the weather forecast were performed using simulated observations of only one component of the horizontal wind vector. The second component is provided by the numerical weather forecast model during the assimilation of the observed component. It is debatable and depends on the assimilation experiment, numerical model or geographical region whether one component of the two-dimensional wind vector provides less, equal or even more than 50 % of the information content of the full wind vector.

49.2 Aeolus Mission and ALADIN Instrument

While the satellite mission is named after Aeolus—the Greek father of the 12 gods of the winds—the instrument on board the satellite is called ALADIN (*Atmospheric Laser Doppler Instrument*). To achieve global coverage a polar orbit with a high inclination angle of 97° has been chosen. The period of one satellite orbit is 93 min and an orbit repeat cycle of 7 day was selected, which is the duration after which the satellite ground tracks on the earth are repeated. As a compromise between minimizing the atmospheric drag of the satellite and maximizing the backscattered lidar signal, a rather low orbital height with a mean altitude of 408 km was chosen. This is significantly lower than the usual altitudes of 700–800 km for polar-orbiting satellites.

In order to simplify the instrument concept, only one component of the wind vector will be sensed, resulting in one LOS direction. The LOS direction is perpendicular to the satellite velocity in order to nullify the contribution of the satellite ground speed to the Doppler frequency shift. Additionally, yaw steering of the complete satellite is applied to compensate for Earth rotation. Only the component of the horizontal and vertical wind speed is obtained, which is projected onto the LOS direction. As a compromise between high horizontal wind contribution ($\sim \sin(\alpha)$), low vertical wind contribution ($\sim \cos(\alpha)$) and low distance from satellite to atmospheric target ($\sim 1/\cos(\alpha)$), an off-nadir angle α of 35° was chosen (Fig. 49.1).

The ALADIN instrument is mounted on top of the satellite structure (Fig. 49.2), which contains a power system including lithium-ion batteries and gallium-arsenide solar arrays, a control and data management computer, the Attitude and Orbit Control System (AOCS), and X-band and S-band antennas (frequencies and wavelengths of 8.040 GHz and 3.73 cm for X, and 2.030 GHz and 14.8 cm for S, respectively) for data downlink and telemetry data link from the ground stations. The AOCS itself consists of several sensors (e.g., star tracker, inertial measurement unit, global positioning sensor GPS, coarse earth-sun sensor) and orbit controls (e.g., thrusters, reaction wheels, magnetic torquers). The most striking part of the ALADIN instrument is its telescope, with a diameter for the primary mirror of 1.5 m. The other instrument parts are mounted below the telescope and consist of two power lasers and two reference lasers for redundancy, the optical receiver, and the electronic detection unit.

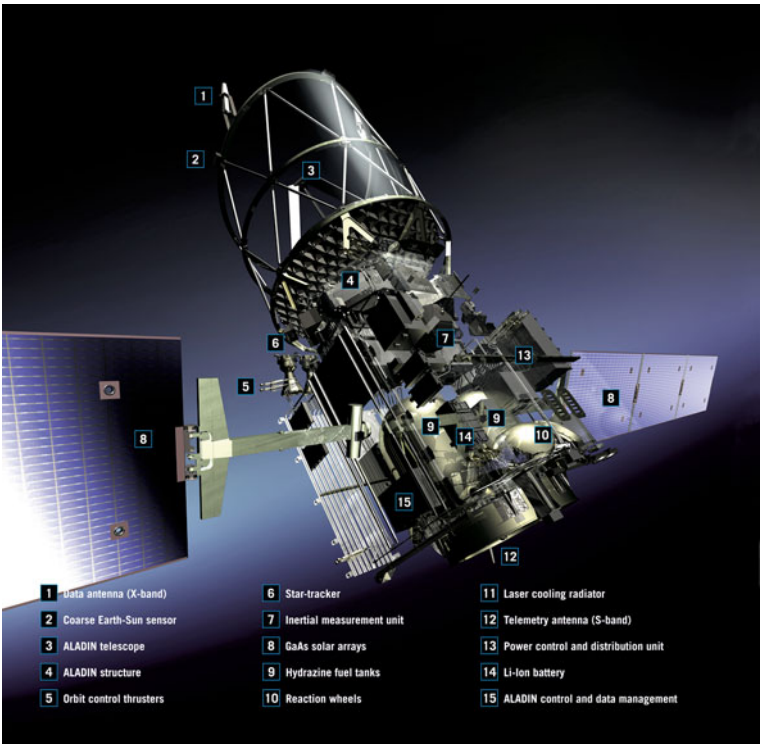

Fig. 49.2 Aeolus satellite including the key subsystems and the ALADIN telescope (*3*) and structure (*4*) where the ALADIN instrument is mounted (from European Space Agency ESA (2008), Credits: ESA/AOES Medialab)

The ALADIN instrument is a direct-detection DWL operating at the ultraviolet (UV) wavelength of 354.9 nm. Due to the need for wind observations from high altitudes in the atmosphere with low aerosol content, a space-borne DWL must mainly rely on Rayleigh backscatter from molecules rather than Mie backscatter from aerosols and cloud particles. A wavelength in the UV spectral region is beneficial in order to increase molecular Rayleigh backscatter, which is inversely dependent on the fourth power of the wavelength. The specific wavelength of 354.9 nm was chosen for technical reasons, because it can be generated with a Nd:YAG laser (Neodymium doped Yttrium–Aluminum Garnet) by tripling its frequency with non-linear crystals and lower wavelengths in the UV spectral region are absorbed by stratospheric ozone. A number of 50 laser pulses, each with a duration of 30 ns and an energy of 120 mJ, are generated per second by the power laser (Fig. 49.3) and transmitted and expanded via the telescope to a diameter of 1.5 m. Due to the very low laser beam divergence of 12 μrad, the laser spot has a diameter of only 6–8 m on the ground for a range of 500 km. In order to achieve high frequency stability and low spectral bandwidth of the laser pulse, which is needed for a precise

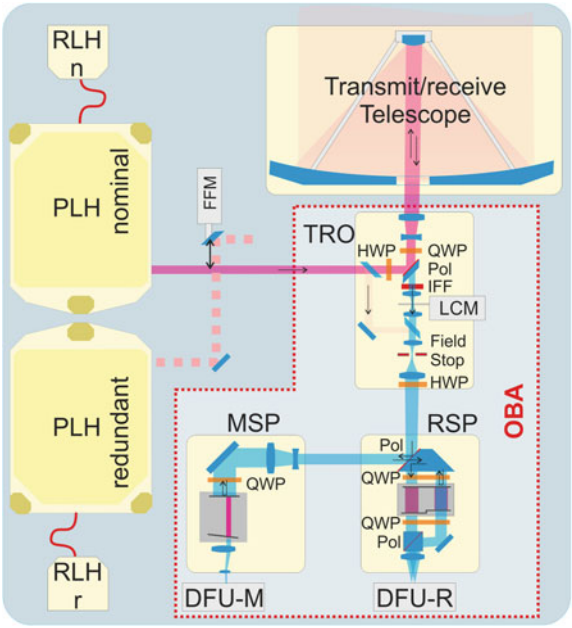

Fig. 49.3 Optical layout of the ALADIN instrument (adapted from ESA, 2008, Credits: ESA/Aeolus Project): *PLH* power laser head, *RLH* reference laser head, *n* nominal; *r* redundant, *FFM* flip-flop mirror, *TRO* transmit-receive optics, *HWP* half-wave plate; *QWP* quarter-wave plate; *Pol* Polarising Beam Splitter, *IFF* interference filter, *LCM* laser chopper mechanism, *MSP* Mie Spectrometer, *RSP* Rayleigh spectrometer, *DFU-M/-R* Detection front-end unit Mie/Rayleigh, *OBA* optical bench assembly

Doppler frequency measurement, laser light from a very frequency-stable reference laser is injected into the power laser. The qualification program for the ALADIN laser was supported by optical testing for laser induced contamination and damage by the DLR Institute of Technical Physics (Riede et al. 2011).

The backscatter signal from molecules, aerosols, and cloud particles is collected by the same telescope used for laser pulse transmission and directed to the transmit-receive optics. In order to suppress unwanted solar background light, which is one of the noise sources for the lidar, the received light is spectrally filtered with a bandwidth of 1 nm and spatially filtered by a field stop, which limits the field-of-view of the telescope to only 19 μrad. As the same telescope is used for transmission of the laser pulse towards the atmosphere and reception of the backscattered signal, the two beam paths have to be separated, which is achieved by use of the *polarization* of the light through several polarization-dependent optics. Thus the depolarized portion of the backscattered light, which is significant for such strongly aspherical particles as *desert dust*, *volcanic ash*, or *cirrus cloud* particles, is discarded and not detected by the ALADIN instrument. In order to protect the very sensitive detectors from the high-energy laser pulses, the optical path is blocked mechanically by a mechanism during emission of the pulse.

The heart of the optical receiver consists of two spectrometers: a Rayleigh spectrometer measures the Doppler frequency shift from molecular backscatter and a Mie spectrometer from aerosol and cloud particles backscatter. Separation between molecular and aerosol backscatter requires significantly different spectral bandwidths of the two signals, the broad bandwidth signal from the molecule

scattering in the atmosphere and the narrow bandwidth from aerosol scattering. Spectral bandwidth is commonly characterized by the full-width at half maximum (FWHM; for a Gaussian distribution, FWHM is 2.35 times the standard deviation). The Rayleigh-*Brillouin scattering* on molecules causes a broad bandwidth signal with a FWHM of 1.6 pm (or 3.8 GHz at 355 nm for 15 °C atmospheric temperature) due to thermal motions of the molecules (Witschas et al. 2010). The aerosol Mie scattering has a narrow bandwidth with a factor of almost 80 lower FWHM of 21 fm (or 50 MHz at 355 nm).

For both the broad and narrow bandwidth signals a direct-detection method is applied to derive the Doppler frequency shift, but the principle of the two spectrometers is different. The double-edge technique is used for the molecular backscattered signal and the fringe imaging technique is used for the aerosol return. The Rayleigh spectrometer is based on the well-established double-edge technique using two Fabry–Perot interferometers as spectral filters. A novel sequential implementation of the two filters was developed, where the reflected signal from one filter is directed towards the second filter in order to efficiently use the signal. The Mie spectrometer is based on a Fizeau interferometer, where the spatial location of the interference pattern (called "fringe") is directly proportional to the Doppler frequency shift. Thus an imaging detector is needed for the Mie spectrometer, which is realized by a charge-coupled device (CCD). The CCD allows accumulation of backscattered signal from a number of laser pulse returns directly on the CCD itself, which is a novel concept for lidar instruments. Two identical CCDs are used for detecting both the Mie spectrometer and the Rayleigh spectrometer signal.

The minimum vertical resolution of the ALADIN instrument is limited by the CCD detector to 250 m and only a number of 24 atmospheric altitude ranges can be acquired. On the other hand the vertical resolution can be varied from 250 to 2 000 m even within one vertical profile. The number of laser pulses accumulated directly on the detector is a setable parameter and determines the minimum along-track horizontal resolution for the raw data, which is 3–5 km for a typical number of 20–35 accumulated laser pulses. Both the Mie and the Rayleigh spectrometer data are stored with this temporal resolution and transmitted to the ground via the X-band downlink every orbit. Further averaging of 600–650 laser pulses (12–13 s data, 90–100 km along-track resolution) is needed to achieve the random error requirement of 1 m/s or 2 m/s. Algorithms for Level 2 data products (derived geophysical wind and aerosol variables) use a horizontal and vertical grouping of raw data in cases of cloudy atmospheric scenes (Tan et al. 2008).

The raw data are processed on-ground to obtain several levels of data products. For Level 1 data products (processed instrument data) only information from the Aeolus satellite with the ALADIN instrument and AOCS system is used. For Level 2 products additional information from other sources is used, mainly the atmospheric temperature and pressure profile for correcting the winds derived with the Rayleigh spectrometer from the temperature and pressure dependent signal bandwidth of the Rayleigh-Brillouin scattering (Dabas et al. 2008). As ALADIN is not only a wind lidar but also a *high spectral resolution lidar* with a separation of aerosol and molecular backscatter, quantitative aerosol optical parameters such as

backscatter and extinction coefficient profiles for these constituents can be derived (Level 2a data products, Ansmann et al. (2007); Flamant et al. (2008). The DLR Institutes of Atmospheric Physics (IPA) and Remote Sensing Technology (IMF) are responsible for developing algorithms and operational processors for Level 1b and Level 2a data products in cooperation with European partners. The operational Level 2b and Level 2c processor is being developed and implemented at ECMWF, while the archive and re-processing facility for Aeolus data will be integrated at the German Remote Sensing Data Center (DFD).

49.3 Pre-launch Validation with an Airborne Demonstrator by DLR

A pre-launch campaign program for ADM-Aeolus was established for several reasons. First, the satellite instrument will be tested and characterized on ground in a clean-room environment by space industry, but without illuminating it with atmospheric signals before launch. Second, direct-detection DWL were operated in the past from the ground, but no *direct-detection DWL* was deployed on an airborne platform in a downward looking geometry, as from space; likewise *coherent DWLs*. Third, ALADIN combines new techniques not yet implemented in a wind lidar, like a novel combination of molecular and aerosol spectrometers, the use of a Fizeau interferometer for the narrow bandwidth aerosol return, sequential implementation of the double-edge Fabry–Perot interferometer for molecular return, and the use of accumulation CCDs as detectors.

The development of an airborne instrument demonstrator was started by the European Aeronautic Defence and Space Company (EADS-Astrium) and DLR in 2003—called the ALADIN airborne demonstrator A2D (Reitebuch et al. 2009). The main objective of the airborne instrument demonstrator is to validate the ALADIN instrument and performance models from ground and from an airborne platform. Furthermore, atmospheric observations with an ALADIN type instrument of various atmospheric scenes (e.g., clear air without clouds, different cloud types or aerosol loadings, sea surface and land returns) were obtained to test, validate and optimize the retrieval and related quality-control algorithms (Tan et al. 2008; Reitebuch et al. 2012), as well as the calibration schemes for the space instruments (Dabas et al. 2008). An end-to-end instrument simulator including the retrieval algorithms was developed for the airborne demonstrator in order to assess radiometric and wind measurement performance (Paffrath 2006; Paffrath et al. 2009). In addition, sea surface reflectance under different incidence angles at 355 nm wavelength was studied using airborne observations from the A2D (Li et al. 2010), which is relevant for satellite calibration.

