Institute for Geophysics, Astrophysics, and Meteorology
University of Graz

Atmospheric Remote Sensing and Climate System Research Group
ARSCliSys – on the art of understanding the climate system

IGAM/UniGraz Technical Report for ASA No. 3/2004

ASA study:

CHAMPCLIM – Radio Occultation Data Analysis Advancement and Climate Change Monitoring based on the CHAMP/GPS Experiment

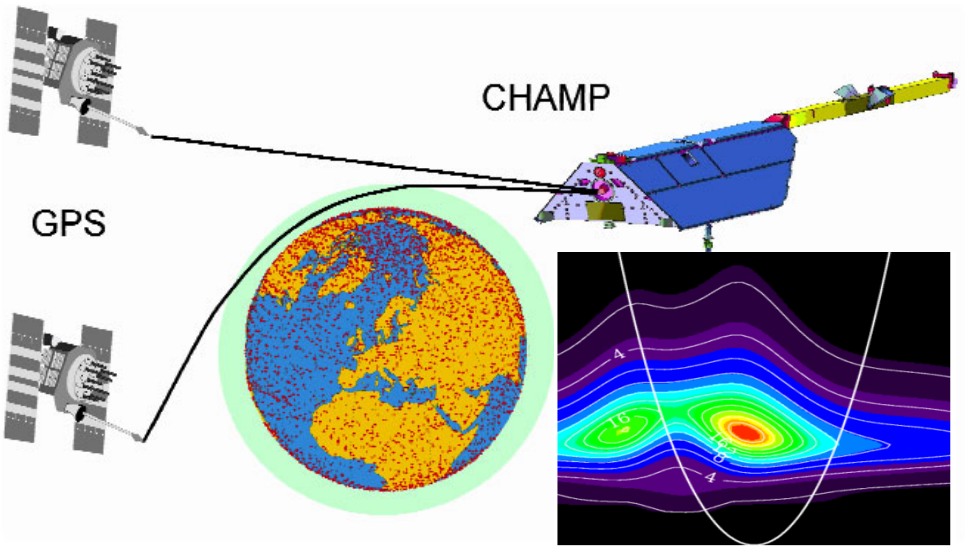
[Contract No: ASAP-CO-004/03 – July 2003]

ADVANCEMENT REPORT

[WP3: ADVANCE]

Radio Occultation Data Processing Advancements for Optimizing Climate Utility

A.K. Steiner, A. Gobiet, U. Foelsche, and G. Kirchengast
Institute for Geophysics, Astrophysics, and Meteorology (IGAM),
University of Graz, Graz, Austria



Final Version — Oktober 2004

(intentionally left blank, back page if double-sided print)

Table of Contents

List of Acronyms

1D-Var	One-Dimensional Variational data assimilation
AR	Advanced receiver
ASA	Austrian Space Agency
CCD	Charge-coupled device
CHAMP	CHAllenging Minisatellite Payload (German research satellite)
CHAMPCLIM	Radio Occultation Data Analysis Advancement and Climate Change Monitoring based on the CHAMP/GPS Experiment – research project in the ASAP program managed by IGAM
CIRA-86	Cospar International Reference Atmosphere 1986 [Fleming, 1988]
CNRS	Centre National de la Recherche Scientifique (France)
COSMIC	Constellation Observing System for Meteorology, Ionosphere, and Climate
DMI	Danish Meteorological Institute (Copenhagen)
ECHAM4	Fourth-generation atmospheric general circulation model developed from the ECMWF atmospheric model (EC) at the Max Planck Institute for Meteorology, Hamburg (HAM)
ECMWF	European Centre for Medium-Range Weather Forecasts
EGOPS	End-to-end GNSS Occultation Performance Simulator
EGOPS/CCR	CHAMPCLIM Retrieval – software tool for the retrieval of CHAMP RO data based on EGOPS
ENVISAT	ENVIronment SATellite (ESA)
EQUARS	Equatorial Atmosphere Research Satellite
ESA	European Space Agency
EUMETSAT	EUropean organisation for the exploitation of METeorological SATellites
GALILEO	European future global navigation satellite system
GCM	General Circulation Model
GCOS	Global Climate Observing System
GFZ	GeoForschungsZentrum Potsdam
GLONASS	Russian global navigation satellite system
GNSS	Global Navigation Satellite System (generic term for GPS, GALILEO, GLONASS)
GPS	Global Positioning System (USA)
GPS/MET	GPS METeorology – "proof-of-concept" experiment for the GPS RO technique on MicroLab 1 in 1995
GRACE	Gravity Recovery and Climate Experiment
GRAS	Global navigation satellite system Receiver for Atmospheric Sounding (on MetOp)
IGAM	Institute for Geophysics, Astrophysics, and Meteorology, University of Graz, Austria
IGAM/ECMWF	RO retrieval scheme using operational ECMWF analyses as background information
IGAM/MSIS	RO retrieval scheme using the MSISE-90 climatology as background information
IMK	Institut für Meteorologie und Klimaforschung, University Karlsruhe / Forschungszentrum Karlsruhe, Germany

IPCC	Intergovernmental Panel on Climate Change
JPL	Jet Propulsion Laboratory
LEO	Low Earth Orbit
LIDAR	Light Detecting and Ranging
MetOp	Meteorological Operational satellite (of future EUMETSAT polar system)
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding (on ENVISAT)
MSISE-90	Mass-Spectrometer-Incoherent-Scatter model, extended version 1990 [*Hedin*, 1991]
NASA	National Aeronautics and Space Administration
NeUoG	The "Electron Density, University of Graz" empirical 3D ionosphere model
NWP	Numerical Weather Prediction
OPAC	Occultations for Probing Atmosphere and Climate
OSSE	Observing System Simulation Experiment
QF	Quality Flag
RO	Radio Occultation
SAC-C	Satelite de Aplicaciones Cientificas-C (Argentinia)
SNR	Signal to Noise Ratio
SR	Standard receiver
TEC	Total Electron Content
TerraSAR-X	Im Bau befindlicher deutscher Satellit mit X-Band Synthetic Aperture Radar (SAR) für die Erfassung der Oberflächenbedeckung
UBH	Upper Boundary Height
UCAR	University Corporation for Atmospheric Research (Boulder, Colo., U.S.)
UniGraz	University of Graz (Austria)
U.S.	United States of America

List of Symbols

α_b	Background bending angle profile
α_o	Observed bending angle profile
α_{opt}	Best linear unbiased estimator of a bending angle profile
α	Total bending angle of a RO ray
α_{LC}	Ionosphere-corrected bending angle
σ	Standard deviation
a	Impact parameter of a RO ray
B	Background error covariance matrix
b	Bias profile
$\mathbf{b}_{rel}$	Relative bias profile
ΔL	LC phase difference profile
f_1	Carrier frequency of L1 (1575.42 MHz)
f_2	Carrier frequency of L2 (1227.60 MHz)
g	Gravitational acceleration
g_0	Standard acceleration of gravity (9.80665 ms^{-2})
H_{strat}	Scale height of the refractivity error increase in the stratosphere
k_1, k_2	Constants in refractivity formula ($k_1 = 77.60\ K\ hPa^{-1}$, $k_2 = 3.73*10^5\ K^2\ hPa^{-1}$)
L	Vertical error correlation length
L1, L2	GPS L-band signals
LC	Ionosphere corrected phase delay
M_d	Molecular mass of dry air (28.964 kg $kmol^{-1}$)
N	Refractivity ($N = 10^6(n - 1)$)
n	Refractive index
O	Observation error covariance matrix
p	Pressure
p_d	Pressure of dry air
p_w,	Partial pressure of water vapor in the atmosphere
Φ_1, Φ_2	Phase delays of GPS signals L1,L2
q	Specific humidity
R	Error correlation matrix
R	Universal gas constant ($8.3145*10^3\ J\ K^{-1}\ kg^{-1}$)
R_C	Radius of curvature of the earth's ellipsoid at the occultation location
rms	Root mean square profile
S	Error covariance matrix
s	Standard deviation profile
$\mathbf{s}_{rel}$	Relative standard deviation profile
s_{TP}	Relative standard deviation of refractivity at tropopause level
s_0	Relative standard deviation of refractivity at 1 km
σ_{bi}	Standard deviation of background bending angle
σ_{oi}	Standard deviation of observed bending angle
T	Temperature
T_d	Dry temperature
z	Geometric height

z_p	Pressure height
z_{TP}	Tropopause height
Z	Geopotential height

1 Introduction

The provision of accurate, long-term, consistent data to sustain and expand the observational foundation for climate studies is one of the high priority areas for action to improve the ability to detect, attribute and understand climate variability and changes as stated by the Intergovernmental Panel on Climate Change [*IPCC*, 2001].

A promising climate monitoring tool for meeting these requirements is the Global Navigation Satellite System (GNSS) Radio Occultation (RO) technique. It is based on a satellite-to-satellite limb sounding concept using GNSS microwave signals to probe the Earth's atmosphere. The propagation of the signals is influenced by the atmospheric refractivity field resulting in deceleration and bending of the signal (Figure 2.1). Measured atmospheric phase delays are the basis for high-quality retrievals of atmospheric key variables, such as pressure, geopotential height, temperature, and humidity profiles. Highest accuracy (e.g., temperature < 1 K) and high vertical resolution (0.5 to 1.5 km) are obtained in the upper troposphere and lower stratosphere, this domain being a sensitive indicator of climate change [*IPCC*, 2001]. The GNSS RO technique provides essential properties for climate studies comprising global coverage, all-weather capability, and long-term stability (< 0.1 K per decade is expected). The latter derives from the self-calibrating nature of the RO technique and is of particular importance for long-term monitoring of the Earth's climate. [*Kirchengast et al*., 2000; *Anthes et al*., 2000; *Steiner et al*., 2001]

After successfully demonstrating the proof of concept with the GPS/MET mission [e.g., *Ware et al.*, 1996; *Kursinski et a*l., 1996; *Rocken et al*., 1997; *Steiner et al*., 1999], the German/US research satellite CHAMP (CHAllenging Minisatellite Payload) now continuously records RO observation and thus provides the first opportunity to create RO based climatologies on a longer term. CHAMP was launched in July 2000 receiving radio signals from Global Positioning System (GPS) satellites. Since 2001 it globally records about 280 RO events per day yielding about 170 atmospheric profiles per day [*Wickert et al*., 2001; 2004].

The overall aim of the CHAMPCLIM project is to contribute that the CHAMP RO data are exploited in the best possible manner for climate monitoring by addressing three areas: 1) CHAMP RO data and algorithms validation, 2) CHAMP/GPS-based RO data processing advancements in order to optimize the climate utility of the data, and 3) CHAMP/GPS-based monitoring of climate variability and change, respectively.

Founded on the investigations and results of work package 2 regarding the validation of CHAMP/GPS RO retrievals [*Gobiet et al.*, 2004a] this report is a documentation of the work carried out during the second year of the project (work package 3 of the CHAMPCLIM Part I project plan) addressing the advancement of the RO data retrieval. Our primary aim was to advance various critical performance issues of retrieval processing in order to improve the maturity and utility of the data products especially for climatological purposes. In particular, three main topics are addressed including retrieval advancement at high altitudes, lower-to mid tropospheric processing and error assessment, and empirical error analysis of the retrieval products. Each of these topics has dedicated one main section of this report and parts of it have already been or will be subject to publication in scientific journals and/or books.

Section 2 gives a brief overview on the state-of-the-art RO retrieval technique and describes advancements achieved with focus on the retrieval in high altitudes. In this respect, section 2.1 presents results of a study using simulated RO data focusing on the ionospheric correction and statistical optimization part in the processing chain. This section is based on a manuscript by A. Gobiet and G. Kirchengast, which in similar form was submitted to the *Journal of Geophysical Research* [*Gobiet and Kirchengast*, 2004b] and is now at the stage of review before publication. Based on these results, the creation of a robust retrieval scheme for CHAMP RO data (IGAM/CCR scheme) is described in section 2.2 including the topics quality control, specific CHAMP retrieval advancements and retrieval performance.

Section 3 addresses the advancement in RO data processing and the assessment of RO retrieval errors in the troposphere region. Results of a simulation study are presented on testing the retrieval sensitivity with respect to atmospheric horizontal variability errors and the angle of incidence as well as the geometry of reference profiles. This section is based on a paper very recently published in *Occulations for Probing Atmosphere and Climate* [*Foelsche and Kirchengast*, 2004b], a book following the 1st International Workshop on Occultations for Probing Atmosphere and Climate (OPAC-1) held late 2002 in Graz.

Section 4 is dedicated to the advancement in error characterization of atmospheric profiles retrieved from RO observations. In section 4.1 results of an empirical error analysis for GNSS RO data based on end-to-end simulations are presented. Complete error information on bending angle, refractivity, pressure, geopotential height, temperature, and specific humidity retrievals is provided in form of bias profiles and error covariance matrices. This section is based on a manuscript, which in similar form war recently submitted by A.K. Steiner and G. Kirchengast to the *Journal of Geophysical Research* [*Steiner and Kirchengast*, 2004b] and is currently in the review process. The statistical methods used in the simulation study were then applied to real data. In section 4.2 results of this error analysis of CHAMP RO refractivity retrievals processed with the IGAM/CCR scheme are presented.

Finally, in section 5, a short summary is given, conclusions are drawn and a brief outlook regarding future work is presented.

2 Advancement of High Altitude Radio Occultation Retrieval

2.1 High Altitude RO Retrieval: State-of-the-Art and Advancement

Executive Summary. Global Navigation Satellite System (GNSS) based radio occultation (RO) measurements promise to become a valuable observational basis for climate research, numerical weather prediction, and atmospheric process studies, thanks to their unique combination of global coverage, high accuracy and vertical resolution, and long-term stability. Concerning the quality of the RO-derived atmospheric profiles, performance simulation studies and first data analysis results from the current CHAMP satellite mission are encouraging but also reveal weaknesses of present RO retrieval chains at high altitudes (above 30 km). This study aims at providing, first, understanding of these weaknesses and, second, mitigation by an advanced retrieval scheme. We first evaluated present state-of-the-art high altitude RO retrieval algorithms. We found clear superiority for using statistical optimization involving background information over using exponential extrapolation and a sensitivity of RO retrieval products to biases in background information calling for sensible bias mitigation. Exploiting these findings, an advanced high altitude retrieval scheme is presented, which focuses on minimizing residual retrieval biases in the upper stratosphere and thereby on optimizing the climate monitoring utility of retrieved profiles. Applied to a large ensemble of simulated occultation events, the advanced scheme proved to be effective. The scheme is currently under evaluation with real data from the CHAMP mission. It will then serve as part of a processing chain generating RO-based global climatologies of refractivity, geopotential height, temperature, and humidity based on RO data from CHAMP, SAC-C, GRACE and other future satellites carrying RO instruments.

2.1.1 Introduction

Detecting and observing changes and variability in the global climate is one of the most important challenges in atmospheric sciences over the coming decades since there exist evidence and concern that the Earth's climate is increasingly influenced by human activities [e.g., *IPCC*, 2001]. Global warming and climate change in general is not only a scientific but also a societal-political topic, which makes it especially important to establish reliable and stable long-term records of key climate variables such as atmospheric temperature. Since long-term in-situ measurements face, especially on global scale, many difficulties like calibration, temporal and spatial coverage, and high costs, satellite borne remote sensing techniques are desirable. However, many satellite derived data records are degraded by problems like instrument and orbit changes, calibration problems, instrument drifts, and lacking vertical resolution [*Anthes et al*., 2000]. The Global Navigation Satellite System (GNSS) radio occultation (RO) technique solves many of these problems. It was originally developed and successfully applied to study planetary atmospheres [e.g. *Fjeldbo et al*., 1971]. The application to Earth's atmosphere became possible with the advent of the U.S. Global Positioning System (GPS) in the early 1980's [*Yunck et al*., 2000]. Together with the Russian

GLONASS system and the emerging European GALILEO system there is now a multitude of transmitter platforms available to be used for sounding Earth's atmosphere with high temporal and spatial resolution. Figure 2.1 depicts schematically the GNSS RO geometry on Earth. The first suggestions to apply the RO technique to the Earth's atmosphere go back to the 1960's [*Fishbach*, 1965; *Lusignan et al*., 1969], but the potential of the method was not clearly recognized until the late 1980's [*Gurvich and Krasil'nikova*, 1987; *Yunck et al*., 1988]. The successful "proof-of-concept" experiment GPS/MET (1995–1997) demonstrated the unique properties and performance of RO measurements applied to the Earth's atmosphere [*Ware et al*., 1996; *Kursinski et al*., 1996; *Rocken et al*., 1997; *Steiner et al*., 1999]: high vertical resolution (0.5–1.5 km), high accuracy (< 1 K in the upper troposphere and lower stratosphere), all-weather capability, and the global coverage. This confirmed that RO data can play a significant role in atmospheric process studies and numerical weather prediction (NWP) in the near future. Additionally, and potentially most valuable, the expected long-term stability of the method (< 0.1 K drift per decade) suggests it as an ideal tool for global climate monitoring and the improvement of climate models [e.g., *Anthes et al*. 2000; *Kirchengast et al*., 2000; *Leroy and North*, 2000; *Steiner et al*., 2001].

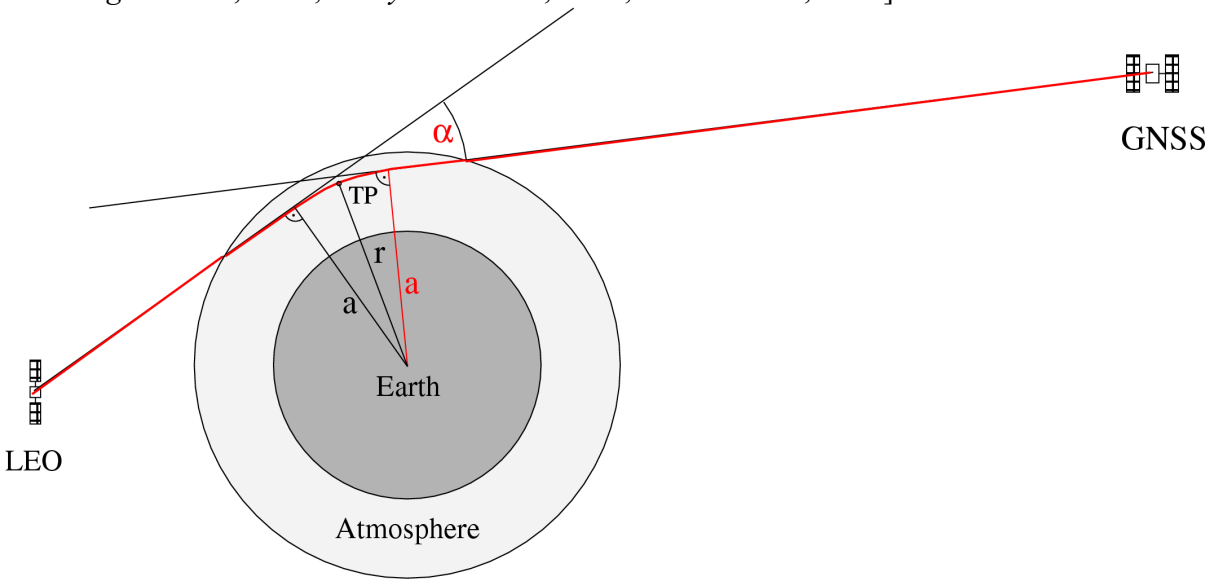


Figure 2.1. Radio occultation geometry (after *Foelsche* [1999]).

The first opportunity for realizing multi-year global RO based climatologies is provided by the German/U.S. CHAMP research satellite, which was launched on July 15, 2000, into low earth orbit (LEO) and almost continuously provides ~250 globally distributed occultation events per day from which ~170 quality approved atmospheric profiles can be derived since late 2001 [*Wickert et al.*, 2001; *Wickert et al.*, 2004]. Sampling error studies by *Foelsche et al*. [2003] showed that even with one single GPS RO receiver such as on CHAMP, accurate global season-to-season temperature climatologies resolving large horizontal scales >1000 km can be achieved. Higher temporal and spatial resolution can be reached by adding RO data from the Argentine/U.S. SAC-C satellite, providing data since July 2001, and the U.S./German twin-satellites GRACE, which are expected to provide RO data from mid 2004. Further future perspectives still within this decade include the MetOp weather satellite series, which will carry the GRAS receiver (1st launch scheduled 2005; *Edwards and Pawlak* [2000]; *Loiselet et al.* [2000]; *Silvestrin et al*. [2000]), the German TerraSAR-X satellite (launch scheduled 2005), the U.S./Taiwan COSMIC mission, consisting of six micro-satellites

carrying GPS RO receivers (launch scheduled 2005; *Rocken et al*. [2000]), and the Brazilian EQUARS satellite (launch scheduled 2006).

We aim at using the complete CHAMP and SAC-C RO data flow (and GRACE and later sensor data, as soon as available) for global climate monitoring via refractivity, geopotential height, temperature, and humidity fields. The respective RO data processing system for climate monitoring is currently under development at IGAM.

First retrieval results from the CHAMP mission [*Wickert et al*., 2001] show a good agreement with co-located ECMWF analyses in the lower stratosphere (temperature bias < 0.5 K between 12–20 km). However, above 20 km the errors increase significantly. *Wickert et al.* [2001] state that for applications such as climate monitoring the retrieval scheme has to be improved, and the occurrence of biases at stratospheric heights has to be investigated. In the most recent published version (v04) of the operational CHAMP retrieval, the middle/upper stratosphere temperature differences compared to ECMWF analysis and radiosonde data were further reduced by introducing ECMWF data as background information into the retrieval [*Wickert et al.*, 2004].

First results from a currently ongoing RO climate monitoring system simulation project designed to test the climate change detection capability of GNSS occultation sensors (GNSS-CLIMATCH; *Kirchengast et al.* [2000]; *Steiner et al*. [2001]; *Foelsche et al*. [2003]), indicate the potential but also the weaknesses of state-of-the-art RO retrieval algorithms. The seasonal mean dry temperature bias (not including the sampling error) of an assumed RO monitoring system consisting of 6 LEO satellites would be below 0.2 K at most latitudes and altitudes between 8–30 km. Above 30 km the bias increases drastically, and in the high latitude winter region an additional bias in the order of 1 K exists.

Both, the first GNSS-CLIMATCH and CHAMP results show the challenging areas of RO retrieval (see also *Kursinski et al.* [1997]; *Steiner et al.* [2001]; *Rieder and Kirchengast* [2001b]): The upper stratosphere, where the signal-to-noise ratio gradually becomes smaller as height increases towards the stratopause and into the mesosphere and where thus thermal noise, residual ionospheric error, and other small errors become increasingly important, and the lower-to-mid troposphere where moisture and strong horizontal gradients complicate the interpretation of RO data. Currently, extensive research is on-going on the latter topic [e.g., *Gorbunov*, 2002a; *Gorbunov*, 2002b; *Jensen et al.*, 2003; *Ao et al.*, 2003; *Sokolovskiy*, 2003; *Beyerle et al*., 2004], which is outside the scope of this work. This paper focuses on the former topic, the advancement of RO retrieval in the upper stratosphere, which is based on the evaluation of the performance of present state-of-the-art schemes.

In section 2.1.2 the relevant methodology is described detail. The results on the evaluation of present schemes and on the performance of a newly developed advanced scheme are discussed in section 2.1.3. Section 2.1.4, finally, provides a summary, the main conclusions, and an outlook.

2.1.2 Methodology

2.1.2.1 High Altitude RO Retrieval Techniques

a) RO Retrieval Overview. The primary observables of RO measurements are the phase delays of GNSS signals (GPS: Φ_1, f_1 = 1575.42 MHz; Φ_2, f_2 = 1227.60 MHz), i.e., the consequences of deceleration of the electromagnetic wave's phase velocity by the atmosphere. Doppler shifts and subsequently total bending angles α of the rays are deduced from phase

delays. The refractive index n can then be derived via the inverse Abel transform F*jeldbo et al*. [1971]),

$$n(a) = \exp\left[\frac{1}{\pi}\int_a^\infty \frac{\alpha(a')}{\sqrt{a'^2 - a^2}}\,\mathrm{d}a'\right], \qquad (2.1)$$

from which the refractivity as a function of height, $N(z)$, is obtained via $N(a) = 10^6(n(a)–1)$ and $z(a) = a/n(a)–R_C$. R_C is the radius of curvature of the ellipsoidal Earth at the occultation location along the occultation plane (~ Earth's radius), a is the ray's impact parameter (see Figure 2.1). Refractivity is related to total atmospheric pressure p, temperature, T, and the partial pressure of water vapor, p_w, via

$$N = k_1\frac{p}{T} + k_2\frac{p_w}{T^2}, \qquad (2.2)$$

where k_1 and k_2 are constants ($k_1 = 77.60$ K/hPa, $k_2 = 3.73\cdot10^5$ K^2/hPa; e.g., *Bevis et al.* [1994]).

Using the refractivity equation (2.2), the hydrostatic equation, the equation of state, and the gravity formula the atmospheric parameters can be derived. For example, dry pressure $p_d(z)$ (which equals the total air pressure $p(z)$ if humidity can be neglected, i.e., above the middle troposphere) is obtained via hydrostatic integration,

$$p_d(z) = \frac{M_d}{k_1 R}\int_z^\infty g(z')N(z')\mathrm{d}z' \qquad (2.3)$$

where R is the universal gas constant ($8.3145\cdot10^3$ J K^{-1} kg^{-1}), M_d is the molar mass of dry air (28.964 kg kmol^{-1}), and $g(z)$ is the acceleration of gravity. Dry temperature is then $T_d(z) = k_1\,p_d(z)/N(z)$.

For the retrieval of temperature and humidity in the lower and middle troposphere, a priori information on at least one of the two parameters is necessary. A detailed treatment of RO retrieval techniques is given by the *Kursinski et al*. [1997] review.

The integration formulae, equations (2.1) and (2.3), are crucial links in the RO retrieval chain. Equation (2.1) indicates that the inversion of bending angles leads to downward propagation of high altitude bending angle errors. Due to the localized kernel $(a'^2–a^2)^{-1/2}$ of the inverse Abel transform, the vertical correlation of errors is still limited in $N(z)$ but further error propagation occurs via the hydrostatic integration, equation (2.3). It is thus vital to use adequate bending angles also at altitudes above any height of interest. A detailed discussion of RO retrieval error propagation in dry air is provided by *Rieder and Kirchengast* [2001b].

b) Ionospheric Correction. Ionospheric effects dominate the RO bending angles at heights above ~45 km [e.g., *Hocke*, 1997]. Since the ionosphere, as a dispersive medium, leads to different phase delays Φ_1 and Φ_2, and subsequently bending angles α_1 and α_2, these effects can be removed to first order by linear combination of the two signals. The traditional approach to do so is the linear correction of phase delays [e.g., *Spilker*, 1980],

$$\Phi_{LC}(t) = \frac{f_1^2 \Phi_1(t) - f_2^2 \Phi_2(t)}{f_1^2 - f_2^2}, \qquad (2.4)$$

where Φ_{LC} is corrected phase delay and t is the time.

In recent applications, the method of linear correction of bending angles has been applied more successfully [*Vorob'ev and Krasil'nikova*, 1994],

$$\alpha_{LC}(a) = \frac{f_1^2 \alpha_1(a) - f_2^2 \alpha_2(a)}{f_1^2 - f_2^2}, \qquad (2.5)$$

where α_{LC} is the corrected bending angle, and α_1, α_2 are the uncorrected bending angles of the two signals at impact parameter a. As shown in several theoretical and simulation studies, the linear correction of bending angles provides better results, since it accounts for different ray paths of the two signals and exploits the fact that most of the total bending angle is accumulated near the ray perigee [e.g., *Vorob'ev and Krasil'nikova*, 1994; *Ladreiter and Kirchengast*, 1996].

There are some methods that account for higher order effects of the ionosphere [e.g., *Gorbunov et al.*, 1996; *Syndergaard*, 2000], but these rely on additional a priori knowledge and on the assumption of spherical symmetry in the ionosphere. In section 2.1.2.2 we present a high altitude RO retrieval evaluation study that, among other matters, systematically tested the ionospheric correction of bending angles under different large-scale ionospheric asymmetry conditions and different ionization levels. Effects of ionospheric small-scale variations ("ionospheric scintillations"), which act to increase random noise on top of receiver thermal noise and clock noise and which are amplified by the linear signal combination [e.g., *Kursinski et al.*, 1997], are not dealt with in this study; these need careful filtering in the processing of real data [e.g., *Gorbunov*, 2002c; and references therein].

c) Review of High Altitude Initialization/Statistical Optimization Methods. Even after ionospheric correction, retrieval results at heights above 20–30 km are sensitive to residual ionospheric noise (higher order terms) and other noise (e.g., receiver thermal noise). This calls for sensible use of RO data at high altitudes, i.e., in the stratopause region and above. Additionally, since its upper integration limit ranges to infinity, the inverse Able transform needs some kind of high altitude initialization of the bending angle profile in any case. The simplest way to do so is to select an upper boundary height (UBH) at altitudes between 50–70 km above which an extrapolated exponential profile is used [e.g., *Kursinski et al.*, 1997]. This approach ("exponential extrapolation" or "no optimization" hereafter) features several weaknesses, the most important ones being the sensitivity of retrievals to the exact UBH setting and to extrapolation quality due to measurement noise as well as the intrinsic assumption of an isothermal (constant scale height) atmosphere above the UBH. The latter assumption leads to a systematic bias in any ensemble of retrieved profiles, which unfavorably is also UBH-dependent, since isothermality above 50–70 km insensibly ignores the actual mesosphere/lower thermosphere temperature structure.

A more robust and sensible approach is statistical optimization, a form of optimal estimation [*Turchin and Nozik*, 1969; *Rodgers*, 2000]. It finds the most probable bending angle profile α_{opt} by combining the observed profile with a background profile in a statistically optimal way [*Sokolovskiy and Hunt*, 1996; *Rodgers*, 2000]. The general effect of statistical optimization is that at higher altitudes, where the observation error exceeds the error of the background, the

background determines α_{opt}. At lower altitudes, where the background error becomes dominant, the observed data determine α_{opt}. The transition zone between these two regimes lies typically between 45 and 60 km, depending on the error characteristics of the observed and background data. Note that statistical optimization does not improve the quality of observed profiles themselves at high altitudes but rather delivers an improved combined profile thanks to the sensible inclusion of background information. The most important effect is that downward propagation of errors via the inverse Abel transform and the hydrostatic integral is reduced.

Background profiles can be obtained from climatological models or profile datasets or from NWP analyses or forecasts. Since the quality of NWP analyses at high altitudes (above ~30 km) is still poor, due to the sparseness of adequate assimilated data at these heights, and since the RO-retrieved profiles may later be evaluated using these analyses or even be used in the data assimilation process producing them (which would result in "incest" problems), we presently prefer the use of climatological models like MSISE-90 [*Hedin*, 1991] or CIRA-86 [*Fleming*, 1988]. Previous work cited below also used either of these two standard climatologies, which are closely similar in the height range 30–120 km.

The full optimization formula which optimally combines, in a least-squares sense, observations and background reads ("inverse covariance weighting" optimization hereafter),

$$\boldsymbol{\alpha}_{opt} = \boldsymbol{\alpha}_b + \mathbf{B}(\mathbf{B}+\mathbf{O})^{-1}(\boldsymbol{\alpha}_o - \boldsymbol{\alpha}_b), \qquad (2.6)$$

where α_b and α_o are the background and observed bending angle profiles, and **B** and **O** are the background and observation error covariance matrices, respectively. The assumptions of this approach are that unbiased (Gaussian) errors and a linear problem are assumed. Linearity is fulfilled in the present case, and the unbiased error assumption holds reasonably well for the observation error due to the self-calibrating nature of RO measurements. The background errors are likely to involve biases in addition to their random component, however. Consequently, biases in background data, which could partially leak into retrievals and thereby degrade the climate monitoring utility of retrieved profiles, are an important problem in the practical application of statistical optimization and are thus one key matter of this paper.