The A2D is the first airborne direct-detection DWL worldwide and was deployed on the DLR Falcon aircraft in October 2005 for its maiden flight. The core of the A2D is based on the ALADIN receiver and transmitter from the pre-development program of ESA, and is therefore representative of the actual satellite instrument. The main differences of the A2D to the satellite instrument are the

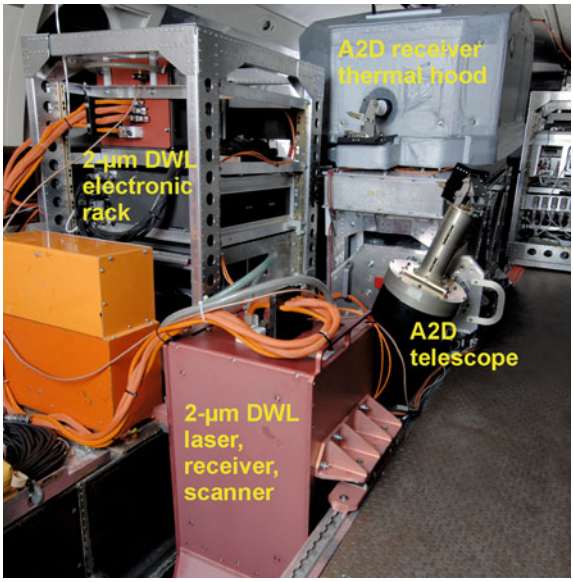

Fig. 49.4 Photo of the 2-µm DWL and A2D installed in the DLR Falcon aircraft for Aeolus pre-launch validation campaigns in 2007, 2008 and 2009

telescope, the transmit/receive optics and the laser transmitter implementation (Schröder et al. 2007). A Cassegrain telescope with Ø 0.2 m is used for the receive path with a co-axial but separate laser transmit path, whereas a Ø 1.5 m telescope in a transceiver configuration is used for the satellite instrument. The airborne front optics include an electro-optical modulator for suppression of the near-field signal, a CCD camera for co-alignment of the transmit-receive path, and a fiber coupling unit for the internal reference signal. The Mie and Rayleigh spectrometers, the accumulation CCD detectors, and the detection electronics differ only in minor details in the airborne demonstrator and the satellite instrument.

The A2D and the 2-µm DWL were deployed on the Falcon aircraft during airborne campaigns in 2007, 2008 and 2009 (Fig. 49.4), which were the first times a coherent and a direction-detection wind lidar were operated on the same aircraft. LOS wind measurements from the well-established 2-µm DWL were used as a reference for the winds retrieved from the A2D Rayleigh and Mie spectrometers. In September 2009 the Falcon aircraft was based in Keflavik, Iceland to perform flights in high-wind conditions and over the Greenland ice shield in order to simulate the ALADIN calibration scheme over ice. An example of a wind measurement of the 2-µm DWL and the A2D along a 30 min flight track (corresponding to 368 km) along the east coast of Greenland is provided in Fig. 49.5, showing LOS wind speeds of the A2D with an off-nadir angle of 20° (Marksteiner et al. 2011). Strong winds are present in the jet-stream region between 6 and 9 km in the second half of the flight track. Due to the absence of thick clouds in a clean atmosphere, solely the Rayleigh channel of the A2D, sensitive to molecular backscatter, is able to provide full vertical wind profiles. Close to the sea surface a distinctive region of up to 12 m/s (in LOS direction) is related to a catabatic flow

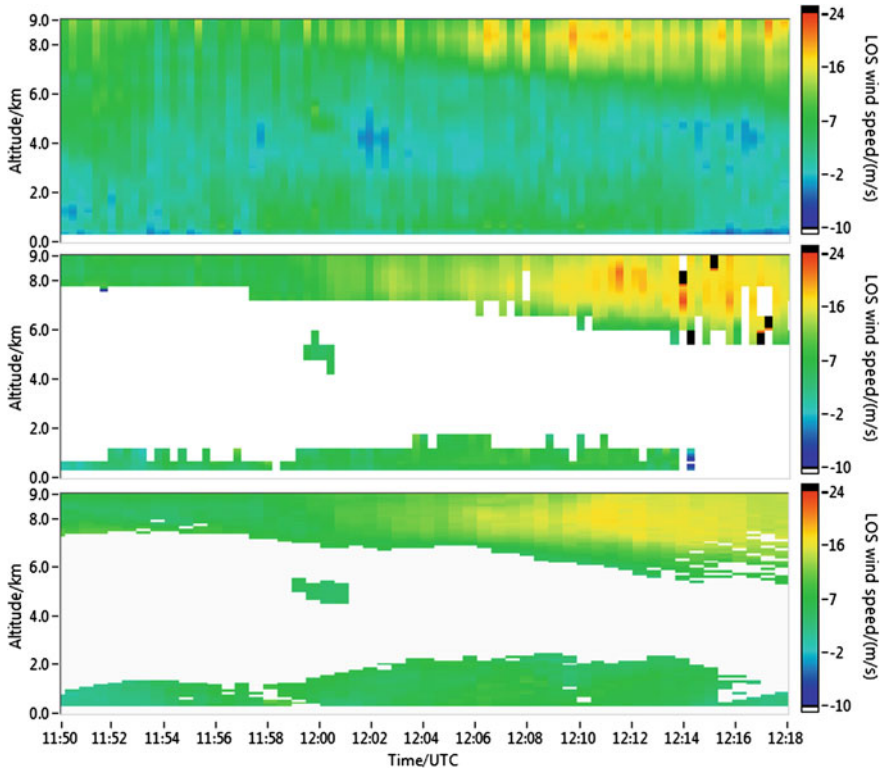

Fig. 49.5 Measurement of the LOS wind speed from a flight of the Falcon aircraft along the east coast of Greenland on 26 September 2009 with the A2D Rayleigh channel (*top*), the A2D Mie channel (*middle*), and the 2-μm DWL (*bottom*); *white* colors indicate nonvalid winds after quality control; length of the flight track is 368 km. Adapted from Marksteiner et al. (2011)

from the Greenland ice shield. Both the A2D Rayleigh and Mie winds are in good correspondence to the 2-μm DWL observations and a statistical comparison between A2D and 2-μm winds yields standard deviations of 1.9 m/s (Rayleigh) and 1.2 m/s (Mie), correlation coefficients of 0.85 (Rayleigh) and 0.95 (Mie), and systematic differences below 1 m/s (Marksteiner et al. 2011).The pre-launch validation program with the A2D was successfully completed in 2012. It provided the proof that wind observations are in principle possible using the novel ALADIN instrument with its double-edge, sequential Fabry–Perot interferometer for molecular backscatter and the Fizeau interferometer for aerosol and cloud backscatter. Even more important for the Aeolus mission was the formulation of about 100 recommendations for testing and operating the ALADIN instrument on-ground and in-orbit and for refinement of the retrieval algorithms and calibration strategy. In contrast to the space industry, which follows an engineering-focused approach, the validation campaign program by DLR was also motivated by scientific curiosity about the response of a rather complex and novel instrument to

atmospheric signals. This revealed some unexpected instrumental characteristics and behavior observed with "real" (instead of "simulated") atmospheric signals. Thus the pre-launch validation program with the airborne prototype significantly reduces some of the risks in developing and launching a novel space-borne instrument.

References

Ansmann, A., Wandinger, U., Le Rille, O., Lajas, D., Straume, A.G.: Particle backscatter and extinction profiling with the spaceborne high-spectral-resolution Doppler lidar ALADIN: methodology and simulations. Appl. Opt. **46**, 6606–6622 (2007). doi:10.1364/AO.46.006606

Dabas, A., Denneulin, M.L., Flamant, P., Loth, C., Garnier, A., Dolfi-Bouteyre, A.: Correcting winds measured with a Rayleigh Doppler lidar from pressure and temperature effects. Tellus A **60**, 206–215 (2008). doi:10.1111/j.1600-0870.2007.00284.x

European Space Agency ESA, ADM-Aeolus Science Report. ESA SP-1311, (2008) 121 pp

Flamant, P.H., Cuesta J., Denneulin M.L., Dabas A, Huber D., ADM-Aeolus retrieval algorithms for aerosol and cloud products. Tellus **60 A**, 273–286 (2008). doi: 10.1111/j.1600-0870.2007.00287.x

Leike, I., Streicher, J., Werner, C., Banakh, V., Smalikho, I.: Virtual doppler lidar instrument. J. Atmos. Oceanic Technol. **18**, 1447–1456 (2001). doi:10.1175/1520-0426(2001)018<1447:VDLI>2.0.CO;2

Li, Z., Lemmerz,C., Paffrath, U., Reitebuch, O., Witschas, B.: Airborne Doppler lidar investigation of sea surface reflectance at a 355-nm ultraviolet wavelength. J. Atmos. Oceanic Technol. **27**, 693–704 (2010). doi: 10.1175/2009JTECHA1302.1

Marksteiner, U., Reitebuch, O., Rahm, S., Nikolaus, I., Lemmerz, C., Witschas, B.: Airborne direct-detection and coherent wind lidar measurements along the east coast of Greenland in 2009 supporting ESA's Aeolus mission. Proc. SPIE **8182**, 81820J-1–81820J-8 (2011). doi: 10.1117/12.897528

Paffrath, U.: Performance assessment of the Aeolus Doppler wind lidar prototype. Dissertation, TU Munich DLR FB 2006-12 (2006) 137 pp

Paffrath, U., Lemmerz, C., Reitebuch, O., Witschas, B., Nikolaus, I., Freudenthaler, V.: The airborne demonstrator for the direct-detection Doppler wind lidar ALADIN on ADM-Aeolus: II. Simulations and Rayleigh receiver radiometric performance. J. Atmos. Oceanic Technol. **26**, 2516–2530 (2009). doi:10.1175/2009JTECHA1314.1

Reitebuch, O., Lemmerz, C., Nagel, E., Paffrath, U., Durand, Y., Endemann, M., Fabre, F., Chaloupy, M.: The airborne demonstrator for the direct-detection Doppler wind lidar ALADIN on ADM-Aeolus: I. Instrument design and comparison to satellite instrument. J. Atmos. Oceanic Technol. **26**, 2501–2515 (2009). doi:10.1175/2009JTECHA1309.1

Reitebuch, O., Huber, D., Nikolaus, I. Algorithm theoretical basis document ATBD: ADM-Aeolus Level 1B Products, 4.0, 76 pp (2012)

Riede, W., Schroeder, H., Bataviciute, G., Wernham, D., Tighe, A., Pettazzi, F., Alves, J.: Laser-induced contamination on space optics. Proc. SPIE **8190**, (2011) doi: 10.1117/12.899190

Schröder, T., Lemmerz, C., Reitebuch, O., Wirth, M., Wührer, C., Treichel, R.: Frequency jitter and spectral width of an injection-seeded Q-switched Nd:YAG laser for a Doppler wind lidar. Appl. Phys. B **87**, 437–444 (2007). doi:10.1007/s00340-007-2627-5

Stoffelen, A., Pailleux, J., Källen, E., Vaughan, J.M., Isaksen, L., Flamant, P., Wergen, W., Andersson, E., Schyberg, H., Culoma, A., et al.: The Atmospheric Dynamics Mission for global wind field measurement. Bull. Am. Meteorol. Soc. **86**, 73–87 (2005). doi: 10.1175/BAMS-86-1-73

Stoffelen, A., Marseille, G.J., Bouttier, F., Vasiljevic, D., de Haan, S., Cardinali, C.: ADM-Aeolus doppler wind lidar observation system simulation experiment. Q. J. R. Meteorol. Soc. **132**, 1927–1947 (2006). doi:10.1256/qj.05.83

Tan, D., Andersson, E., Fisher, M., Isaksen, L.: Observing-system impact assessment using a data assimilation ensemble technique: application to the ADM-Aeolus wind profiling mission. Q. J. Roy. Meteorol. Soc. **133**, 381–390 (2007). doi:10.1002/qj.43

Tan, D., Andersson, E., Kloe, J., Marseille, G.-J., Stoffelen, A., Poli, P., Denneulin, M.-L., Dabas, A., Huber, D., Reitebuch, O., et al.: The ADM-Aeolus wind retrieval algorithms. Tellus A **60**, 191–205 (2008). doi:10.1111/j.1600-0870.2007.00285.x

Witschas, B., Vieitez, M.O., van Duijn, E.-J., Reitebuch, O., van de Water, W., Ubachs, W.: Spontaneous Rayleigh-Brillouin scattering of ultraviolet light in nitrogen, dry air, and moist air. Appl. Opt. **49**, 4217–4227 (2010). doi:10.1364/AO.49.004217

WMO World Meteorological Organization: Statement of guidance for global numerical weather prediction, Version 10 May 2011, approved June 2011. Available online at http://www.wmo.int/pages/prog/www/OSY/GOS-RRR.html

Zagar, N.: Assimilation of equatorial waves by line of sight wind observations. J. Atmos. Sci. **61**, 1877–1893 (2004). doi:10.1175/1520-0469(2004)061<1877:AOEWBL>2.0.CO;2

Cloud-Aerosol-Radiation Interaction: Towards the EarthCARE Satellite Mission

50

Bernhard Mayer, Robert Buras, Gerhard Ehret, Martin Hagen, Andreas Petzold and Bernadett Weinzierl

Abstract

The EarthCARE satellite mission, to be launched by the European Space Agency (ESA) in 2015, is expected to provide new insights into aerosol-cloud-radiation interactions thanks to simultaneous active and passive observations of Earth's atmosphere. Here the contributions of the Institute of Atmospheric Physics (IPA) and Ludwig-Maximilians-University (LMU) Munich to mission preparation are presented, including the provision of atmospheric profiles, airborne demonstrators, and instrument simulators, all essential prerequisites for a successful mission and data exploitation.

B. Mayer (✉) · R. Buras · B. Weinzierl
Ludwig-Maximilians-Universität München (LMU),
Meteorological Institute Munich (MIM),
Theresienstraße 37, 80333 München, Germany
e-mail: bernhard.mayer@lmu.de

R. Buras
e-mail: robert.buras@lmu.de

B. Weinzierl
e-mail: Bernadett.Weinzierl@dlr.de

B. Mayer · G. Ehret · M. Hagen · A. Petzold · B. Weinzierl
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Gerhard.Ehret@dlr.de

M. Hagen
e-mail: Martin.Hagen@dlr.de

A. Petzold
e-mail: a.petzold@fz-juelich.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_50,

50.1 Introduction

Solar radiation drives *atmospheric circulation* and hence *weather* and *climate*. Tropospheric and stratospheric chemistry are controlled by *photochemical reactions* and thus by *shortwave radiation*. Accurate knowledge about shortwave and *longwave radiation* and their interaction with *clouds*, *aerosol* particles, and trace gases is therefore required for all fields of atmospheric science, in particular for determining *Earth's radiation budget*. Considerable efforts are currently expended on aerosol-cloud interactions, such as ship exhausts, which modify marine stratocumulus by increasing its reflectivity for solar radiation and by potentially suppressing *precipitation*.