Since it is difficult to obtain accurate covariance matrices, *Sokolovskiy and Hunt* [1996] used a simpler form assuming vertically uncorrelated errors. This allows to calculate the optimized bending angles ray-per-ray ("inverse variance weighting" optimization hereafter),

$$\alpha_{opti} = \alpha_{bi} + \frac{\sigma_{bi}^2}{\sigma_{bi}^2 + \sigma_{oi}^2}(\alpha_{oi} - \alpha_{bi}), \qquad (2.7)$$

where i is a height index, and σ_{bi} and σ_{oi} are the estimated standard deviations of the background and the observation data, respectively. σ_{bi} is typically assumed to be a constant fraction of the background bending angle profile (5%–20%) and σ_{oi} is estimated, generally as a constant σ_o, from the root-mean-square (RMS) error of the observed profile compared to the background, or from the high frequency variations of the observed signal, within a high altitude range of 10–15 km above 65 km, where noise dominates the signal [*Gorbunov et al.,* 1996; *Gorbunov and Gurvich*, 1998; *Hajj et al.*, 2002].

A modification of this "inverse variance weighting" is the "heuristic weighting" optimization [*Hocke et al.*, 1997; *Hocke*, 1997; *Steiner et al*, 1999],

$$\alpha_{opti} = \alpha_{bi} + \frac{\sigma_{bi}}{\sigma_{bi} + \sigma_{oi}} (\alpha_{oi} - \alpha_{bi}), \qquad (2.8)$$

where σ_{bi} is again assumed to be a constant fraction of the background bending angle (typically 20%) but σ_{oi} is defined to be the absolute difference between the observed profile and the background at each height i, $\sigma_{oi} = |\alpha_{oi} - \alpha_{bi}|$. This approach effectively down-weights "outliers" and smoothes the profile at high altitudes, but is non-optimal in a formal statistical sense and tends to produce retrievals biased towards the background (see also *Healy* [2001]).

Healy [2001] suggested to use the full inverse covariance weighting approach, equation (2.6), and demonstrated it with a simplified analytical background error covariance matrix of Gaussian shape,

$$B_{ij} = \sigma_{bi}\sigma_{bj} \exp\left[-\frac{(a_i - a_j)^2}{L^2} \right], \qquad (2.9)$$

where B_{ij} are the elements of the error covariance matrix **B**, σ_{bi} and σ_{bj} are background standard errors at heights i and j, a_i and a_j are the impact parameters, and L is the error correlation length. L was set to 6 km and σ_{bi} again to 20% of the background value. Regarding observation errors, a constant vertically uncorrelated error σ_o set to 5 μrad was assumed.

A somewhat similar (applying inverse covariance weighting as well) but more general concept was introduced and discussed by *Rieder and Kirchengast* [2001b], who generalized the treatment of the optimization problem to employ the full breath of the optimal estimation methodology detailed by *Rodgers* [2000]. In demonstrating the concept, the background bending angle errors σ_{bi} were set to linearly increase from 6% to 18% of the background profile between 30 and 120 km; the background covariances were analytically derived according to equation (2.9). Though the demonstration started with phase delay observations the concept can be equally well used starting with bending angles.

Gorbunov [2002c] proposed a combined ionospheric correction and statistical optimization algorithm where the background is linearly fitted to the observations between 40 and 60 km to reduce the background bias, observation and background errors are assumed to be vertically uncorrelated, and both, σ_o and σ_b are dynamically estimated: σ_o is estimated from the variance of the observation above 50 km and σ_b is estimated from the RMS deviation between background and observations between 12 and 35 km.

An alternative to the inversion approach described in section 2.1.2.1.a is the assimilation of observed (ionosphere-corrected) bending angle profiles with forward-modeled background data using one-dimensional variational (1DVAR) data assimilation as proposed and described, e.g., by *Eyre* [1994], *Palmer et al.* [2000], and *von Engeln et al.* [2003]. Using this approach, the observed data is only processed to the bending angle level and no further inversion is applied. The main challenges using this approach are the adequate definition of the forward modeling operators and background error covariance matrices as well as the lack of a simple method to cope with biases in the background data.

d) The Basic IGAM High Altitude Retrieval Scheme. The basic high altitude retrieval scheme used at IGAM applies the linear correction of bending angles, equation (2.5), followed by inverse covariance weighting, equation (2.6), using MSISE-90 bending angle profiles as background together with an analytical background error covariance matrix with an error

correlation length of L = 6 km. Different from *Healy* [2001], exponential decay of covariances was used,

$$B_{ij} = \sigma_{bi}\sigma_{bj} \exp\left[-\frac{|a_i - a_j|}{L} \right], \qquad (2.10)$$

which was found more adequate in an empirical error study by *Steiner and Kirchengast* [2004a]. Observation errors O_{ij} were specified in the same form as B_{ij}, but with a shorter correlation length (L = 1 km) based as well on results by *Steiner and Kirchengast* [2004a]. This short-range correlation accounts for the smoothing of the phase delays in an earlier step of the retrieval. The background standard deviation σ_b is assumed to amount to 20% of the background bending angle, which is a frequently used estimate of the stratopause/mesosphere RMS uncertainty of the MSISE-90 climatology and is also based on retrieval performance studies conducted at IGAM (not presented here). The observation errors σ_o are estimated from the RMS deviation of the observed bending angle profile compared to the background between 70 and 80 km, where the bending angle signal is at the ≤ 1 μrad level and ionospheric residual and measurement noise dominate. The optimization, equation (2.6), is applied from 120 km down to 30 km height, below which the signal-to-noise ratio is so strong that it is safely no longer required [e.g., *Rieder and Kirchengast*, 2001b].

Syndergaard (Inst. of Atmos. Physics, Univ. of Arizona, Tucson, pers. communications, 1999) suggested to perform background search prior to statistical optimization, i.e., to fit the available ensemble of background profiles to the observed profile, and use the best-fit, instead of the co-located profile, as background in the statistical optimization process. We adopted this approach in the IGAM retrieval scheme, searching the global MSISE-90 climatology on a 5 deg latitude x 15 deg longitude grid through all months from January to December and using a least squares fit in the 45–65 km height interval to determine the best-fit profile. Besides this statistical optimization of bending angles no further background information is used in the IGAM scheme, different from most other retrieval chains, which also initialize the hydrostatic integral, equation (2.3), using data from NWP analyses or climatologies.

The scheme, which was tested in a systematic high altitude retrieval evaluation study [*Gobiet and Kirchengast*, 2002] and applied in the GNSS-CLIMATCH study (see section 2.1.1), already proved to be effective. However, it cannot cope with a situation where no unbiased data is available in the background climatology. As will be shown in section 2.1.3, for the MSISE-90 climatology this situation frequently occurs in the high-latitude winter region. Similar results were found by *Randel et al*. [2003] for the CIRA-86 climatology, which provided the basis for the MSISE-90 climatology below the thermosphere.

e) The Enhanced IGAM High Altitude Retrieval Scheme. Considering the results from CHAMP, the GNSS-CLIMATCH study, and high altitude retrieval evaluation studies presented in section 2.1.3, the most promising pathway to advance high altitude RO retrieval is to pay more attention to biases in the background information used in the statistical optimization approach. Since presently no better mesospheric climatology than MSISE-90 (or CIRA-86) is available and since direct use of NWP analyses as background is not desirable (see section 2.1.2.1.c), an extension to the background bias correction ability of the basic IGAM scheme was developed.

The main components of the enhanced algorithm are linear ionospheric correction of bending angles, equation (2.5), empirical background bias correction of bending angles

(explanation below), and inverse covariance weighting statistical optimization of bending angles, equation (2.6). These main components involve the following ingredients:

Background information. Bending angle profiles derived from the MSISE-90 climatology. (Alternatively, the use of more diverse libraries of upper stratosphere/mesosphere bending angle profiles is studied, e.g., based on middle atmosphere LIDAR profile or ENVISAT/-MIPAS limb sounder profile databases; also ECMWF analyses may be useful, which currently do not reach beyond ~60 km, however.)

Empirical background bias correction, step 1. Search the best-fit bending angle profile in the climatology using a least-squares criteria in the 45–65 km interval (section 2.1.2.1.d). The optimal height interval depends on the noise level of the observations. For testing this scheme we used the GRAS error specifications [e.g., *GRAS-SAG*, 1998]. For more noisy observations (e.g., GPS/MET, CHAMP) it is necessary to place the fit interval at lower altitudes.

Empirical background bias correction, step 2. Adjustment of the background profile by multiplying it with a fitting coefficient – a factor usually close to unity – derived from regression with respect to the observed profile at high altitudes (least-squares adjustment). A similar approach was suggested by *Gorbunov* [2002c], though that approach uses lower altitudes and was not combined with background search. Tests showed that the 55–75 km interval is especially sensitive for detecting remaining background biases. The optimal height interval depends on the observation noise level and the error characteristics of the background. Regarding the resulting optimized profile it represents the region just above the transition zone from background-dominated to observation-dominated. GRAS error specifications were used when testing the correction. For more noisy observations (e.g., GPS/MET, CHAMP) it is necessary to use a lower altitude range.

Background error. Standard deviation 15% (instead of 20%) of the background bending angle profile (accounting for the bias reduction and determined in empirical tests); vertically correlated over about a scale height.

Observation error. Standard deviation estimated from the RMS deviation of the observed bending angle profile compared to the background (after bias correction) within 70–80 km; weekly vertically correlated, equation (2.10).

The key advancement of this scheme is its bias correction before the optimization, which is done in an empirical way by comparing the background to the observations within suitable altitude ranges. It combines the advantages of the near bias-free observations and the small statistical error of the background. Step 1 means that the climatology is used as a library of representative bending angle profiles and the best-fit profile is chosen as background. Step 2 means that if there is no well fitting profile in the library, there is the possibility to adjust the background profile towards the observed profile.

2.1.2.2 High Altitude RO Retrieval Evaluation Studies

The performance of the high altitude RO retrieval schemes was evaluated in a twofold way: The first part was a systematic case study, which focused on individual occultation events and their error characteristics under different ionospheric conditions. The second part was a large-sample study focusing on sample error statistics and spatial (latitudinal) distribution of errors. Both parts were based on end-to-end simulations of RO events, which provides the advantage that the "true" state of the atmosphere is known and a detailed error analysis is possible. The complete study, beginning with satellite geometry simulation and proceeding with modeling of GNSS signal propagation through the atmosphere/ionosphere, simulation of the observation system and the observables, and retrieval of the atmospheric profiles was performed by means

of a study-tailored version of the EGOPS software tool (End-to-end GNSS Occultation Performance Simulator; *Kirchengast et al*. [2002]).

a) Systematic Case Study Setup. The setup of this study is described in some detail by *Gobiet and Kirchengast* [2002]. In the framework of this paper we expanded it by two additional retrieval schemes.

In order to study the effects of statistical optimization as a function of ionospheric state, we choose three occultation events (out of a sample consisting of ~2100 events) as base scenarios, each of them being representative for one type of symmetry or asymmetry of electron density distribution in the ionosphere (Figure 2.2a–c). For the sake of brevity we refer to them as "Nice", "Nasty 1", and "Nasty 2" event hereafter. We simulated the propagation of radio signals through the atmosphere without ionosphere ("*no ionosphere*" reference case) as well as at three different ionization levels, represented by the radio flux at 10.7 cm ($F_{10.7}$ index) ranging from $F_{10.7} = 70$ via $F_{10.7} = 140$ to $F_{10.7} = 210$ (low, mid, high solar activity).

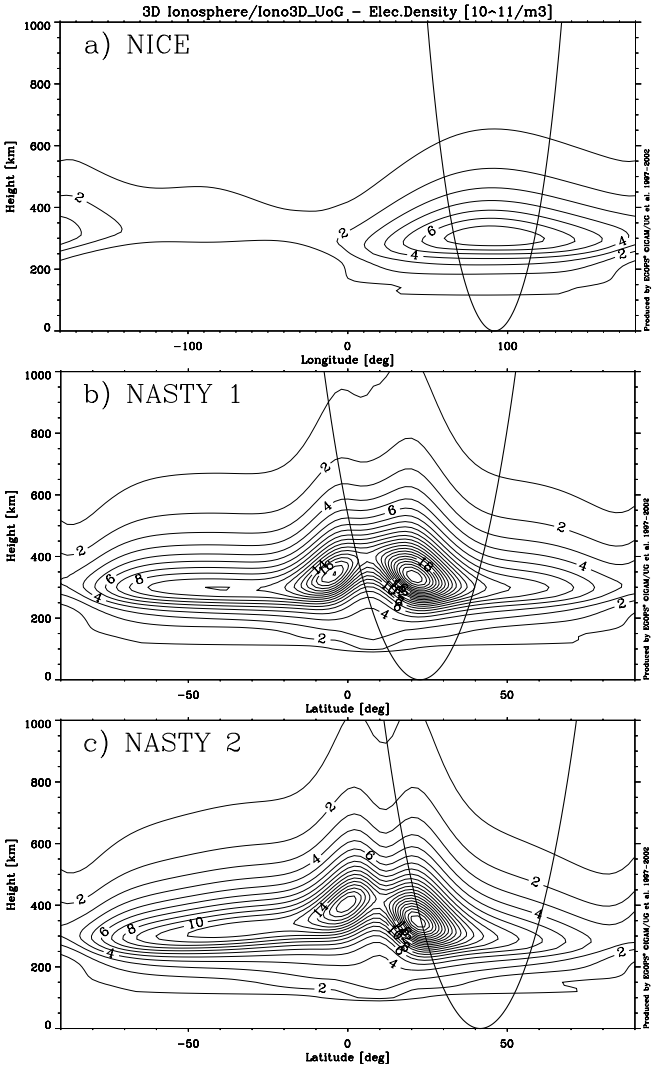


Figure 2.2. Ionospheric conditions for the three representative occultation events (for $F_{10.7} = 140$). "Nice" event (a): vertical total electron content (TEC) above tangent point ~20×10^{16} m^{-2}; electron density varies < 1.5×10^{11} m^{-3} between inbound and outbound of lowermost ray. "Nasty 1" event (b): vertical TEC above tangent point ~60×10^{16} m^{-2}; electron density varies ~12×10^{11} m^{-3} between inbound and outbound of lowermost ray and is maximal above tangent point. "Nasty 2" event (c): vertical TEC above tangent point ~25×10^{16} m^{-2}; electron density varies ~18×10^{11} m^{-3} between inbound and outbound of lowermost ray and is maximal at outbound.

The ionosphere model utilized in the simulation was the NeUoG model [*Leitinger and Kirchengast*, 1997a]. NeUoG provides global 3D electron density distributions depending on local time, season, and solar activity. It features reasonably realistic climatological ionization conditions and has already proven to be useful in the past in occultation-related studies [e.g., *Leitinger and Kirchengast*, 1997b] but does not feature small-scale structures. This latter restriction should be kept in mind when interpreting the results of this study.

The "Nice" event (Figure 2.2a) was chosen to be zonally oriented. By this and due to careful selection of the occultation event location, the spherical symmetry assumption is met as good as it is possible in a realistic ionosphere during daytime conditions. The two "Nasty" events (Figures 2.2b and 2.2c) were chosen to be meridionally oriented and probe through severe electron density gradients associated with the equatorial anomaly, each violating the spherical symmetry assumption in a specific distinct manner. The "Nasty 1" event exhibits relatively low electron densities at the in- and outbound of the occultation rays and a maximum in electron density above the occultation tangent point. The "Nasty 2" event exhibits low electron densities at the inbound and high electron densities at the outbound of the occultation rays.

Since the aim of the study is to describe the effects of the ionospheric conditions and the method of high-altitude initialization on the RO retrieval products, rather than effects of error sources internal to the neutral atmosphere, we supplied the forward model with a simple neutral atmosphere in order to ensure that the errors in the subsequently retrieved atmospheric parameters would be fully traceable to ionospheric residuals and the method of statistical optimization. We used a dry atmosphere, local spherical symmetry, and identical conditions for all three simulated occultation events. The neutral atmosphere employed this way was one selected vertical profile out of the MSISE-90 climatology, i.e., the same climatology we used as background in the statistical optimization process. This adds some interesting aspects to the interpretation of the results (section 2.1.3.1). We chose the profile co-located with the "Nice" event (63°N, 93°E; month September).

Additionally, two different receiving systems were modeled: an idealized, where observation system related errors were neglected, and a "realistic" one, which was based on the error specifications of the GRAS receiver [e.g., *GRAS-SAG*, 1998] and included the main receiving system related errors such as orbit uncertainties, receiver noise, local multipath errors, and clock errors. Subsequently, we performed the retrieval using five different retrieval scheme settings.

We show results for *no optimization* (with observed profile exponentially extrapolated; section 2.1.2.1.c), *heuristic weighting* optimization, equation (2.8), inverse covariance weighting optimization without (*basic IGAM-no search*) and with (*basic IGAM-search*) background search (section 2.1.2.1d), and inverse covariance weighting with enhanced background bias correction (*enhanced IGAM*; section 2.1.2.1.e). In total, 24 different scenarios were modeled (3 representative events x 4 ionization levels x 2 receiving system types), and each of these 24 scenarios evaluated using the 5 different retrieval schemes. Results of this study are discussed in section 3.1.

b) Large Ensemble Study Setup. Selected RO retrieval schemes were also tested using a large ensemble of quasi-realistically simulated occultation events. For allowing comparison to earlier results, we used a simulated occultation event ensemble, which is globally distributed and was originally produced in the framework of the GNSS-CLIMATCH study (section 2.1.1). It involves the simulation of a six-satellite constellation, each satellite equipped with a GRAS-type receiver. The neutral atmosphere modeling used a middle atmosphere-extended version (T42L39 resolution) of the ECHAM4 General Circulation

Model [*Roeckner et al*., 1999] with a horizontal resolution of ~2.8° and 39 vertical levels up to 0.01 hPa (~80 km). Above about 80 km, the neutral atmosphere was extrapolated using the MSISE-90 climatology. While typical current GCM fields, including ECMWF analyses, do not reach beyond 0.1 hPa, the extend up to near the mesopause is important for the present assessment of retrieval schemes as it ensures that any search and best-fit estimates are performed at height ranges governed by the more variable GCM conditions rather than the climatological conditions. The ionosphere was modeled using the NeUoG model. For more details see *Foelsche et al.* [2003] and references therein.

The ensemble consists of ~1000 occultation events that characterize the northern summer/southern winter season 1997 (June, July, and August) and are evenly distributed in time and space. It was partitioned into 17 equal area latitude bins of 10° width (equatorial bin: 10° lat x 15° lon), each containing 50 – 60 occultation events [*Foelsche et al*., 2003], and enables to analyze error statistics and the latitudinal dependency of errors. We used this ensemble to systematically test the performance of three different high-altitude retrieval schemes (*no optimization*, *basic IGAM-search*, and *enhanced IGAM*; see section 2.1.2.2.a), with focus on the latitude-height distribution of biases. Results of this study are discussed in section 2.1.3.2.

2.1.3 Results and Discussion

2.1.3.1 The Systematic Case Study

Figure 2.3 gives an overview on the performance of the *no optimization* retrieval and the inverse covariance weighting optimization retrieval with search in MSISE-90 applied to the 24 simulated scenarios. Each bar in Figure 2.3 represents the mean temperature error between 35 and 45 km ("upper stratosphere bias" hereafter) of one specific scenario. At these altitudes, the RO-retrieved temperatures are heavily influenced by background information, thus the upper stratosphere bias indicates the bias in the background, or in the exponentially extrapolated bending angle profile, respectively.

The most salient result is the strong mitigation of the upper stratosphere bias by statistical optimization. Without optimization, the behavior of the retrieval above ~30 km depends heavily on noise in the data. One and the same simulated occultation event can give different upper stratosphere biases of up to 20 K, depending on the starting value of the random number generator used in the simulation of receiver noise. Applying statistical optimization (in this case the *basic IGAM* scheme) strongly stabilizes the retrieval, makes it robust against random noise, and residual ionospheric errors and retains the upper stratosphere bias below 1 K in most cases. Additionally, it is shown for the linear ionospheric correction of bending angles that neither highly asymmetric ionospheric conditions nor high ionization levels do significantly degrade the retrieval performance, especially when statistical optimization is applied, as one should expect as long as noise due to small-scale ionospheric structures is small [e.g., Gorbunov [2002c]; and references therein].

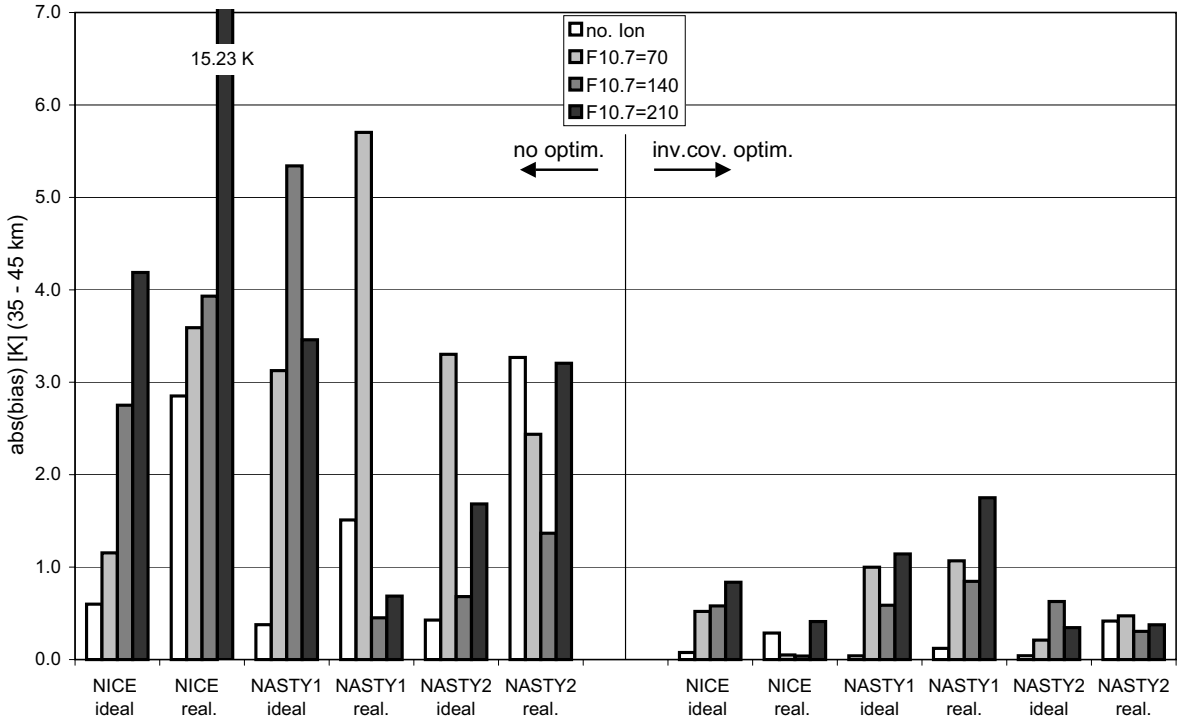


Figure 2.3. Mean error of the retrieved temperature between 35 and 45 km ("upper stratosphere bias") for all simulated occultation scenarios (3 events, 4 ionization levels, ideal and realistic receiving system). Left: *no optimization* retrieval. Right: *basic IGAM-search* retrieval.

Figure 2.4 shows the error profiles of temperature for one example occultation scenario ("Nasty 1", $F_{10.7} = 70$, realistic receiving system) retrieved with the five different high altitude retrieval schemes introduced in section 2.1.2.2.a.

In this scenario, the co-located background profile from MSISE-90 is biased, since we used a different neutral atmosphere profile (from the "Nice" event) in the forward modeling (see section 2.1.2.2.a). However, inspection shows that the background error is < 8% in the strato- and mesosphere, which is clearly within the assumed uncertainty of 20%. The example scenario can thus be considered representative for a real RO retrieval situation.

Figure 2.4, panel a, shows that errors due to deficient high altitude initialization can propagate down to below 20 km in the temperature profile. While this may be considered a "bad case", error propagation down to < 30 km will be common. *Heuristic weighting* optimization (panel b) stabilizes the retrieval, but it is strongly biased towards the background. The inverse covariance weighting optimization without search leads to much better results (panel c), but the temperature stays biased towards the background. The retrieval with search (panel d) is able to account for the biased background and manages to find a un-biased background profile. The retrieval results stay virtually unbiased up to 40 km. This demonstrates the situation, when a search library is adequate, i.e., when an unbiased profile is available in the library. In such a situation a second step of background bias reduction is not necessary and the *enhanced IGAM* retrieval (panel e) gives essentially the same result as the *basic IGAM-search* retrieval. In the example shown here (panel e), the *enhanced IGAM* retrieval is still slightly better than the basic one with search, while in other cases it can also be slightly worse or closely the same, but never far away from the basic retrieval (difference of upper stratosphere bias < 0.5 K), as one expects for an adequate search library.

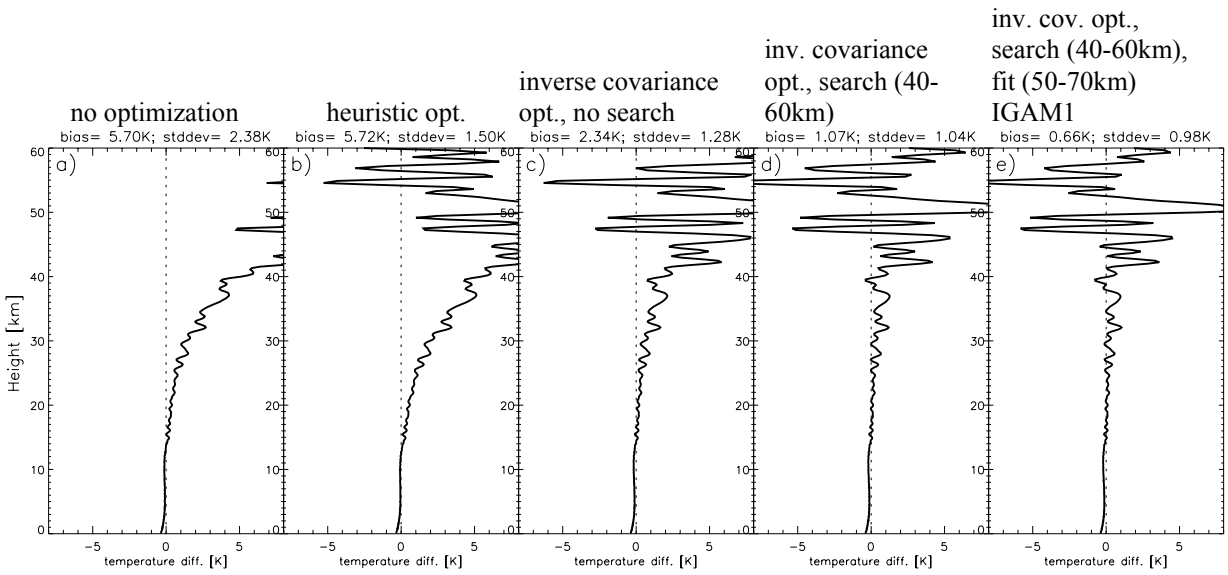


Figure 2.4. Temperature error profiles for the "Nasty 1" occultation event (asymmetric conditions, $F_{10.7} = 70$, realistic receiver) retrieved with different high altitude retrieval schemes: *no optimization* (a), *heuristic weighting* optimization (b), *basic IGAM-no search* c), *basic IGAM-search* (d), and *enhanced IGAM* (e).

2.1.3.2 The Large Ensemble Study

The ECHAM4 model was used to create neutral atmospheric fields representing the summer season 1997 for the simulation of a large ensemble of occultation events (~1000 events; see section 2.1.2.2.b). In contrary to the systematic case study, this setup does not automatically ensure that an unbiased background profile is available for each retrieval, since the forward model atmosphere is different from the MSISE-90 climatology used for background information. The MSISE-90 search library may thus be partially inadequate, which is a situation more realistic but also more demanding for the statistical optimization scheme.

Figure 2.5 shows latitude versus height slices of the dry temperature retrieval seasonal mean bias and standard deviation for the *no optimization*, *basic IGAM-search*, and *enhanced IGAM* retrieval schemes.

The *no optimization* case shows again that the high altitude retrieval performance is very unstable if the bending angle profile is simply exponentially extrapolated. The > 1 K standard deviation region starts at ~25 km, which is about 10 km lower than in the two optimization cases. The most salient improvement of the *enhanced IGAM* algorithm over the *basic IGAM-search* algorithm is the drastic error reduction in the southern high latitude region, where the MSISE-90 search library is most inadequate. Less salient, but robustly evident, is furthermore a general improvement of the retrieval quality above 30 km.

The additional bias correction step of the *enhanced IGAM* scheme exhibits the behavior of an "emergency reserve": In most cases, where the background library is adequate, the background bending angles are only slightly modified (~1 %) and the effect on the retrieved temperatures below 40 km is small. In the cases, however, when no unbiased background profile is found in the library, the modification reaches up to 15%, and significant improvement is achieved. For these cases the MSISE-90 climatology is inadequate to supply profiles unbiased against the ECHAM4 model atmosphere used. The capability of the *enhanced IGAM* scheme to handle these conditions is important since also in case of real data existing climatological libraries will not be able to supply adequate background profiles under all conditions.

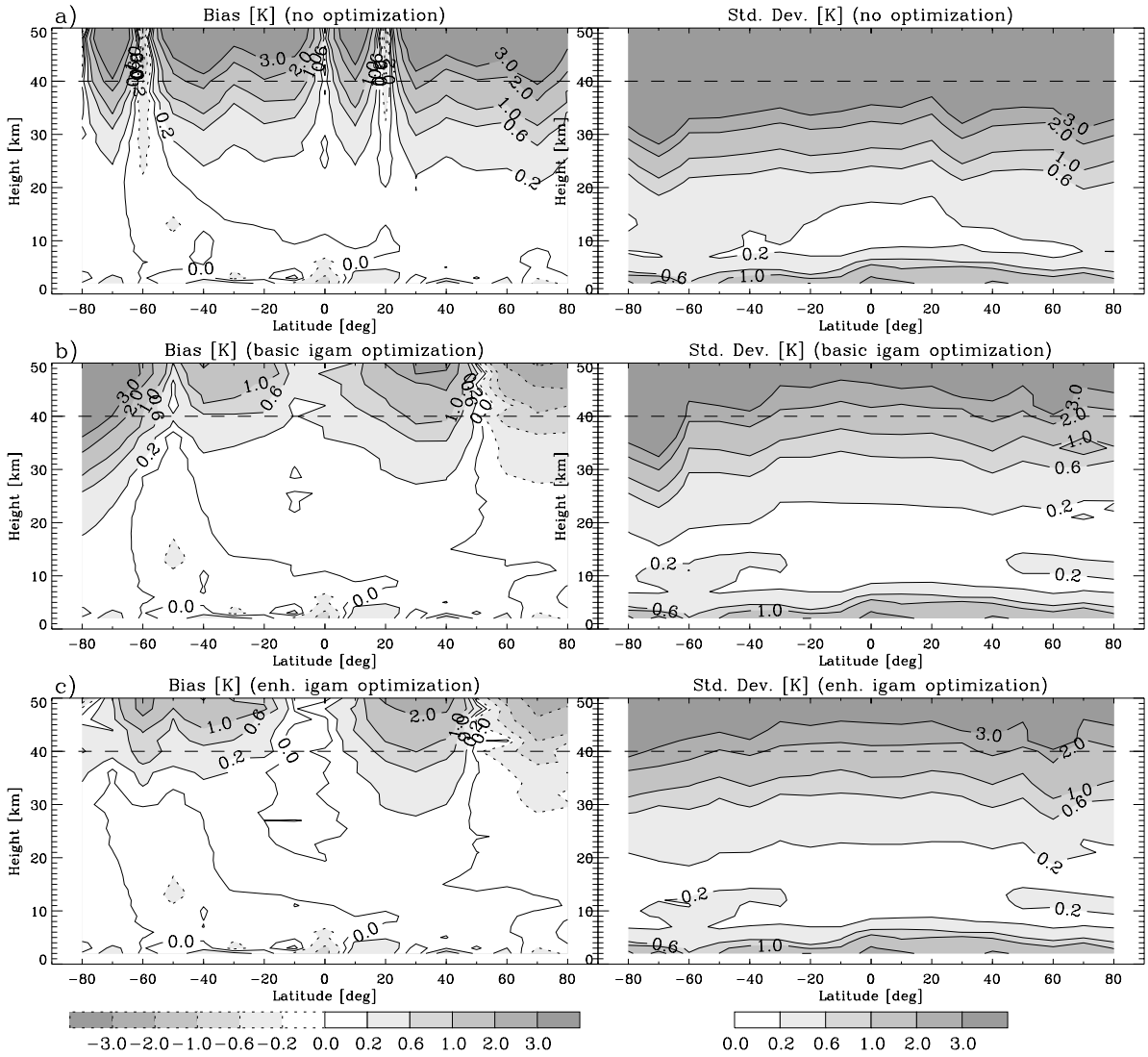


Figure 2.5. Seasonal mean bias (left) and standard deviation (right) of temperature retrieved from an ensemble of ~1000 simulated occultation events. High altitude retrieval schemes: *no optimization* (a), *basic IGAM-search* (b), and *enhanced IGAM* (c).