To better understand the role of clouds and aerosols in climate, observations of cloud and aerosol microphysical, *optical*, and chemical properties are crucial. While ground-based or airborne instruments provide high-resolution observations with limited spatial and temporal coverage, *satellite* sensors provide data covering large temporal and spatial ranges, but they do not give detailed microphysical or optical properties. Benefits for climate research emerge through combining ground-based, airborne and satellite observations together with model *simulations*.

From 2015 on, the joint European-Japanese Earth Clouds, Aerosol and Radiation Explorer (EarthCARE) mission has the objective to provide collocated observations of clouds, aerosols and radiation with global coverage. EarthCARE will deliver accurate global observations of top of atmosphere (TOA) fluxes with coincident cloud and aerosol properties with the aim to improve the understanding and prediction of climate, to characterize the impact of aerosol and clouds on radiation, and to investigate *feedback* processes in the Earth-atmosphere system.

To achieve these scientific goals, EarthCARE will orbit Earth at an altitude of 390 km and carry an unprecedented collection of active and passive cloud and aerosol instruments on one single satellite platform (Fig. 50.1):

- The ATmospheric LIDar ATLID, a *high spectral resolution lidar* operated at 355 nm with 10 m footprint and *polarization* channels;
- The 94 GHz Doppler Cloud Profiling *Radar* CPR with 500 m sampling distance and 400 m vertical resolution;
- The Multi-Spectral Imager MSI with seven channels in the solar and thermal spectral range, 500 m nadir spatial resolution and a swath width of 150 km;
- The Broad Band *Radiometer* BBR with shortwave and longwave channels, three along-track views and 10 km ground pixel size.

Products from the four EarthCARE instruments (either individually or in synergy) comprise optical, micro- and macro-physical properties of clouds (e.g., ice and liquid water content, *effective cloud particle radius*, as well as cloud top temperature and *height*), vertical velocities (*convection* and/or *sedimentation* processes), *rain* rates (rainfall, drizzle), aerosol properties, as well as narrow band and broadband TOA radiation fluxes. The MSI imager will enable detailed characterization of the BBR footprint. This is particularly important for correctly taking into account the type and spectral signatures of the observed targets because

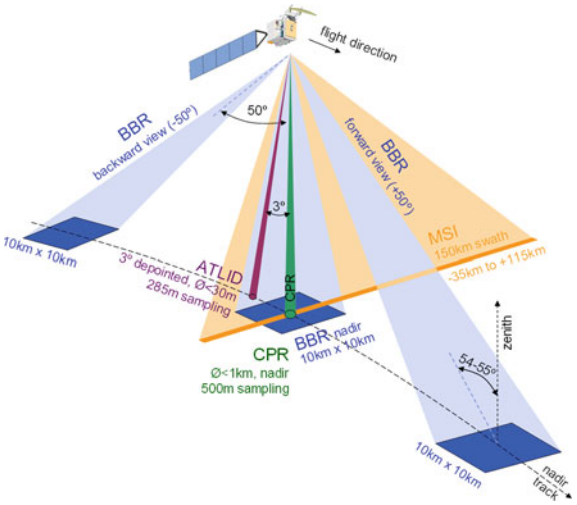

Fig. 50.1 EarthCARE and its four instruments ATLID, CPR, MSI, and BBR. *Source* ESA

they greatly influence the determination of solar and thermal broadband fluxes at top of atmosphere. Furthermore, ATLID and CPR, with their profiling capabilities, will provide detailed knowledge of the vertical structure of the atmosphere along the track. In particular, a lidar is the most powerful tool to obtain range resolved aerosol and *cirrus* cloud information, while radar is sensitive to optically thicker clouds and rain and provides complementary information.

NASA's A-Train (from Afternoon Train), a constellation of satellites spaced a few minutes apart, has a suite of instruments (CALIOP, CPR, MODIS and CERES) similar to EarthCARE's (Anderson et al. 2005). However, the instruments are distributed over several satellites, which prohibits the observation of the same atmospheric volume at the same time.

EarthCARE is a dedicated research-oriented Earth Explorer Core mission of ESA's Living Planet Program. Based on its high scientific value, this mission was selected by ESA for feasibility studies on pre-Phase A and Phase A levels in response to its second call for ideas for Earth Explorer core missions in 2000. In 2004, during a user consultation meeting at ESRIN, Frascati, Italy, EarthCARE received strong support from the earth observation research community. In the end, ESA decided to continue with the implementation phases (Ingmann 2004). Thanks to a strong Mission Advisory Group (MAG) with members from the UK (Slingo 1990; Hogan and Illingworth 2003), France (Pelon et al. 2000), Netherlands (Donovan 2003), Germany (Ansmann et al. 2003; Wandinger et al. 2010), Japan (Nakajima et al. 2001), and others, EarthCARE is well prepared to make a vital contribution to the climate focus with respect to aerosols and clouds. In the following, IPA and LMU activities in connection with preparation, validation and use of EarthCARE data products are described. In addition, ideas are proposed for the next steps.

In 1986, IPA conducted a Phase A study for a Mie *backscatter Lidar* ALEXIS for the German Space Shuttle Mission D2 (Mörl et al. 1988). ALEXIS was designed to measure cloud top heights, the height of the planetary *boundary layer*, *optical thickness*, cloud base height of thin and medium thick clouds, ice/water phase discrimination in clouds, *tropopause* height, tropospheric aerosols, and stratospheric aerosols. Although the mission was not selected (D2 launch 26 April 1993), the ALEXIS study contributed to prepare for lidar in space.

The expertise of IPA and LMU with active and passive remote sensing, radiative transfer, and aircraft experiments is ideally suited for preparing the EarthCARE mission. IPA and LMU will contribute at various ends, including the development of EarthCARE-like equipment for the *HALO* aircraft; organization and analysis of preparatory missions using combinations of instruments such as the SAMUM and EUCAARI field campaigns; the development of an end-to-end simulator with special emphasis on the realism of the scenes with respect to cloud geometry and microphysics as well as aerosol; *retrieval* development; and finally, exploratory studies for a possible Post-EarthCARE mission carrying a multi-wavelength lidar.

50.2 Past and Future Airborne Campaigns

50.2.1 Airborne Demonstrators

Demonstration of the instrument and mission ideas for an airborne platform are very beneficial for a space-borne mission. This is true not only from a technological point of view in order to reach the needed level of maturity, but also with respect to retrieval issues associated with the envisaged data products. In particular, the multi-payload sensor package foreseen for EarthCARE requires advanced algorithm development where the use of aerial test data sets is essential. In the context of the ESA Earth Explorer missions that incorporate lidar observations, two prominent airborne *demonstrator* instruments have been developed at IPA in recent years. The first one is the ALADIN Airborne Demonstrator (A2D) that supports the development of a retrieval algorithm for the *wind* profile of the *Atmospheric Dynamics Mission* ADM-Aeolus. The second one is the Water Vapor Lidar Experiment in Space (WALES) airborne demonstrator that supports future mission ideas for lidar-based observations of the atmospheric water vapor profile or missions like EarthCARE for the measurement of aerosols and clouds. While the ADM-Aeolus mission is scheduled for launch in 2013, the WALES mission, which was studied at the Pre-Phase 0/A level at ESA in 2001–2002 (Ingmann 2001), was finally not selected because of the low level of maturity of major technical subsystems, like the four-wavelength *laser* transmitter that is needed to cover the large dynamic range of atmospheric water vapor concentrations. The system concept and instrument performance of the WALES demonstrator is described in detail by Wirth et al. (2009). Beside its multispectral and *polarimetric* capabilities, a special feature of this instrument is the

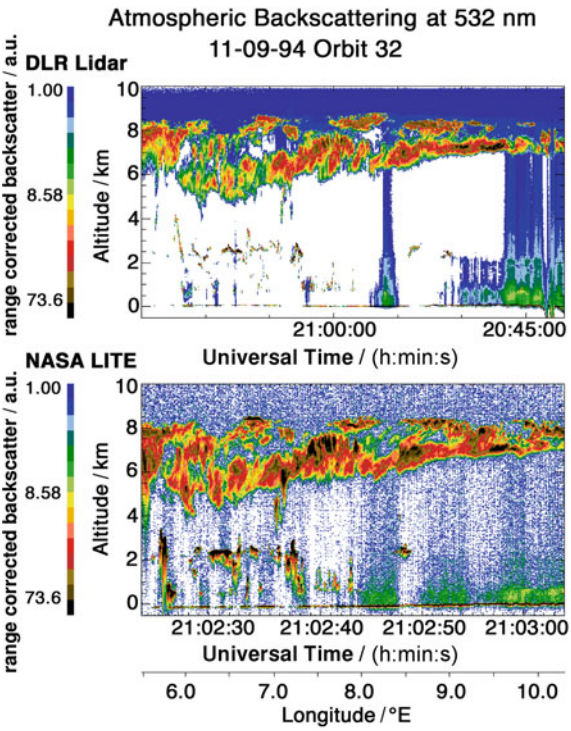

Fig. 50.2 Color-coded time-height cross sections of the range-corrected backscatter signal in arbitrary units for orbit 32 during the shuttle lidar experiment LITE. The lower panel is the shuttle lidar (NASA LITE) result, the upper panel the airborne IPA lidar ALEX-F measurements

high spectral resolution (HSR) lidar option, which is also a key feature of ATLID on EarthCARE. It enables the direct measurement of the particle extinction profile without relying on further assumptions about the particle extinction to *backscatter ratio*, commonly referred to as the *lidar ratio*. It is worth noting that another airborne HSR lidar is operated by NASA in the USA (Hair et al. 2001).

IPA's current airborne lidar instruments have their origin in IPA's ALEX-F (Airborne Aerosol Lidar Experiment) system (Mörl et al. 1981). This instrument was developed more than 30 years ago for deployment on the *Falcon* aircraft. In 1994, it was used to validate the first Lidar-In-Space Technology Experiment (LITE) that was flown on NASA's space shuttle. The measurements were conducted as part of the European contribution to the LITE project (ELITE) that was coordinated by ESA at that time. Figure 50.2 shows an example from correlated aerial and space lidar measurements of a cirrus cloud with a cloud top of about 8 km. Both cross sections agree remarkably well. Differences can be seen at the cloud base, which can be attributed to multiple *scattering* in the space shuttle lidar measurements.

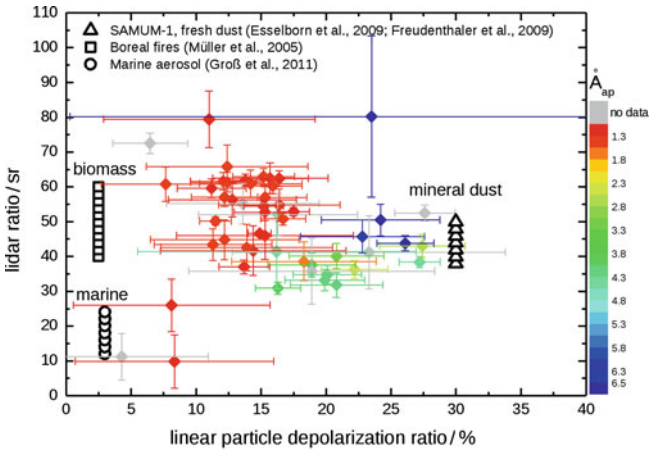

Fig. 50.3 Classification of different aerosol types by their lidar ratio and depolarization ratio. Both are available from high spectral resolution lidar observations with polarization information. For the SAMUM-2 data set, the Ångstrøm exponent of absorption $Å_{ap}$, as derived from in-situ absorption measurements, is given by color-coded values. The black triangles, circles and squares are data from earlier studies without absorption data (Müller et al. 2005; Esselborn et al. 2009; Freudenthaler et al. 2009; Groß et al. 2011). Figure adapted from Weinzierl et al. (2011)

50.2.2 Field Campaigns

The great potential of the EarthCARE concept for aerosols was demonstrated during three major field campaigns, namely the two Saharan Mineral Dust Experiments (SAMUM-1 and SAMUM-2) and the European Integrated Project on Aerosol, Cloud, Climate, and Air quality Interactions (EUCAARI). For these experiments, IPA's airborne HSR lidar WALES was used to measure the aerosol extinction profile, while the *size distribution* and other microphysical and optical properties were obtained from an extensive set of IPA's aerosol in situ probing instruments. The entire airborne instrumentation was operated aboard the IPA's Falcon aircraft.

The goal of the SAMUM-1 campaign was to obtain vertical profiles of natural and anthropogenic aerosols on a regional scale over the Saharan desert, over the North Atlantic area in the outflow of Africa and/or Europe, and over Portugal in the case of *biomass burning* events. Since profiles of aerosol properties are required for calculating aerosol *forcing*, key parameters such as the *extinction* and *absorption coefficients*, the large-scale structure of elevated *desert dust* and biomass burning aerosol layers, and particle size distribution, shape and chemical composition were measured. Although the specific target wavelength of the ATLID instrument (355 nm) was not directly matched by IPA's HSR lidar that operates at 532 nm, the airborne HSR lidar measurements were validated by the ground-based Raman lidar of the Institute for Tropospheric Research in Leipzig which operated simultaneously at both 532 and 355 nm. SAMUM-2 focused on

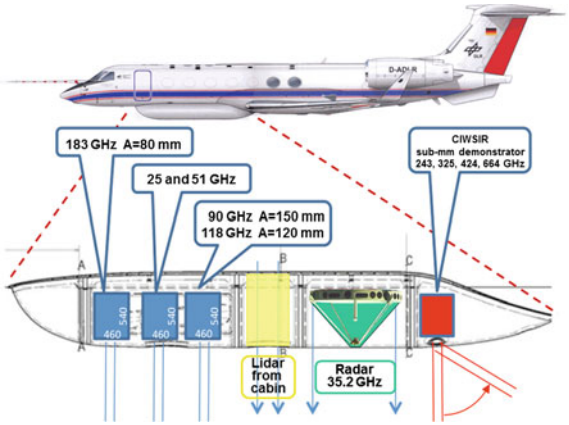

Fig. 50.4 Schematic layout of the installation of the HAMP radar and DLR lidar on HALO. *Blue* denotes the three microwave radiometer racks, *yellow* is the lidar, *green* the radar and *red* the CIWSIR (Cloud Ice Water Submillimeter Imaging Radiometer) instrument package. A defines the diameter of the apertures and the radiometer, radar and CIWSIR frequencies are as given. The numbers in the *blue* boxes specify the radiometer dimensions in mm

the observation of aged Saharan dust and biomass burning aerosols from the African tropical regions, and the objective of the EUCAARI field campaign was aerosol-cloud-climate and air quality interactions. One important finding of SAMUM and EUCAARI was that the measurement of two aerosol-specific quantities, the lidar ratio and the linear particle *depolarization* ratio, can be used to distinguish different aerosol types (Groß et al. 2011; Weinzierl et al. 2011). For example, a depolarization ratio of (31 ± 3) % at 532 nm is characteristic for pure Saharan dust (Freudenthaler et al. 2009), whereas aged Siberian and Canadian forest *fire* smoke layers typically show depolarization ratios of 2–3 % (Müller et al. 2005). In the case of mixtures, the *Ångstrøm* exponent of absorption, which describes the wavelength dependence of the absorption coefficient, can be useful for distinguishing different aerosol types. Figure 50.3 shows the results of the aerosol classification on the basis of HSR lidar observations and in-situ measurements during these three campaigns..