2.1.4 Summary and Conclusions

An enhanced statistical optimization scheme for GNSS RO retrievals was presented. This and other retrieval schemes were tested in two retrieval performance studies, one based on representative profile scenarios, the other on a large ensemble of profiles.

Sensible statistical optimization of bending angles is found vital to obtain results of good quality in the upper stratosphere. In this context, it is shown for the linear ionospheric correction of bending angles that neither highly asymmetric ionospheric conditions nor high ionization levels do significantly degrade the retrieval performance, as one should expect as long as noise due to small-scale ionospheric structures ("ionospheric scintillations") is small. The latter has been disregarded in this investigation; in the processing of real data it needs separate care in filtering before correction (see section 2.2.5).

Using statistical optimization, it is found that the retrieval performance can be significantly improved further if biases in the background information are corrected. Since the statistical optimization method fundamentally assumes unbiased errors, potential biases need to be removed before the optimization. The "enhanced IGAM scheme" presented incorporates a 2-step empirical background bias correction: a first step searches and extracts a best-fit profile (fitting over the stratopause height range) from a suitable library of profiles instead of using the co-located profile, a second step adjusts this best-fit profile towards the observed profile (by fit over the lower mesosphere height range), further compensating potential bias in the best-fit profile if existing. The exact height ranges for fitting depend on the observational noise in the RO data, the higher the noise, the lower down these height ranges (the present study used the European GRAS receiver as baseline). The bias correction is particularly effective when background data are systematically inadequate, as was found in this work in the high-latitude winter atmosphere.

Use of an independent search library for best-fit selection of background profiles (e.g., the MSISE-90 climatology) rather than co-located NWP analysis profiles (e.g., from ECMWF) avoids introducing potential analysis biases into RO climatologies. Moreover, validation of RO retrievals relative to analysis fields needs independency of the RO data from these reference data. Improved search libraries than the currently available MSISE-90 and CIRA-86 climatologies should be built in the future (e.g., from middle atmosphere LIDAR data or ENVISAT/MIPAS data) in order to ensure that adequate background profiles are available for all the diverse strato- and mesosphere conditions of different locations and seasons.

2.2 Creation of a robust retrieval scheme for CHAMP data - CCR

2.2.1 Introduction

The retrieval schemes described in the preceding section were so far applied to simulated data in the framework of a modified version of the End-to-end GNSS Occultation Performance Simulator (EGOPS) [*Kirchengast et al*., 2002]. Starting from this point, we developed a robust stand-alone retrieval system suited for application in an operational processing system for CHAMP data: the CHAMPCLIM Retrieval (CCR). The CCR-chain starts with atmospheric phase delay profiles from CHAMP (level 2 data provided by GeoForschungsZentrum Potsdam, GFZ) and returns profiles of the atmospheric parameters refractivity, density, pressure, geopotential height, dry temperature, temperature, and humidity.

The following sections describe the developments leading to the current stable version 2 of CCR (CCRv2). In this report only a few CCRv2 results are shown for demonstrational purposes. A more comprehensive discussion of retrieval results and several validation studies are described in the CHAMPCLIM validation report [*Gobiet et al.*, 2004a].

2.2.2 Interface to GFZ Level 2 Data and Technical Robustness

The creation of a flexible interface to GFZ level 2 data included data format conversion (GFZ to CCR/EGOPS format) and the modification of the retrieval algorithm's control structure. To reach maximum flexibility, the retrieval sequence is now controlled by an occultation table containing the filenames of all GFZ level 2 data files to be processed. This enables to flexibly process any period or randomly selected level 2 files regardless of temporal or any other order. In consequence to the drop of a monotonous timeline inside the algorithm the handling of date/time variables and the interface to background data had to be adapted.

After building up the interface to GFZ level 2 data, CCR was principally capable to retrieve CHAMP data but in practice only few profiles could be processed without "repairing" input data manually in advance. Since this is not acceptable for a future operational retrieval system, much afford was made to assure technical robustness of the algorithm. During this development phase atmospheric profiles from more than 4500 different CHAMP phase delay profiles were processed to assure robustness against any kind of input data including completely unphysical data, data gaps, outliers and the like. At that stage, data was not modified in any way and we did not regard the quality of the retrieval results but only assured that CCR would not collapse from a technical point of view.

2.2.3 Treatment of Noisy, Corrupted, or Missing High Altitude Data

After technical robustness was reached, CCR's performance was still not comparable with results of leading RO retrieval centers like NASA's Jet Propulsion Laboratory (JPL), EUMETSAT's GRAS Satellite Application Facility, GFZ, and the University Corporation for Atmospheric Research (UCAR). Detailed inspection showed that most profiles could be retrieved properly by CCR but some were very far away from any reference data used for validation. These outliers are generally caused by corrupted or very noisy input data, most of the problems stemming from the uppermost part of the phase delay profiles. These data often led to failure of the ionospheric correction algorithm (equation 2.5) yielding negative (unphysical) ionosphere corrected bending angles and many subsequent problems in the retrieval chain like unrealistic observation error estimates and the downward propagation of

high-altitude bending angle errors in the atmospheric profiles. For this reason, a combined strategy of making the retrieval more robust against noisy data and quality control effectively sorting out unrecoverable corrupted data was developed:

(1) Implicit quality control focusing on high altitude data including rejection of single data points (outliers) and removal of unphysical data (see section 2.2.4.1).
(2) Enhanced ionospheric correction less sensible to very noisy high altitude data (see section 2.2.5).
(3) More stable estimation of the observation error used in statistical optimization by dynamically regarding the quality of high altitude data (see section 2.2.6).
(4) Selection of CHAMP-tailored height intervals used for background bias correction in the statistical optimization process (see sections 2.1.2.1 e and 2.2.6).
(5) Implementation of an alternative background information source (numerical weather prediction (NWP) fields) in the statistical optimization process (see section 2.2.7).

2.2.4 Quality Control

Quality control implemented in the CCR algorithm can be subdivided into implicit and explicit quality control. While explicit quality control is only applied to the retrieval results, implicit quality control modifies or rejects data during the retrieval process. From the data user's point of view it is indispensable that quality control criteria are transparent and do not modify statistical properties of the 'good' part of the data. Hence, it seems desirable to apply as few implicit modifications of data as possible. Nevertheless, some data modifications are necessary to assure high delivery rate of the algorithm and to prevent the retrieval results from being biased towards atmospheric conditions favorable for RO retrieval performance.

In order to make the quality control mechanisms transparent for the user, suspect data is not simply rejected but flagged by a 2-digit integer value. Implicit quality control affects the 1-decimal place (1, 2, ..., 9) of the quality flag while explicit quality control affects the 10-decimal place (10, 20, ..., 90). The overall quality flag is obtained by adding the two numbers.

2.2.4.1 Implicit Quality Control

CCR applies the following implicit quality control mechanisms. If data is rejected due to implicit quality control it is flagged according to Table 2.1.

(1) **Enforcement of positive phase delays**: Negative phase delays are often not related to the state of the atmosphere, but due to technical aspects of the RO receiver ('unphysical data'). In the case of CHAMP, they often indicate a known receiver clock malfunction. Therefore, negative values are replaced by 'pseudo-zero' (10^{-12}) and subsequent quality control ensures that profiles severely disturbed by the enforcement of positive phase delays are sorted out. However, negative phase delays are not necessarily unphysical but can also be related to layers in the ionosphere that can be corrected for. In that case, it would be preferable not to artificially introduce pseudo-zero. For other RO experiments than CHAMP it might be possible to drop the enforcement of positive phase delays.
(2) **3-σ outlier rejection of phase delays**: Each one-second interval (50 data points) of the phase delay profiles is checked for data points being 3 times the standard deviation away from the mean. If so, the data point is replaced by the interval's mean.

(3) **Impact parameter consistency 1:** If impact parameters are rising with falling altitude for more than one second in the uppermost part of the profile, the entire profile is regarded as corrupt and rejected (quality flag 9).

(4) **Impact parameter consistency 2:** If large impact parameter ambiguities occur below 15 km impact height, i.e., if the impact parameters rise by more than 0.2 km from one data point to the next downwards (for 50 Hz sampling rate), the profile is cut off. Data downwards are regarded as non-retrievable because of receiver tracking failure or multipath conditions in the atmosphere. This check is necessary in a geometric optical retrieval as applied here. However, different lower tropospheric retrieval schemes based on wave optics exist which are capable of dealing with bending angle ambiguities to some degree [e.g., *Gorbunov*, 2002a; *Gorbunov*, 2002b; *Jensen et al.*, 2003; *Ao et al.*, 2003; *Sokolovskiy*, 2003; *Beyerle et al*., 2004]. Our future plans are to combine CCR with one of these methods to obtain a lower mean cut-off altitude.

(5) **Bending angle consistency:** If negative bending angles occur after ionospheric correction, these data are regarded to be unphysical and set to pseudo-zero. Pseudo-zero data above 50 km impact height and all data above are removed. Subsequent quality control assures that data severely degraded by this procedure are rejected entirely.

(6) **Bending angle availability**: If the ionosphere corrected bending angle profile contains no data above 35 km or below 20 km impact height the entire profile is regarded as non-retrievable and rejected (quality flag 6).

(7) **Consistency of estimated observation error 1:** The error of the observed bending angles is estimated from the variance of observed profile in the impact height-interval from 65 to 80 km. If this estimation is not possible using at least 25 data points or if the estimated observation error is unrealistically small (< 0.5 μrad), the observation error is set to a very high value (50 μrad) and the retrieval results are flagged to be 'suspect above 25 km' (quality flag 2).

(8) **Consistency of estimated observation error 2:** If the estimated observation error is larger than the bias of the observed profile compared to the background profile in the 65 to 85 km impact height interval or if the estimated observation error is larger than 50 μrad, the data is regarded to be corrupt and is rejected (quality flags 7 and 8, respectively).

Table 2.1. Meaning of quality flags (QF) for implicit quality control (EGOPS/CCR Version 2, June 2004).

Meaning	QF implicit
The observation error could not be estimated from data and was set conservatively to 50 μrad. Retrieval results may not be used above 25 km.	2
Not enough observed bending angle data available to perform statistical optimization. Retrieval results may not be used.	6
The bias (background compared to observation) is larger than estimated observation error. Retrieval results may not be used.	7
The estimated observation error is huge (> 50 μrad). Retrieval results may not be used.	8
The impact parameters are irrecoverably inconsistent at high altitudes. Retrieval results may not be used.	9

2.2.4.2 Explicit Quality Control

CCR's explicit quality control compares retrieval products to colocated atmospheric profiles extracted from operational analyses of the European Centre for Medium-Range Weather Forecasts (ECMWF) in T42L60 resolution. This comparison is performed in height intervals where RO retrieval errors as well as the ECMWF analysis errors are regarded to be small and only retrieval results far away from the analysis (outliers) are rejected to assure that eventual errors in the analysis are not introduced to the statistical properties of the retrieval results. The explicit quality control criteria are listed in Table 2.2. Retrieved data not meeting the criteria are flagged as listed in Table 2.3.

The implicit and explicit quality control described here is the most essential advancement compared to the retrieval schemes described in section 2.1. It is strongly related to all further developments described in the next sections and enables the IGAM retrieval schemes for the first time to operationally cope with CHAMP data.

Table 2.2. CCR's explicit quality control. Retrieval results containing data further away from the reference (operational ECMWF analyses, T42L60) than indicated by the threshold value are rejected (EGOPS/CCR Version 2, June 2004).

Atmospheric Parameter	Height Interval	Threshold
Temperature	8 – 25 km	20 K
Refractivity	5 – 35 km	10 %

Table 2.3. Meaning of quality flags (QF) for explicit quality control (EGOPS/CCR Version 2, June 2004).

Meaning	QF explicit
The temperature criterion is not met. Retrieval results may not be used.	10
The refractivity criterion is not met. Retrieval results may not be used.	20
The temperature and refractivity criteria are not met. Retrieval results may not be used.	30
No reference data is available. Retrieval results may not be used.	40
Not enough retrieved data is available to perform explicit quality control. Retrieval results may not be used.	50

2.2.5 Ionospheric Correction

The linear ionospheric correction of bending angles (equation 2.5) cannot correct for effects of small scale structures in the ionosphere and higher order terms in the formula describing ionospheric refractivity [e.g., *Bassiri and Hajj*, 1993]. These effects, together with measurement noise and technically corrupted data (leisurely called 'noise' thereafter), are amplified by the original linear correction of bending angles by the factor of 3 if the noise-contribution in the L1 and L2 signals are comparable or by the factor of 1.5 if L2 noise is dominant. In any way, the corrected signal carries at least 1.5 times the noise of the signal with the lower signal-to-noise ratio (SNR) which is not satisfying. More than that, technically corrupted data often lead to unphysical (negative) bending angles after correction causing the subsequent retrieval steps to fail. *Hocke et al.* [2003] describe a modification of the linear ionospheric correction of bending angles aiming at more stability and at reducing the noise-amplification (equation 2.11). This modified scheme uses smoothed bending angle profiles $\overline{\alpha}(a)$ assuming that highly varying signatures can not be corrected for anyway since they are either not present in both signals (due to different ray paths) or represent higher order terms not covered by the linear correction.

$$\alpha_{LC}(a) = \frac{f_1^2 \overline{\alpha}_1(a) - f_2^2 \overline{\alpha}_2(a)}{f_1^2 - f_2^2} + \delta\alpha_1(a), \quad \delta\alpha_1(a) = \alpha_1(a) - \overline{\alpha}_1(a) \qquad (2.11)$$

The advantage of this scheme is that noise amplification is reduced and that the formula is more robust against technically corrupted data. Since some of the small scale features filtered out can also be related to structures in the neutral atmosphere, the high-frequency part is added to the signal again after correction ($\delta\alpha_1(a)$) disregarding the more noisy L2 signal. This has the advantage that the ionospheric corrected bending angle profile is less sensitive to the low L2 SNR without loosing information about the non-dispersive neutral atmosphere.

Hocke et al. [2003] proposed that bending angle fluctuations in the scale of < 1 km should be disregarded in the smoothed bending angle profiles. Tests showed that this procedure can reduce the variance of errors in an sample of profiles, but that the filter width has to be carefully selected since too broad filters introduce positive temperature biases above 30 km. Applied to CHAMP data we found, in agreement with *Hocke et al.* [2003], a 1 km boxcar filter applied to bending angles and impact parameters giving a good performance.

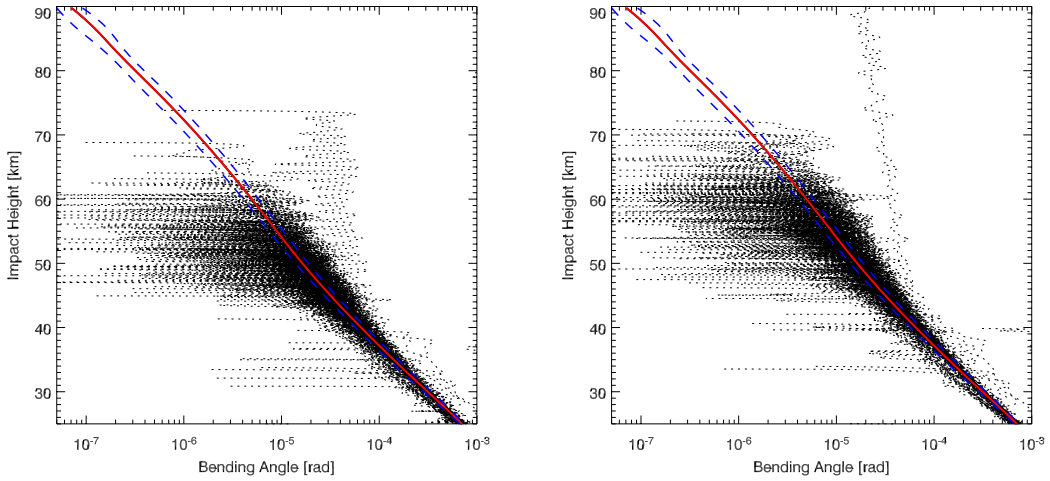


Figure 2.6. Ionospheric corrected bending angle profiles measured by CHAMP on Jan. 1, 2003. Sample-size: 206 events, no outliers were removed. The red and blue lines demonstrate the climatological mean and standard deviation, respectively (derived from CIRA86). Left panel: Standard ionospheric correction of bending angles. Right panel: Smoothed ionospheric correction of bending angles (1 km boxcar filter).

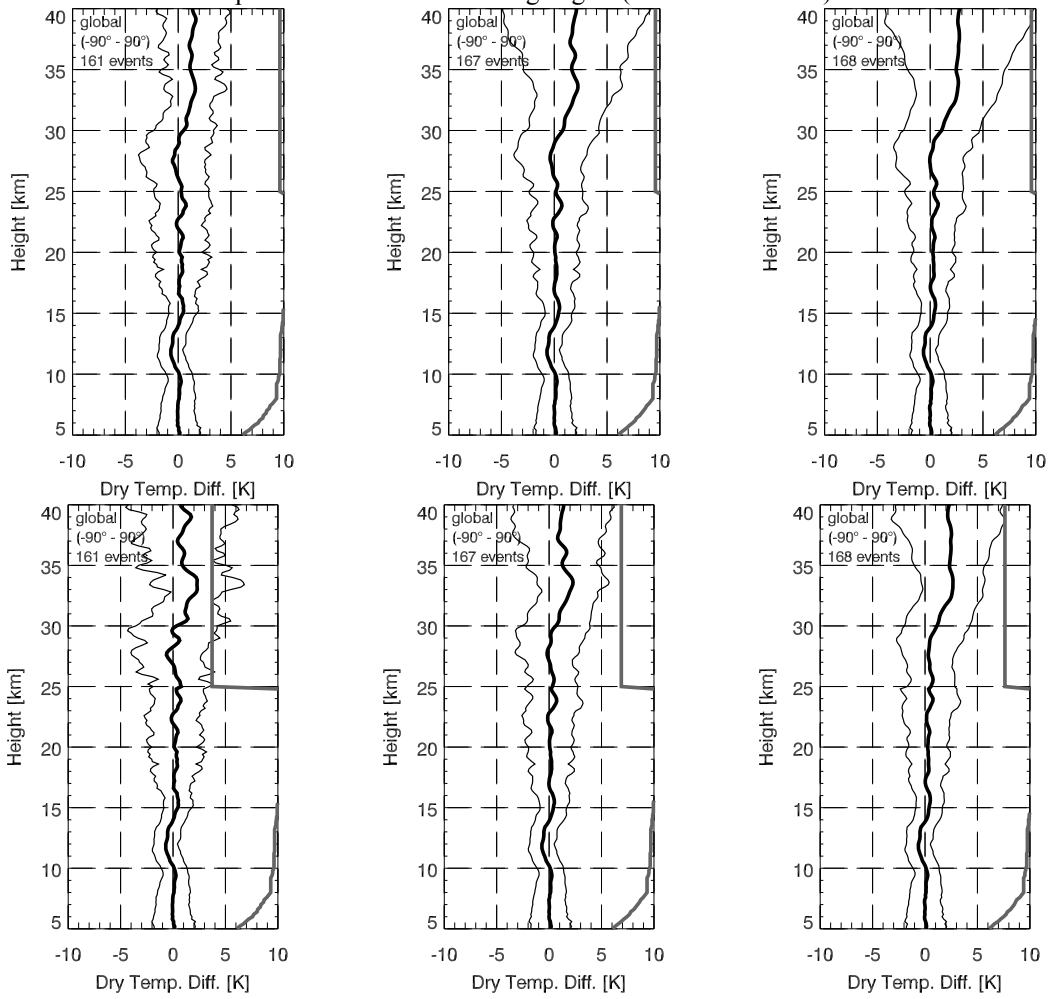


Figure 2.7. CHAMP dry temperature error statistics (Jan. 1, 2003, ~170 profiles) compared to ECMWF analyses. Bold line: bias. Thin line: standard deviation. Grey line: percent number of data points contributing to each altitude's error statistic (10 K corresponds to 100 %). Upper left panel: standard ionospheric correction. Upper middle and upper right panels: smoothed ionospheric correction with 1 km and 2 km boxcar filter, respectively. Lower panels: same ionosphere correction schemes as above with stricter implicit quality control for high altitude data.

The direct effect of the smoothed ionospheric correction is demonstrated in Figure 2.6 which shows the ionospheric corrected bending angles of 206 CHAMP profiles (measured on Jan. 1, 2003) including all outliers. For comparison, the climatological mean and standard deviation derived from CIRA86 [*Barnett and Corney*, 1985; *Fleming et al.*, 1988] is displayed. After the standard ionospheric correction hardly any profile contains physical information (positive bending angles) above 60 km (Figure 2.6, left panel). After smoothed correction the mean maximum altitude could be raised (Figure 2.6, right panel) which is favorable for the subsequent statistical optimization process.

The effect of smoothed ionospheric correction on retrieved dry temperature profiles of the same CHAMP sample (Jan. 1, 2003) is demonstrated in Figure 2.7. Dry temperatures retrieved with standard (upper left panel) and smoothed ionospheric correction (upper middle and right panels, 1 km and 2 km boxcar filter, respectively) compared to operational ECMWF analyses are displayed showing mean (bold line) and standard deviation (thin lines). In exchange for a raised variance above 30 km, the variance in the core RO altitude region between 15 and 25 km is reduced and the delivery rate is raised by ~4 % (167 instead of 161 profiles) by applying the 1 km filter. Further broadening of the filter width to 2 km results in slightly further reduction of variance in the core region and yields one additional profile, but also modifies the bias above 30 km. Since IGAM is aiming to produce high-quality data for climate applications where bias-free data is most important, we decided to accept the slightly more pronounced variance of the 1 km filter in order to avoid high-altitude biases.

Besides the positive aspect of reducing the error standard deviation in the core RO altitude region, the main effect is a trade-off between a high delivery rate (including the danger to include some low-quality measurements resulting in high variance and potential biases at high altitudes) and restrictive quality control. If a high delivery rate is not necessary, broader filters combined with stricter implicit quality control could be applied which would result in combining low variance in the core-region with low variance at high altitudes. However, filters broader than 1 km are not guaranteed to operate bias-free since they interfere with scales covered by the ionospheric correction and are not recommended.

Figure 2.7 (lower panels) shows the results of the same ionosphere correction schemes as in the upper panels and stricter implicit quality control of high altitude data resulting in ~25 % – 60 % (depending on the filter width) of the profiles sorted out above 25 km (quality flag 2) as indicated by the gray line. The implicit quality control only depends on technical quality of the data, but not on atmospheric background information about the atmosphere. It reduces the high altitude variance but does not modify the mean of the data. If stricter explicit quality control would be applied, a bias towards the reference dataset (ECMWF) could not be avoided.

The current CCRv2 schemes apply moderately smoothed ionospheric correction (1 km boxcar) without strict implicit high altitude quality control (Figure 2.7, upper middle panel). In applications not crucially dependent on high data density the same scheme with stricter implicit quality control (Figure 2.7, lower middle panel) is recommended.

2.2.6 Observation Error Estimation and Background Bias Correction

The statistical optimization process (Section 2.1.2.1) needs an estimate about the error characteristics of the observed data. This estimate is obtained by analyzing the altitudinal variance of the observed ionosphere corrected bending angle profile at high altitudes where the neutral atmospheric contribution to the signal is negligible. The appropriate height interval for estimating the observation error should be above the altitude where the expected bending angle noise of CHAMP equals the bending angle caused by the neutral atmosphere. For a first

estimate of the lowermost altitude usable for this purpose, CHAMP-derived bending angle profiles are compared with profiles derived from a climatology in Figure 2.6 which shows the mean and the standard deviation of 1130 bending angle profiles derived from CIRA86 (globally equal distributed). Above ~70 km the neutral atmospheric bending angles amount less than $\sim 10^{-6}$ rad which is in the order of the lower bound for the CHAMP error. Our original interval for observation error estimation is 70–80 km (see section 2.1.2.1 e). However, from CHAMP data hardly ever useful ionosphere corrected bending angles can be derived above 70 km which makes it necessary to set the interval as low as possible. Extensive studies showed that an interval above 65 km still yields good results and drastically increases the number of profiles that allow observation error estimation. For CHAMP data we use the error estimation interval between 65 and 80 km. The main criteria to accept the error estimation is an upper (50 μrad) and lower (0.5 μrad) threshold value, a consistency check comparing the variance to the bias relative to the background data, and a minimum of 25 data points available for estimation (see also section 2.2.4.1).

A further effect of the low quality of CHAMP high-altitude data is that on the background bias correction scheme as described in section 2.1.2.1 e. The intervals found for searching and adjusting the background data in an optimal way (45–65 km and 55–75 km) are not applicable to CHAMP data and had to be displaced to 35–55 km and 45–65 km, respectively. This displacement has a negative effect on the performance of the background bias correction scheme and no significant performance-gain compared to the usage of non-modified colocated profiles from climatology can be reached. However, the scheme still easily competes with other schemes using climatological background data [*Gobiet et al.*, 2004a] and has a good potential for RO missions using further developed RO receivers yielding better data at high altitudes like GRACE, GRAS/MetOp, or COSMIC.

2.2.7 ECMWF Background Data

Depending on the data user's intention it might be desirable to have available RO-data independent from NWP-models which is the case for the IGAM/MSIS retrieval scheme described so far. On the other hand, it might be favorable to use the best-possible temperature data available and to accept a slight dependency from NWP-models at high altitudes. For these users, we developed a second retrieval mode using ECMWF operational analyses as background information in the statistical optimization process. The scheme is very similar to the IGAM/MSIS scheme with the main difference of making direct use of ECMWF data without any modification (no background bias correction). The main characteristics of the retrieved atmospheric profiles with respect to both schemes are discussed in section 2.2.9, the main technical features of the retrieval schemes are listed in Table 2.4.

2.2.8 CCR Version 2 Overview

In order to give an overview, the main features of the two CCRv2 retrieval schemes are listed in Table 2.4. It provides some additional information on aspects of the retrieval scheme not discussed in this report since they cannot be regarded as advancements but rather as state-of-the-art procedures and are equally implemented in the EGOPS software [*Kirchengast et al.*, 2002] which is the basis for CCR.

Table 2.4. Overview of the IGAM CHAMP-RO retrieval schemes (EGOPS/CCR Version 2, June 2004).

	IGAM/MSIS	IGAM/ECMWF
Outlier Rejection and Smoothing	"3σ" outlier rejection on phase delays and smoothing using regularization (third order norm, regularization parameter = $10^{(\text{sampling rate}/10)}$ [*Syndergaard*, 1999]).	Like IGAM/MSIS
Bending Angle Retrieval	Geometric optics retrieval.	Like IGAM/MSIS
Ionospheric Correction	Linear combination of bending angles [*Vorob'ev and Krasil'nikova*, 1994]. Correction is applied to low-pass filtered bending angles (1 km sliding average), L1 high-pass contribution is added after correction [*Hocke et al*., 2003]. L2 bending angles < 15 km derived via L1-L2 extrapolation.	Like IGAM/MSIS
Bending Angle Initialization	Statistical optimization of bending angles 30 – 120 km. Vertical correlated background (corr. length L = 6 km) and observation (L = 1 km) errors. Obs. error estimated from observed profile >65 km. Background error: 15 %. Backg. information: MSISE-90 [*Hedin*, 1991] best fit-profile, bias corrected [*Gobiet and Kirchengast*, 2004].	Like IGAM/MSIS, but co-located bending angle profile derived from ECMWF operational analysis (T42L60) as backg. information (above ~60 km: MSISE-90). No further pre-processing.
Hydrostat. Integral Init.	At 120 km: pressure = pressure(MSISE-90).	Like IGAM/MSIS
Humidity Retrieval	Optional: 1DVar using ECMWF analyses as background.	Like IGAM/MSIS
Quality Control	Refractivity 5 – 35 km: $\Delta N/N < 10$ %; Temperature 8 – 25 km: $\Delta T < 20$ K. Reference: ECMWF operational analysis (T42L60).	Like IGAM/MSIS

2.2.9 Performance of CCRv2

2.2.9.1 Overall Performance

The overall performance of the IGAM/ECMWF retrieval is demonstrated in Figure 2.8 which shows the CHAMP retrieval results of 20 days of observations in Sept. and Oct. 2002 (details are published in the CHAMPCLIM validation report [*Gobiet et al.*, 2004a]). The dry temperatures are virtually unbiased against ECMWF analyses (left panel) below 30 km and slightly warm biased above that. One interesting feature of this comparison is that though ECMWF analyses were used as background information in the retrieval, the results differ from

ECMWF at high altitudes where the influence of the background is strongest. This indicates that the retrieval is not dominated by the background at least down to 40 km which makes it a potentially valuable observation up to this altitude. The discrepancy between IGAM/ECMWF temperatures and ECMWF analysis temperatures is not necessarily a deficiency of the observation since ECMWF is known to be cold biased above 30 km since it is hardly constrained by observations at this altitude.

This interpretation is supported by the comparison of CHAMP-results with coinciding data from the MIPAS instrument on ENVISAT [*v. Clarmann et al*., 2003a; 2003b] shown in the right panel of Figure 2.8. The most interesting result of this comparison is the fact that the IGAM/ECMWF temperatures are not warm biased against MIPAS temperatures above 30 km as it is the case in the comparisons with ECMWF analyses. Since MIPAS is virtually independent from ECMWF this is a strong indication for the reliability of the CHAMP derived temperatures of the IGAM/ECMWF-scheme above 30 km as well.

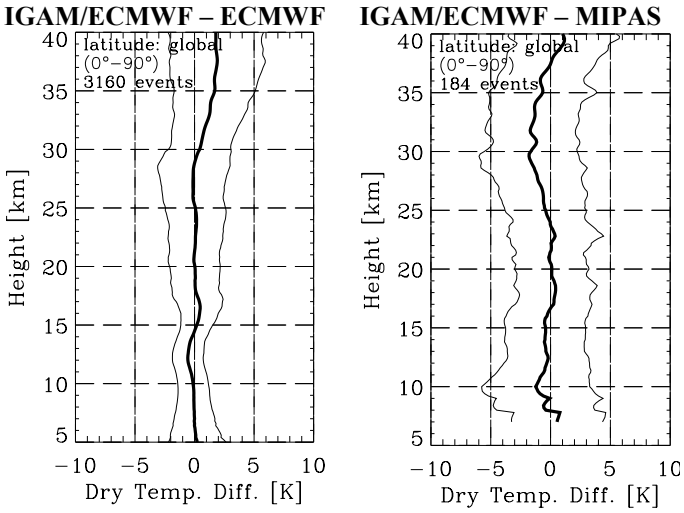


Figure 2.8. CHAMP dry temperature error statistics. Period Sep. 8, 12, 14 – 28, Oct. 11 – 13 2002; 20 days. Bold line: bias. Thin line: standard deviation. Left panel: CHAMP (IGAM/ECMWF retrieval) compared to ECMWF analyses. Right panel: CHAMP (IGAM/ECMWF retrieval) compared to coinciding MIPAS events. Coincidence criteria: 300 km and 3 h max. spatial and temporal displacement, respectively.