50.2.3 Future Demonstration and Validation with HALO

The new research aircraft HALO provides an ideal platform for the demonstration and validation of EarthCARE measurements and retrieved products. HALO can fly above clouds with downward looking remote sensing instruments. HALO will fly the HAMP (HALO *microwave* package), a combination of a highly-sensitive 35.2 GHz cloud radar and five microwave radiometers covering different frequency bands (Fig. 50.4).

The HAMP instruments will be located in the protective *belly pod* beneath the aircraft fuselage. The belly pod will also have an additional lidar window so that it

will be possible to fly HAMP in combination with the WALES lidar, thus providing a synergetic combination similar to the EarthCARE instruments. The lower frequency of the radar (35 compared to the 94 GHz CPR) allows for better penetration of the radar beam even through convective systems in the *tropics*, compared to higher frequencies. The radar polarization provides additional information about cloud particle properties, for example, whether particles are spherical or have irregular shapes. Due to Doppler signal processing the radar can estimate the vertical velocities of cloud particles. A beam-steering technique for the radar will allow in the future the retrieval of the wind profile along the flight path. The passive microwave radiometers will operate in the *oxygen* absorption bands at 60 and 118 GHz as well as in the water vapor absorption bands near 22 and 183 GHz. This combination of passive microwave sensors will allow for application of a number of precipitation and cloud microphysics retrieval algorithms to extract, among other information, *liquid* and *ice water paths*, and water vapor and temperature profiles. The combination with the water vapor lidar WALES gives additional retrieval possibilities: the lidar water vapor data can be related to the water vapor profile estimated from the microwave radiometers; thin ice layers can be identified better with lidar than with radar; cloud particle identification algorithms can be implemented. The first flights are scheduled for the NARVAL (Next-generation Aircraft Remote-sensing for Validation Studies) campaign planned for 2013. Even though the instruments are not identical to those proposed for EarthCARE, HALO can provide valuable experience in adaptation and application of algorithms originally developed for upward looking ground based instruments. Once EarthCARE is launched, the HAMP radar and lidar as well as the microwave radiometers will allow for validation of the EarthCARE products by independent measurements when HALO is flying beneath the satellite path.

50.3 End-to-End Simulations

For the development and testing of new remote sensing instruments as well as for validating retrievals, *end-to-end simulation* of the sensor output is an indispensable tool. Simulated observations complement "traditional" methods, that is, comparison of retrieval results with independent observations—if such are available at all. Starting from given cloud and aerosol distributions the satellite observation is simulated to produce data sets where radiation as well as cloud properties are fully known, in contrast to using satellite observations alone where only the *radiances* are available and the accuracy of the derived cloud information cannot be assessed because the real cloud properties are not known. On this basis, retrieval algorithms can be tested and optimized by comparing the retrieved properties with the initial cloud properties.

For such studies it is mandatory to use input data sets which are as realistic as possible, because otherwise the performance of the retrieval under real-world conditions is not properly assessed. Here, consistent, accurate, and bias-free 3D

simulations of two of the four EarthCARE instruments, MSI and ATLID, are shown. The results of our calculations have been compared with those of the EarthCARE simulator developed by ESA (Voors et al. 2007).

50.3.1 Models

The simulations are based on the output of the *numerical weather prediction* model COSMO-DE of the German Weather Service (DWD) and take into account all known characteristics of the EarthCARE payload. This work is based on many years of experience with 3D radiative transfer and with the generation of realistic cloud and aerosol data. COSMO-DE is a high-resolution nonhydrostatic model with a horizontal mesh size of 2.8 km on a 421 × 461 horizontal grid. The model domain encompasses essentially the area of Germany, which is large enough to include a considerable variability of cloud as well as surface properties.

COSMO-DE uses generalized terrain following vertical coordinates which divide the atmosphere into 50 layers from the surface up to 21 km. The prognostic model variables are the wind vector, temperature, pressure, specific *humidity*, cloud liquid and ice water, rain and snow water. For the radiative transfer simulations vertical profiles of pressure, temperature, specific humidity, cloud liquid water and cloud ice are used together with surface skin temperature, orography, and a land-sea mask. As an example of EarthCARE simulation, a scene from 12 UTC 31 July 2011 has been selected. An extended water cloud was present over Germany, partially overlaid by a cirrus cloud, which is one the most complicated retrieval situations (thin cirrus, multilayer clouds). In order to obtain realistic surface properties, MODIS observations of the surface *albedo* were used because the surface albedo in COSMO-DE is monochromatic. The optical properties of the liquid water clouds were calculated using Mie theory, while for ice clouds the *parameterization* by Baum et al. (2005) was used, which includes a realistic mixture of particle shapes.

With the radiative transfer package *libRadtran* (Mayer and Kylling 2005) and the *Monte Carlo* code for the physically correct tracing of *photons* in cloudy atmospheres MYSTIC (Mayer 2009) the tools are available to accurately simulate all four EarthCARE instruments, including all possible effects such as three-dimensional photon transport (relevant for the imaging instruments), multiple scattering (relevant for the space borne lidar), and *coherent* backscatter (relevant for the radar).

50.3.2 MSI Simulations

First, in Fig. 50.5 a simulation of a *geostationary* observation is shown in order to visualize the COSMO-DE domain and cloud conditions. Results for the visible spectral channel at 0.6 μm are presented. The simulations take into account the

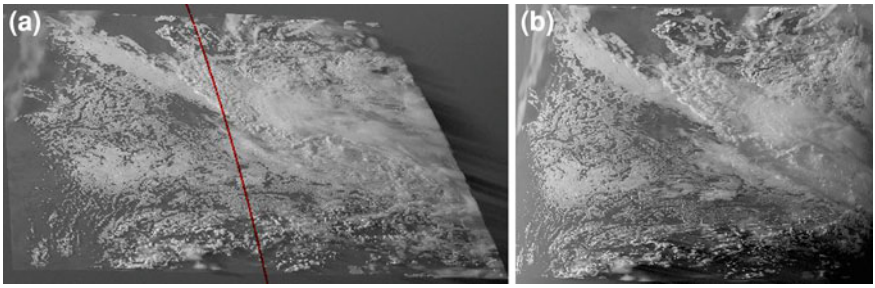

Fig. 50.5 **a** Simulated geostationary Meteosat Second Generation (MSG/SEVIRI) visible channel observation for the COSMO-DE domain. The *red* line indicates the flight path of the EarthCARE satellite for the MSI and ATLID simulations. **b** Simulated MSI (EarthCARE Multi-Spectral Imager) visible channel *image*

sphericity of Earth and the variation of the solar zenith and satellite viewing angle across the scene. Figure 50.5a shows a simulation of the *SEVIRI* sensor aboard the geostationary satellite *Meteosat* Second Generation for the entire COSMO-DE area. For illustrative purposes, in order to emphasize 3D radiative transfer effects, the sun was moved to a position corresponding to 18:00 UTC. The average solar zenith angle in the scene is 80°. MSG/SEVIRI views the full area. On the right the clouds end because of the boundary of the COSMO-DE viewing area. The mentioned cloud deck is clearly visible, as well as the higher cirrus clouds, which appear as fluffy patches. For the following MSI and ATLID simulations, EarthCARE flew along the red line.

Figure 50.5b shows the corresponding MSI scan. The position of the sun corresponds to 18:00 UTC again; the sun has already set in the southeastern part of the scene. The figure looks similar to the MSG/SEVIRI observation, but the viewing geometry is that of a cross-track scanner with nadir view in the center and increasing viewing angle towards the left and right edges. Note that the viewing angle was extended to ±55°, whereas MSI is planned to cover only a viewing angle of ±30°.

50.3.3 ATLID Simulations

The HSR lidar aboard the EarthCARE satellite will allow a distinction to be made between Rayleigh and Mie backscatter. The MYSTIC lidar simulator can also simulate in HSR mode for an ideal instrument (cross-talk between the Rayleigh and Mie channels has not yet been included in the model). The flight altitude as well as the opening angle of ATLID are correctly considered. With the help of variance reduction techniques (Buras and Mayer 2011) the lidar simulator in MYSTIC can quickly calculate the lidar signal for the EarthCARE satellite to about 1 % accuracy, including all orders of multiple scattering and providing also the Jacobian matrix needed for retrieval algorithms. Figure 50.6 shows a

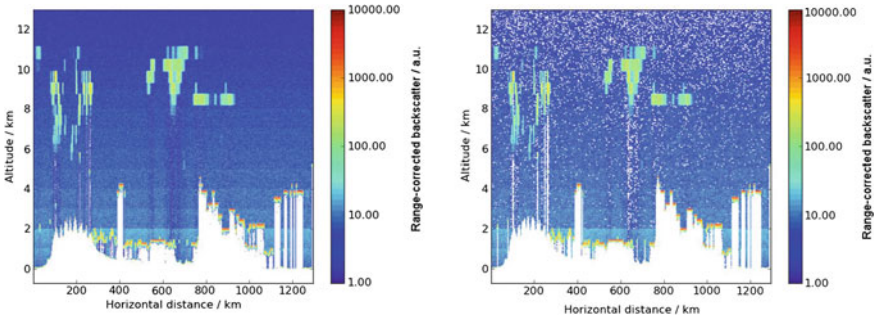

Fig. 50.6 Simulated ATLID signal: range-corrected backscatter; high accuracy MYSTIC simulation (*left*); realistic instrument noise added (*right*)

simulation of the lidar signals measured by EarthCARE passing over the cloud scene from the previous section. The left plot shows a simulation without noise; in the right plot observational photon noise was added corresponding to a laser shot containing 3×10^6 photons. The optically thick lower clouds and the optically thin cirrus clouds are clearly visible.

50.4 Post-EarthCARE Mission Development

Future space-borne aerosol lidar missions must allow an accurate characterization of the aerosol in terms of type (e.g., dust, maritime, fire smoke, urban haze), vertical layering, light extinction and absorption properties, and microphysical properties. The separation of anthropogenic from natural aerosols is essential for assessing the human impact on climate. For reaching that goal, the lidar instruments and algorithms should be able to report the identification of aerosol layers in the *troposphere* and *stratosphere* and vertical profiles of optical and microphysical properties with known and specific precision. These include backscatter and extinction coefficients at selected wavelengths, lidar ratio, Ångstrøm coefficients, aerosol type, volume and surface concentrations, size distribution parameters, and *refractive index*.

High-level aerosol microphysical properties necessary for evaluating the *climate impact* of anthropogenic aerosols are the aerosol optical depth and *single-scattering albedo* in the visible spectral range (typically at 550 nm), the effective particle diameter, and the hemispheric backscatter fraction. For achieving an accuracy of 1 W m^{-2} in TOA *radiative forcing*, the required accuracy of the measured aerosol optical depth should be 0.01–0.02 at 500 nm, and the single scattering albedo should be constrained to 0.02 over land. Threshold values for *backscatter coefficients* are 0.0005 km^{-1} sr^{-1}, corresponding to 0.02 km^{-1} for extinction coefficients.

Current space-borne high-spectral resolution lidar concepts like ATLID for EarthCARE or ALADIN for ADM-Aeolus are based on a single-wavelength HSR channel and on one backscatter channel. While this is extremely valuable information compared to passive remote sensing, more detailed information would be desirable for fully quantifying the aerosol radiative forcing and atmospheric response to the forcing. EarthCARE and ALADIN data have limited spatial and temporal coverage. Furthermore, current aerial lidar measurements are not very sensitive to aerosol absorption. The IPA-coordinated ESA study ICAROHS (Inter-Comparison of Aerosol Retrievals and Observational Requirements for Multi-wavelength HSRL Systems) investigated the potential improvements and benefits of novel multiwavelength HSR lidar technology combined with innovative retrieval methods for future satellite missions for meeting these requirements (Petzold 2011). In fact, apart from the aerosol optical depth AOD the high-level aerosol parameters cannot be measured directly by lidar. Instead, they must be derived from a suitable set of multiwavelength aerosol backscatter and extinction measurements. ICAROHS demonstrated that a data set consisting of backscatter coefficients β at 355, 532, and 1 064 nm and extinction coefficients α at 355 and 532 nm measured with a multiwavelength aerosol lidar (denoted as 3 β + 2 α in the following) allows the retrieval of microphysical properties and of the refractive index characteristics from the measured spectrally resolved optical properties with the required accuracy. Such lidars are currently only operated from the ground, e.g., the MULIS lidar of LMU. Extinction and backscatter coefficients must be known with an uncertainty (standard deviation) of the order of 10 % to allow successful inversion of the optical data into microphysical properties and other important aerosol parameters. An additional depolarization channel provides a clear indication of the presence of desert dust, the most important natural continental aerosol component and of *ash* in forest fire smoke and aerosol intrusions in the stratospheric layer.

The configuration of the space-borne 3 β + 2 α system with a passive remote sensor that detects aerosol properties at several visible and near-infrared wavelengths could further improve the aerosol characterization. It would be desirable for this passive sensor to be a multiangle instrument such as NASA's MISR satellite. Such a sensor could provide information about the *phase function* of the observed particles. The combination of data from such a passive remote sensor with the multiwavelength lidar data could for the first time provide height-resolved information on the phase function (respectively asymmetry parameters) of particles at several wavelengths, which goes far beyond any currently applied aerosol remote sensing technology.

In another step the passive remote sensor should be equipped with polarization-sensitive channels. These channels would allow for better separation of the satellite signals from aerosols from those from Earth's surface. Polarization-dependent information will lead to a better characterization of nonspherical particles like mineral dust. In addition, such polarization signals could lead to better insight into aerosols which consist of internal mixtures of different aerosol types.