2.2.9.2 Effects Due to Background Information Source

The direct comparison of the two CCRv2 retrieval schemes (IGAM/MSIS and IGAM/ECMWF) in Figure 2.9 drastically demonstrates the effect of different sources of background information. On refractivity level, the background influence is limited to higher altitudes (< 1 % below 35 km) than on temperature level. The hydrostatic integration used to derive pressure and temperature from refractivity (equation 2.3) causes a downward propagation of information resulting in lower-reaching influence of background information. For this reason and due to the high reliability of the IGAM/ECMWF retrieval it is recommended to use the IGAM/MSIS retrieval results only on refractivity level and below 35 km. If strict independence from ECMWF analyses is not required and/or if temperature data is needed we recommend to use the IGAM/ECMWF retrieval results.

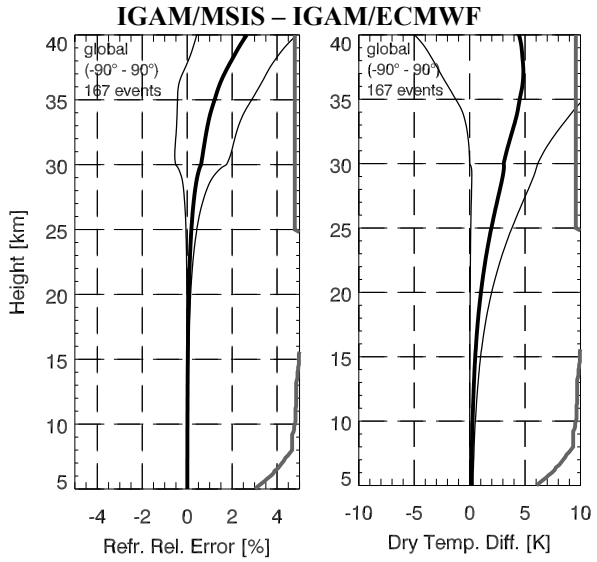


Figure 2.9. Intercomparison of the two CCRv2 schemes IGAM/MSIS and IGAM/ECMWF (Jan. 1, 2003, 167 profiles). Bold line: bias. Thin line: standard deviation. Grey line: percent number of data points contributing to each altitude's error statistic (10 K corresponds to 100 %). Left panel: IGAM/MSIS – IGAM/ECMWF refractivity error statistics. Right panel: IGAM/MSIS – IGAM/ECMWF dry temperature error statistics.

2.2.9.3 Effects Due to Statistical Optimization Methodology

As described in section 2.1.2.1 e, the statistical optimization schemes implemented in CCRv2 are based on the aim of optimally combining observed and background data under consideration of their error characteristics. This combination is accomplished in an early stage of the retrieval process to prevent erroneous data from disturbing the performance of the retrieval algorithm and to prevent error propagation. Since the degree of reliability of background information about the ionosphere is much smaller than of the background information about the neutral atmosphere, background information is integrated directly after ionospheric correction at the bending angle level when no ionospheric background information is required any more. After statistical optimization, the data can be regarded as our best knowledge about the state of the atmosphere. From this stage the IGAM retrieval schemes perform the inverse Abel transform (equation 2.1) to derive refractivity and the hydrostatic integral (equation 2.3) to derive pressure and temperature without further inclusion of background information.

This is a major difference to most other RO retrieval schemes which generally, in addition to statistical optimization, initialize the hydrostatic integral at a height of ~40–50 km with some pressure value derived from NWP short range forecasts or analyses temperatures [e.g., *Kursinski et al.*, 1997; *Wickert et al.*, 2004; *Hajj et al.*, 2004]. This approach has some disadvantages: It makes it impossible to trace the error characteristics of the retrieval results directly from the error characteristics of the measurement and the background information, it completely disregards any information in the refractivity profile above the height of hydrostatic integral initialization, it produces physically inconsistent refractivity and temperature profiles, and it gives too much weight to background information, i.e. potentially produces biased results (see also discussion in section 4.1 on this topic).

As an example for the effect of these two different high altitude initialization strategies, Figure 2.10 depicts a comparison of the CCRv2 schemes with the operational CHAMP retrieval products provided by GFZ. The GFZ retrieval scheme is based on statistical optimization of bending angles using the MSISE-90 climatology as background information

similar to the 'basic IGAM scheme' (Section 2.1.2.1 d) and additionally initializes the hydrostatic integral at 43 km with data from ECMWF analyses (J. Wickert, GFZ Potsdam, personal communication, 2004). The left two panels (IGAM/MSIS – GFZ) demonstrate the refractivity-temperature inconsistency: While refractivity shows no bias between the IGAM/MSIS and the GFZ retrievals, dry temperatures are severely biased at high altitudes though both retrieval schemes rely on the same formulas to relate atmospheric refractivity to temperature (equations 2.2 and 2.3). The two right panels show the IGAM/ECMWF – GFZ comparison. The temperatures are slightly biased above 30 km in a qualitatively similar way as the IGAM/ECMWF temperatures relative to ECMWF analyses (Figure 2.8, left panel). This means that the GFZ retrieved temperatures are much closer attached to ECMWF which increases the risk of introducing eventual high altitude biases from ECMWF into the retrieval results.

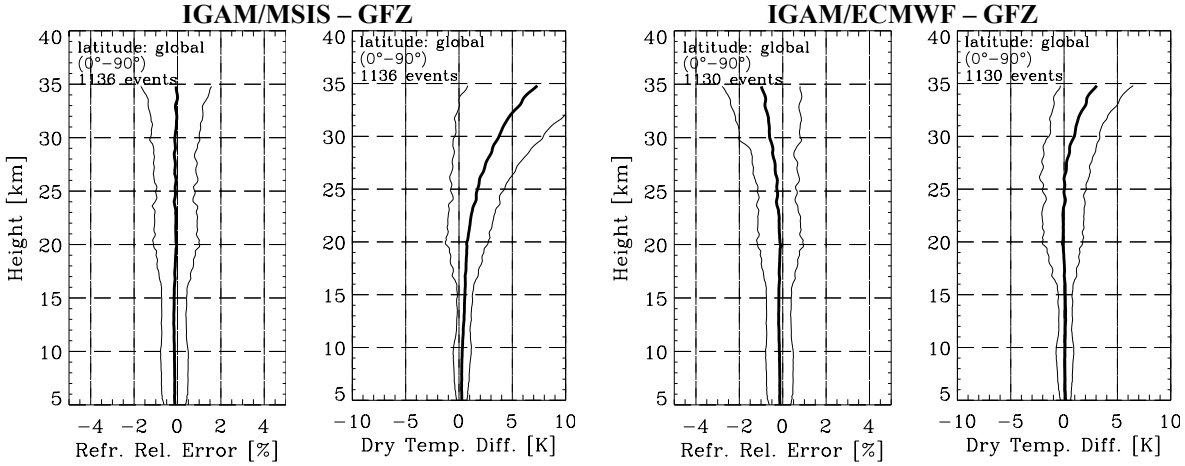


Figure 2.10. Comparison of the two CCRv2 schemes IGAM/MSIS and IGAM/ECMWF with the GFZ operational retrieval (Jan. 1 – 7, 2003, 1136 profiles). Bold line: bias. Thin line: standard deviation. Left panels: IGAM/MSIS – GFZ refractivity and dry temperature error statistics. Right panels: IGAM/ECMWF – GFZ refractivity and dry temperature error statistics.

2.2.10 Conclusions

The preceding subsections described the development of a stable retrieval scheme for CHAMP data. Quality control mechanisms were developed enabling the IGAM retrieval schemes to operationally cope with CHAMP data for the first time. A robust ionospheric correction scheme based on smoothed bending angles and impact parameters (1 km boxcar filter) was implemented yielding a higher delivery rate and less variance in the retrieval results in the 15–25 km altitude region without removing effects from small-scale structures in the neutral atmosphere. Due to the low quality of CHAMP high-altitude data, the observation error estimation and background bias correction procedures had to be adjusted resulting in reduced quality of the 'enhanced IGAM scheme' using climatological background information (IGAM/MSIS scheme). Still this scheme, which has the advantage to be independent from NWP background data, shows a performance similar to other state-of-the-art retrieval schemes and has a big potential for RO missions using further developed RO receivers yielding better data at high altitudes like GRACE, GRAS/MetOp, or COSMIC.

Since it is acceptable for many applications to use slightly NWP-dependent data at high altitudes, we developed a second retrieval mode using ECMWF operational analyses as background information (IGAM/ECMWF scheme). It shows relative independency from the

background up to at least 40 km whilst the retrieval performance, especially for temperature profiles, is drastically improved.

It is recommended to use IGAM/MSIS retrieval results on refractivity level and below 35 km if strict independence from ECMWF analyses is required. If strict independence is not required and/or if temperature data is needed we recommend to use results from the IGAM/ECMWF retrieval scheme.

(intentionally left blank, back page if double-sided print)

3 Advancement in Assessing Radio Occultation Retrieval Errors in the Troposphere

Executive Summary. We investigated the sensitivity of atmospheric profiles retrieved from Global Navigation Satellite System (GNSS) radio occultation data to atmospheric horizontal variability errors. First, the errors in a quasi-realistic horizontally variable atmosphere relative to errors in a spherically symmetric atmosphere were quantified based on an ensemble of about 300 occultation events. This investigation was based on simulated data using a representative European Centre for Medium-Range Weather Forecasts (ECMWF) T511L60 analysis field with and without horizontal variability. Biases and standard deviations are, below 20 km, significantly smaller under a spherical symmetry assumption than corresponding errors in an atmosphere with horizontal variability. The differences are most pronounced below ~7 km height. Second, we assessed the relevance of either assuming the "true" profile vertically at a mean event location (the common practice) or along the actual 3D tangent point trajectory. Standard deviation and bias errors decrease significantly if the data are exploited along the tangent point trajectory. Third, the sensitivity of retrieval products to the angle-of-incidence of occultation rays relative to the boresight direction of the receiving antenna (aligned with the orbit plane of the Low Earth Orbit satellite) was analyzed based on the same ensemble of events for three different angle-of-incidence classes (0–10 deg, 20–30 deg, 40–50 deg; ensembles of about 100 events in each class). Below about 7 km, most errors were found to increase with increasing angle of incidence. Dry temperature biases between 7 km and 20 km exhibit no relevant increase with increasing angle of incidence, which is favorable regarding the climate monitoring utility of the data.

3.1 Introduction

The EGOPS4 software tool (End-to-end GNSS Occultation Performance Simulator, version 4) was used to generate simulated phase measurements and retrievals of the observables bending angle, refractivity, total air pressure, geopotential height, and (dry) temperature. For a detailed description of EGOPS see *Kirchengast* [1998] and *Kirchengast et al.* [2002]. Section 3.2 gives an overview of the experimental setup. Results on the sensitivity to horizontal variability are presented in section 3.3 and the relevance of the geometry of reference profiles is discussed in section 3.4. The results on sensitivity to angle-of-incidence are shown in section 3.5. Conclusions and an outlook are provided in section 3.6.

3.2 Experimental Setup

3.2.1 Geometry

We assumed a full constellation of 24 Global Positioning System (GPS) satellites as transmitters and a GRAS (GNSS Receiver for Atmospheric Sounding) sensor onboard the METOP satellite with a nominal orbit altitude of ~830 km [*Silvestrin et al.*, 2000]. With such a constellation, about 500 rising and setting occultations per day can be obtained. We simulated measurements over a 24 hour period on September 15, 2001, the date of the ECMWF analysis field used in the forward modeling.

We collected occultation events in three different azimuth sectors relative to the boresight direction of the receiving antenna. A schematic illustration of this division into several

"angle-of-incidence" sectors is given in the left panel of Figure 3.1, while Table 3.1 summarizes the simulation design in terms of numbers of events simulated per sub-sector defined. With restriction to the described azimuth sectors we obtained a total of 306 occultation events during the selected 24 hour period. The geographic distribution is shown in the right panel of Figure 3.1. We obtained uniform distribution in latitude as well as equal density over oceans and over continents in each sector.

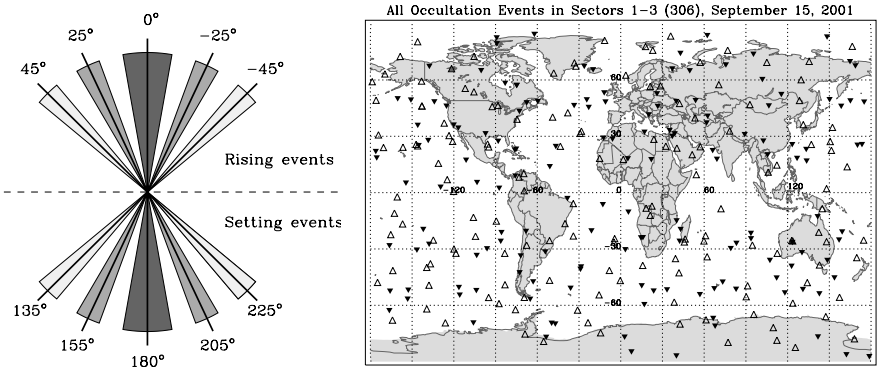


Figure 3.1. Left panel: Schematic illustration of azimuth sectors used in the study: sector 1 (dark gray), sector 2 (medium gray), and sector 3 (light gray). Right panel: Locations of occultation events during one day in all three azimuth sectors. Upright open triangles denote rising occultations while upside-down filled triangles denote setting occultations.

Table 3.1. Number of occultation events in the three azimuth sectors used in this study.

	Sector 1	**Sector 2**	**Sector 3**	No. of events
Rising events	0° to –10°	–20° to –30°	–40° to –50°	76
	0° to +10°	+20° to +30°	+40° to +50°	76
Setting events	170° to 180°	150° to 160°	130° to 140°	77
	180° to 190°	200° to 210°	220° to 230°	77
No. of events	105	114	87	**306**

3.2.2 Forward Modeling

High resolution (T511L60) analysis fields from the European Centre for Medium-Range Weather Forecasts (ECMWF) for September 15, 2001, 12 UT, were used to generate quasi-realistic atmospheric phase delays. The horizontal resolution (T511) corresponds to 512 x 1024 points in latitude and longitude, respectively, and thus furnishes about eight or more grid points within the typical horizontal resolution of an occultation event of ~300 km [e.g., *Kursinski et al.*, 1997]. This dense sampling is important to have a sufficient representation of horizontal variability errors in occultation measurements. In the vertical, 60 levels (L60, hybrid pressure coordinates) extend from the surface to 0.1 hPa, being most closely spaced in the troposphere, which represents good vertical resolution. In order to illustrate the resolution of the T511L60 fields utilized, typical slices of temperature and specific humidity are displayed in Figure 3.2. The MSIS climatological model [*Hedin*, 1991] was used, with a smooth transition, above the vertical domain of the ECMWF analysis field (from ~60 km upwards).

As we focused on the troposphere, we made the reasonable assumption that ionospheric residual errors can be neglected below 20 km [*Steiner et al.*, 1999]. Forward modeling was

thus employed without the ionosphere, which corresponds to considerable savings in computational expenses. We performed high-precision 3D ray tracing with sub-millimeter accuracy and a sampling rate of 10 Hz for all forward modeled events through the ECMWF analysis (refractivity) field. In order to address the effects of horizontal variability, two separate ensembles of 306 events were forward modeled: One employing the analysis field with its 3D structure as is, the other by artificially enforcing spherical symmetry for each event. The latter case was obtained by using the refractivity profile at the mean tangent point of an occultation event (estimated between 12 and 15 km altitude) over the entire domain probed. As in the real atmosphere, occultation events over oceans at low latitudes occasionally failed to penetrate the lowest ~2–4 km of the troposphere, if the ray tracer encountered super-refraction or closely such conditions.

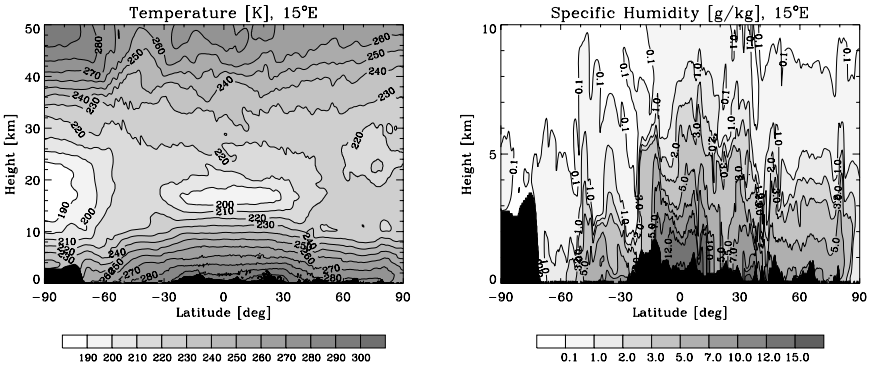


Figure 3.2. Latitude-versus-Height slices at 15° eastern longitude; Sept. 15, 2001, 12 UT (T511L60 ECMWF analysis fields). Left panel: temperature [K], right panel (different height range): specific humidity [g/kg].

3.2.3 Observation System Modeling and Retrieval Processing

Realistic errors (including error sources like orbit uncertainties, receiver noise, local multipath errors and clock errors) have been superimposed on the obtained simulated phase measurements. For this receiving system simulation, we used (conservatively) the specifications and error characteristics of the GRAS instrument [e.g., *Silvestrin et al*., 2000].

Regarding retrieval processing, we applied a geometric optics bending angle retrieval scheme. The core of this algorithm, transforming phase delays to bending angles, is the algorithm described by *Syndergaard* [1999], which was enhanced to include inverse covariance weighted statistical optimization (with prior best-fit a priori profile search) as described by Gobiet and Kirchengast (2002). Since the forward modeling has been performed without an ionosphere, ionospheric correction was omitted. Refractivity profiles have been computed using a standard Abel transform retrieval employing the algorithm of *Syndergaard* [1999]. Profiles of total air pressure and of temperature have been obtained using a standard dry air retrieval algorithm as again developed by *Syndergaard* [1999]. Geopotential height profiles where obtained by converting geometrical heights z of pressure levels via the standard relation $dZ = (g(z,\varphi)/g_0)dz$ [e.g., *Salby*, 1996] to geopotential heights Z, where $g(z,\varphi)$ invokes the international gravity formula [e.g., *Landolt-Börnstein*, 1984] and $g_0 = 9.80665$ ms^{-2} is the standard acceleration of gravity. We did not undertake to separately analyze temperature and humidity. For this baseline analysis of horizontal variability errors we decided to inspect variables such as refractivity and dry temperature, which do not require prior information.

3.2.4 Reference Profiles

All retrieved profiles have been differenced against the corresponding "true" ECMWF vertical profiles at the mean tangent point locations. The differences incurred by either assuming the reference profile vertically at a mean event location (the common practice) or more precisely along the estimated 3D tangent point trajectory has been assessed as well.

We compare to "true" (dry) geopotential height and dry temperature profiles from the ECMWF fields. This implies that the temperature profiles have an increasing moisture effect below 10 km. The dependence of dry temperature on actual temperature and humidity is accurately known, however, so that one can always determine the influence of moisture if desired.

3.3 Sensitivity to Horizontal Variability

While the analyses have been carried out for all described parameters we will focus here on discussing geopotential height and temperature errors only. In Figure 3.3, the results for the "real" atmosphere with horizontal variability (top panels) are compared with the results for the artificial spherically symmetric atmosphere (middle panels). Furthermore, we computed errors by comparing with the "true" profiles extracted along the 3D tangent point trajectories (bottom panels).

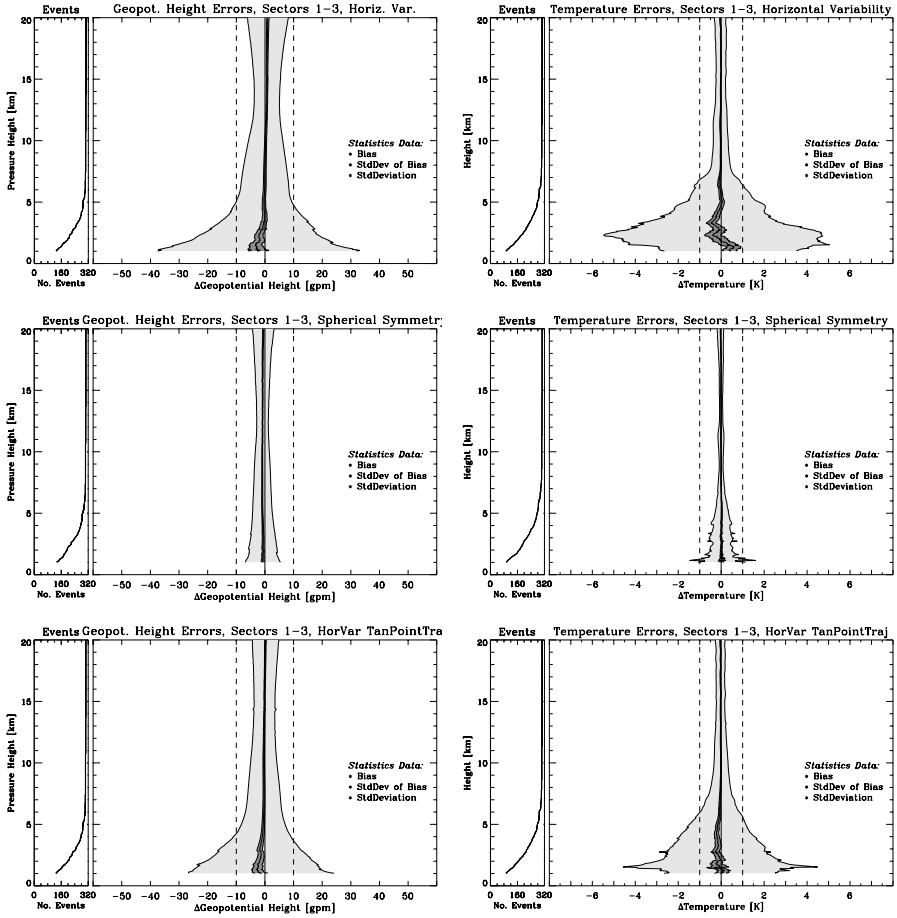


Figure 3.3. Geopotential height (left) and Temperature (right) error statistics for the ensemble of all 306 occultation events. Top panels: atmosphere with horizontal variability; middle panels: atmosphere with spherical symmetry applied; bottom panels: horizontal variability with profile along 3D tangent point trajectory as reference. Sub-panels: number of events entering the statistics at a given height versus height.

The statistical results for the full ensemble of 306 events are illustrated in each panel. The bias errors and standard deviations have been empirically estimated by differencing of retrieved profiles with co-located reference profiles. The gradual decrease in the number of events towards lower tropospheric levels (small left-hand-side subpanels) is due to the different minimum heights reached by individual events.

3.3.1 Geopotential Height Errors

Errors in the geopotential height of pressure surfaces are shown in the left panels of Figure 3.3 as function of pressure height z_p (defined as $z_p = -7 \cdot \ln(p[\text{hPa}]/1013.25)$, which is closely aligned with geometrical height z. We note that geopotential height errors mirror pressure errors (not shown), i.e., positive biases in geopotential height correspond to negative biases in pressure [see, e.g., *Syndergaard*, 1999].

In the horizontally variable atmosphere (top panel) standard deviations above 5 km are smaller than 10 gpm (geopotential meters) while reaching 35 gpm at 1 km. Biases are generally negligible and bias values of more than 1 gpm are only found below 3 km. With reference profiles along the 3D tangent point trajectory (bottom panel), standard deviations remain smaller than 10 gpm down to 4 km and do not exceed 25 gpm at 1 km; biases are similarly small as in the top panel case. Under spherical symmetry (middle panel), biases never exceed 1 gpm, and standard deviations reach 6 gpm only at the lowest height levels.

3.3.2 Temperature Errors

The dry temperature errors are depicted in the right panels of Figure 3.3. In the scenario with horizontal variability (top panel), the errors are significantly larger than the corresponding errors under spherical symmetry.

Under horizontal variability, the bias reaches –0.5 K near 3.3 km and +0.6 K near 1.4 km, respectively, while standard deviations exceed 5 K near 2.3 km height. Comparing with reference profiles along the 3D tangent point trajectory (bottom panel) reduces the maximum bias by a factor of ~2 and the height where a standard deviation of 1 K is reached is lowered by about 1 km. Under spherical symmetry, the bias is smaller than 0.1 K everywhere, standard deviations remain smaller than 1 K with the exception of a small height interval below 1.5 km near the lower bound. Between 7 km and 20 km there is essentially no temperature bias in all scenarios (i.e., always smaller than 0.1 K).

3.4 Dependence on Geometry of Reference Profiles

Figure 3.4 (left panel) illustrates that differences between "true" vertical profiles at mean tangent point locations and "true" ones along actual 3D tangent point trajectories are fairly small at tropopause/lower stratosphere heights, since the EGOPS mean location estimate is designed to fit best near 12-15 km. Below 7 km, however, the differences are comparable to the errors estimated under horizontal variability (top left panel in Figure 3.3). This implies that the geometrical mis-alignment of the actual tangent point trajectory with the mean-vertical contributes as an important source to horizontal variability errors. Overall, the results show that the performance in the horizontally variable troposphere is markedly improved if measured against the actual tangent point trajectory.

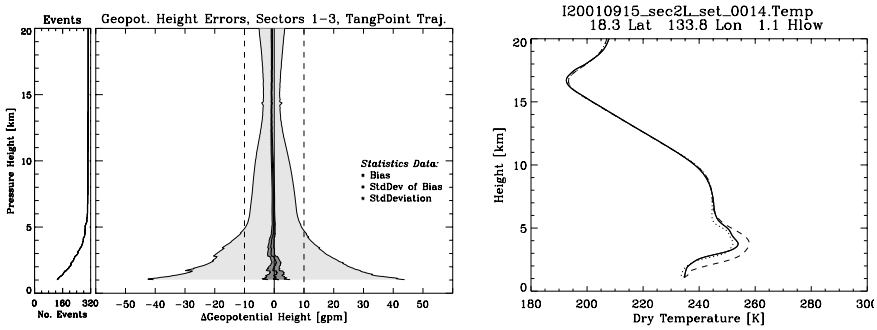


Figure 3.4. Left panel: Differences between vertical reference profiles at the mean tangent point location and profiles along the 3D tangent point trajectory. Sub-panel: number of events versus height. Right panel: Typical dry temperature profile under high horizontal variability: retrieved profile (solid), reference profile at mean tangent point location (dashed), and reference profile along the 3D tangent point trajectory (dotted), respectively.

It should be noted that the exact trajectory cannot be determined without performing ray tracing. To a high degree of accuracy, however, it can be estimated from observed data (GNSS and LEO satellite positions and bending angles). Future work will thus investigate by how much the applicable standard deviations and bias errors decrease if the data are exploited along a tangent point trajectory deduced purely from observed data.
The right panel of Figure 3.4 shows a typical temperature retrieval in an area with high horizontal variability. In this cases, the two types of reference temperature profiles differ by several Kelvins and the retrieved profile is clearly much closer to the 3D tangent point reference profile than to the vertical one.

3.5 Sensitivity to the Angle-of-Incidence

In this section the sensitivity of retrieval products to the angle-of-incidence of occultation rays relative to the boresight direction of the receiving antenna (aligned with the LEO orbit plane) is analyzed. Error analyses have been performed for each azimuth sector defined in section 3.2.1 (ensembles of 105, 114, and 87 events, respectively), for every atmospheric parameter under study. We show again results for geopotential height and temperature errors only.

Events in sector 1 are associated with almost co-planar GNSS and LEO satellites, which should lead to the most-vertical and best-quality occultation events. We would thus expect an increase of errors with increasing angle of incidence. Figure 3.5 shows the results for sector 1 in the top panels, for sector 2 in the middle panels, and for sector 3 in the bottom panels, respectively.

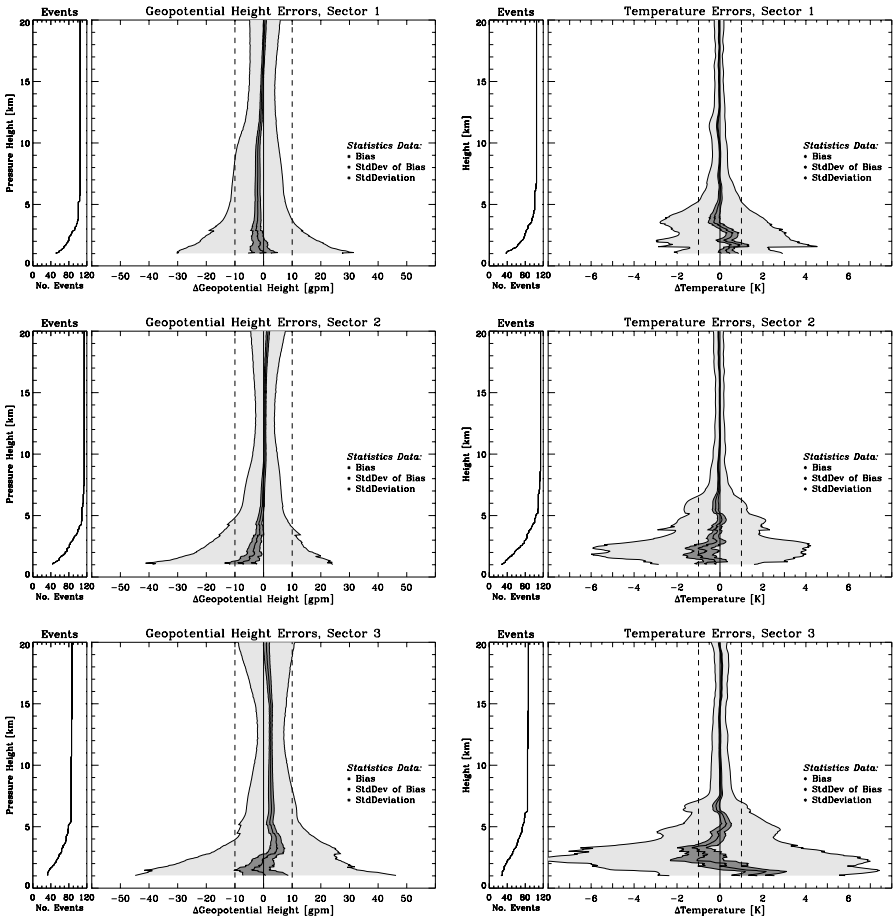


Figure 3.5. Geopotential height (left) and Temperature (right) error statistics for the different azimuth sectors. Top panels: sector 1; middle panels: sector 2; bottom panels: sector 3. Sub-panels: number of events versus height.

3.5.1 Geopotential Height Errors

Errors in the geopotential height of pressure surfaces are shown in the left panels of Figure 3.5, again as function of pressure height (cf. section 3.3.1). While standard deviations generally increase with increasing angle of incidence, the behavior of the bias profiles is more complex and interesting. Between ~3 and ~18 km we find the smallest biases in sector 2, events in this sector are close to bias-free between 5 and 18 km altitude (bias < 1 gpm).

In sector 1, on the one hand, there is an almost constant negative bias of ~2 gpm between ~2 and ~10 km, while in sector 3, on the other hand, we find an almost constant positive bias of ~2.5 gpm between ~5 and ~15 km. In the ensemble of all events (see Sect. 3.1) this adds (together with a smaller number of events in sector 3) to very small biases down to about 3 km. Though biases as small as up to 2.5 gpm are not a real concern, we are currently investigating these datasets closer in order to understand the causes of this subtle behavior.

3.5.2 Temperature Errors

The dry temperature errors are shown in the right panels of Figure 3.5. Their general behavior is more in line with expectations than that of geopotential height errors.
Below 7 km, biases and standard deviations increase significantly with increasing angle of incidence. Standard deviations in sector 1 (azimuth 0-10°) have maximum values of 3.7 K,

those in sector 2 (azimuth 20-30°) reach 5.0 K, and those in sector 3 (azimuth 40-50°) even 8.0 K. Biases in sectors 1 and 2 remain smaller than 0.9 K and 1.1 K, respectively, while they reach about 2 K in sector 3.