References

Anderson, T.L., Charlson, R.J., Bellouin, N., Boucher, O., Chin, M., Christopher, S.A., Haywood, J., Kaufman, Y.J., Kinne, S., Ogren, J.A., et al.: 2005: An "A-Train" strategy for quantifying direct climate forcing by anthropogenic aerosols. Bull. Amer. Meteorol. Soc. **86**, 1795–1809 (2005). doi:10.1175/BAMS-86-12-1795

Ansmann, A., Bösenberg, J., Chaikovsky, A., Comerón, A., Eckhardt, S., Eixmann, R., Freudenthaler, V., Ginoux, P., Komguem, L., Linné, H., et al.: Long-range transport of Saharan dust to northern Europe: The 11–16 October 2001 outbreak observed with EARLINET. J. Geophys. Res. **108**, 4783 (2003). doi:10.1029/2003JD003757

Baum, B.A., Heymsfield, A.J., Yang, P., Bedka, S.T.: Bulk scattering properties for the remote sensing of ice clouds. Part I: Microphysical data and models. J. Appl. Meteor. **44**, 1885–1895 (2005). doi: 10.1175/JAM2308.1 http://dx.doi.org/10.1175/JAM2308.1

Buras, R., Mayer, B.: Efficient unbiased variance reduction techniques for Monte Carlo simulations of radiative transfer in cloudy atmospheres: The solution. J. Quant. Spectrosc. Radiat. Transf. **112**, 434–447 (2011). doi:10.1016/j.jqsrt.2010.10.005

Donovan, D.P.: Ice-cloud effective particle size parameterization based on combined lidar, radar reflectivity, and mean Doppler velocity measurements. J. Geophys. Res. **108**, 4573 (2003). doi:10.1029/2003JD003469

Esselborn, M., Wirth, M., Fix, A., Weinzierl, B., Rasp, K., Tesche, M., Petzold, A.: Spatial distribution and optical properties of Saharan dust observed by airborne high spectral resolution lidar during SAMUM 2006. Tellus. **61B**(1), 131–143 (2009). doi: 10.1111/j.1600-0889.2008.00394.x

Freudenthaler, V., Esselborn, M., Wiegner, M., Heese, B., Tesche, M., Ansmann, A., Müller, D., Althausen, D., Wirth, M., Fix, A., et al.: Depolarization ratio profiling at several wavelengths in pure Saharan dust during samum 2006. Tellus **61B**, 165–179 (2009). doi:10.1111/j.1600-0889.2008.00396.x

Groß, S., Tesche, M., Freudenthaler, V., Toledano, C., Wiegner, M., Ansmann, A., Althausen, D., Seefeldner, M.: Characterization of Saharan dust, marine aerosols and mixtures of biomass burning aerosols and dust by means of multi-wavelength depolarization- and Raman measurements during SAMUM-2. Tellus. **63B**(4), 706–724 (2011). doi: 10.1111/j.1600-0889.2011.00556.x

Hair, J.W., Caldwell, L.M., Krueger, D.A., She, C.-Y.: High-spectral-resolution lidar with iodine-vapor filters: measurement of atmospheric-state and aerosol profiles. Appl. Opt. **40**, 5280–5294 (2001). doi:10.1364/AO.40.005280

Hogan, R.J., Illingworth, A.J.: Parameterizing ice cloud inhomogeneity and the overlap of inhomogeneities using cloud radar data. J. Atmos. Sci. **60**, 756–767 (2003). doi: 10.1175/1520-0469(2003)060<0756:PICIAT>2.0.CO;2

Ingmann, P.: WALES-water vapour lidar experiment in space, ESA, SP-1257 (2001)

Ingmann, P.: EarthCARE—earth clouds, aersols, and radiation explorer, Tech. Rep., ESA, Sp-1279(1) (2004)

Mayer, B., Kylling, A.: The libRadtran software package for radiative transfer calculations: Description and examples of use. Atmos. Chem. Phys. **5**, 1855–1877 (2005). doi: 10.5194/acp-5-1855-2005

Mayer, B.: Radiative transfer in the cloudy atmosphere. European Phys. J. Confer. **1**, 75–99 (2009). doi:10.1140/epjconf/e2009-00912-1

Mörl, P., Reinhardt, M.E., Renger, W.: The use of the airborne lidar ALEX-F for aerosol tracing in the lower troposphere. Beitr. Phys Atmos. (Contr. Atm. Phys.), **45**, 403–410 (1981)

Mörl, P., Quenzel, H., Renger, W., Schmitz-Peiffer, A., Schumann, U., Weber, E., Wiegner, M.: ALEXIS—atmospheric LIDAR experiment in space—phase-A study report, DFVLR-Mitt. 88–10, pp. 84 (1988)

Müller, D., Mattis, I., Wandinger, U., Ansmann, A., Althausen, D., Stohl, A.: Raman lidar observations of aged Siberian and Canadian forest fire smoke in the free troposphere over

Germany in 2003: microphysical particle characterization. J. Geophys. Res. **110**, D17201 (2005). doi:10.1029/2004JD005756

Nakajima, T., Higurashi, A., Kawamoto, K., Penner, J.E.: A possible correlation between satellite-derived cloud and aerosol microphysical parameters. Geophys. Res. Lett. **28**, 1171–1174 (2001). doi:10.1029/2000GL012186

Pelon, J., Flamant, C., Trouillet, V., Flamant, P.H.: Optical and microphysical parameters of dense stratoculus clouds during mission 206 of EUCREX'94 as retrieved from measurements made with the airborne lidar LEANDRE 1. Atmos. Res. **55**, 47–64 (2000)

Petzold, A.: STSE—ICAROHS—inter-comparison of aerosol retrievals and observational requirements for multi-wavelength HSRL systems, Final Report. Tech. rep., ESA (2011)

Slingo, A.: Sensitivity of the Earth's radiation budget to changes in low clouds. Nature **343**, 49–51 (1990). doi:10.1038/343049a0

Voors, R., Donovan, D., Acarreta, J., Eisinger, M., Franco, R., Lajas, D., Moyano, F.P.R., Ramos, J., Wehr, T.: ECSIM: The simulator framework for EarthCARE. Proceedings of SPIE, 67441Y (2007)

Wandinger, U., Tesche, M., Seifert, P., Ansmann, A., Müller, D., Althausen, D.: Size matters: Influence of multiple scattering on CALIPSO light-extinction profiling in desert dust. Geophys. Res. Lett. **37**, L10801 (2010). doi:10.1029/2010GL042815

Weinzierl, B., Sauer, D., Esselborn, M., Petzold, A., Veira, A., Rose, M., Mund, S., Wirth, M., Ansmann, A., Tesche, M., et al.: Microphysical and optical properties of dust and tropical biomass burning aerosol layers in the Cape Verde region—an overview of the airborne in situ and lidar measurements during SAMUM-2. Tellus. **63B**, 589–618 (2011). doi:10.1111/j.1600-0889.2011.00566.x

Wirth, M., Fix, A., Mahnke, P., Schwarzer, H., Schrandt, F., Ehret, G.: The airborne multi-wavelength water vapor differential absorption LIDAR WALES: system design and performance. Appl. Phys. B Lasers Opt. **96**, 201–213 (2009). doi:10.1007/s00340-009-3365-7

Roots, Foundation, and Achievements of the "Institut für Physik der Atmosphäre"

51

Hans Volkert and Dania Achermann

Abstract

This chapter collects some basic ingredients of the research carried out at the Institute of Atmospheric Physics (Institut für Physik der Atmosphäre-IPA) during the five decades 1962–2012. A few older roots are mentioned and the initial vision of the founding director is recalled. A march through five decades presents important achievements and illustrates the contributions which IPA made to the steady development of the atmospheric sciences in experiment and simulation during the past half century. A vision into the decade ahead and a look at partner institutes of similar age conclude the chapter.

51.1 Institutional Roots of Atmospheric Physics in Germany

Around 1890, Wilhelm von Bezold, director of the Königlich Preußisches Meteorologisches Institut in Berlin, coined the term *Physik der Atmosphäre* for a general research program concerning the upper levels of the atmosphere (Höhler 2007). The foundation of the Aeronautisches Observatorium Lindenberg, today an observatory of Deutscher Wetterdienst (DWD), in 1905 and the inauguration of the

H. Volkert (✉) · D. Achermann
DLR, Institute of Atmospheric Physics (IPA), Münchner Straße 20,
82234 Oberpfaffenhofen, Germany
e-mail: Hans.Volkert@dlr.de

D. Achermann
e-mail: Dania.Achermann@dlr.de

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4_51,

Deutsche Versuchsanstalt für Luftfahrt (DVL) in 1912 constituted important seeds for the interlinked development of *aviation* and related meteorological research around Berlin. The Modellversuchsanstalt der Motorluftschiff-Studiengesellschaft, founded in Göttingen in 1907 and renamed the Aerodynamische Versuchsanstalt (AVA) in 1920, formed another root with the emphasis on engineering aspects.

The combination of sportive competition between pilots and atmospheric research prevailed from 1924 in the *Rhön-Rossitten-Gesellschaft*, the forerunner of the Deutsche Forschungsanstalt für Segelflug (DFS; 1933–1945). The latter contained a research unit named Institut für Physik der Atmosphäre, which was led by Fritz Höhndorf (Georgii 1969). After the Second World War, German research activities in aviation could be resumed after 1954. Harald *Koschmieder* (1897–1966), head of the new DFS unit Institut für Flugmeteorologie (aviation meteorology) advocated a German Forschungszentrum Physik der Atmosphäre in 1959, in analogy to the plans to create a National Center for Atmospheric Research (NCAR) in the United States. Walter *Georgii* (1888–1968) led the DFS unit Institut für Flugraumforschung (air-space research) at München-Riem airport and combined broad aerodynamic expertise in soaring flight with meteorological research interests. Both institute directors retired in 1961, during a period when plans gained momentum to overcome the fragmentation of aviation related research and to secure steady institutional funding from federal ministries.

Extended negotiations with Hans Gerhard *Müller* (1905–1978), director of the aviation meteorology and radiosonde teams of DWD at München-Riem airport, resulted in the formal merger of both institutes by 1 July 1962 under the more general name, which also referred to a tradition: Institut für Physik der Atmosphäre. Müller was well connected, also internationally (e.g., he reported to the Radiation Commission of the International Association of Meteorology and Atmospheric Physics; Ohring et al. (2009); person #102 on the unique group photograph). He had obtained a leave of absence from his tenured position as a senior civil servant with DWD by November 1961 and continued to lecture as *professor extraordinarius* at the meteorological institute of Munich university.

These early institutional roots and their duration are summarized in Fig. 51.1. A subtle analogy becomes apparent between the atmosphere and research aiming at improving knowledge and reducing forecasting uncertainties. The atmospheric continuum is generally divided into scales (e.g., *global*, *synoptic*, *mesoscale*, *microscale*) and comprises many physical processes all acting at the same time with variable intensity (e.g., dynamics, radiation, phase changes of water). Likewise the research organization tends to be clustered in larger units with embedded science institutes and highly dependent on the prevailing political situation, while the research interests combine more basic foci (e.g., atmospheric processes for weather forecasting or climate studies) with applied ones (e.g., atmospheric influences on aircraft performance). Currently, the research portfolio of DLR embraces the wide field spanned by aviation and atmosphere in a number of research institutes. This overview concentrates on the atmospheric physics achievements of IPA scientists.

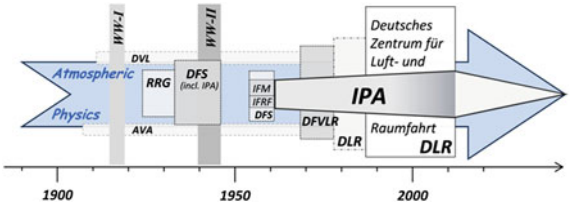

Fig. 51.1 Institutional roots and embedding of IPA since Wilhelm von Bezold defined Atmospheric Physics in 1890. Abbreviations: AVA—Aerodynamische Versuchsanstalt; DFS—Deutsche Forschungsanstalt für Segelflug; DFVLR—Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt; DVL—Deutsche Forschungsanstalt für Luftfahrt; IFM—Institut für Flugmeteorologie; IFRF—Institut für Flugraumforschung; IPA—Institut für Physik der Atmosphäre; RRG—Rhön-Rossitten-Gesellschaft; WW—World War

51.2 Purpose and Tasks of an Aerological Observatory Called "Institut für Physik der Atmosphäre"

On 25 April 1962, during the mentioned transition period, Hans Gerhard Müller (Fig. 51.2) presented his vision of the function and tasks of the united institute under his direction to the scientific advisory committee of DWD using as a concise title: *Aerologisches Observatorium* (aerological observatory; Müller (1962)). The content of his 13-page typescript is used here as a 'verification tool' to judge the coming achievements.

First, he distinguished tasks of routine observations and special measurements within meteorological services and explained the nature of scientific observatories, including the development of innovative instrumentation in specialized laboratories. Research interests of university institutes, weather service offices and aviation authorities are mentioned, including their varying emphases regarding a more basic or a more applied orientation. Müller referred to the need for a new aerological observatory since the traditional prototype in Lindenberg near Berlin was meanwhile situated in East Germany. He called the intended fusion of the smaller DFS institutes under his direction a pragmatic development for the benefit of atmospheric research in West Germany. *Aerology*, i.e., measurements of the vertical structure above the atmospheric boundary layer, should make especial use of aircraft rather than mainly balloons, not least in order to become again competitive in the international arena.

In the main part of his presentation, Müller provided a selection of several broad research areas (translated and put in italics here). *Flight physics* denotes exchange processes at altitudes which are relevant for aviation and during the start phase of large rockets. Also mentioned are problems of air pollution by radioactive and other substances. In *cloud physics*, where previous pioneering achievements by German scientists were supplemented by laboratory studies abroad, a more meteorological perspective of condensation and sublimation processes was deemed necessary,

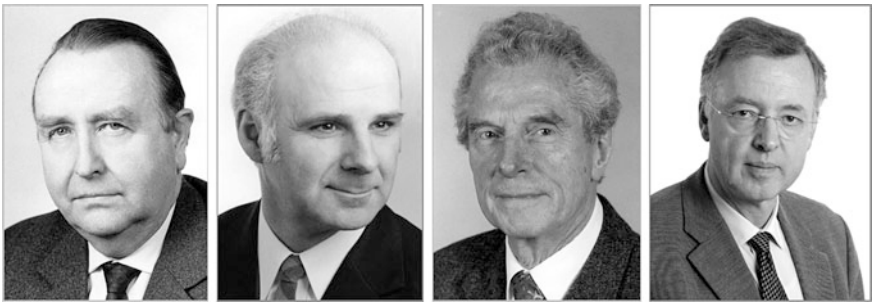

Fig. 51.2 IPA directors with periods of office and birth/death years (from *left*): Hans Gerhard Müller (1962–1972; *1905, +1978); Heinz Fortak (1973–1976; *1926); Manfred E. Reinhardt (1974–1992; *1927); Ulrich Schumann (1982–2012; *1945)

including topographic triggering and mesoscale structures of precipitation. Within the *stratosphere* a better knowledge of horizontal structures was sought to complement the sparsely placed radio soundings as air traffic was seen to increase in the region above the tropopause. In the context of international cooperation, *high atmosphere* problems in the range from 35 to 60 km were seen to be relevant, with reference to plans for a European rocket. Studies regarding the *radiation budget* were to comprise measurements of brightness temperatures at inhomogeneous surfaces and of infrared fluxes within the free atmosphere, including automatic retrievals by an electronic computer. Evidentially, the established field of *atmospheric optics* ought to provide practical applications soon, e.g., for robust visibility measurements at airports. More generally, the optical determination of the size and number density of aerosol particles was regarded as an important field of basic research.