Biases between 7 km and 20 km are smaller than ~0.1 K and exhibit no relevant increase with increasing angle of incidence. Standard deviations in the same height interval remain smaller than 0.5 K (except close to 7 km), those in sector 3 are only slightly larger than respective errors in sector 1 and sector 2.

3.6 Summary and Conclusions

We investigated the sensitivity of atmospheric profiles retrieved from GNSS radio occultation data to atmospheric horizontal variability errors, based on an ensemble of about 300 simulated events. Biases and standard deviations are, below 20 km, significantly smaller under a spherical symmetry assumption than corresponding errors in a quasi-realistic atmosphere with horizontal variability. The differences are most pronounced below ~7 km height. Temperature standard deviations, for example, remain smaller than 1 K in a spherically symmetric atmosphere, while they reach values of about 5 K in the horizontally variable atmosphere. This confirms earlier results based on a more simplified estimation by *Kursinski et al*., [1997], that horizontal variability is an important error source in the troposphere. Dry temperature profiles between 7 km and 20 km were found to be essentially bias-free in both the horizontal variability and spherical-symmetry scenarios (biases smaller than 0.1 K), which confirms the unique climate monitoring utility of GNSS occultation data.

A significant part of the total error below ~7 km can be attributed to adopting reference profiles vertically at mean tangent point locations instead of extracting them along actual 3D tangent point trajectories through the troposphere. Future work will investigate by how much the standard deviation and bias errors decrease if a tangent point trajectory deduced purely from observed data is used instead.

Below about 7 km most errors were found to increase with increasing angle of incidence. This is in line with the hypothesis that larger angles of incidence lead to more sensitivity to horizontal variability. Geopotential height biases in the 20–30 deg azimuth sector above ~3 km, however, are smaller than corresponding biases in the 0–10 deg sector, which merits further investigation. In general, the sensitivity of bias errors to increases of the angle of incidence has been found to be relatively small, which is favorable regarding the climate monitoring utility of the data. For example, dry temperature biases between 7 km and 20 km exhibit no relevant increase with increasing angle of incidence. Current cautionary approaches restricting the events used in climate studies to small angles of incidence (such as < 15 deg; [*Steiner et al*., 2001]) appear thus to be overly conservative and can be safely relaxed to using the data at least up to 30 deg.

4 Advancement in Error Characterization of Radio Occultation Retrievals

4.1 Empirical Error Analysis for GNSS Radio Occultation Data based on End-to-end Simulations

Executive Summary. Radio occultation (RO) observations using the Global Navigation Satellite System (GNSS) globally provide high quality atmospheric data which can support the advancement of climate monitoring and modeling as well as the improvement of numerical weather prediction. In order to make optimal use of the data, e.g., via data assimilation systems, the characterization of measurement errors is of importance. Within this context we present results of an empirical error analysis based on quasi-realistically simulated GNSS RO data. The study is based on an end-to-end forward-inverse simulation involving (i) modeling of the neutral atmosphere and ionosphere, (ii) simulation of RO observations, (iii) forward modeling of excess phase observables including realistic observation system error modeling, and (iv) retrieval of atmospheric parameters. Occultation observations were simulated for one day from which an ensemble of 300 occultation events was chosen, with 100 events in each of three latitude bands (low, middle, high). Phase path profiles were computed showing a realistic rms error of the ionosphere corrected phase paths of 2–3 mm at meso- and stratospheric heights at 10 Hz sampling rate. Atmospheric profiles were retrieved by applying a dry air retrieval in the stratosphere and an optimal estimation retrieval in the troposphere. The retrieved profiles were referenced to the "true" co-located ones of the analysis field of the European Centre for Medium-range Weather Forecast (ECMWF) used as atmospheric model. We empirically estimated bias profiles and covariance matrices (standard deviations and correlation functions) for the retrieval products bending angle, refractivity, pressure, geopotential height, temperature, and specific humidity. Specific results include: Refractivity exhibits a relative standard deviation of 0.1–0.75% and a relative bias of <0.1% at 5–40 km height. Temperature shows a standard deviation of 0.2–1 K at 3–31 km height and a bias of <0.1–0.5 K below 33 km and of <0.1 K below 20 km. Simple analytical error covariance formulations are presented for refractivity, as deduced from the empirically estimated covariance matrices, and their applicability to the other retrieval products is discussed. The reasonably realistic error estimates presented are a valuable basis for further retrieval algorithm improvements and for proper specification of observational errors in data assimilation systems.

4.1.1 Introduction

The GNSS RO technique is based on active limb sounding using a satellite-to-satellite radio link to probe the Earth's atmosphere. The GNSS-transmitted radio signals are influenced by the atmospheric and ionospheric refractivity field during their propagation to a receiver on a Low Earth Orbit (LEO) satellite. Observed excess phases (relative to propagation in vacuum) are the basis for high quality retrievals of atmospheric variables such as bending angle, refractivity, density, pressure, geopotential height, temperature, and water vapor as well as of ionospheric electron density [e.g., *Melbourne et al*., 1994; *Kursinski et al.*, 1997].

Originating from planetary radio science [e.g., *Fjeldbo and Eshleman*, 1965; 1969] the first practical application of the RO method to the terrestrial atmosphere was successfully demonstrated with the U.S. Global Positioning System Meteorology (GPS/MET) experiment [*Ware et al*., 1996]. The strengths of the technique include all-weather capability, long-term stability, global coverage, and high accuracy and vertical resolution. Analysis and validation of GPS/MET data confirmed that the highest quality of the retrieved variables is achieved in the upper troposphere and lower stratosphere region with an (upper bound) accuracy in refractivity of <0.4%, in temperature of <1 K [*Kursinski et al*., 1996; 1997; *Rocken et al.*, 1997; *Steiner et al.,* 1999], and in geopotential height of <20 m [*Leroy*, 1997]. The accuracy in specific humidity in the lower to middle troposphere is of the order 0.2–0.5 g/kg [*Kursinski and Hajj*, 2001].

Ongoing GPS RO experiments are placed on board the German CHAMP satellite [*Reigber et al.*, 2000; *Wickert et al.*, 2001] and the Argentine SAC-C satellite [e.g., *Hajj et al.*, 2004]. The validation of CHAMP RO data with corresponding weather analyses profiles revealed an ensemble mean agreement of <0.4 K between 10–35 km height with a standard deviation of ~1 K at 10 km increasing to ~2 K at 30 km height [*Wickert et al.*, 2004]. The precision of RO data was investigated by *Hajj et al*. [2004], who performed an inter-comparison of nearby CHAMP and SAC-C temperature retrievals, which were found to be consistent to <0.1 K in the mean, and to 0.5 K in the standard deviation, between 5–20 km height.

These results demonstrated the suitability of RO data for global climate monitoring and modeling and indicated their potential to make a vital contribution to the Global Climate Observing System [*GCOS*, 2003]. Such a contribution is studied by GNSS RO observing system simulation experiments (OSSE) [*Steiner et al.*, 2001; *Foelsche et al.*, 2003] and by climatological exploitation of RO data from GPS/MET [*Schroeder et al.*, 2003] and CHAMP [*Schmidt et al.*, 2004; *Steiner et al.*, 2004]. It is planned to be fully realized by future RO constellation missions such as COSMIC [*Rocken et al*., 2000; *Lee et al.*, 2001] and potentially ACE+ (Atmosphere and Climate Explorer) [*Hoeg and Kirchengast*, 2002; *Kirchengast and Hoeg*, 2004]. Regarding weather forecasting, *Kuo et al*. [2000] confirmed that the assimilation of RO data has the potential to significantly improve the accuracy of global and regional analysis and weather prediction, which is subject to investigation in assimilation studies [e.g., *Eyre*, 1994; *Zou et al.*, 2000; *Collard and Healy*, 2003; *Healy et al*., 2003].

One important issue in this respect is the knowledge on RO observation errors. *Syndergaard* [1999] investigated in a theoretical analysis the step-by-step covariance propagation from excess phase profiles via Doppler shift, bending angle, refractivity and pressure to temperature profiles. *Palmer et al*. [2000] presented an optimal estimation inverse method for the simultaneous retrieval of temperature, specific humidity and surface pressure from bending angle data with an error analysis emerging from the optimal estimation theory [*Rodgers*, 2000]. The practical application of this method to GPS/MET bending angles was carried out in a follow-on study by *Palmer and Barnett* [2001] including assumptions of simple background and measurement error covariances with correlations being neglected. A complete theoretical error analysis and characterization, based on optimal estimation methodology and accounting for error correlations, was performed by *Rieder and Kirchengast* [2001b]. Starting with excess phase error profiles, the properties of bending angle, refractivity, pressure, and dry temperature profiles were assessed and their covariance structure and correlation functions were analyzed.

Within this context we present results of an ensemble-based empirical error analysis, which involves quasi-realistically simulated GNSS RO data and provides a complete error characterization for each retrieval product, including profiles of bending angle, refractivity,

pressure, geopotential height, temperature, and specific humidity. The purpose of the work is to provide reasonably realistic error estimates of covariance matrices and, in addition to the above described previous studies, on bias profiles. Estimations of the latter are in-accessible by the classical BLUE (Best Linear *Un-biased* Estimation) theory [*Rodgers*, 2000], whilst the present empirical approach, estimating the error statistics from Monte Carlo-type "retrieved-minus-true" profile ensembles, fully allows for non-linearities and biases.

The study is based on an end-to-end forward-inverse simulation involving modeling of the neutral atmosphere and ionosphere and simulation of radio occultation events (section 4.1.2), forward modeling of excess phase observables including realistic observation system error modeling (section 4.1.3), and retrieval of atmospheric parameters (section 4.1.4). The error analysis scheme is described in section 4.1.5. Results of the empirical error statistics for each retrieval product, comprising bias profiles, standard deviation profiles and correlation functions, are presented and discussed in section 4.1.6. One practical result of the study is the specification of simple analytical observation error covariance matrices for refractivity, which is described in section 4.1.7, where also the applicability to other parameters beyond refractivity is discussed. Section 4.1.8 finally summarizes the findings and conclusions of the study and provides a brief outlook to further investigations.

4.1.2 Ensemble Design and Simulations

The simulation of the data ensemble, from occultation event simulations via excess phases to atmospheric retrieval products, was performed with the End-to-end GNSS Occultation Performance Simulator (EGOPS) [*Kirchengast et al*., 2002]. Quasi-realistic atmospheric and ionospheric models as well as realistic error characteristics were involved in order to obtain reliable simulated occultation data.

Occultation events were simulated for one day, September 15, 1999, adopting the planned European Meteorological Operational satellite (METOP) [*Edwards and Pawlak*, 2000] as Low Earth Orbit (LEO) satellite and its GNSS Receiver for Atmospheric Sounding (GRAS) [*ESA/EUMETSAT*, 1998; *Silvestrin et al.*, 2000] as sensor. Rising and setting occultation events were taken into account employing the GPS satellites as transmitter system. The number of simulated events for this day was 574 in total. In order to save computation time with respect to the time consuming forward modeling, we chose an ensemble of 300 occultation events equally distributed in space and time with 100 events in each of three latitude bands, low (-30° to +30°), middle (±30° to ±60°), and high (±60° to ±90°) latitudes. Figure 4.1 illustrates the distribution of these 300 profiles, which were used for the following simulations and error analysis.

As atmospheric model input an ECMWF analysis field was used (Sept. 15, 1999, 12UT), with resolution T213L50 and the highest model level at 0.1 hPa (~60 km x 60 km grid spacing, 50 vertical levels up to ~65 km height). This model horizontal sampling is about 5 times higher than the along-ray horizontal resolution of RO data of about 300 km [e.g., *Kursinski et al.*, 1997] and is capable of representing weather features such as frontal systems, but smaller scale horizontal variations are no longer resolved. Figure 4.2 shows exemplary slices of temperature and specific humidity, extracted from the analysis field, which indicate the reasonably realistic nature of the model for the purpose of the study. The ionosphere was prescribed with the NeUoG model, which is a global empirical solar-activity and local-time dependent 3D climatological model of the ionospheric electron density field developed by *Leitinger et al.* [1996]; a moderate solar activity index $F_{10.7}$ = 130 was supplied to the model

for the present study. NeUoG was of good use already in other occultation-related studies [e.g., *Leitinger and Kirchengast*, 1997; *Steiner et al.*, 2001; *Gobiet and Kirchengast*, 2004a].

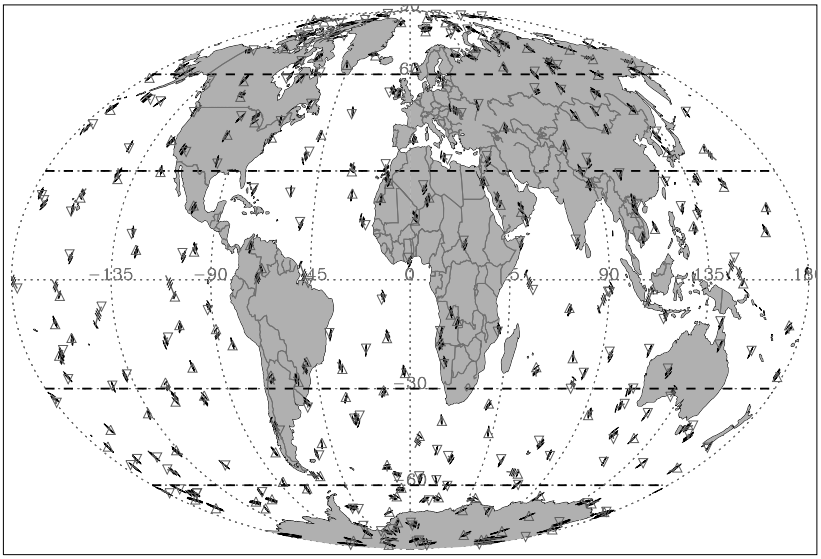


Figure 4.1. Distribution of the 300 occultation events used in the study; rising occultations (triangles), setting occultations (upside-down triangles), latitude circles delimiting selected latitude bands (heavy dashed lines). For each individual event (inspect, e.g., the event at the equator at about –110 deg longitude as a clear example), small lines give an indication of the "verticality" of the event (the shorter the tangent point trajectory/heavy line, the more vertical) and of the occultation plane orientation (orientation of the parallel light solid lines), respectively.

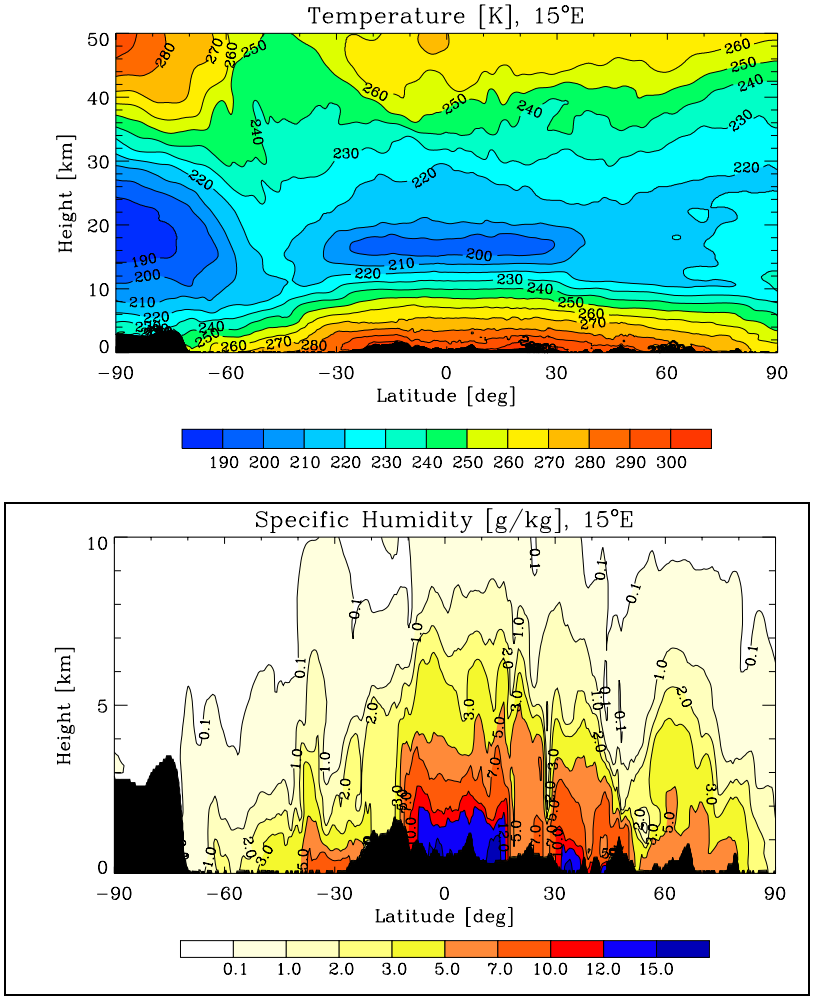


Figure 4.2. ECMWF analysis field of September 15, 1999, 12 UT, at 15°E: temperature slice (upper panel) and specific humidity slice (lower panel).

4.1.3 Phase and Doppler Observables

The geometry for each occultation event was calculated based on Keplerian orbits. Forward modeling of the signal propagation through the atmosphere-ionosphere system was performed with a sub-millimetric precision 3D ray tracer providing signal profiles as seen by the sensor. Since ray tracing stops at multipath situations in the lower troposphere in case of sharp vertical gradients, this study does not include multipath and diffraction effects, which has to be kept in mind when interpreting the results at heights below 5 km in the lower troposphere.

Observation system modeling was subsequently performed to superpose instrumental and raw processing system effects on the forward modeled signal. The results are fairly realistic excess phase path profiles reflecting the GRAS receiving system performance [*Silvestrin et al.*, 2000; *Ramsauer and Kirchengast*, 2001]. Modeling of the observation system error characteristics comprises precise orbit determination (POD) errors, antenna pattern, local multipath, receiver thermal noise, and clock instabilities. The POD error model contains satellite positioning and velocity errors with the along-ray velocity error of 0.05 mm/s being the dominant error source. The radial position errors of the GPS and the LEO satellites were fixed at 0.2 m and at 0.4 m, respectively, a conservative bound for modern POD performance. Receiver noise was modeled as white Gaussian noise for a LEO antennae noise temperature of 150 K and a loop bandwidth (single-side) of 10 Hz. Local multipath effects were modeled using a sinusoidal shaped function with the multipath phase error amplitude set to 0.5 mm and the multipath phase error period set to 100 sec. The modeling of clock errors, representative of the performance of high-quality ultra-stable oscillators (USOs), was based on a random walk model and a ground-based single-differencing clock correction method with the relative stability of the ground clock set to a 1-sec Allan deviation of 1×10^{-13}. *Ramsauer and Kirchengast* [2001] performed a thorough error budget analysis for a GRAS-type receiving system and more details can be found therein.

Figure 4.3 (left column) shows the ionosphere-corrected excess phase path LC for the global ensemble (panel a) and for the three latitudinal data sets (panels b–d) with the mean LC excess phase path indicated. An occultation event starting/ending at 90 km (0 sec) lasts 1–2 min with a (neutral-gas) excess phase path of about 2 mm near the mesopause (~80 km/~4 sec; in Figure 4.3, left column, involving ionospheric residuals), ~20 cm near the stratopause (~50 km/~16 sec), and >20 m below the tropopause level (~15 km/~30 sec). The near surface excess phase path reaches ~0.7–2 km depending on the water vapor content of the atmosphere. The ionospheric residual after linear ionospheric correction of phases is of the order of a few centimeters, which can be seen in the phase path profiles in the first few seconds of the occultation, especially in the mean profiles of the ensembles. The different latitudinal ensembles reflect the varying influence of the ionosphere, showing ionospheric residuals of up to ~1 cm at high latitudes increasing to up to ~10 cm at low latitudes.

Figure 4.3, furthermore, illustrates the phase errors (middle column) and associated Doppler errors (right column), which are shown without the ionospheric errors in order to explicitly depict the errors due to the GRAS receiving system specifications. The error statistics for the LC phase is based on difference profiles, $\mathbf{\Delta L}_k$, each of which was computed by subtracting the raw phase profile without observational error from the phase profile including observational error.

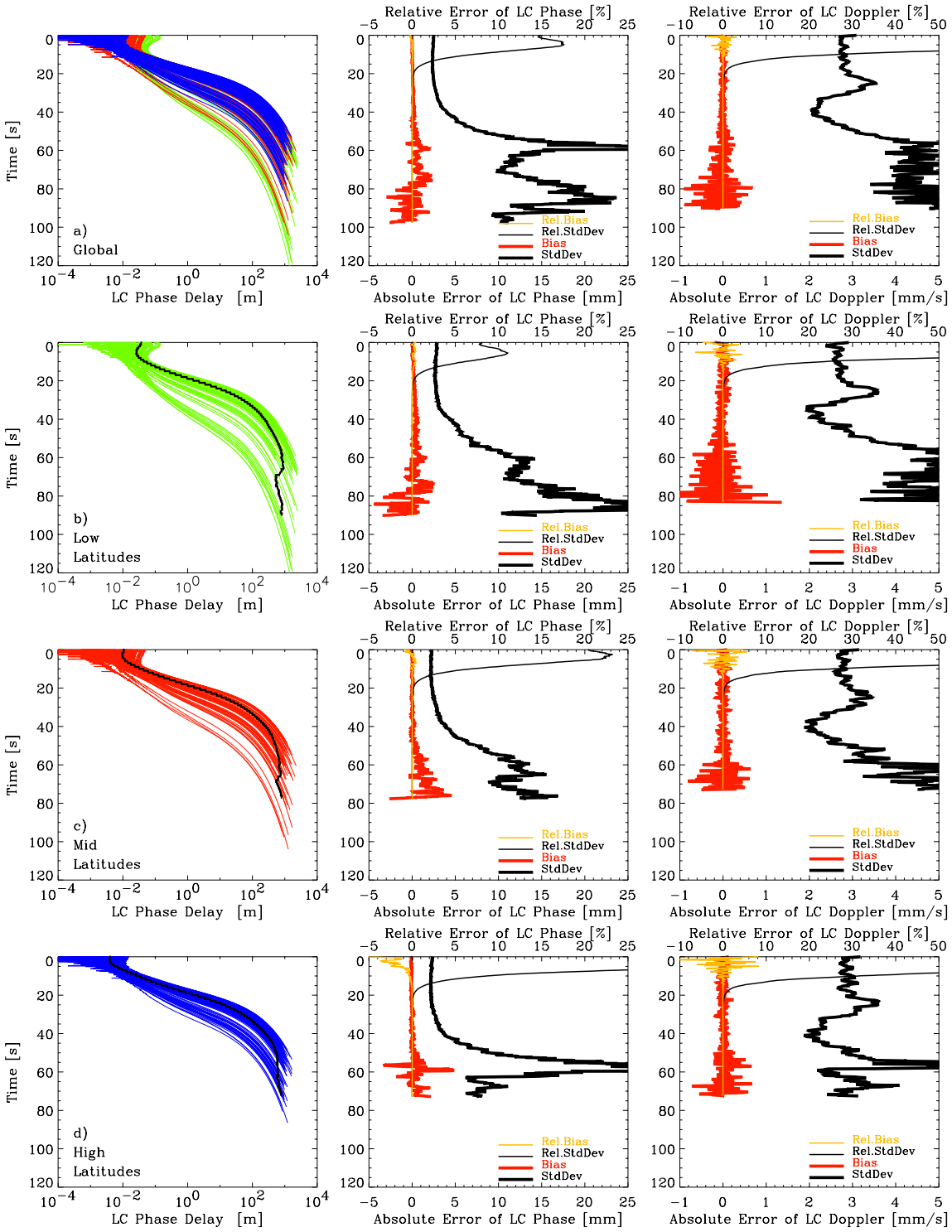


Figure 4.3. The left panels show the ionosphere corrected LC excess phase path for the global ensemble (a) and the latitudinal ensembles (b–d); as well as the mean profile of each ensemble (heavy black line). The middle and right panels show the associated absolute (heavy) and relative (light) bias (red/orange) and standard deviation (black) of the LC phase delay (middle) and Doppler shift (right), respectively.

Calculation of the mean of the difference profiles led to the bias profile, $\mathbf{b}_{\Delta L}$,

$$\mathbf{b}_{\Delta L} = \frac{1}{n}\sum_{k=1}^{k=n} \Delta \mathbf{L}_k \text{ , with } n = \text{number of events in ensemble.} \qquad (4.1)$$

Furthermore, profiles of the standard deviation, $\mathbf{s}_{\Delta L}$,

$$\mathbf{s}_{\Delta L} = \sqrt{\frac{1}{n-1}\sum_{k=1}^{k=n} \left(\Delta \mathbf{L}_k - \mathbf{b}_{\Delta L}\right)^2} \,, \qquad (4.2)$$

and of root mean square (rms) error, $\mathbf{rms}_{\Delta L} = \sqrt{\mathbf{s}_{\Delta L}^2 + \mathbf{b}_{\Delta L}^2}$, were calculated (the latter not shown in Figure 4.3, since essentially identical with the standard deviation). Relative error quantities (percentage errors) were derived by relating the absolute error quantities to the mean phase profile (Figure 4.3, left column) and by multiplying it with the factor 100. The errors in the Doppler shift, the time derivative of the phase (see section 4.1.4 below), were computed in the same manner.

The standard deviation (and also the rms error, since the bias contribution is negligible) of the LC phase sampled at 10 Hz is found to be 2–3 mm at <30 sec (above the tropopause), which reasonably reflects the METOP/GRAS receiving system performance [*Silvestrin et al.*, 2000]. While absolute errors increase into the troposphere to 25 mm, the relative standard deviation is found to be <0.02% at these low heights. The Doppler shift exhibits a standard deviation and rms of ~3 mm/sec above the tropopause. Also for the Doppler shift, relative errors are found <0.02% in the troposphere. Biases in both, excess phase and Doppler shift, are negligible, reflecting the self-calibrated nature of these basic observables.

4.1.4 Retrieval of Atmospheric Profiles

The calculation of atmospheric profiles from excess phase paths included filtering of the phases with a regularization method resembling a low pass filter to eliminate high frequency noise [*Syndergaard*, 1999]. Time-derivation of the filtered phase path data led to the Doppler shift, from which bending angles were calculated. We applied the ionospheric correction of bending angles, which leads to a considerably smaller residual bias in derived bending angles than the phase correction method [*Vorob'ev and Krasil'nikova*, 1994; *Hocke et al*., 1997].

The processing involved an initialization of retrieved bending angles with background bending angles derived from the MSISE-90 climatology [*Hedin*, 1991]. A background search in the height interval of 40–55 km was performed to find the best-fit background bending angle profile by using a least squares criterion. An inverse-covariance-weighting statistical optimization [*Healy*, 2001] was applied to combine observed and background bending angle profiles as described in detail by *Gobiet and Kirchengast* [2004a].

Briefly, for the specification of the background error covariance matrix a relative standard deviation of 20% of the background bending angle was assumed, and the variances constituting the diagonal elements of the covariance matrix set accordingly. The off-diagonal elements representing covariances were calculated by assuming an exponential correlation decay with a correlation length of 6 km, representing the length scale at which the inter-level correlation is decayed to 1/e [*Rodgers*, 2000]. The observation error was estimated as the standard deviation of the observed bending angle profile from the background profile between 70–80 km, where the bending angle signal is small and ionospheric residual and measurement

noise dominate. The average standard error of the 300 retrieved bending angles was found to be ~1.2 μrad, fairly consistent with the accuracy requirement specification of 1 μrad of the METOP/GRAS operational RO receiving system [*ESA/EUMETSAT*, 1998]. For reference and scaling, bending angle accuracy from present experimental RO systems such as CHAMP/GPS is found more at the 2–3 μrad level (A. Gobiet, Univ. of Graz, Austria, priv. communications, 2004). The off-diagonal observation error covariance elements were set to follow an exponential decay with a correlation length of 1 km. All these settings are based on fairly extensive adequacy and sensitivity tests, partly inherited from earlier studies [*Hocke et al.*, 1997; *Healy*, 2001; *Steiner et al.*, 2001].

Refractivity, pressure, geopotential height, temperature, and specific humidity profiles were calculated from the bending angle profiles by using, from bending angle via refractivity towards dry temperature, a classical GPS RO retrieval chain summarized by *Gobiet and Kirchengast* [2004a]. A detailed review of the algorithms involved was given by *Kursinski et al.* [1997]. In the troposphere, an optimal estimation temperature and humidity retrieval algorithm was applied to refractivity, using the ECMWF 24-hour short-range forecast for September 15, 1999, 12 UT (the time of the analysis field used) as background field for temperature and humidity. This algorithm is an implementation closely following the variational retrieval (1D-Var) algorithm of *Healy and Eyre* [2000]. It was implemented for a latitude-dependent tropospheric top height defined to decrease linearly with latitude from 15 km at the equator to 9 km at the poles. Above this height, standard dry air retrieval was applied, with a smooth (half-Gaussian-weighted) transition to the 1D-Var solution over a width scale of ~2 km below this height.

For the background error, covariance matrices following an exponential decay with a correlation length of 6 km for temperature and of 3 km for specific humidity were adopted. The standard deviation of the background specific humidity was assumed 20% at 0 km linearly increasing to 50% at 10 km and kept constant further up. The standard deviation of the background temperature was taken constant with height but dependent on latitude with 2.5 K at high, 2.0 K at middle, and 1.5 K at low latitudes, respectively. These simple settings, checked in various sensitivity tests, were found to reasonably reflect the uncertainty of the short-term forecast fields used as background. Further incremental improvement can be achieved by using more sophisticated forecast error covariance matrices. The refractivity error covariance matrix was specified following an exponential decay with a correlation length of 3 km, and a latitude and height dependent standard deviation. The standard deviation of refractivity was 0.8%, 1.4%, 2% at 0 km linearly decreasing to 0.2%, 0.2%, 0.1% at 10 km and staying constant further up for high, middle, and low latitudes, respectively. These settings were based on the refractivity error analysis results of this study itself.

4.1.5 Error Analysis Formulation

For each atmospheric retrieval product an empirical error analysis was performed as follows. We computed, for each retrieved profile, the co-located ECMWF analysis vertical profile at the mean tangent point location of the event to serve as the "true" profile. This implies that our error estimates represent an upper bound error estimate including both the observational error (due to the measurement and retrieval process) and the representativeness error (due to atmospheric horizontal variability), the latter becoming important in the lower troposphere due to increased horizontal variations [*Foelsche and Kirchengast*, 2004a, 2004b; *Syndergaard et al*., 2004].

The difference profiles $\Delta\mathbf{x}_k$ ($\Delta\mathbf{x}_k = (\Delta x_1, \Delta x_2, \ldots, \Delta x_i)^T{}_k$, with i denoting the height levels, k the profile number, and ^{T} the matrix transpose) of the retrieved profiles, $\mathbf{x}_k^{retr}$, and the "true" profiles, $\mathbf{x}_k^{true}$, were calculated at an ECMWF-type L60 vertical grid in the form

$$\Delta\mathbf{x}_k = \left(\mathbf{x}_k^{retr} - \mathbf{x}_k^{true}\right). \tag{4.3}$$

Calculation of the mean of the difference profiles led to the bias profile, **b**,

$$\mathbf{b} = \frac{1}{n}\sum_{k=1}^{k=n} \Delta\mathbf{x}_k \text{ , with } n = \text{number of events in ensemble.} \tag{4.4}$$

Next, the bias was subtracted from each profile giving bias-free profiles, $\mathbf{\Delta x}^{\text{biasfree}}$,

$$\Delta\mathbf{x}_k^{biasfree} = \Delta\mathbf{x}_k - \mathbf{b}\,. \tag{4.5}$$

From these bias-free profiles we calculated the error covariance matrix, **S**,

$$\mathbf{S} = \left[\frac{1}{n-1}\sum_{k=1}^{k=n}\left(\Delta\mathbf{x}_k^{biasfree}\right)\left(\Delta\mathbf{x}_k^{biasfree}\right)^T\right], \tag{4.6}$$

with its diagonal elements representing the variances (S_{ii}) at height level i and with its non-diagonal elements representing the covariances (S_{ij}) between height levels i and j. The square root of its diagonal gives the standard deviation profile, **s**,

$$\mathbf{s} \text{ with } s_i = \sqrt{S_{ii}}\,. \tag{4.7}$$

The root-mean-square error profile, **rms**, then reads

$$\mathbf{rms} = \sqrt{\mathbf{b}^2 + \mathbf{s}^2}. \tag{4.8}$$

The error correlation matrix, **R**, with elements R_{ij}, denotes the error correlation between errors Δx_i at height i and Δx_j at height j. It is computed by dividing the covariances S_{ij} by the square root of the product of variances S_{ii} and S_{jj},

$$\mathbf{R} \text{ with } R_{ij} = \frac{S_{ij}}{\sqrt{S_{ii}\,S_{jj}}}\,. \tag{4.9}$$

For details on these statistics estimators see, e.g., *Mardia et al.* [1979], *Anderson* [1984], *Von Storch and Zwiers* [1999], or similar textbooks. The relative error quantities (percentage errors) were computed by dividing the absolute error quantities by the mean profile of the "true" profiles and multiplying this ratio with 100.