Finally, the research areas were also viewed from the perspective of improvement of current *measuring techniques*. Müller explicitly referred to the development of improved sensors for temperature above the 30 km level, including campaigns for testing and calibration, new approaches towards more reliable humidity measurements, and the application of the recently invented laser technique for concentration measurements of atmospheric aerosol and visibility. In combination with more complex observational tools such as precipitation *radar* and instrumented *aircraft*, a broad experimental basis was sought to the mutual benefit of aviation research and the meteorological service.

From today's perspective it is interesting to note that four of the five current IPA departments are concerned with experimental tasks mentioned by Müller back in 1962 and that they are named accordingly*: Atmosphärische Spurenstoffe* (trace species), *Fernerkundung der Atmosphäre* (*remote sensing* including *radiative transport*), *Wolkenphysik und Verkehrsmeteorologie* (cloud physics and *transport meteorology*), *Lidar* (including laser development). The overarching aim of all these techniques lies in a thorough investigation of physical and, increasingly, chemical processes in the *atmosphere*.

51.3 Self-Determined Versus Programmatic Research Funding

Publicly financed research activities generally take place following two different principles set by the funding agency: On the one hand, there is a fully self-determined style where resources are given to a creative group around an academic leader, trusting that relevant results will be achieved sooner or later. In the programmatic approach, on the other hand, governmental institutions define the tasks for which they expect solutions, sooner rather than later. In Germany, fundamental research activities at Max-Planck-institutes follow the former model, while research and development within a federal institution such as DWD are closely tied to service targets, partly officially defined by legislation. Research establishments such as DLR and its predecessors are situated in between, and the concrete orientation of a research institute like IPA depends on a number of factors, which are not constant over time. Their combination has a considerable influence on the prevailing research perspectives.

Müller (1962) explicitly declared that IPA's activities should be partly determined by practical demands expressed and paid by federal ministries, e.g., those relating to transport and defence. In the beginning, IPA was still part of the Flugwissenschaftliche Forschungsanstalt München (FFM), financed by the state of Bavaria. In 1963, it was transferred to the mainly federally funded Deutsche Versuchs-anstalt für Luft- und Raumfahrt (DVL), which in turn was merged in 1969 with other institutions to become the Deutsche Forschungs- und Versuchs für Luft- und Raumfahrt (DFVLR; Trischler (1992)). While the name and the status of the parenting institution changed two more times in 1989 and 1997 (Fig. 51.1), the role of IPA as a self-contained research institute remained the same.

During all five decades the IPA budget has been composed of a significant base part provided by the parenting institution plus the revenue from research contracts from numerous sources. The proportion of these 'third party funds' has grown significantly over time. During the past two decades all funding became increasingly project-oriented. The growing importance of DLR-wide projects, of national and European research initiatives (e.g., priority programs of DFG and BMBF, and various environmental research themes supported by the EU research directorate since 1984) induced useful networks with numerous cooperation partners, but also led to a higher administrative load for project acquisition and monitoring. Broad initiatives to monitor the environmental impact of transports systems (e.g., aviation, shipping) and to compare anthropogenic emissions with natural sources (e.g., nitric oxides by lightning) stem from seminal publications by IPA scientists. Since about 2000, institutions within the Helmholtz-Gemeinschaft (HGF), as DLR, have been receiving large parts of their basic funding through competitive 'programmatic steering' mechanisms.

Conducting and completing a dissertation project in cooperation with an appropriate university institute enhances the fundamental research component, even within projects of a more applied character. Until 1982 eight staff scientists at

Year	Politics / Economics	DVL / DFVLR / DLR	Nature / Science / Technology
62		IPA within FFM	
63	Federal govt. under Ludwig Erhard	FFM into DVL	Atmospheric chaos (E. Lorenz)
64			1st Chinese atomic bomb test
65			
66	Federal govt. under Kurt-G. Kiesinger		
67		IPA to Oberpfaffenhofen	
68			
69	Federal govt. under Willy Brandt	DVL into DFVLR	1st flight of B747 aircraft
70	1st nuclear power protest in Germany		Unix operating system
71			
72	Employee participation (federal law)		Interactive terminal available
73	Oil crisis / car-free Sundays	Hermann Jordan chairs DFVLR	
74	Federal govt. under Helmut Schmidt		
75			ECMWF and ESA established
76		Delivery of Falcon aircraft	
77			Launch of Meteosat
78			
79			1st operational ECMWF forecast
80	*Die Grünen* – „Green" political party		
81			1st PCs (IBM) / launch of AVHRR
82	Federal govt. under Helmut Kohl	Installation of Cray supercomputer	
83			
84	Start of European research funding		Munich hailstorm / 1st E-Mail
85			Ozone hole article (in *Nature*)
86	Federal environment ministry	Installation of POLDIRAD radar	Tschernobyl accident
87		Walter Kröll chairs DFVLR	1st Sun-4 workstation released
88		Workstations installed at IPA	IPCC launched by WMO & UNEP
89	Fall of Berlin Wall	DFVLR renamed, acronym: DLR	
90	German unification		IPCC: 1st Assessment Report
91			Oil fires in Kuwait
92	Maastricht-treaty of European Union	In-situ chemistry on Falcon	1st ensemble weather forecasts
93			
94		Direct data link from Meteosat	1st SMS from a mobile phone
95	Helmholtz Association (HGF) inaugurated		Launch of ERS-2 satellite with GOME
96			
97		DLR also German space agency	Kyoto protocol signed
98	Federal govt. under Gerhard Schröder		
99		IPA building renovated	
00			
01	Programmatic funding by HGF		IPCC: 3rd Assessment Report
02	Euro as common currency	Sigmar Wittig chairs DLR	Launch of Envisat satellite
03			
04		Portable lightning stations LINET	Meteosat 2nd generation
05	Federal govt. under Angela Merkel		1st flight of A380 aircraft
06			
07		Johann-Dietrich Wörner chairs DLR	Nobel Peace Prize for IPCC
08	Financial crisis („Lehmann collapse")	Acquisition of HALO instruments	
09	New federal govt. under Angela Merkel	Delivery of HALO aircraft	
10			Eruption of Eyjafjalla (Iceland)
11	Nuclear power moratorium in Germany	MERLIN satellite, phase A	Earthquake / tsunami (Japan)
12		100th anniversary of DVL	

Fig. 51.3 Selection of external events with relevance for IPA work and major investments during the five decades 1962–2012

Experimental milestone *(aircraft* or *campaign* name)	Year	Simulation milestone *(model code, software package)*
Airborne observations in foehn flows	**62**	Meridional circulation in stratosphere and mesosphere
Turbulence data at Atlantic tropopause (*Lufthansa*)	**63**	
Mesospheric rocket observations (from Sardinia)	**64**	
Visibility registration for marine applications	**65**	
Receiver construction for high altitude laser returns	**66**	
Review of weather modification experiments in Bavaria	**67**	Optical retrievals of aerosol size spectra
Precipitation intensity inferred from radar reflectivity	**68**	Review of polar light spectra
Airborne sampling of radionucleids	**69**	
Ground-based laser observations of stratospheric structures	**70**	
Fog reduction through aerosol emission	**71**	
Evaluation of hail suppression experiments	**72**	DFVLR concept for „environmental research“
Vertical looking cloud cover indicator (laser method)	**73**	Conception of a Raman lidar system
Airborne probing of cooling tower emissions	**74**	Simulation of microwave transmittance through clouds
Airborne aerosol measurements (over southern Bavaria)	**75**	Simulation concept for urban pollution
Low level maximum wind statistics (Northern/Southern Germany)	**76**	Thermal impact of power plant parks
Four-channel radiative fluxes in the tropics (*GATE-74*)	**77**	Low-resolution global simulations
Airborne backscatter lidar missions (over Po valley)	**78**	Spectral weighting scheme for atmospheric visibility
Stratospheric ozone missions (North Atlantic, Iceland)	**79**	Concept study for „cloud physical research“
Foehn characteristics aloft from balloon soundings	**80**	Linear theory for mesoscale orographic flows
Radiation budget of the tropical atmosphere	**81**	Implementation of *Pielke*-model for Alpine flows
Momentum flux over mountains (*Falcon* at *ALPEX*)	**82**	Temperature profiles from satellite data (*TOVS*)
High-frequency turbulence data (from *Falcon*)	**83**	Bi-spectral satellite data analysis of ice cloud heights
Weather and hailstone analyses of Munich storm	**84**	Elliptic solver for 3-dim. terrain following coordinates
Falcon observations for Alpine heat low study	**85**	Direct simulation of homogenous turbulent shear flows
Aircraft icing experiments	**86**	Automatic cloud retrievals from satellite data (*APOLLO*)
Cold fronts experiment (in Germany, Austria, Switzerland)	**87**	Multi-purpose mesoscale model validated (*MESOSCOP*)
Trans-border pollution transport (data from *Queenair*)	**88**	Large eddy simulations with simple chemistry
Airborne calibration of Meteosat (data from *Falcon*)	**89**	3-dim. simulation of convectively forced gravity waves
Trans-Alpine pollution transport (*MEMOSA*)	**90**	Large eddy simulations of convective boundary layer
Falcon flights in aircraft exhaust plumes	**91**	Basic 3-dimensional turbulence studies (*DISTUF*)
Ozone lidar OLEX on *Transall* aircraft	**92**	First global climate model simulations (*ECHAM*)
In-situ observations across North Atlantic flight corridor	**93**	3-dimensional turbulence over wavy terrain
Fuel sulfur impact on aircraft emissions (*SULFUR-1*)	**94**	Large eddy simulations of aircraft wake vortices
Lidar observations of Artic ozone depletion (*SESAME-2*)	**95**	Revised Schmidt-Appleman criteria for contrail generation
NO_x produced by lightning over Germany (*LiNOx*)	**96**	Detailed condensation trail simulation
Model guided lidar observations of stratospheric ice-clouds	**97**	Mesoscale stratospheric temperature forecasts
Multisensor aerosol observations (*LACE-98*)	**98**	Radiative forcing by condensation trails
Aerial wind and water vapor lidar missions (*MAP-SOP*)	**99**	Noise propagation in realistic atmospheric states
Cirrus formation in both hemispheres (*INCA*)	**00**	Numerical rain forecasts checked with radar data
Pollution and transport (over southern France, *ESCOMPTE*)	**01**	Inertia-gravity waves confirmed as cause for PSC
Tropopause sharpness from radiosonde climatology	**02**	Global ozone-climate interaction simulations
Condensation trails from satellite data (for ESA)	**03**	Studies of past and future ozone layer development
Airborne wake vortex observations by lidar (*I-WAKE*)	**04**	Aerosol-cirrus interaction parameterized (*ECHAM*)
NO_x produced by tropical lightning (*TROCCINOx-2*)	**05**	*libRadtran* for radiative transfer calculations
Air-chemistry and dust in Africa (*AMMA*, *SAMUM*)	**06**	Ice-supersaturation in ECMWF forecast model
Airborne H_2O observation by 3rd generation DIAL (WALES)	**07**	Indirect aerosol effect determined for global ship traffic
Lidar observations around typhoons (*T-PaRC* for WMO)	**08**	Postulated link ocean temperature / stratospheric dynamics
Polar ADM-Aeolus pre-launch validation (for ESA)	**09**	Explanation of aerodynamic contrails
17 volcanic ash missions (lidar & in-situ observations, *Falcon*)	**10**	Realistic large-eddy simulations of aircraft wake vortices
Optical depth of condensation trails (*CONCERT-1*)	**11**	Contrail cirrus global climated impact estimated
Chemistry and transport in thunderstorm anvils (*DC3*)	**12**	Contrail prediction tool (*CoCiP*) tested

Fig. 51.4 Selected experimental and simulation achievements and conceptual milestones at IPA during the five decades 1962–2012

IPA extended their project results to PhD theses. Since 1983 more than 90 scientists completed PhD dissertations while working at IPA, mostly financed by dedicated PhD positions over 3 years. Thereby the high ranking aim expressed in DLR's statutes ("further training of junior scientists" [Sect. 1, 2d)]) was continuously met and academically oriented creativity inspired in a variety of disciplines (e.g., mathematics, physics, meteorology, chemistry, mechanical engineering). In 2012, some 25 of the former PhD students are employed at either DLR in Oberpfaffenhofen (18 at IPA) or at both Munich universities, while others work at research institutes and meteorological services in Germany and abroad (Austria, Switzerland, Italy, UK, USA) as well as for industrial partners (e.g., Lufthansa).

51.4 Experimental and Simulation Achievements Over Five Decades

While all previous chapters in this book provide detailed descriptions of specific fields of IPA expertise with some historical background and cross-links, the overall achievements are sketched here in broad steps by decade. Some emphasis is placed on the four directors (Fig. 51.2) whose periods of office tended to follow a 10-year rhythm. Some relevant external events are recalled year-by-year in Fig. 51.3 in the category *Politics* and *Economics*, changes within the overarching organizations, and developments under *Nature, Science* and *Technology*. Fig. 51.4 contains an analogous attempt to select single achievements for every year, one each in what evolved as the main pillars of IPA work: *experiment* and *simulation*. More detailed collections of experimental and simulation achievements are available for the first two decades in two volumes of assembled annual reports (Reinhardt 2005; Reinhardt 2010) and in five institute reports produced for regular evaluations during the past three decades at roughly 5-year intervals (Reinhardt and Schumann 1987; Schumann 1992, 1997, 2003, 2007).