4.1.6 Analysis Results and Discussion

Figures 4.5 to 4.10 present the error analysis results of each retrieval product from bending angle, via refractivity, pressure, geopotential height, to temperature and specific humidity. The results are shown for the global ensemble (panel row a) and for the latitudinal data sets (panel rows b–d) up to 50 km height. The left panels show the number of events entering the error statistics at a given height. The middle panels illustrate the error characteristics in terms of absolute and/or relative quantities of bias, standard deviation, and rms profiles. We refer primarily to relative error quantities for the atmospheric parameters with near-exponential height decrease and to absolute error quantities for geopotential height and temperature. These primarily-referred-to error quantities are denoted by heavy lines, solid black for the standard deviation, solid red for the bias, and dashed blue for the rms error, respectively. Mean atmospheric profiles are depicted by light dotted black lines. The right panels display error correlation functions, which are defined as the rows of the error correlation matrix **R** (equation (4.9)). They are displayed for three different height levels (~40 km, ~20 km, ~5 km) representative of upper stratosphere, lower stratosphere, and troposphere. For humidity, three representative tropospheric levels are shown (~7 km, ~5 km, ~3 km). The error correlation functions express the correlation of errors at these specific heights with the errors in the remainder of the profile.

Bending Angle. Figure 4.4 displays the error statistics for the bending angle profiles showing the mean bending angle profile to range from ~20 mrad at near 2 km to ~15 μrad at 50 km impact height. The global data set shows a relative bias of the bending angle of <0.1% at 5–40 km with a maximum of 0.5% outside this height interval. The 0.1% bias height interval shrinks to ~7–30 km at mid latitudes. Maximum bias values of up to 1% occur below 5 km impact height at all latitudes. The relative standard deviation (Rel.StdDev) of the bending angle is 0.2–1% between ~8–35 km, increasing to about 5% at 50 km at all latitudes. In the troposphere, the Rel.StdDev increases to ~5% at high latitudes and to ~8% at low and mid latitudes.

Comparison of these empirical bending angle errors with the findings in theoretical studies [*Syndergaard*, 1999; *Rieder and Kirchengast*, 2001b] shows that the theoretical estimates of the standard deviation appear somewhat smaller than the empirical ones. *Rieder and Kirchengast* [2001b] report a Rel.StdDev of <1% below ~37 km for a standard receiver scenario (SR; GPS/MET receiver type) and below ~43 km for an advanced receiver scenario (AR; GRAS receiver type). The difference of that GRAS receiver type AR scenario 1%-height at ~43 km to the present empirical 1%-height at ~35 km is explained by the lower vertical resolution, leading to higher accuracy, assumed in that work (~4 km resolution near the stratopause instead of ~2 km here), the absence of ionospheric errors in that work, the lower background bending angle uncertainty assumed (~10% near the stratopause instead of 20% here), and the somewhat higher excess phase accuracy (2 mm instead of ~2.5 mm here).

Accounting for the differences, the stratospheric results of that theoretical error analysis are found consistent with the present empirical analysis results. This underlines the relevance of one of the conclusions drawn by *Rieder and Kirchengast* [2001b], which stressed that "though the discussed results furnish a good baseline idea, a dedicated error analysis should be performed for each specific processing setup. The formalism proposed here is one excellent candidate; a complementary one is ensemble-based empirical error analysis" (end of quote). Since the present empirical analysis involves significantly less restrictive assumptions (no linearity, un-biasedness, dry air, spherical atmosphere assumptions) the results obtained here are evidently more realistic. Still it is to be kept in mind, not only for the bending angle

results but also for all results following below, that data products from different measurement and retrieval processes than here (i.e., other than GRAS-type receiving system performance, different retrieval algorithms) will need at least a scaling of the present results to those different conditions, if not repeat of the error analysis itself, in case of severe differences.

Deriving from uncorrelated measurements at the phase level, the bending angle error correlation functions show sharp structures with small negative correlation of the given height levels with their lower and upper neighboring levels. These negative correlations stem from derivative operations at the step from phase to Doppler shift [*Syndergaard*, 1999; *Rieder and Kirchengast*, 2001b] and are most pronounced in the low latitude ensemble at lower to middle stratospheric heights where they reach values of about –0.4.

When interpreting and using the bending angle error results it has to be taken into account that the "true" bending angle was computed here by Abel transform from the "true" refractivity assuming spherical symmetry. While this still leads to reasonable results in the stratosphere and upper troposphere it does not represent with full adequacy the errors under significant horizontal variations below ~7 km [*Foelsche and Kirchengast*, 2004b].

Refractivity. Figure 4.5 displays the error statistics for the refractivity profiles with the mean refractivity profile ranging from ~280 [N units] at 1 km to ~0.2 at 50 km height. Refractivity exhibits a Rel.StdDev of 0.1–0.75% and a relative bias of <0.1% at 5–40 km height in all data sets. Outside this "core" domain, the relative bias reaches 0.3% in the lower troposphere at all latitudes and up to 0.7% at 50 km height at high latitudes. The relative standard deviation stays below 2% outside the "core" domain, except for the low latitude ensemble, where it reaches ~2.3% at the bottom [*Steiner and Kirchengast*, 2004a].

In the *Rieder and Kirchengast* [2001b] analysis, the upper height limit for a Rel.StdDev of <1% is 43 km for the SR scenario and 49 km for the AR scenario, respectively. Our empirically estimated height limit lies in between these heights at 45 km. *Kursinski et al.* [1997] estimated an rms error of refractivity of <0.4% between 5–30/40 km for a bad/good scenario. This agrees well with the empirically derived Rel.StdDev of <0.4% at ~6–36 km height, but at lower tropospheric heights the error derived here is bigger, mainly due to subsuming more horizontal variability error. Our findings are also consistent with the results of a study carried out by *Kuo et al*. [2004], who performed an error analysis of CHAMP and SAC-C data and found the observation error of refractivity to be of the order of 0.3–0.5% at 5–25 km increasing to 3% in the tropical lower troposphere.

The refractivity error correlation functions are broader than the bending angle error correlation functions revealing the effect of the Abelian integration. The broadening in the troposphere seen in the refractivity error correlation functions is mostly a result of the errors due to horizontal variations; it is particularly marked in the mid and high latitude ensembles with negative correlations vanished. Negative correlations are most pronounced in the low latitude ensemble at lower to middle stratospheric heights. Since the error statistics is performed on an L60 height grid, these anti-correlation features are smoothed and can hardly be seen in Figure 4.5 but can be resolved on a finer, e.g., L90, grid (further addressed in section 4.1.7 below). To note one specific feature of interest, the enhanced error correlation seen at mid latitudes (Figure 4.5c, right panel) between the lower stratosphere near 20 km and the upper stratosphere near 30 km (double-peaked structure of the 20 km correlation function) was found robust under sensitivity tests. It seems to indicate an interesting "memory" across the middle stratosphere of the (in general) nearly isothermal mid-latitude temperature profiles in this height domain.

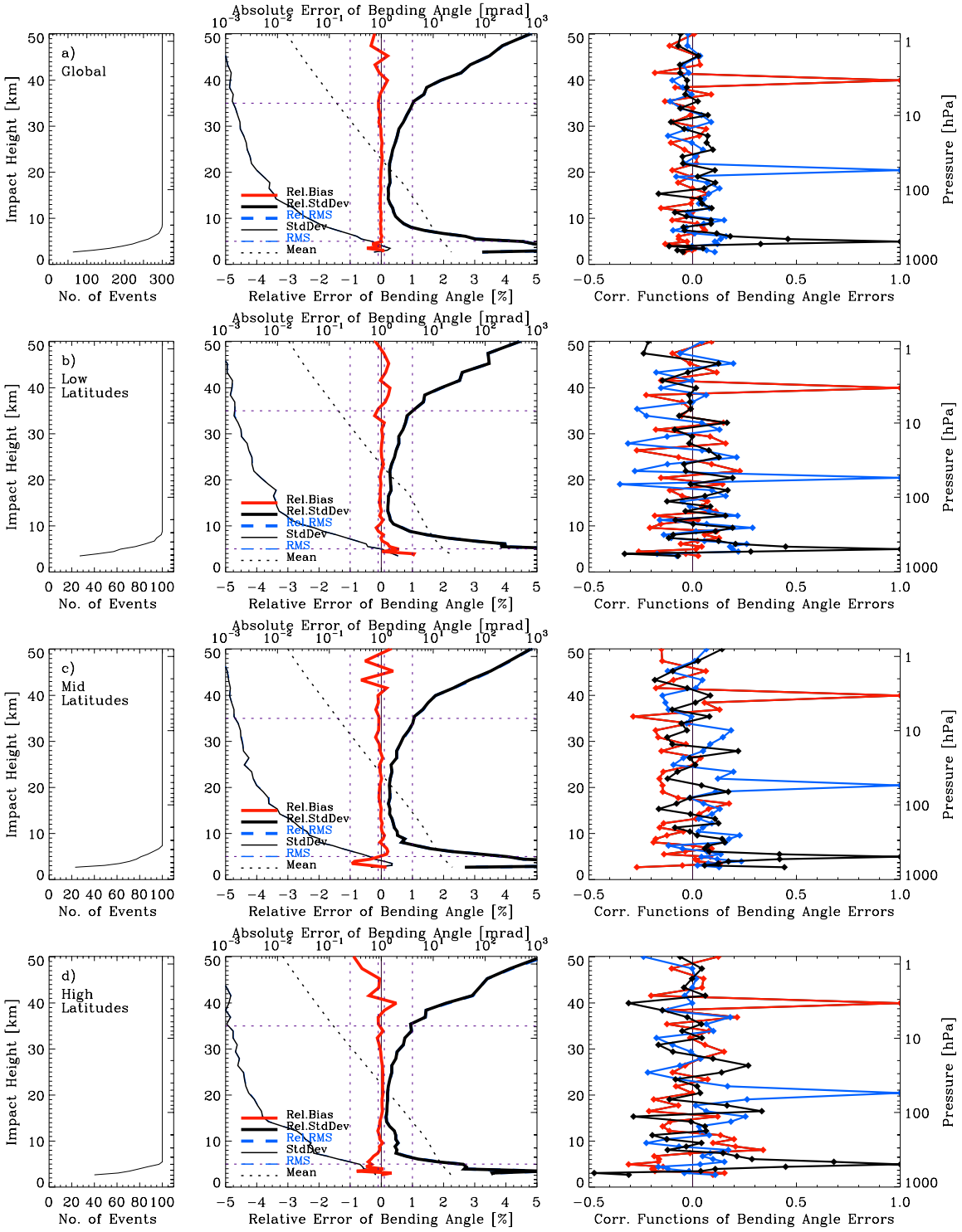


Figure 4.4. Bending angle error analysis results, as a function of impact height, for the global (a) and the latitudinal (b–d) ensembles. Left panels: number of events used for the error statistics calculation at any given height. Middle panels: relative bias (red), relative standard deviation (heavy black), relative rms (heavy blue dashed) as well as the absolute standard deviation (light black), absolute rms (light dashed), and the mean of the "true" profiles (dotted). Right panels: error correlation functions for ~40 km (red), ~20 km (blue), and ~5 km (black) height, representative of upper stratosphere, lower stratosphere, and troposphere, respectively.

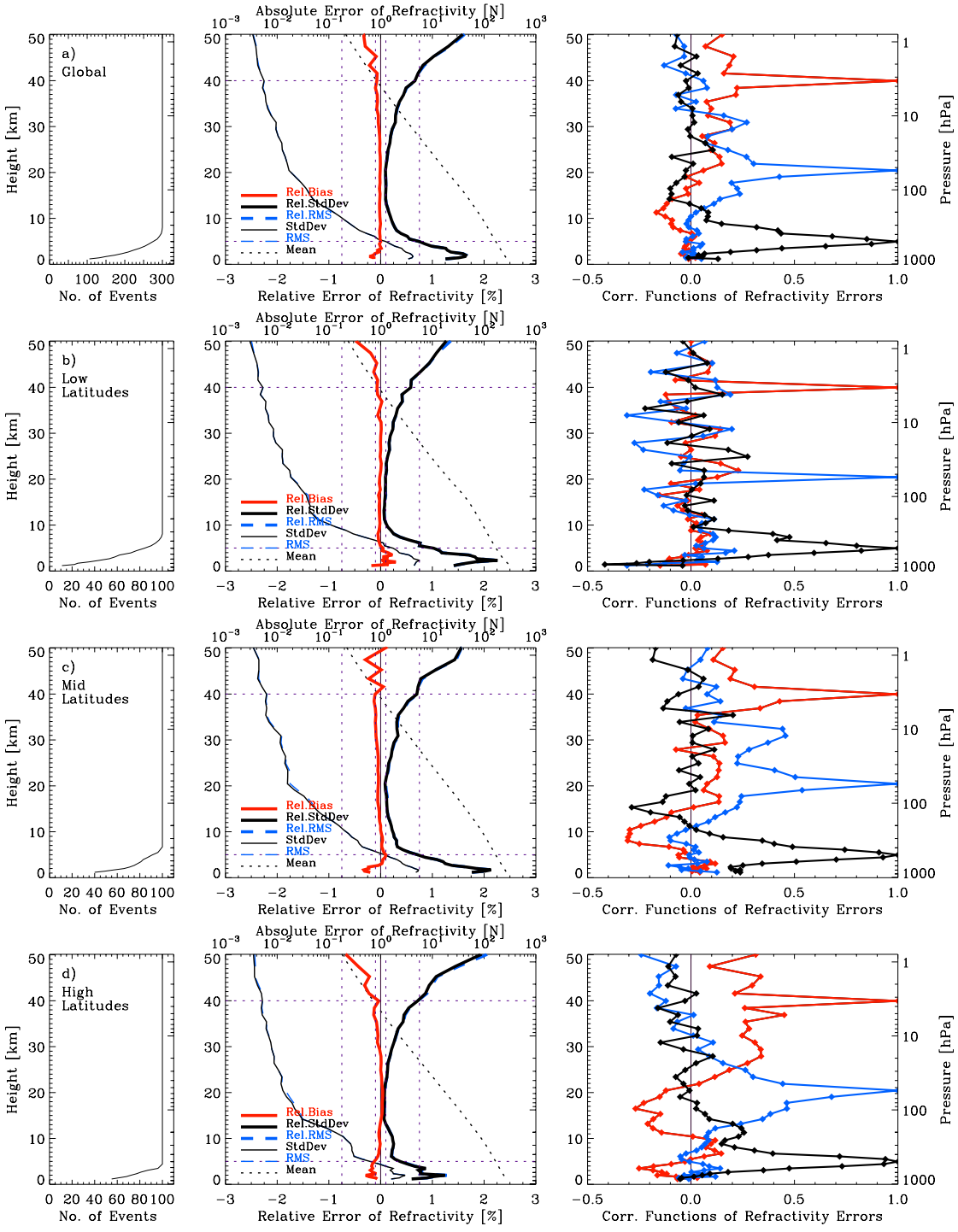


Figure 4.5. Refractivity error analysis results, as a function of height, for the global (a) and the latitudinal (b–d) ensembles. Figure layout and style same as Figure 4.4; see that caption for details.

As refractivity is the RO data product perhaps most convenient for data assimilation [e.g., *Healy and Eyre*, 2000; *Syndergaard et al.*, 2004], we present in section 4.1.7 a simple analytical formulation for refractivity error covariance matrices based on the empirical results obtained.

Pressure: Figure 4.6 presents the error statistics of pressure with the mean profile ranging from ~900 hPa at 1 km to ~0.7 hPa at 50 km height. For all data sets the Rel.StdDev of pressure is 0.25% in the troposphere decreasing to 0.1% above the tropopause and increasing to 1% at 40 km height.

The relative bias is <0.2% below 30 km and <0.5% below 40 km for the global and low latitude ensembles. The smallest biases occur at mid latitudes, the largest ones at high latitudes, with a relative bias of close to –3% and a Rel.StdDev of ~3% at 50 km. These biases are the result of downward hydrostatic integration of refractivity (or, equivalently, density), which carries residual density biases from above the stratopause, where background information dominates [e.g., *Rieder and Kirchengast*, 2001b; *Gobiet and Kirchengast*, 2004a], by about two scale heights downward into the upper stratosphere. This is most readily seen in comparing the high latitude results of Figures 4.5 and 4.6 (weaker, but otherwise similar, also in the low latitude results): Inspecting, for example, the height limit of biases smaller than –0.25%, this lies at ~46 km for refractivity (density) (Figure 4.5d, middle panel) while it lies at ~30 km for pressure (Figure 4.6d, middle panel). At low latitudes, these heights are ~48 km (refractivity, density) and ~36 km (pressure), respectively. As this type of bias characteristics also maps further into geopotential height and temperature, special efforts are required to keep them as low as possible in magnitude, in order to have the upper stratospheric profiles accurate as high up towards the stratopause as possible.

A simple approach commonly taken — mostly regarding temperature, since pressure and geopotential height related studies have been rarely presented in the open literature so far — is to impose, in addition to the upper-boundary initialization/statistical optimization of bending angle profiles, a "2nd initialization" of the hydrostatic integral. The integral is not downward integrated from 120 km across the full statistically optimized refractivity (density) profile but only from some top height chosen between 30 and 50 km, where it is initialized by a model (*a priori*) temperature usually taken from ECMWF [e.g., *Kursinski et al.*, 1997; *Wickert et al.*, 2004; *Hajj et al.*, 2004]. The most complete discussion of this approach, including sensitivity tests, has so far been given by *Hajj et al.* [2004].

Focusing on climatological use of geopotential heights and temperatures, which is our primary interest, this approach is problematic, since it leads to intricate error characteristics and *a priori*-dependence in the stratospheric data down to about 20 km, which threats the crucial aims of un-biasedness and a clear understanding of the degree of residual biasedness. For these reasons, we do not perform any "2nd initialization" but use full downward integration of the hydrostatic integral from 120 km. The initialization at 120 km, whether zero or the MSISE-90 pressure as we adopted, has negligible influence on the results below 50 km [e.g., *Steiner et al.*, 1999]. This climatologically preferable approach needs fair care in performing the bending angle initialization in order to keep the residual refractivity (density) biases above the stratopause, and consequently the downward-propagated biases in pressure, geopotential height, and temperature, as small as possible (as discussed above and illustrated in Figures 4.5 vs. 4.6–4.8, middle panels). *Gobiet and Kirchengast* [2002; 2004a] and A. Gobiet and G. Kirchengast (submitted manuscript to *J. Geophys. Res.*, 2004b) discuss respective optimization algorithm advancements in detail. They report that particularly residual upper stratospheric biases at high latitudes, which are the most salient ones, can be further reduced significantly by bias correction of background bending angle profiles, at least given data of GRAS-type quality (1 μrad bending angle accuracy). *Gobiet et al.* [2004b] validated CHAMP RO data with various sources of correlative data in the stratosphere and found that these less accurate current RO data demand initialization with better background

bending angle profile libraries than from pure MSIS or CIRA climatologies, e.g., from ECMWF analyses (extended above ~60 km by climatology).

Pressure errors, furthermore, exhibit strong correlation due to the hydrostatic integration; see *Rieder and Kirchengast* [2001b] for a theoretical discussion of pressure error correlation properties.

Geopotential height. Figure 4.7 displays the error statistics of geopotential height, which is discussed with respect to absolute error quantities. The absolute standard deviation of geopotential height is 5–20 gpm below 30 km, staying <70 gpm up to 40 km pressure height for all data ensembles (pressure height $z_p = -(7\text{km})\cdot\ln[p(z)/1013.25\text{hPa}]$, for showing pressure levels as heights). The bias of geopotential height is <5 gpm below 30 km and <30 gpm below 40 km except for the high latitude ensemble. The latter shows a bias of 50 gpm at 40 km, bias and standard deviation increase to 200 gpm at 50 km pressure height. The correlation functions of geopotential height errors are basically the same as for pressure (cf. right panels of Figures 4.6 and 4.7). The same holds for the behavior of relative pressure errors vs. absolute geopotential height errors (cf. middle panels of Figures 4.6 and 4.7), a property rooted in the physical relation that fractional pressure is proportional to geopotential height [*Leroy*, 1997; *Syndergaard*, 1999]. The results are in good agreement with the findings of *Kursinski et al.* [1997], who published an rms error of ~10 gpm at 10–20 km and with those of *Leroy* [1997] publishing a geopotential height error of <20 gpm at upper tropospheric and lower stratospheric heights.

Temperature. Temperature error is presented in Figure 4.8 and is also discussed with respect to absolute error quantities. For the global data set, temperature shows a standard deviation of 0.2–1 K at 3–31 km height and a bias of <0.1–0.5 K below 33 km and of <0.1 K below 20 km. In the lower troposphere, the standard deviation does not exceed 1.5 K in all latitude bands; note that it is a result of the 1D-Var retrieval below about 8 km (high latitudes) to 12 km (low latitudes). The temperature bias is lowest at mid latitudes with <0.1–0.2 K at 2–40 km. Large temperature errors occur at high latitudes, revealing a bias of >3 K and a standard deviation of >5 K above 40 km, which is dominating the global error statistics. The reasons for the residual stratospheric biases are the same as discussed above for pressure; the pressure biases transfer to temperature biases via the equation of state for dry air, which expresses that temperature is proportional to pressure divided by refractivity (or density).

The correlation functions of temperature error are reminiscent of both pressure and refractivity correlation. The correlating influence of hydrostatic integration is most pronounced at high latitudes, least at low latitudes. This indicates that temperature profile errors at high latitudes are more vertically coupled, and hence in particular vulnerable to downward propagation of errors by integration, in line with the findings based on independent larger-ensemble climatologies (~1000 profiles in total from a full season) of *Steiner et al.* [2001], *Foelsche et al.* [2003], and *Gobiet and Kirchengast* [2004a]. The relations determining how both refractivity and pressure errors propagate into temperature errors are somewhat involved and not addressed further here; *Rieder and Kirchengast* [2001a, 2001b] provide some insight. The different shapes of errors and correlation functions below and above the tropopause are due to the application of the 1D-Var retrieval algorithm in the troposphere (see section 4.1.4) [*Steiner and Kirchengast*, 2004a].

Kursinski et al. [1997] found the rms error of temperature to be <1 K at 8/2–30/40 km for a bad/good scenario, which is in good agreement with our empirical estimates of sub-Kelvin accuracy at 3–31 km for the global data set and at 3–36 km for the low latitudes. In fact the error estimates in this study can be viewed in several respects as an advancement and

generalization of the *Kursinski et al.* [1997] results under more realistic conditions as possible due to the ensemble-based empirical error analysis approach used here. *Palmer et al*. [2000] applied a 1D-Var retrieval and showed higher temperature errors of 1.5 K below ~30 km and up to 5–8 K at 40–50 km height, mainly due to different assumptions on error covariances.

Specific Humidity. The error statistics for specific humidity is presented in Figure 4.9 and is discussed with respect to relative error quantities again. The global mean profile of specific humidity reaches from ~4–0.1 g/kg at 1–10 km height. In the lower troposphere below 2 km, mean specific humidity exhibits values of ~10 g/kg at low latitudes and of ~2 g/kg at high latitudes, in the upper troposphere near 8 km the mean values range from ~0.6 g/kg at low latitudes to below 0.1 g/kg at high latitudes. Note that all humidity retrieval results are based on the optimal estimation algorithm applied (section 4.1.4).

The Rel.StdDev of specific humidity in the global ensemble is 20–25% below 5.5 km increasing to 40% at 10 km height. The low latitude data set reveals a smaller Rel.StdDev of 15–30 % at 1–10 km. At mid latitudes the Rel.StdDev increases to 50% at 8 km height, the increase in error above about 7 km indicating the height where the background profile dominates the retrieved profile. At high latitudes, a priori knowledge plays a role already in the lower troposphere in the 1D-Var as used here, since, for example, at 3 km humidity values are smaller than 1 g/kg already and the a priori uncertainty is assumed no more than about 30% at this height.

A very encouraging result from the climatological point of view is the low bias in all humidity datasets: Humidity biases are found within 5% at all latitudes and heights except at mid latitudes above 6 km, where biases reach up to 10%, despite the quite simplified error covariance specifications used in the optimal estimation (section 4.1.4). This finding, together with the small biases also found in temperature below 10 km (Figure 4.8), indicates that 1D-Var retrieval based on ECMWF short-term forecasts as background can lead to tropospheric humidity and temperature retrievals clearly of climatological utility. Our plans for climatologies consequently include the full range of products from refractivity via geopotential height to temperature and humidity, where the latter parameters will, in the troposphere, be retrieved from refractivity by the optimal estimation approach [*Steiner et al.*, 2004].

The correlation functions for specific humidity errors are displayed for three tropospheric height levels showing that the statistical errors in humidity are correlated over several kilometers in the troposphere, comparable to the correlation range of refractivity and temperature profiles rather than to that of pressure profiles.

Theoretical estimates by *Kursinski et al.* [1997] on a specific humidity rms error of 10–20% in the lower to middle troposphere are comparable to our empirical estimates in low latitude regions, generally our rms error estimates are more conservative. Good general agreement exists with the results of *Palmer et al.* [2000], who show a 25–35% error at <9 km in specific humidity. Humidity rms errors will always significantly depend on the algorithm used, in particular on the refractivity errors and the error characteristics of the background or other external information used to extract humidity and temperature from refractivity. Optimal estimation is rated the method of choice for this purpose [cf. also *Healy and Eyre*, 2000].

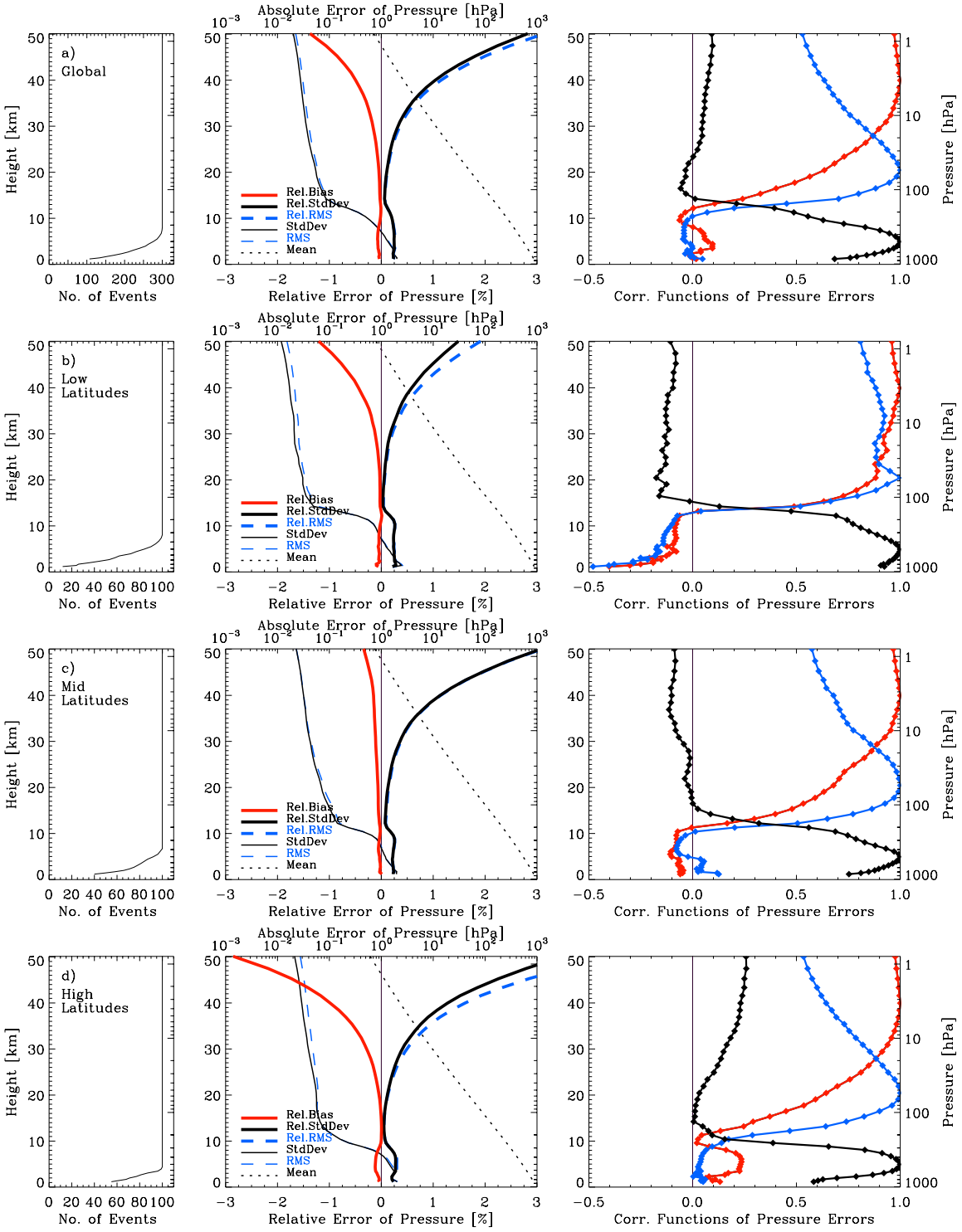


Figure 4.6. Pressure error analysis results, as a function of height, for the global (a) and the latitudinal (b–d) ensembles. Figure layout and style same as Figure 4.4; see that caption for details.

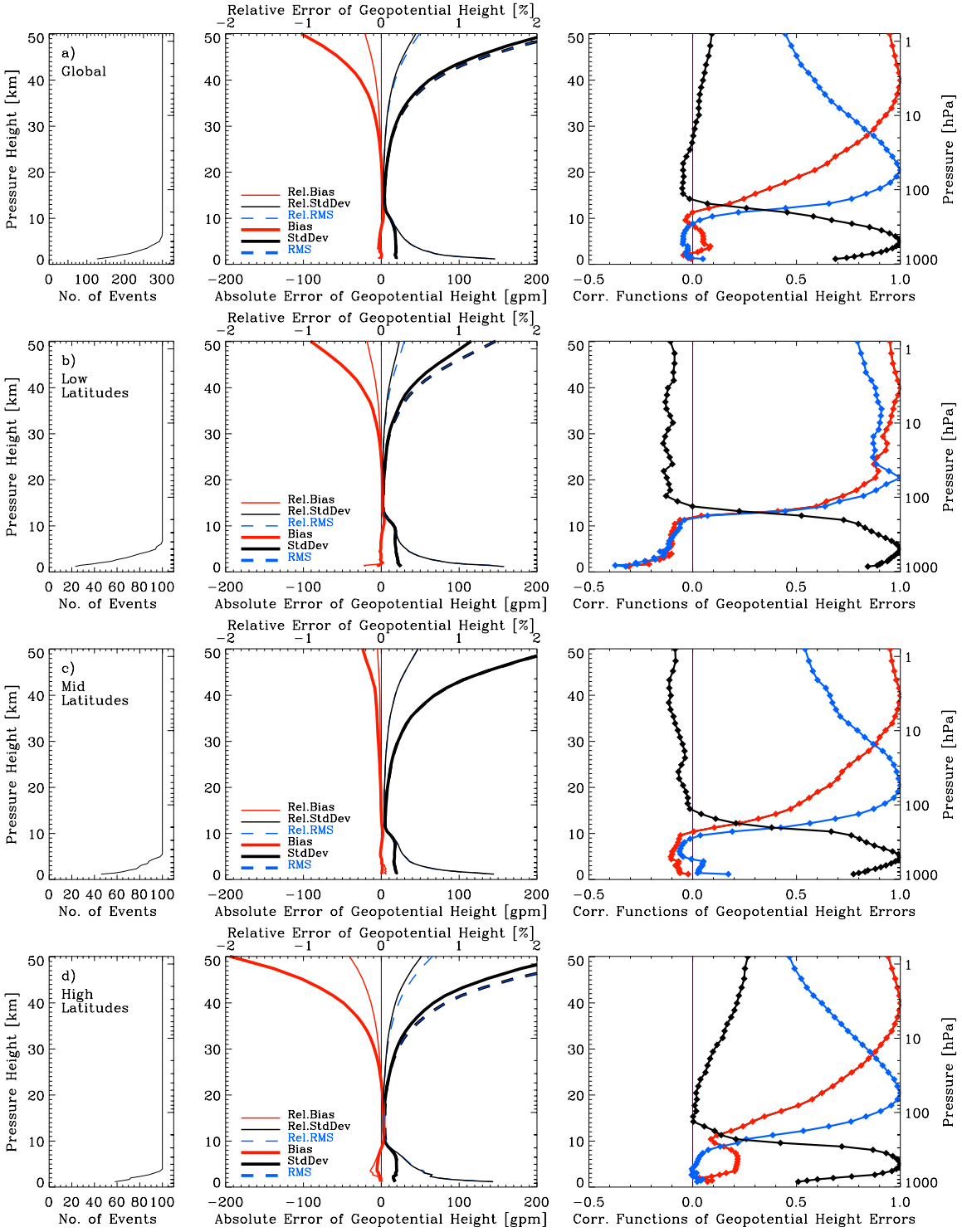


Figure 4.7. Geopotential height error analysis results, as a function of pressure height, for the global (a) and the latitudinal (b–d) ensembles. Left panels: number of events used for the error statistics calculation at any given height. Middle panels: absolute bias (heavy red), absolute standard deviation (heavy black), and absolute rms (heavy blue dashed) as well as relative bias (light red), relative standard deviation (light black), and relative rms (light blue dashed). Right panels: error correlation functions for ~40 km (red), ~20 km (blue), and ~5 km (black), representative of upper stratosphere, lower stratosphere, and troposphere, respectively.

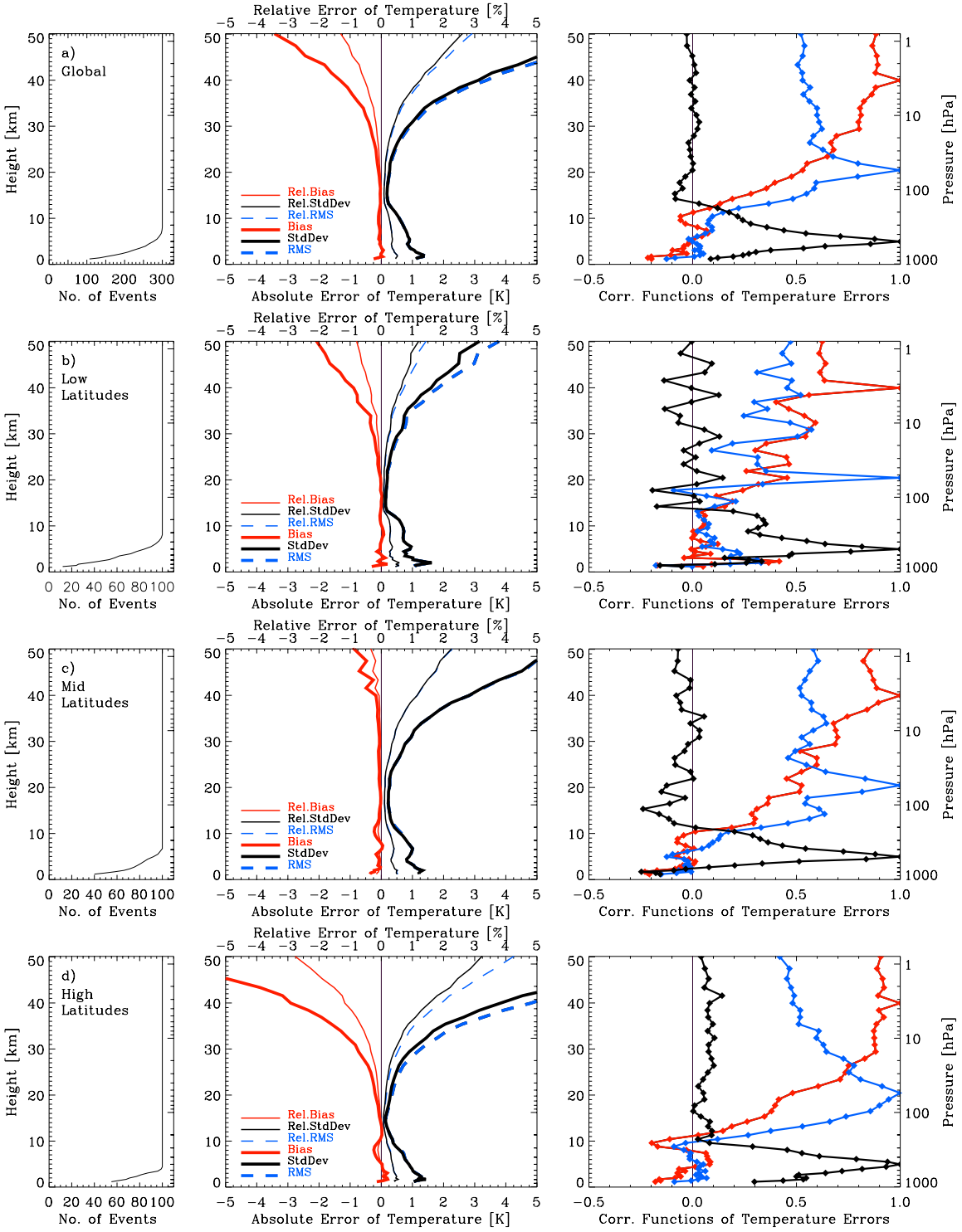


Figure 4.8. Temperature error analysis results, as a function of height, for the global (a) and the latitudinal (b–d) ensembles. Figure layout and style same as Figure 4.7; see that caption for details.

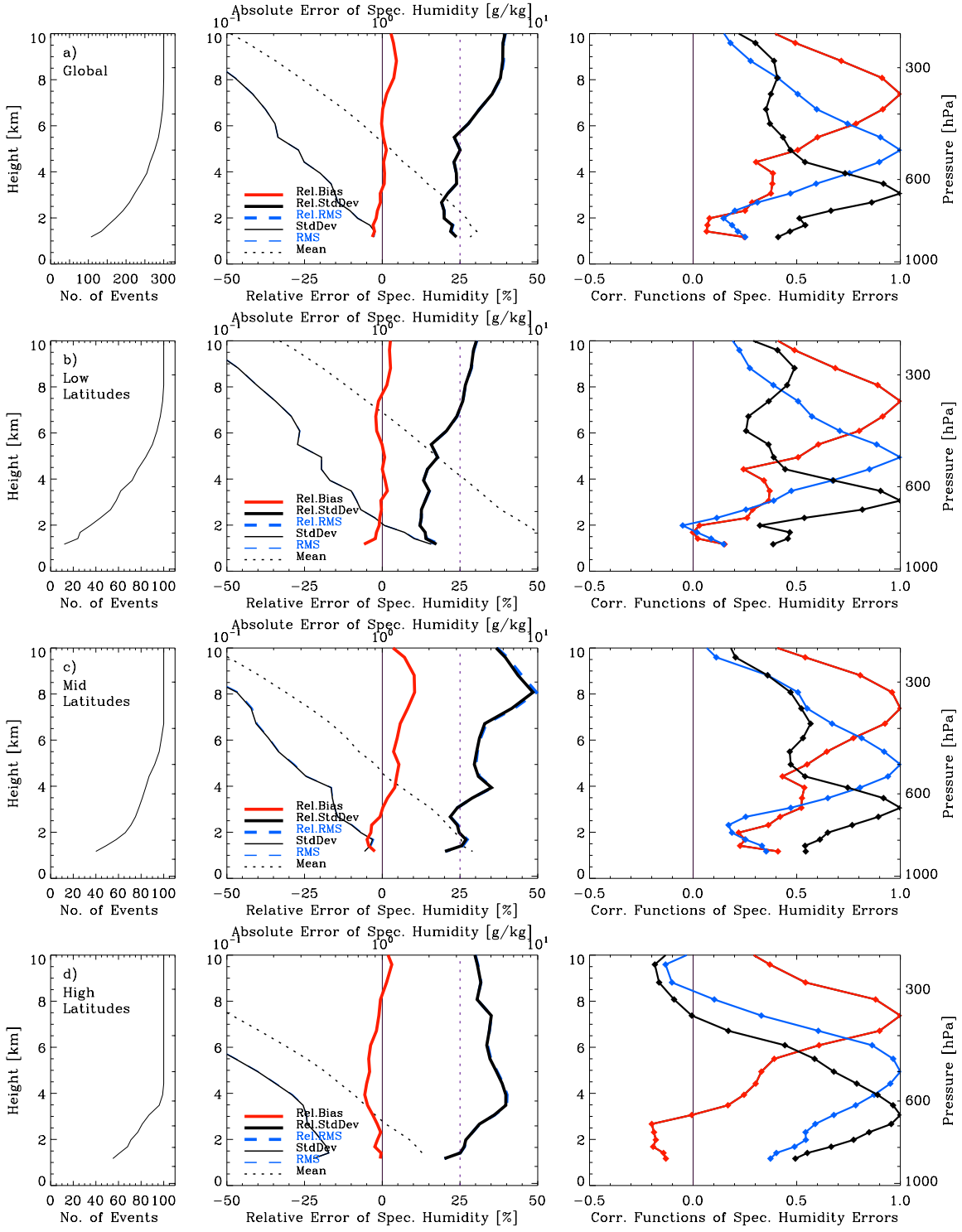


Figure 4.9. Specific humidity error analysis results, as a function of height, for the global (a) and the latitudinal (b–d) ensembles. Left panels: number of events used for the error statistics calculation at any given height. Middle panels: relative bias (red), relative standard deviation (heavy black), relative rms (heavy blue dashed) as well as the absolute standard deviation (light black), absolute rms (light blue dashed), and the mean of the "true" profiles (dotted). Right panels: error correlation functions for tropospheric height levels at ~7 km (red), ~5 km (blue), and ~3 km (black) height.

4.1.7 Error Covariance Matrices for Refractivity and other Data Products

Given that refractivity is particularly useful for data assimilation [e.g., *Healy and Eyre*, 2000; *Syndergaard et al.*, 2004], we deduced a simple observation error covariance matrix formulation for refractivity from the empirical estimation discussed above, for convenient use in data assimilation systems. We performed this in fitting simple analytical functions to the relative standard deviation profile and to the error correlation functions of refractivity. We also comment on the suitability of this model for retrieval products other than refractivity.

We used a least-squares fit of analytical functions to the empirical relative standard deviation, which exhibits two different types of height dependence, one above and one below the tropopause, respectively. Above the tropopause, the relative standard deviation of refractivity errors well follows an exponential increase (Figure 4.10, top; cf. Figure 4.5), at least up to the stratopause, which roots in the downward exponential decrease of upper boundary initialization errors [e.g., *Kursinski et al.*, 1997; *Rieder and Kirchengast*, 2001b]. The standard deviation model can thus be formulated as

$$s(z) = s_{TP} \cdot \exp\left[\left(z - z_{TP}\right) / H_{strat}\right], \quad \text{for } z_{TP} \le z < 50\,\text{km}, \qquad (4.10)$$

where z_{TP} is the tropopause level, s_{TP} is the relative (fractional or percentage) error at height z_{TP}, and H_{strat} is the scale height of the error increase over the stratosphere. In fitting our global refractivity results, we found z_{TP} = 15 km and s_{TP} = 0.075% adequate values at the tropopause level and, from a fit within 15–50 km, a best-fit value H_{strat} = 11.9 km for the error scale height. Figure 4.10, top-middle panel, illustrates this fit. The model may also be valid over some height domain beyond the stratopause, the vertical extend of which will depend on the exact handling of the upper boundary initialization. Furthermore, this type of error model over the stratosphere also holds reasonably well for the other retrieval products (see Figures 4.4 and 4.6 to 4.8, middle panels), with s_{TP} adjusted and H_{strat} fit to the empirical data of those products.

From the tropopause level downwards, the relative refractivity error is found to increase closely proportional to an inverse-height law ("$1/z^p$" law, with p = 1), which captures the increasing influence of horizontal variability errors [e.g., *Kursinski et al.*, 1997; *Foelsche and Kirchengast*, 2004a]. For this domain, the standard deviation model thus reads

$$s(z) = s_{TP} + s_0\left[\left(1/z^p\right) - \left(1/z_{TP}^p\right)\right], \quad \text{for } 2\,\text{km} < z \le z_{TP}, \qquad (4.11)$$

where z_{TP}, s_{TP} are the same as in equation (4.10), p is the order of the inverse-height dependence (not necessarily integer), and s_0 is the error at ~1 km ($z << z_{TP}$). In fitting our global refractivity results with p = 1 ("$1/z$") within 2–15 km, using the same z_{TP} and s_{TP} as for equation (4.10), which is mandatory at this joint boundary, we found a best-fit value s_0 = 4.46% for the ~1 km error, corresponding to an error of ~2% at 2 km. Figure 4.10, top-middle panel, illustrates the fit. From our fit we can recommend use of equation (4.11) down to 2 km, the physics-based limit we see is that it should be used only within the free troposphere above the planetary boundary layer, i.e., in general above 1–2 km only. Regarding other retrieval products, this "$1/z^p$" model is applicable also to bending angle (Figure 4.4), where equation (4.11) with p = 3 ("$1/z^3$" dependence) fits well down to an impact height of 4 km. For the

other parameters (after hydrostatic integration, optimal estimation) the structure is less simple. It is anyway error models for refractivity and bending angle, however, which are of main interest for data assimilation purposes.

Equations (4.10) and (4.11) can be used, via re-adjusting/-scaling/-fitting the parameters z_{TP}, s_{TP}, H_{strat}, s_0, p, to similarly approximate any other error profile data. An example is scaling to current RO data such as CHAMP data, which furnish somewhat less accuracy than the GRAS-type system baselined here. For this case we suggest that adjusting s_{TP} to 0.2–0.4% and re-scaling or re-fitting H_{strat} to near 15 km, keeping the other parameters the same, captures the main differences in refractivity errors between GRAS-type and CHAMP-type systems.

The second aspect of the error covariance matrices is the correlation structure. We approximated the shape of the refractivity error correlation functions at an L60 height grid by an exponential drop-off with a correlation length L = 2 km in the troposphere (up to 15 km), linearly decreasing to L = 1 km at 50 km height. Figure 4.10, right panel, illustrates that a reasonable fit of correlation functions is achieved by this exponential drop-off model with height-varying L. Combining the standard deviations from the model above with this error correlation modeling, we can write the global refractivity error covariance matrix, **S**, as

$$\mathbf{S} = S_{ij} = s_i s_j * \exp\left(-\frac{|z_i - z_j|}{L(z)} \right), \qquad (4.12)$$

where $L(z)$ is the (linearly) height dependent correlation length, z_i and z_j are the height levels between which the correlation is measured, s_i and s_j are the standard deviations (from equations (4.10) and (4.11)) at z_i and z_j, respectively, and S_{ij} are the covariance elements of **S** obtained. Considering use of equation (4.12) also for error covariance matrices of other retrieval products, it is found that it also reasonably applies to fit bending angle covariance matrices at the L60 grid, whereby in this case correlation lengths appear to be as small as 0.5–1 km (see Figure 4.4). For the parameters further retrieved from refractivity, correlation structure is more involved, as already noted above also for the standard deviation structure in the troposphere (Figures 4.6 to 4.8).

As described in section 4.1.6, negative correlations in refractivity can hardly be seen at an L60 height grid. They can be resolved at a finer grid, such as an L90 grid. Since these negative correlations are most pronounced in the low latitude ensemble at lower to middle stratospheric heights we also tried to model error correlation functions at a higher vertical resolution, comparable to an L90 grid, including the modeling of anti-correlation features. Figure 4.10, bottom panels, illustrates the results for the low latitude (middle panel) and global ensemble (right panel), respectively.

For this alternative modeling of **S**, we used a Mexican Hat function of the form (see, e.g., *Bronshtein* [2004] for details on this type of functions),

$$\mathbf{S} = S_{ij} = s_i s_j \cdot \left(1 - \frac{(z_i - z_j)^2}{(c \cdot L)^2} \right) \cdot \exp\left(-\frac{(z_i - z_j)^2}{L^2} \right), \qquad (4.13)$$

where, for the results of Figure 4.10, L was used as above ($L = L(z)$). For the global ensemble the stretching factor c was set to $c = 2$, and for the low latitude ensemble it was set to $c = 2$ below and to c = 0.8 above the tropopause level (15 km), respectively. Since the Mexican Hat function implies a Gaussian shape of the correlation structure, equation (4.13) is not generally invertible. This can be overcome by approximating the Gaussian factor, $f = \exp\left(-\frac{(z_i - z_j)^2}{L^2}\right)$, in equation (4.13) with a 5th order polynomial function as described by *Gaspari and Cohn* [1999]. Using a re-scaled height z defined by

$$z = \left|\frac{(z_i - z_j)\cdot\sqrt{0.3}\cdot\sqrt{2}}{L}\right|, \tag{4.14}$$

an approximate function $\tilde{f}$ can be written in the form,

$$\begin{aligned}\tilde{f} &= -\frac{z^5}{4} + \frac{z^4}{2} + \frac{5z^3}{8} - \frac{5z^2}{2} + 1, \quad 0 \le z \le 1, \\ \tilde{f} &= \frac{z^5}{12} - \frac{z^4}{2} + \frac{5z^3}{8} + \frac{5z^2}{3} - 5z + 4 - \frac{2}{3z}, \quad 1 < z \le 2.\end{aligned} \tag{4.15}$$

Substituting the Gaussian factor f with this approximate factor $\tilde{f}$, we can write **S** as

$$\mathbf{S} = S_{ij} = s_i s_j \cdot \left(1 - \frac{(z_i - z_j)^2}{(c\cdot L)^2}\right)\cdot \tilde{f}, \tag{4.16}$$

which is robustly invertible and delivers closely the same result as equation (4.13). When aiming at modeling the bending angle error correlation structure, which is for physical reasons showing more anti-correlation than refractivity (see discussion in 4.1.6), use of equation (4.16) instead of equation (4.12) will often also lead to superior results.

As can be seen in Figure 4.10, lower panels, the empirical error correlation values at the higher vertical resolution are approximated quite well with the Mexican Hat model. For the low latitude ensemble (middle panel) the anti-correlation structures are reasonably captured, while in the global ensemble (right panel) the negative correlations are smoothed already and consequently not modeled. This finding points also to the fact that use of global covariance matrices, and thereby disregarding latitudinal, and other, dependences of errors, is a rather crude approximation, in particular so at high vertical resolution. On the other hand, data assimilation systems are often not that sensitive to the precise formulation of the observation error covariance matrices employed, as long as it is reasonably adequate, but the exact sensitivities have to be individually investigated for each specific system (e.g., C. Marquardt and S.B. Healy, Met Office, Exeter, UK, priv. communications, 2003).

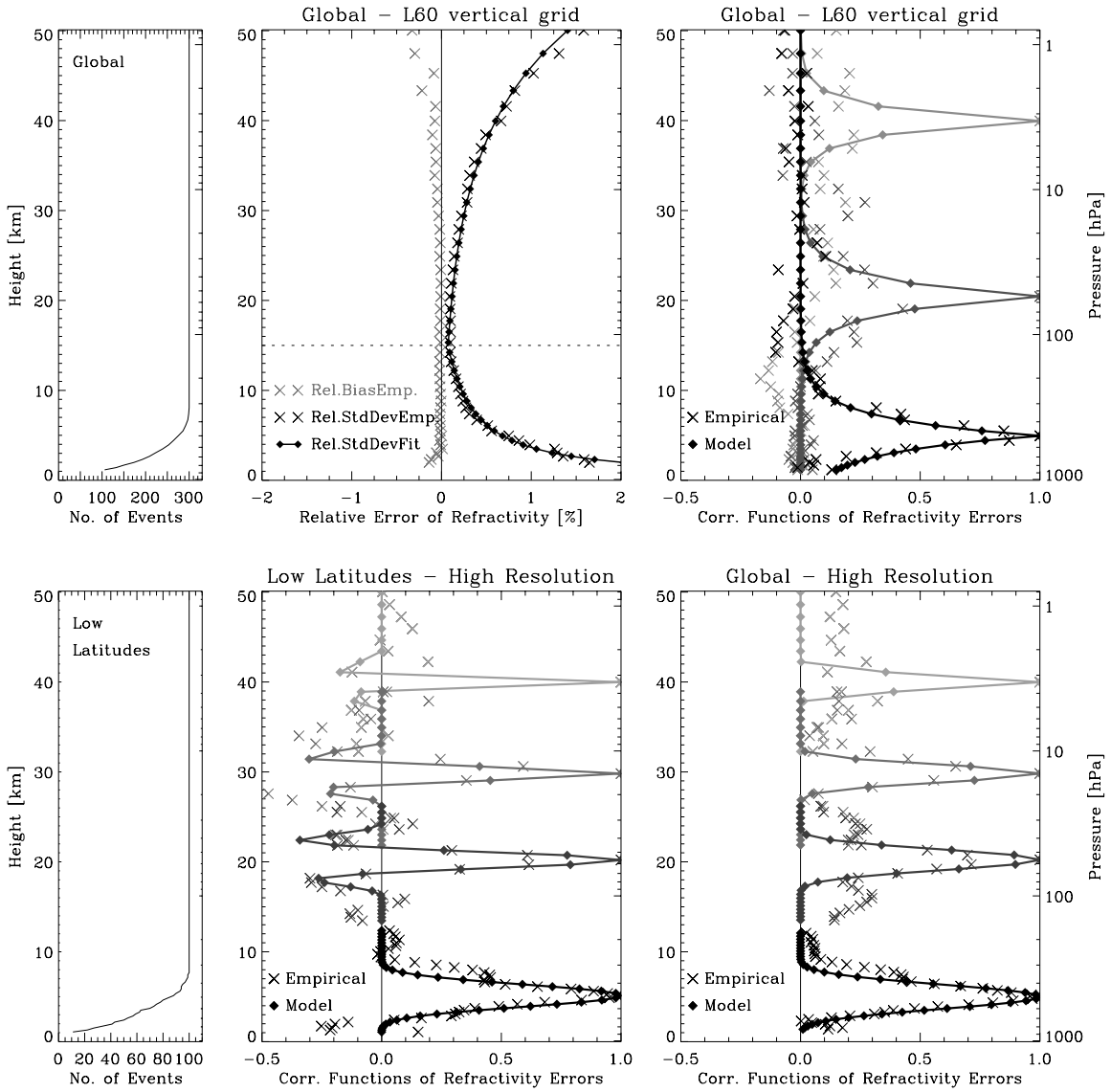


Figure 4.10. Analytical error models (connected diamonds) fitted to empirical values ("x" symbols) for the specification of refractivity error covariance matrices at an L60-type grid (upper panels) and an L90-type grid (bottom panels), respectively. For the standard L60-type grid results for the global ensemble are shown in terms of relative standard deviation (upper middle; also bias values indicated as grey "x" symbols) and error correlation functions (upper right; for ~40 km, ~20 km, ~5 km height). For the L90-type grid error correlation functions are shown for the low latitude ensemble (lower middle) and for the global ensemble (lower right), respectively; the functions shown for ~40 km, ~30 km, ~20 km, and ~5 km height. Left panels: number of events used for the error statistics calculation at any given height.

4.1.8 Summary, Conclusions, and Outlook

We presented results of an empirical error analysis for GNSS radio occultation (RO) data based on end-to-end forward-inverse simulations of an ensemble of RO events under fairly realistic modeling conditions. Starting with information on excess phase and Doppler errors, which are representative of the expected performance of the upcoming METOP/GRAS GPS

RO data, we provided a comprehensive error estimation for the atmospheric retrieval products bending angle, refractivity, pressure, geopotential height, temperature, and specific humidity.

The ensemble-based empirical error analysis approach allowed to avoid making assumptions such as linearity, un-biasedness, dry troposphere, a few standard atmospheric conditions only, spherically symmetric atmosphere, and no ionosphere, which were to varying degrees involved in previous error studies [e.g., *Kursinski et al.*,1997; *Syndergaard et al.*, 1999; *Palmer et al.*, 2000; *Healy and Eyre*, 2000; *Rieder and Kirchengast*, 2001b]. One limitation of realism also in this study is the use of geometric optics ray-tracing in the forward modeling, which implies that multipath and super-refraction conditions in the lower troposphere are not accounted for; all simulated profiles in the statistical ensembles stop above such conditions if they encounter them. Future consolidation of estimates will thus be valuable in particular below about 3 km at low latitudes, where multipath and super-refraction effects can play the largest role.

The present study advanced the earlier results in that it provided, based on the quasi-realistic datasets, both complete error covariance estimates (standard deviation profiles and correlation functions) and bias profile estimates, for all parameters from bending angle to humidity. In particular the bias estimation was an aspect not systematically addressed before, partly due to the fact that error analyses applying linear un-biased optimal estimation methodology [*Rodgers*, 2000], used by most previous studies, cannot deal with biases. Understanding of residual biases in retrievals from the self-calibrated and thus nominally un-biased raw RO data, and subsequent efforts to mitigate them, is crucial for climate applications.

Where comparable, the results of this study are consistent with the findings of theoretical studies [*Syndergaard*, 1999; *Rieder and Kirchengast*, 2001b]. When comparing and interpreting them in relation to the findings of the theoretical studies, one has to be aware of the differences between the studies, however, especially of the limiting assumptions in the theoretical studies like un-biasedness and dry air assumptions, but also of retrieval algorithm differences such as different resolution-accuracy trade-offs and background error specifications. This way the theoretical results obtained under more simplified conditions proved a valuable aid in the interpretation of the empirical results under more realistic conditions.

The results, furthermore, expanded and consolidated the rms error estimates obtained by *Kursinski et al.* [1997] based on a more simplified analysis. The empirically derived standard deviations of refractivity, geopotential height and temperature agree well with the rms errors estimated by *Kursinki et al.* [1997] in the upper troposphere and lower stratosphere; in lower troposphere refractivity we found a twice as large error due to the more realistic horizontally variable fields used in the present study. *Kursinski et al.* [1997] also looked into error budget components not addressed in this study on total retrieval errors. An error budget analysis for a GRAS-type receiving system was recently performed by *Ramsauer and Kirchengast* [2001].

The estimated bias profiles confirmed the high climate monitoring utility of RO data but also pointed to further improvement potential in the retrieval algorithms, especially related to the high-altitude initialization. In general, residual biases are found negligible below about 30 km in all data products and increasingly relevant throughout the upper stratosphere. More specifically, bending angle and refractivity biases start to become appreciable (systematically >0.25%) close to the stratopause only. Appreciable pressure, geopotential height and temperature biases are found to extend about 2 scale heights (~12–15 km) lower into the stratosphere than those of refractivity, being smallest at mid latitudes and reaching down to below 30 km into the lower stratosphere at high latitudes. These biases are caused by residual

mesospheric refractivity (equivalently, density) biases carried downward by the hydrostatic integration in the retrieval of pressure from refractivity. The density biases are, in turn, caused by residual mesospheric bending angle biases carried via the Abelian integration from bending angle to refractivity.

This insight implies that the key to further mitigation of upper stratospheric biases is further improvements in the statistical optimization of bending angles, i.e., in their optimal estimation from retrieved and background bending angles. This finding has already prompted further improvement of the bending angle retrieval, focused on minimizing biases in background bending angle profiles before the linear un-biased optimal combination with retrieved bending angle profiles is done, which led to significant bias reduction, especially at high latitudes where it is most needed (see section 2) [*Gobiet and Kirchengast*, 2004a; *Gobiet et al.*, 2004b]. The frequently used alternative way of pressure, geopotential height, and temperature bias mitigation by a "2nd initialization" at pressure retrieval — of the hydrostatic integral at a height within 30–50 km by an *a priori* model temperature [e.g., *Kursinski et al.*, 1997; *Wickert et al.*, 2004; *Hajj et al.*, 2004] — is discouraged if the retrieved profiles target climatological use. The approach is problematic, since it leads to intricate error characteristics and *a priori* dependence in the stratospheric profiles down to about 20 km, which threats the crucial climate requirements of un-biasedness and of a clear understanding of the systematic and statistical error properties.

An encouraging result from the climatological point of view is, furthermore, that biases in specific humidity are found small within 5% at all latitudes and heights, except at mid latitudes above 6 km, where biases reach up to 10%. This finding, together with the small biases also found in temperature below 10 km, indicates that optimal estimation (1D-Var) retrieval based on ECMWF short-term forecasts as background can lead to tropospheric humidity and temperature retrievals clearly of climatological utility. While there is other approaches still in use, usually based on *a priori* temperature aiding observed refractivity for humidity retrieval [e.g., *Kursinski et al.*, 1997], we thus rate 1D-Var retrieval the method of choice for humidity and temperature retrieval from GNSS RO refractivity in the troposphere [cf. *Healy and Eyre*, 2000].

Regarding error correlation functions, the results provide insight into covariance propagation under quasi-realistic conditions. The main differences between the empirically estimated correlations and theoretical ones are seen in the troposphere, including broader refractivity error correlation functions due to horizontal variability. The 1 D-Var retrieval does not significantly broaden the correlation functions of temperature and humidity compared to the ones of refractivity. Negative correlation minima adjacent to the main correlation function peaks of bending angle and refractivity can be hardly resolved on the L60 vertical grid. Nevertheless, they are present and found resolved on an L90 grid; they are found most pronounced at stratospheric heights at low latitudes, where they are comparable to the error correlation features in theoretical studies [*Syndergaard*, 1999; *Rieder and Kirchengast*, 2001]. The hydrostatic integration process prevents any appreciable negative correlations in pressure, geopotential height, temperature, and humidity error profiles.

We derived simple analytical formulations of refractivity error covariance matrices by closely fitting the empirically estimated matrices. The functional formulations depend on a few sensibly selected adjustment/fitting parameters only. While we provided values for these parameters based on our empirical datasets reflecting GRAS-type performance, the formulations can be used, via re-adjusting/re-scaling/re-fitting the parameters, to approximate any other error estimation dataset; as an example, scaling to CHAMP RO data was addressed. Advice on how to potentially use the formulations as well for retrieval products other than

refractivity, in particular for bending angles, was also provided. Comparison of the formulations with results of an error analysis of CHAMP RO refractivity profiles confirmed their practical utility [*Steiner*, 2004].

The error covariance matrices, especially those for refractivity and bending angle, should be useful for evaluation and proper specification of observational errors in data assimilation systems as well as in optimal estimation parts of retrieval algorithms, i.e., the optimal bending angle estimation part and the 1D-Var temperature and humidity retrieval part. Variants of the simple analytical formulations might be of particular convenience for use in large-scale operational data assimilation systems. In utilizing the results it is to be kept in mind that retrieval products from different measurement and retrieval processes than used in this study, i.e., other than GRAS-type performance and different retrieval algorithms, will usually need scaling of the results to those different conditions, if not repeat of the error analysis itself, in case of severe differences.