51.4.1 The First Decade: Consolidation and Move to Oberpfaffenhofen

Hans Gerhard Müller systematically followed his agenda as described in Sect. 51.2 (Müller 1964). The traditional link to aviation remained strong. Campaigns were for instance undertaken on regular Lufthansa airliners to document temperature and humidity at stratospheric levels, including turbulence as a severe hazard. The construction of 'visibility meters' eventually resulted in regular deployments along motorways and runways at airports. Considerable effort went into the administrative integration of IPA into DVL and the 40-km move to the growing research center at Oberpfaffenhofen, where a large building was completed in 1967 (Fig. 51.5, Müller (1968)).

Fig. 51.5 IPA building at DVL/DLR research center in Oberpfaffenhofen. *Left*: spring 1968 shortly after completion with 1.4-m dish of a precipitation radar on the roof, also containing offices of other DVL branches. *Right*: January 2012, renovated with new entry pavilion, extended 5th floor and 5.5-m dish of POLDIRAD on the roof, occupied solely by IPA staff

Weather modification experiments were generally en vogue in the 1960s with fog dissolution at airports being of practical importance, especially in relation to the planned transfer of the Munich airport from Riem to the moist fields between Erding and Freising. The new 'measurement field' was equipped, e.g., with machinery to eject aerosol particles for fog reduction experiments (Fig. 51.6). Moreover, aerial environmental monitoring of various pollutants was important. The determination of radionuclide concentrations after the atomic bomb tests in the atmosphere as far away as China (May 1966) gained political relevance; a maximum was found in a height of 13 km 11 days after the first explosion from systematic air sampling by an ad-hoc equipped military aircraft (Lockheed T33). The twin-engine *Queenair* (Fig. 51.7), purchased by DVL in 1962, was converted to an airborne laboratory using different equipment for a variety of missions. It was regularly used by IPA staff. Their expertise proved to be essential for the successful operation of a jet aircraft dedicated to atmospheric research from 1976 onwards.

At the end of 1972, Müller retired from his position as IPA director, but continued his research dealing with the dissipation of super-cooled fog. He remained a charismatic figure, highly respected by IPA staff and honored with a dedicated colloquium on the occasion of his 70th birthday in November 1975. He kept contact with his former colleagues until he died in December 1978.

51.4.2 The Second Decade: Experiment and Simulation Guided by Two Directors

During the social-liberal coalition of the federal government (1969–1982), a number of novel trends gained widespread acceptance in West Germany. Co-management and participation (*Mitbestimmung*), improved *environmental protection* (*Umwelt-schutz*), and technological impact assessment (*Technikfolgenabschätzung*) were among the keywords. DFVLR advocated the co-management of research institutes by two directors. For IPA they were to establish a closer link between traditional

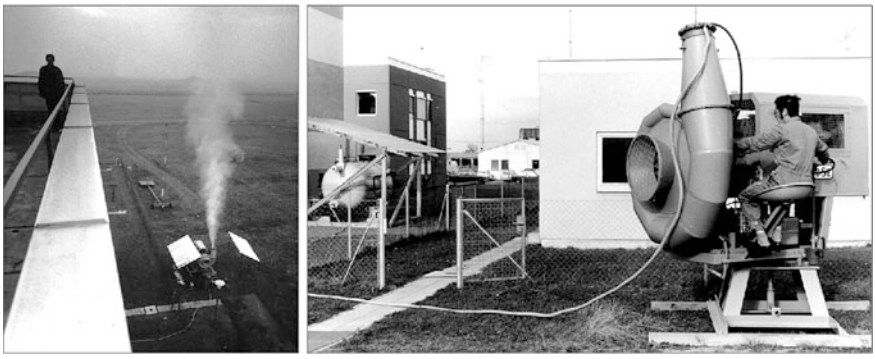

Fig. 51.6 Application test of condensation nuclei generator viewed from the 5th floor roof terrace (*left*) and detailed view of the equipment at the 'measurement field' behind the IPA building (1971; *right*)

experimental tasks and theoretically based simulations aiming at possible scenarios or even concrete forecasts.

In December 1973, Heinz Fortak (Fig. 51.2) began as director at IPA while maintaining his position as professor of 'theoretical meteorology' at the Free University in Berlin, which he had held since 1961. Like Müller, he was also a pilot and highly interested in measurements as well as simulations, e.g., of plume dispersion. The exhausts of the large power station in Neurath were systematically probed with airborne and ground-based equipment in spring and autumn of 1974. On a global scale, the development of a low-resolution spectral model was started. Although Fortak resigned his position in 1976, the concept of a co-directorship, which closely links experiment with simulation, prevailed.

From January 1974 onwards, Manfred E. Reinhardt (Fig. 51.2) became the director at IPA responsible for the experimental tasks. He was a trained pilot, had joined the IPA forerunner in 1957 and had acted as deputy-director under Müller since 1962. Starting in 1973, Reinhardt, Fortak, and many meteorological colleagues in Germany advocated and organized the acquisition of a twin-engine business-jet solely dedicated to atmospheric research. Three years later, the Dassault *Falcon* 20E aircraft (Fig. 51.7) was delivered to the DFVLR flight facility. After numerous adjustments (e.g., nose-boom for turbulence probe, air sampling inlets, optical windows at the bottom and top of the fuselage, wing stations for outside instrumentation) it gradually became an important flagship for airborne research in Germany and beyond. In parallel, the Queenair instrumentation was upgraded to an environmental monitoring platform to detect trace gas concentrations. Moreover, three motorized gliders were equipped with miniaturized equipment for cloud-free boundary layer studies.

More and more, the Falcon became an integrative platform carrying variable sets of instruments during dedicated missions. The atmospheric field experiments JASIN (1978 around Scotland) and ALPEX (1982 over the Alps and the

Fig. 51.7 Main research aircraft of DVL/DFVLR/DLR used for numerous IPA missions (from *left*): Beechcraft 65 Queenair (1962–1991) above power-station near Ingolstadt; Dassault Falcon 20E (since 1976) in home hangar; Gulfstream G550 HALO (since 2010) landing near IPA

neighboring Pyrenees) were the first examples in a long series of mostly international cooperation, which is ongoing. Also of technological importance was the production of the first airborne data curtains from collected backscattered laser pulses emitted by the ALEX-system on the Falcon. By the end of the 1970s, atmospheric systems of mesoscale extent were declared to be the combined focus of measurements and simulations (with modeling codes imported from U.S. colleagues).

51.4.3 The Third Decade: Expertise Expanding to Microscale and Global Extent

On 1 October 1982, Ulrich Schumann (Fig. 51.2) became a director at IPA, for one decade in co-operation with Reinhardt and from 1992 with sole responsibility. He was trained as a mechanical engineer in Berlin and gained broad simulation experience with turbulence codes applied for the safety testing of power station cooling systems and for atmospheric boundary layer investigations. Originally affiliated with Karlsruhe University, Schumann qualified for a senior position in meteorology at Munich University on the basis of a post-doctoral thesis (habilitation) and joined the physics department as a *professor extraordinarius* in 1987. At IPA, he additionally led for 10 years the meteorological model department, later called the atmospheric dynamics department.

In accordance with a general trend in academic research, mesoscale modeling was pursued along several tracks: (i) hydrostatic airflow across mountain ranges in two dimensions inspired by airborne measurements (e.g., during ALPEX), (ii) regional climate as well as low-level wind simulations contracted by state agencies, (iii) combined thermal and dynamical impact of mountain blocks like the Alps on atmospheric flows, and (iv) the development of the nonhydrostatic Mesoscale Flow and Cloud Model Oberpfaffenhofen (MESOSCOP). A technical prerequisite was the installation of a Cray-1 as the first high-performance computer at the neighboring computer center in 1982 and regular upgrades during the following years. A number of PhD students profited from Schumann's expertise in fundamental modeling studies of turbulent flows, which were also applied to explain the statistics of vertical velocities as observed by glider pilots.

In 1985, the installation of an innovative *polarimetric Doppler radar* (POLDIRAD) changed the silhouette of the IPA building (Fig. 51.5). It was a joint acquisition with the neighboring DLR Microwaves and Radar Institute and served as the key device for systematic observation of hail-producing thunderstorms as well as the progression of cold fronts over the Alpine foreland. The development of airborne lidar techniques was augmented by the DIAL method to obtain absolute values for water vapour concentrations. Satellite data started to be archived on a routine basis and a software packet (APOLLO) was developed to differentiate cloud types and to infer physical quantities such as vertically integrated cloud water content (liquid water path).

Prior and after the unification of western and eastern Germany in 1990, mid-tropospheric transport of pollutants (SO_2 and NO_x), was systematically monitored by Queenair flights along the Bavarian borders and over the 'new' federal states in the east. From 1990 onwards a number of east-German post-doctoral scientists and talented PhD students contributed to IPA's expertise. In anticipation of Reinhardt's retirement in January 1992, the meteorological models department became augmented by a global modeling group with the intention to study the impacts of aircraft and other anthropogenic emissions on global chemistry and climate. At the turn to IPA's fourth decade, this expansion contributed to a gradual generation-change in the science staff.

51.4.4 The Fourth Decade: Fundamental Research Closely Tied to Practical Applications

The pollution transport studies used to be restricted to sources on Earth's surface, e.g., industry and automobiles. From 1988 onwards IPA participated in proposing a number of national and European research programs to investigate analogous effects caused by aviation. This included extended series of missions during which the Falcon closely followed dedicated airliners or crossed highly frequented air corridors at cruising altitudes (e.g., over the North Atlantic), and climate scenario simulations based on average emission inventories. As a technological prerequisite, high precision and fast reacting chemical analysis equipment had to be adapted and certified for airborne missions at altitudes around the tropopause. The annual mean emissions of the trace gas components NO_x were estimated for the global aviation system and set in relation to general features of global atmospheric circulation in the troposphere and the stratosphere (Fig. 51.8). These efforts contributed to intergovernmental assessment reports on the impact of aircraft emissions on atmospheric ozone, contrails and climate as well as on the generation and decay of the Antarctic 'ozone hole.'

For mesoscale studies, modeling systems were implemented which allowed episode-type simulation driven by three-dimensional large-scale data sets obtained from weather forecasting centers (such as DWD or ECMWF). A European aircraft campaign based in Kiruna (Sweden) could therefore be guided by tailored

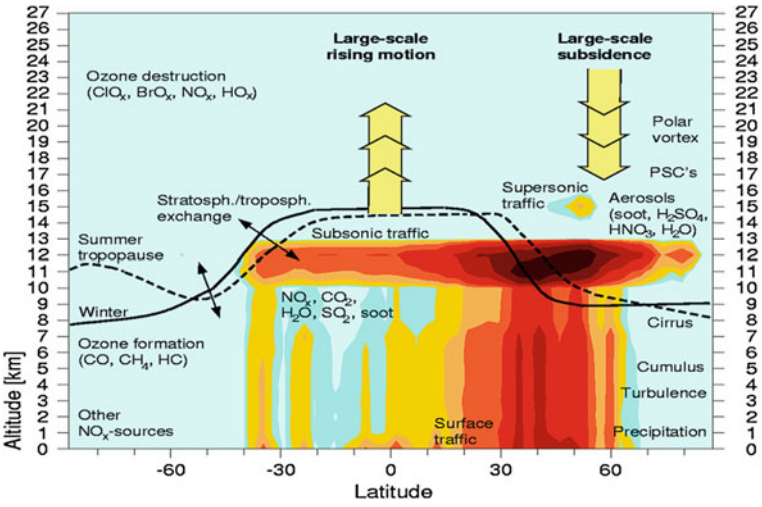

Fig. 51.8 Pole-to-pole section through troposphere and stratosphere with seasonal-mean tropopause altitude. Processes relevant for the impact of aviation emissions on photochemistry and climate are indicated, with the mean zonal NO_x emissions of global aviation as background (Schumann et al. 1997)

stratospheric forecasts of 'cold spots' above 20 km altitude where polar stratospheric ice clouds were suspected to form at temperatures below −85 °C. An upward looking lidar system shuttled with and against the ambient wind direction on board the Falcon and indeed detected ice clouds at three foreseen locations, whereas the first lidar observations some years earlier had been made by chance. This case is regarded as a fruitful outcome of the close link between simulation and observation for both mission planning and subsequent data analysis. For instance, the generalization was recognized that mountain-generated atmospheric waves around the Arctic have a previously overlooked potential for stratospheric ozone destruction. The link to archived and forecast fields of forecasting centers like ECMWF started to become common practice, in particular through the definition of 'special projects.'

In June 1999, the staff moved back to their redesigned 'home' after a thorough renovation of the institute's over 30-year-old building had been completed (Fig. 51.5). Contributions to international assessments about past and future states of the stratospheric ozone layer, based on coupled climate-chemistry simulations, brought numerous contacts with the international research community and enhanced visibility. Successful cross-validation studies between year-long satellite observations of trace gas concentrations and simulation spanning decades raised confidence in both independent tools. During the INCA campaign flights around 55 °S (Chile) and 55 °N (Scotland), the Falcon collected calibration data sets for cirrus studies both in the vicinity of and remote from regular human emissions.

The expertise in idealized turbulence simulations was linked with mesoscale forecast models and applied to the construction of a model chain for weather and

wake vortex warnings at airports. Together with new aircraft icing warning algorithms the model chain was then integrated into the operational procedures of DWD. The wide arc from basic research via applied studies to tailored procedures for routine operations was noted in the votes concluding external reviews of IPA work (in 1997 and 2003).

51.4.5 The Fifth Decade: Widespread Recognition as Attractive Research Partner and Employer

The past decade saw a further increase of simulation activities spanning the full spatial spectrum from, for instance, photons irregularly scattered in a cloudy atmosphere via ice-cloud parameterizations for the operational forecasting model of ECMWF to global assessments for the World Meteorological Organization (WMO). A number of projects were stimulated by research on aviation-produced condensation trails as an important prototype of ice-clouds: large eddy simulation experiments sought to combine turbulent flow configurations with appropriate cloud microphysics, while tailored contrail modules within a global climate model helped to estimate the full climate effect of contrail-cirrus. Global modeling covered numerous components of aircraft climate impact, including aerosol and contrail cirrus effects. Aviation impacts were put into a broader context by extending the simulations to respective impacts caused by road traffic and shipping. Future projections were simulated for the development of the ozone layer throughout the twenty-first century.

During the entire decade, the Falcon aircraft with associated IPA staff and instruments participated in international campaigns on all continents except Antarctica; for instance, in North America (USA, IHOP 2002), over Iceland and Greenland (TReC 2003), in Australia (SCOUT-O_3 2003), in South America (Brazil, TROCCINOX 2005), in Africa (Morocco, SAMUM-1 2006; Burkina Faso, AMMA 2006) over Europe (CONCERT 2008/2011, volcanic ash flights 2010), and in Asia (Japan, T-PaRC 2008; Malaysia, SHIVA 2011). During a number of these campaigns the Falcon served as pathfinder aircraft for the high-altitude *Geophysica* aircraft, the core platform for a European multiagency consortium. In 2001 concrete plans were made to acquire a larger aircraft customized for atmospheric research with a higher ceiling and greater endurance. In 2010, first technology missions with the multiwavelength laser system *WALES* for humidity measurements and an ion trap mass spectrometer for trace gas concentrations were carried out with the new High Altitude and Long-range (*HALO*) *research aircraft* (Fig. 51.7).

As research missions attained worldwide breadth, the IPA staff also became more international. From 2003 to 2009 an IPA department was headed by a Canadian colleague. PhD students and post-doctoral junior scientists came for extended periods of work from Austria, China, France, Italy, Japan, Spain and Switzerland. Three applications at national programmes for scientific excellence

led to young investigator groups (SeaKLIM, AEROTROP, AerCARE). In addition to the engagement of the IPA director at LMU, about ten IPA colleagues regularly held classes at universities (e.g., LMU and TU Munich; Berlin; Bremen; Innsbruck; Vienna). A number of international conferences and workshops were organized by IPA personnel, e.g., a series on Transport, Atmosphere and Climate (TAC; 2006, 2009, 2012), on Chemistry-Climate Model Validation (CCMVal; 2003, 2005, 2007, 2009, 2012), on High Impact Weather (STISS for WMO 2006) and on Severe Storms (ESSC 2009). Since 2007, the general secretariat of the International Association of Meteorology and Atmospheric Sciences (IAMAS) has been hosted by IPA. The number of female scientists grew from one in 1982 to 23 in 2012, 25 % of the science staff. These indicators show the attractiveness of IPA as employer and research partner, also at the international level.

On Wednesday 14 April 2010, the Icelandic *volcano Eyjafjallajökull* began a long period of eruptions which had an unprecedented impact on air-traffic in central Europe. Within a period of only 4 months numerous colleagues within IPA and many research partners demonstrated under the leadership of IPA director Ulrich Schumann what can be achieved by flexible teams in an exceptional situation. Established knowledge about many physical processes was successfully pieced together (cf. Chap. 38), and airborne missions contributed to the lifting of air-space closures and, eventually, to a revision of flight regulations related to volcanic eruptions for international air-traffic.

****

The progress in experimental and modeling studies undertaken at IPA over 50 years and their gradual integration towards a more consistent knowledge of atmospheric processes certainly benefited tremendously from the general progress in measuring techniques and computational power since 1960. This brief description of five decades of achievements also demonstrates how Müller's vision of 1962 identified research areas of lasting relevance. One may ponder whether his prescriptions were sufficiently general to be automatically fulfilled after enough time had elapsed, or whether he inscribed an enduring spirit into "his" IPA, which was then tacitly relayed by the three following directors and through at least three generations of researchers, even though they hardly had explicit knowledge of his original text.

51.5 Trends and Outlook Towards 2022

Many of the previous chapters of this book contain perspectives for future developments. At this point, some general trends are substantiated together with an important caveat.

Satellite missions using active remote sensing with lidar systems are technologically most ambitious. The Atmospheric Dynamics Mission ADM-Aeolus of the European Space Agency receives scientific advice from IPA staff and has a high potential for improved global weather forecasts from consistent wind observations. Similarly the WALES system for the precise height allocations of atmospheric humidity structures awaits the step from aircraft to a satellite application after a convincing technology demonstration. Numerical weather prediction models can be used in research mode to quantify the impact of the additional data. In the context of climate monitoring the French-German MERLIN mission aims at determining methane column data from an orbit elevation of 600 km by using IPA expertise. All these data acquisition projects have to be accompanied by tailored numerical experiments.

Airborne campaigns with the new flagship HALO, a platform for the German research community beyond DLR and an internationally attractive asset, will see a new kind of synergy between instrumentation once the backlog of initial certification procedures have been worked off. Mass spectrometry for low trace gas concentrations and better resolved particle spectra will provide additional insights from in-situ measurements, while the synchronous operation of several laser systems offers a new dimension for the 'data curtains' which were pioneered at IPA. Appropriate instrumentation of high altitude research rockets operated by DLR for international clients has the potential to explore mesoscale details at mesospheric heights, which would resume a research avenue already tried in the 1960s.

Along with the anticipated further growth in computational power, long awaited fundamental studies will become feasible for three-dimensional turbulent flows, including applications for wake-vortex behavior in actual meteorological conditions at airports. Probabilistic aspects of weather prediction can be linked to climate scenarios in a seamless fashion. Joining results from three-dimension radiative transfer measurements with multichannel satellite retrievals from geostationary orbit will enhance the detailed physical interpretation of the spatio-temporal evolution of cloud structures. Efforts in global chemistry-climate model development and evaluation will include more processes and feedbacks within an Earth system model. The link between modeling efforts and data from airborne and spaceborne instruments will be strengthened and synergisms explored, e.g., concerning cirrus cloud development and aerosol-cloud interaction.

A constant challenge is to retain a sufficient influx of talented scientists for projects of longer duration. Well qualified colleagues are ready, but enough long-term positions for them are yet to be made available. By deepening and broadening the link forged between IPA and the meteorological institute (MIM) at Ludwig-Maximilians-Universität in Munich (LMU) the trinity "education–research–application" holds promise for achieving lasting stability for the atmospheric sciences in Oberpfaffenhofen and its environs.

51.6 Concluding Remarks

This book contains a comprehensive collection of highlights about technological developments and physical techniques which are currently available for research at IPA in the continuous endeavor to better understand and more reliably forecast the multitude of physical processes within the vast and shallow atmosphere around the Earth. This chapter added some sidelights on how the current status evolved during the past five decades under the leadership of four directors and their experimental and modelling teams.

If IPA's overall development is regarded as a case study, in analogy to what is observed during an atmospheric *field experiment*, some general characteristics become apparent. The beginning of the 1960s was a fertile period for atmospheric research institutes. In Germany, for instance, the Meteorological Institute at Bonn University was established in 1961 (founded by the former DWD scientist Hermann Flohn) as was the Atmospheric Physics Laboratory at ETH in Zurich, Switzerland. In Britain, Reginald Sutcliffe (former director of research at the UK Meteorological Office) founded the Meteorology Department at Reading University in 1965. In the USA, the formation of the National Center for Atmospheric Research (NCAR) took place in 1961. So, four widely known institutions and cooperation partners of IPA are of similar age. Public investments for atmospheric research and training were apparently available at that time and have grown gradually since.

Specific to IPA is its situation within the German Aerospace Center DLR. As noted in the first section, it originated with the natural interest of aviators in the subtleties of the atmosphere and their interplay. The growing use of satellite information as atmospheric data provider and many studies on two-way atmosphere-aviation feedback since the mid-1980s provided tight and stable links between IPA work and two of the core programmatic aims of DLR, which are expressed by the letters 'L' (*Luft-)* for Aero- and 'R' (*Raumfahrt*) for space.

The numerous subtleties contained in the atmosphere around us certainly contain a multitude of aspects to be discovered and then properly linked to each other, enough to keep the IPA staff busy during the coming decade and beyond. In a wider sense, the past and continuing projects undertaken at IPA also provide rich material for questions like, "How is (atmospheric) knowledge generated?" and "Are experiment and theory the only pillars for knowledge generation, or does simulation provide a third footing?" Historians and philosophers of science have started to tackle such questions with direct reference to the atmospheric sciences (Gramelsberger and Feichter 2011).

References

Georgii, W.: Beitrag zur Geschichte der Deutschen Forschungsanstalt für Segelflug, DVL-Mitteilung 69–04 (DVL-Bericht Nr. 890, aus dem Nachlass des am 24.7.1968 verstorbenen Walter Georgii), 29 pp. Porz-Wahn (1969)

Gramelsberger, G., Feichter, J.: Climate Change and Policy—The Calculability of Climate Change and the Challenge of Uncertainty, 240 pp. Springer, Heidelberg (2011)
Höhler, S.: Ballonfahrer und Meteorologen: Die Frühgeschichte der Luftfahrtforschung. In: Trischler, H., Schrogl, K.-U. (eds.) Ein Jahrhundert im Flug—Luft- und Raumfahrtforschung in Deutschland 1907–2007, pp. 31–50. Campus Verlag, Frankfurt (2007)
Müller, H.G.: Aerologisches Observatorium, internal typescript, 13 pp. (1962)
Müller, H.G.: Das Institut für Physik der Atmosphäre in München, Flughafen Riem. DVL Nachrichten **23**, 43–46 (1964)
Müller, H.G.: Arbeiten zur Erforschung der Atmosphäre. DVL Nachrichten **37**, 374–377 (1968)
Ohring, G., Bojkov, R.D., Bolle, H.-J., Hudson, R.D., Volkert, H.: Radiation and ozone—catalysts for advancing international science programs for over half a century. Bull. Am. Meteorol. Soc. **90**, 1669–1681 (2009). doi:10.1175/2009BAMS2766.1
Reinhardt, M.E., Schumann, U.: Institut für Physik der Atmosphäre, Institutsbericht 1982–1987, 138 pp. DFVLR-IPA, Oberpfaffenhofen (1987)
Reinhardt, M.E.: Tätigkeitsberichte 1954–1974 des DFLVR-Instituts für Physik der Atmosphäre und der Vorgängerinstitute, 339 pp. DLR-IPA, Oberpfaffenhofen (2005)
Reinhardt, M.E.: Ergebnisberichte 1975–1982 des DFLVR-Instituts für Physik der Atmosphäre, 473 pp. DLR-IPA, Oberpfaffenhofen (2010)
Schumann, U.: Institut für Physik der Atmosphäre, Institutsbericht 1987–1992, 130 pp. DLR-IPA, Oberpfaffenhofen (1992)
Schumann, U.: Institut für Physik der Atmosphäre, Statusbericht 1992–1997, 162 pp. DLR-IPA, Oberpfaffenhofen (1997)
Schumann, U., Chlond, A., Ebel, A., Kärcher, B., Pak, H., Schlager, H., Schmitt, A., Wendling, P. (eds.): Pollutants from air traffic—results of atmospheric research 1992–1997, 301 pp. DLR Mitt. 97–04, Köln (1997)
Schumann, U.: Institut für Physik der Atmosphäre, Status Report 1997–2003, 121 pp. DLR-IPA, Oberpfaffenhofen (2003)
Schumann, U.: Institute of Atmospheric Physics, Status Report 2003–2007, 148 pp. DLR-IPA, Oberpfaffenhofen (2007)
Trischler, H.: Luft—und Raumfahrtforschung in Deutschland 1900–1970: Politische Geschichte einer Wissenschaft, 542 pp. Campus Verlag, Frankfurt (1992)

Appendix

Numerical values of physical constants and other data

Symbol	Parameter	Value
c	Speed of light in vacuum	2.99792×10^{8} m s^{-1}
k	Boltzmann constant	1.38×10^{-23} J/K
h	Planck constant	6.626×10^{-34} J s
σ	Stefan-Boltzmann constant	5.670×10^{-8} W m^{-2} K^{-4}
S_0	Average solar constant	1 365 W/m^2
g	Earth gravity acceleration at mean sea level	9.81 m s^{-2}
c_p	Specific heat capacity of dry air at standard conditions	1 004 J kg^{-1} K^{-1}
R	Dry air gas constant	287.04 J kg^{-1} K^{-1}
κ	Ratio of specific heats at constant pressure and constant volume for an ideal gas	1.4
R_v	Water vapor gas constant	461.91 J kg^{-1} K^{-1}
PVU	Potential vorticity unit	10^{-6} m^2 s^{-1} K kg^{-1}
Q	Combustion heat of kerosene	43.2 MJ kg^{-1}
EI_{CO2}	CO_2 emission index for kerosene	3.16
EI_{H2O}	H_2O emission index for kerosene	1.23

Standard units—SI (Système Internationale) units

Quantity	Time by time	Symbol
Length	Meter	m
Time	Second	s
Mass	Kilogram	kg
Temperature	Kelvin	K
Molar amount of a substance	Mole	mol

Further time units: s, min, d, a—for second, minute, day, year (anno)

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4, © Springer-Verlag Berlin Heidelberg 2012

UTC: Coordinated Universal Time (short: universal time) is the primary time standard by which the world regulates clocks and time. It is closely related to the previously used Greenwich Mean Time. Dates are identified by hour:minute UTC day month year, e.g.: 12:45 UTC 2 April 2012.

Julian day is used in the Julian date (JD) system of time measurement for scientific use by the astronomy community, presenting the interval of time in days and fractions of a day since 1 January 4713 BC Greenwich noon.

Derived units

Quantity	Name of unit	Symbol	Definition
Force	Newton	N	kg m s^{-2}
Pressure	Pascal	Pa	N m^{-2}
Energy	Joule	J	N m
Power	Watt	W	J s^{-1}
Frequency	Hertz	Hz	s^{-1}

Mass mixing ratio: kg/kg
Molar mixing ratio: mol/mol

Prefix: Metric prefixes act as a multiplier to the base unit that follows. See below for the values of the prefix (10**X = 10^X)

Factor	In digits	Prefix	Symbol
10** + 18	1 000 000 000 000 000 000	exa	E
10** + 15	1 000 000 000 000 000	peta	P
10** + 12	1 000 000 000 000	tera	T
10** + 9	1 000 000 000	giga	G
10** + 6	1 000 000	mega	M
10** + 3	1 000	kilo	k
10** + 2	100	hecto	h
10** + 1	10	deca	da
10**-1	0.1	deci	d
10**-2	0.01	centi	c
10**-3	0.001	milli	m
10**-6	0.000 001	micro	μ
10**-9	0.000 000 001	nano	n
10**-12	0.000 000 000 001	pico	p
10**-15	0.000 000 000 000 001	femto	f
10**-18	0.000 000 000 000 000 001	atto	a

Further factors: yotta (Y): 10^{24} , zetta (Z): 10^{21} , zepto (z): 10^{-21} , yocto (y): 10^{-24}

Non-SI units often used in international aeronautics and elsewhere
1 foot = 1 ft = 12 inches = 30.48 cm
1 hectofoot = 100 feet = 1 hft = 30.48 m
1 nautical mile = 6 076.12 feet = 1 852 m
1 knot = 1 nautical mile per hour = 1.852 km h^{-1} or 0.5144 m s^{-1}
1 ton = 1 000 kg = 1 Mg; 1 Megaton = 1 Mt = 1 Tg.

Index

U. Schumann (ed.), *Atmospheric Physics*, Research Topics in Aerospace,
DOI: 10.1007/978-3-642-30183-4, © Springer-Verlag Berlin Heidelberg 2012

Q

R

T