(intentionally left blank, back page if double-sided print)

4.2 Error Analyses of Refractivity Profiles Retrieved from CHAMP Radio Occultation Data

4.2.1 Introduction

The assimilation of radio occultation data has the potential to significantly improve the accuracy of global and regional meteorological analysis and weather prediction which has been confirmed by several studies [e.g., *Kuo et al*., 2000; *Healy et al*., 2003]. One important issue in this respect is the knowledge on radio occultation measurement errors in order to formulate proper observation error covariance matrices for data assimilation systems.

Since refractivity seems to be the most appropriate parameter for assimilation purposes we performed an empirical error analysis of a set of refractivity profiles retrieved from CHAMP radio occultation observations building on the heritage of the simulation study (section 4.1). The data comprise one week of observations, January 1–7, 2003, with about 1200 radio occultation profiles equally distributed over the globe. The processing was performed with our CHAMPCLIM Retrieval version 2 (CCRv2) which is described in detail in section 2.2. The error statistics is based on comparison to reference profiles calculated from ECMWF analyses fields. We analyzed refractivity error characteristics comprising bias, standard deviation and error correlation functions by applying the same statistical methods as described in section 4.1. These error characteristics contain both the observational error of the retrieved data and the model error of the ECMWF analyses. For separate specification of the observation error we estimated the ECMWF model error based on the propagation of temperature, humidity, and pressure errors as provided by ECMWF into refractivity. The results of this study are compared to the findings of *Kuo et al.* [2003] who analyzed CHAMP and SAC-C (Satelite de Aplicaciones Cientificas-C) data. We also compare to the results of our simulation study being the precursor of this work as presented in the previous section 4.1 of this report [*Steiner and Kirchengast*, 2004a].

Section 4.2.2 gives a brief description of the data set and the retrieval algorithms. In section 4.2.3 the estimation of the combined error (CHAMP RO + ECMWF) is described. The ECMWF model error is estimated in section 4.2.4 and the results on the observed refractivity error are presented in section 4.2.5. Summary and conclusions are drawn in section 4.2.6.

4.2.2 Description of the Data Set and the Retrieval Scheme

This study is based on a CHAMP level 2 data set comprising one week of radio occultation observations, January 1–7, 2003, with 1253 profiles of atmospheric excess phases in total. Refractivity profiles were calculated from this data base of excess phase profiles by using CCRv2 which is described in detail in section 2.2 of this report.

In brief, the retrieval scheme is based on the standard geometric optics approach for the calculation of bending angle profiles and on the Abel inversion technique for the calculation of refractivity profiles [see e.g., *Hocke*, 1997; *Steiner et al*., 1999; *Gobiet and Kirchengast*, 2004]. Basically, the processing chain from excess phases via bending angles to refractivity involves smoothing of the data by applying a regularization routine [*Syndergaard*, 1999] and an ionospheric correction scheme based on the ionospheric correction of bending angles [*Vorob'ev and Krasil'nikova*, 1994]. The measured bending angle profiles are then initialized

with background bending angle profiles. Since this study deals with the use of refractivity data for assimilation systems we use the MSISE-90 model for the initialization of bending angles (see detailed discussion on different initialization schemes and the data usage in section 2.2.9.2).

Using only data with quality flag 0 (see section 2.2.4 on quality control) a global ensemble of 1136 retrieved refractivity profiles was analyzed. In addition the data sets were separated into three latitude bands, low (–30° to +30°), middle (±30° to ±60°), and high (±60° to ±90°) latitudes which is illustrated by Figure 4.11. The profiles are almost equally distributed over the globe with 356, 410, and 370 profiles in low, mid-, and high latitudes, respectively. The number of profiles for the ensemble statistics decreases with decreasing height giving at the lowest heights 73 profiles globally and only 12 profiles in the low latitude ensemble. The height at which more than 50% of the events are available for the error statistics is varying from ~2.5 km in high latitudes, to ~5 km in mid latitudes and for the global ensemble, and ~6.5 km in low latitudes. Since in addition to this limitation the retrieval is based on geometric optics we will not interpret the results below 5 km height.

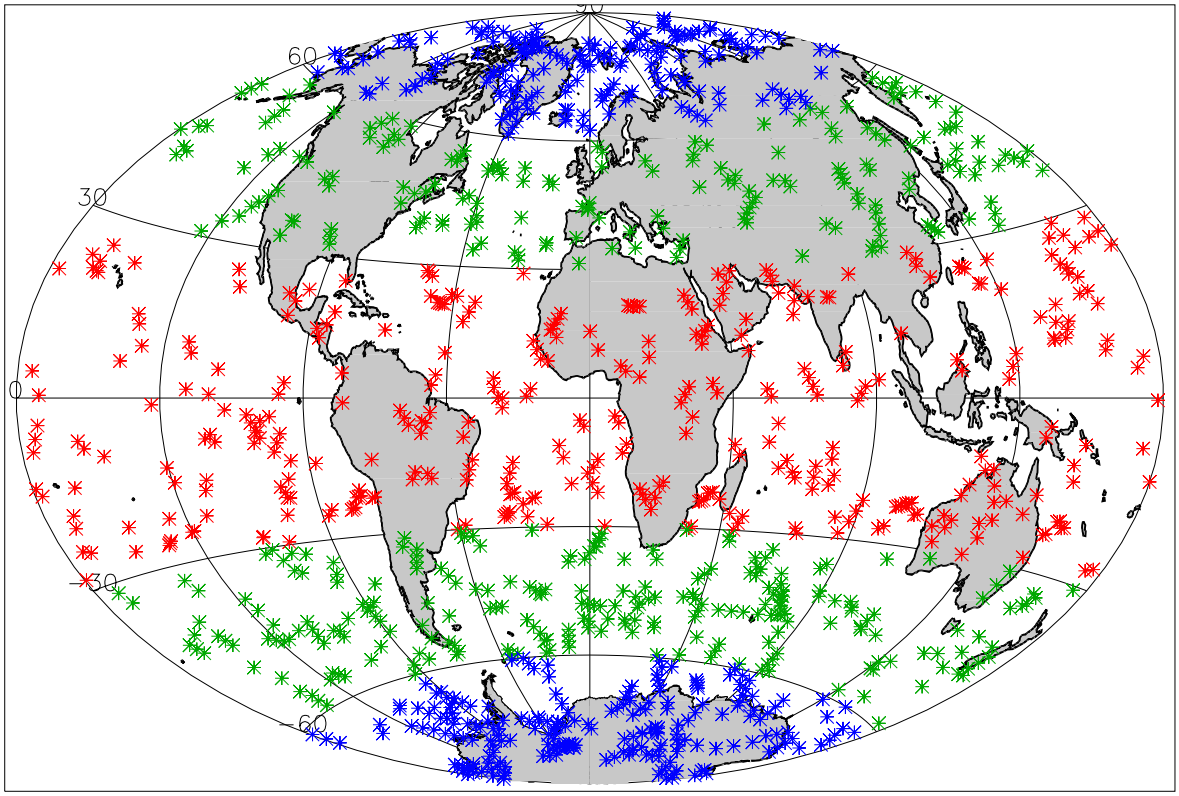


Figure 4.11. Distribution of refractivity profiles from CHAMP occultation observations for January, 1–7, 2003, using the EGOPS/CC retrieval. The latitudinal circles show the separation into three latitude bands, low (–30° to +30°), middle (±30° to ±60°), and high (±60° to ±90°) latitudes.

4.2.3 Estimation of the Combined Error (CHAMP RO + ECMWF)

The error statistics is based on the comparison of the retrieved and smoothed (comparable to ECMWF grid resolution) refractivity profiles with colocated refractivity profiles derived from 6 hourly operational meteorological analyses fields from ECMWF. The colocated vertical ECMWF profiles were calculated at a fixed mean tangent point location. When regarding the ECMWF profiles as the truth this implies that the error estimates represent an upper bound

error estimate including the observation error, the model (ECMWF) error and the representativeness error. The representativeness error stems from the limited spatial and temporal model resolution and from the comparison of the retrieved profiles with vertical reference profiles. This fact becomes important in the lower troposphere, below ~7 km, when high horizontal variability is present (see detailed discussion in section 3 of this report) [*Foelsche and Kirchengast*, 2004b; *Syndergaard et al.*, 2004]. Since we will not interpret results in the lower troposphere the representativeness aspect is negligible in this respect.

The statistical method for calculating the combined (observed error + ECMWF error) error is described in detail in section 4.1.5 (Equation 4.3 to 4.9). Bias profiles and error covariance matrices are provided, the latter separated into standard deviation profiles and error correlation matrices.

The resulting error statistics for the combined refractivity error is shown in Figure 4.12 for the global ensemble (panels row a) and the latitudinal data sets (panels row b–d) up to 50 km height. The left small panels show the number of events entering the statistics at a given height. The middle panels illustrate the error characteristics comprising bias, standard deviation, and rms profiles in terms of absolute quantities shown at the upper x-axis and in terms of relative quantities displayed at the lower x-axis. We discuss relative error quantities for refractivity which are denoted by heavy lines, solid black for the standard deviation, solid red for the bias, and dashed blue for the rms error, respectively. Mean atmospheric profiles are depicted by light dotted black lines. The domain of main interest for the interpretation of the results is 5–35 km height (denoted by dotted lines), the lower boundary being due to the geometric optics retrieval, the upper boundary being due the initialization with background data (see also section 2.2.10). The right panels display the error correlation functions for different height levels (~40 km, ~30 km, ~20 km, ~10 km, ~5 km) representative of upper stratosphere, lower stratosphere, and troposphere.

The relative refractivity bias of CHAMP RO with respect to ECMWF oscillates around –0.4% at 5 –25 km globally and in mid- and high latitudes increasing to 0.5% at 35 km and to 3% at ~40 km. Strong bias oscillations are seen in low latitudes ranging from –0.4% to 0.5% at 5 –35 km. Prominent structures appear at tropopause heights most probably stemming from the higher resolved tropopause in CHAMP RO data than in the ECMWF data.

The combined relative standard deviation (Rel.StdDev) is of the order of 0.7–1% at 5–25 km height globally and in low latitudes while it oscillates around 0.75% at these heights in mid- and high latitudes. It increases to 1.5% at 35 km and to 3% at 40 km in all ensembles.

The refractivity error correlation functions show the same characteristics in the global ensemble as well as in the latitudinal data sets but they show different characteristics in different height domains. In the troposphere, the broadening seen in the refractivity error correlation functions is mostly a result of the errors due to horizontal variations while the broadening in the uppermost error correlation function at 40 km height reveals mainly the influence of the background model which was used for the initialization at bending angle level. In stratospheric heights small correlations are seen with the error correlation functions being relatively narrow. Comparison with the results of the simulation study presented in section 4.1.6 (Figure 4.5) shows that the refractivity error correlation features are quite similar except above 35 km. At these heights the higher bending angle error of the CHAMP data than of the simulated RO data for a METOP/GRAS system (about a factor of 2) results in a stronger impact of the model data in the statistical optimization process.

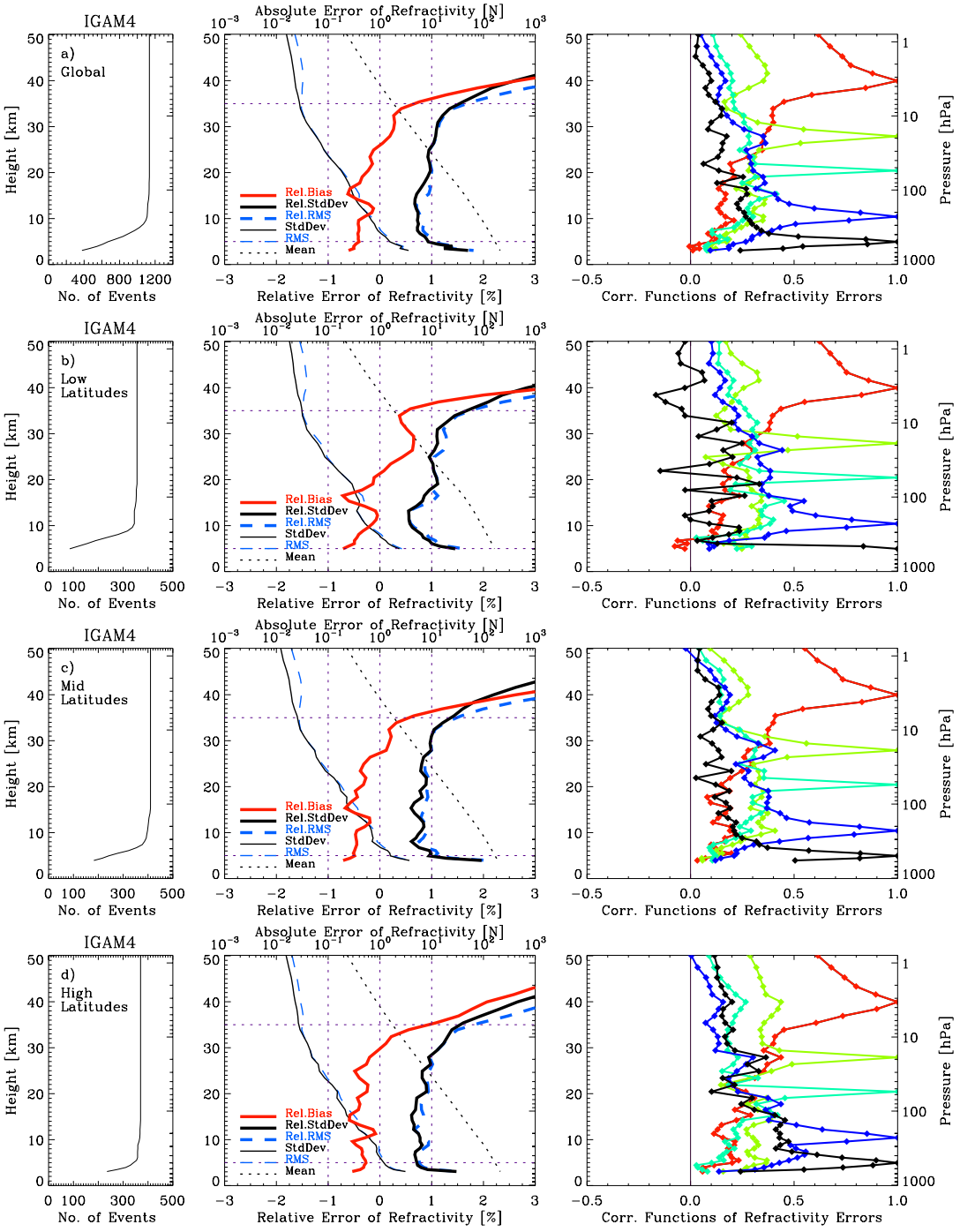


Figure 4.12. Combined refractivity error (CHAMP RO + ECMWF), as a function of height, for the global (a) and the latitudinal (b–d) ensembles. Left panels: number of events used for the error statistics calculation at any given height. Middle panels: relative bias (red), relative standard deviation (heavy black), relative rms (heavy blue dashed) as well as the absolute standard deviation (light black), absolute rms (light dashed), and the mean of the "true" profiles (dotted). Right panels: error correlation functions for ~40 km (red), ~30 km (yellow), ~20 km (green), ~10 km (blue), and ~5 km (black) height, representative of upper stratosphere, lower stratosphere, and troposphere, respectively.

4.2.4 Estimation of the ECMWF Error

In order to separate the observed error of the CHAMP RO refractivity retrievals from the combined error we performed a global estimate of the ECMWF refractivity model error. M. Fisher (ECMWF, Reading, U.K.) provided us with global error estimates of ECMWF analyses in form of standard deviations for temperature, specific humidity, and surface pressure and of vertical error correlations for temperature and specific humidity. Temperature T [K], water vapor pressure e [hPa], and total pressure p [hPa] is related to refractivity N [N units] via the Smith-Weintraub formula [*Smith and Weintraub*, 1953],

$$N = c_1 \frac{p}{T} + c_2 \frac{e}{T^2}, \tag{4.17}$$

with the constants $c_1 = 77.6$ K/hPa and $c_2 = 3.73*10^5$ K^2/hPa. Water vapor pressure in equation 4.17 was substituted for specific humidity q [kg/kg] using the following relation,

$$e = \frac{p \cdot q}{a + bq}, \tag{4.18}$$

with $a = 0.622$ and $b = 0.378$. A simple error propagation based on equation 4.17 then was applied via

$$\Delta N = \sqrt{\left(\frac{\partial N}{\partial T}\Delta T\right)^2 + \left(\frac{\partial N}{\partial q}\Delta q\right)^2 + \left(\frac{\partial N}{\partial p}\Delta p\right)^2} \tag{4.19}$$

in order to calculate the standard deviation of refractivity. The pressure error at a given height was calculated by error propagation using the given standard deviation of surface pressure of 250 Pa. The vertical refractivity error correlations were derived by a weighted combination of temperature error correlations ($w_T = \Delta N(\Delta p, \Delta T) / \Delta N(\Delta p, \Delta q, \Delta T)$) and specific humidity error correlations ($w_q = \Delta N(\Delta q) / \Delta N(\Delta p, \Delta q, \Delta T)$).

Figure 4.13 presents ECMWF error specifications for temperature (left panels), specific humidity (middle panel), and for estimated refractivity (right panels). Standard deviations are shown in the upper panel row and error correlation functions for three different height levels (~10 km, ~20 km, ~30 km) are displayed in the lower panel row.
We tested the sensitivity of the estimated refractivity error with respect to the temperature error input for the four cases displayed in Figure 4.13, where we multiplied the temperature standard deviation by 1 (blue), 1.5 (green), 2 (red), and 2.5 (black). The most reasonable result was found for the case of doubling the temperature standard deviation (red) giving a Rel.StdDev of ECMWF refractivity of the order of 0.5% at 8–15 km increasing to 0.75% at 30 km and to 1% at 35 km. These results are consistent with the findings of *Kuo et al.,* [2004] who performed an estimation of short-range forecast errors using the Hollingsworth-Lönnberg method [*Hollingsworth and Lönnberg*, 1986]. For comparison we included their estimates for low latitudes (dotted) and mid-latitudes (dashed) in Figure 4.13 (right, upper panel).

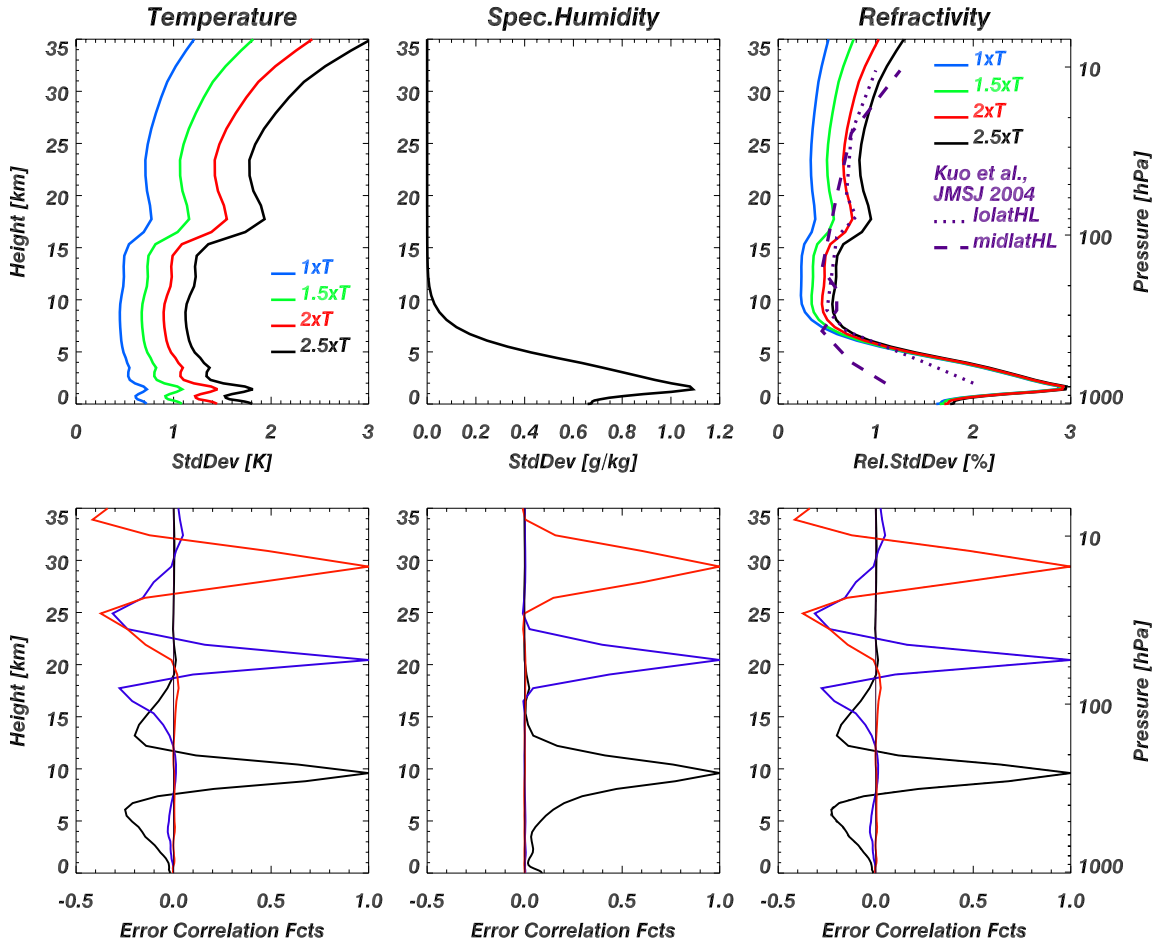


Figure 4.13. ECMWF error for temperature (left) for 4 test cases (1xT, 1.5xT, 2xT, 2.5xT), specific humidity (middle), and corresponding estimated refractivity (right) in terms of standard deviation (upper panels) and error correlation functions (lower panels), the latter shown for three heights (~10 km (black), ~20 km (blue), ~30 km (red)). Estimates of short-range forecast errors for refractivity for low latitudes (dotted) and mid-latitudes (dashed) made by *Kuo et al.* [2004] are shown.

4.2.5 Estimation of the Observation Error

The observed refractivity error was then derived by subtracting the ECMWF error from the combined error in terms of variances s^2,

$$(s_{obs})^2 = (s_{combined})^2 - (s_{ECMWF})^2 . \qquad (4.20)$$

The results are displayed in Figure 4.14 for the global data set with the number of events at a given height shown in the left panel. The right panel shows the combined error (black), the ECMWF error for the 2xT case (blue), and the corresponding observation error (red) in terms of Rel.StdDev, respectively. The relative bias of CHAMP RO with respect to ECMWF (green) has already been discussed in section 4.2.3 (Figure 4.12).

The observed Rel.StdDev for CHAMP refractivity in the 2xT case is of the order of 0.5% at 6–18 km increasing to 0.75% at 30 km and to 1.2% at 35 km. Our observation errors seem to be a more conservative estimate compared to the results of *Kuo et al.*, [2004], who found

the observation error of refractivity to be of the order of 0.3–0.5% at 5–25 km but did not investigate error correlations in their study.

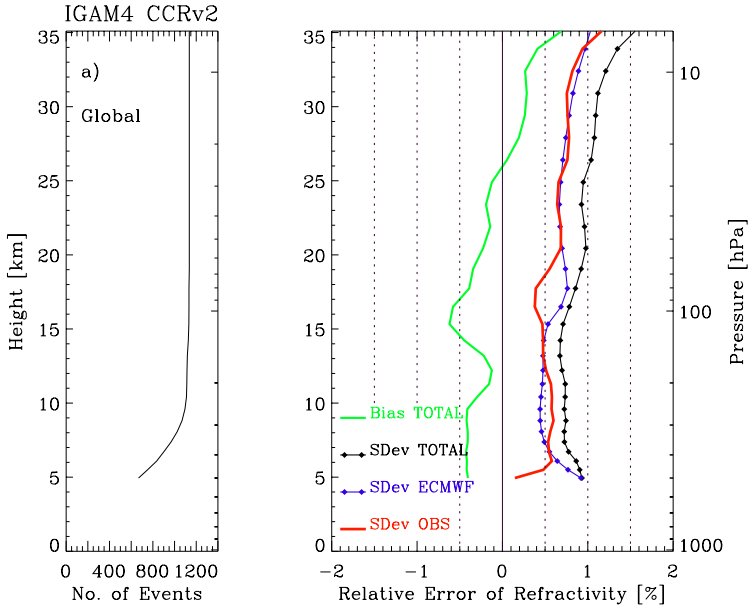


Figure 4.14. Left panel shows the number of events at any given height. Right panel shows the estimated refractivity errors: Relative bias of CHAMP RO with respect to ECMWF (green), relative standard deviation: combined (black), ECMW model error (blue), observation error (red).

As a further result refractivity error correlation functions are displayed in Figure 4.15 for three different heights, ~10 km, ~20 km, ~30 km, representative for troposphere, lower and upper stratosphere. The ECMWF refractivity error correlation functions (dotted) show negative correlation features in the vicinity of the peaks while the correlation functions for the combined error (solid) show a flattening. These features suggest that the correlation wings are dominated by the observed data which seems also to be confirmed when comparing to the results of the simulation study (see Figure 4.5). For the construction of refractivity observation error covariance matrices for data assimilation systems we therefore suggest a combination of the observed Rel.StdDev with the total error correlation matrix.

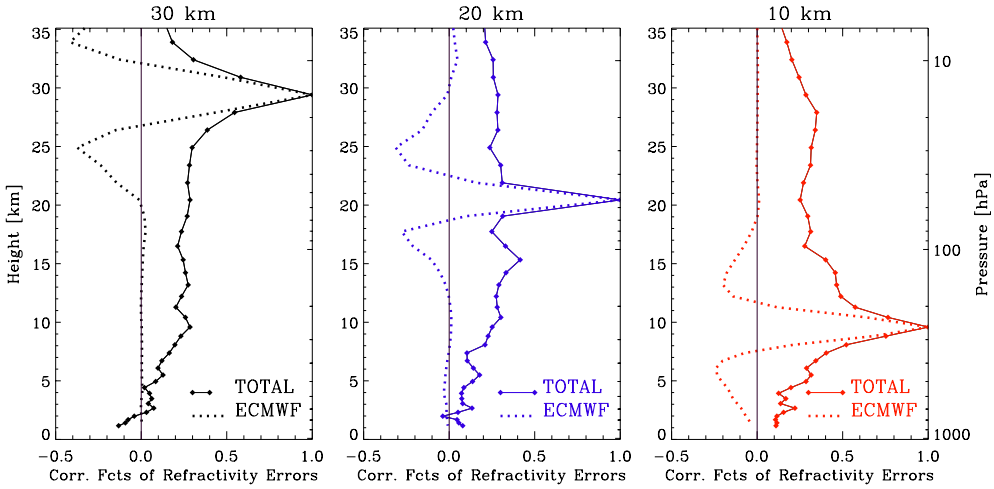


Figure 4.15. Error correlation functions for the combined refractivity error (solid line) and the ECMWF refractivity error (dotted) shown for three heights ~30 km (black), ~20 km (blue), ~10 km (red).

4.2.6 Summary and Conclusions

As a follow-on study to the empirical error analysis of simulated RO data (section 4.1) we performed an error analysis of CHAMP RO refractivity profiles processed with the CCRv2 for one week of observations, the data set consisting of about 1200 profiles. The error statistics was based on a comparison to reference profiles from ECMWF analyses fields implying that the statistics includes both, the observation error and the ECMWF model error. In order to separate the errors we then performed an error estimation of the ECMWF error based on error propagation of temperature, humidity, and pressure error into refractivity error. Finally, the subtraction of the ECMWF error from the combined error allowed an estimation of the global observation error.

The relative refractivity bias of CHAMP RO with respect to ECMWF was found to oscillate around –0.4% at 5–25 km globally and in mid- and high latitudes and to increase to 0.5% at 35 km. Strong bias oscillations in low latitudes range from –0.4% to 0.5% at 5–35 km. Prominent structures appear at tropopause heights most probably stemming from the higher resolved tropopause in CHAMP RO data than in the ECMWF data. The combined Rel.StdDev of refractivity (CHAMP + ECMWF) was found to be 0.7–1% at 5–25 km height increasing to 1.5% at 35 km. The refractivity error correlation functions (CHAMP + ECMWF) show different characteristics in different height domains. In the troposphere, the broadening is mostly a result of the errors due to horizontal variations while in the upper stratosphere the broadening reveals mainly the influence of the background model which was used for the initialization at bending angle level. In stratospheric heights <35 km small correlations are seen resulting in a narrowing of the error correlation functions. The refractivity error correlation features are quite similar to those presented in the simulation study in section 4.1.6 except above 35 km. At these heights the higher bending angle error of the CHAMP data than of the simulated RO data for a METOP/GRAS system (about a factor 2) results in a stronger impact of the model data through the statistical optimization process.

Using reasonable ECMWF temperature errors (2xT case) we found the global estimate of the ECMWF refractivity error to be 0.5% at 8–15 km increasing to 0.75% (1%) at 30 km (35 km), these results being quite consistent with the findings of *Kuo et al.* [2004].

For the observation error we found a Rel.StdDev of refractivity of 0.5% at 6–18 km increasing to 0.75% at 30 km and to 1.2% at 35 km. This is a more conservative estimate but still in line with the results of *Kuo et al.*, [2004] who found an observation error of 0.3–0.5% at 5–25 km.

In terms of error correlations the ECMWF refractivity error correlation functions show negative correlation features in the vicinity of the peaks while the correlation functions for the combined error show a flattening. These features suggest that the correlation wings are dominated by the observed data which may also be confirmed by the similar structure of the correlation functions seen in the previous simulation study. We therefore suggest a combination of the observed Rel.StdDev with the total error correlation matrix for the construction of observation error covariance matrices for data assimilation systems. These observation error covariance matrices can also be approximated with the simple analytical functions presented in section 4.1.7 by using an exponential drop off for the error correlations and by scaling the error magnitude of the Rel.StdDev models at the tropopause.

In currently on-going work the described error analysis is applied to bigger CHAMP RO data sets, where we investigate regional and seasonal aspects in detail. In this respect we also plan further investigations regarding a more advanced estimation of ECMWF errors.

5 Summary, Conclusions, and Outlook

This report on CHAMPCLIM work package 3 "Radio Occultation Data Processing Advancements for Optimizing Climate Utility" presented the achievements obtained with respect to the retrieval advancements and the error assessment of RO observations. Several critical performance issues of occultation data processing in the stratosphere region were investigated and improved, in particular ionospheric correction and statistical optimization of bending angles. Furthermore, the retrieval sensitivity in the troposphere was tested, focusing on atmospheric horizontal variability errors and the angle of incidence as well as the geometry of reference profiles. Finally, a complete empirical error analysis of RO retrieval products was performed giving insight into the propagation of errors from bending angle profiles to final atmospheric data products and resulting in the definition of adequate error covariance matrices for use in assimilation systems. The essential outcome of this work was the installation of a robust CHAMPCLIM Retrieval scheme (IGAM/CCRv2 scheme) and a reasonable error characterization for CHAMP/GPS RO data in fulfillment of the aim to improve the maturity and utility of the data products especially for climatological purposes.

The developed IGAM/CCRv2 is a stable retrieval scheme for CHAMP data with quality control mechanisms included that enable a pre-operational processing of CHAMP data.

A robust ionospheric correction scheme based on smoothed bending angles and impact parameters (1 km boxcar filter) was implemented yielding an improved profile delivery rate and less variance in the retrieval results in the 15–25 km altitude region, without removing effects from small-scale structures in the neutral atmosphere.

The statistical optimization schemes implemented in CCRv2 are based on the aim of optimally combining observed and background bending angle data under consideration of their error characteristics. Further improvement of retrieval performance in the upper stratosphere was obtained through background bias correction. The "enhanced IGAM scheme" incorporates a 2-step empirical background bias correction: a first step searches and extracts a best-fit profile from a suitable library of profiles, a second step adjusts this best-fit profile towards the observed profile. The exact height ranges for fitting depend on the observational noise in the RO data, the higher the noise the lower down the fitting height range. Due to the limited quality of CHAMP high-altitude data, the observation error estimation and background bias correction procedures had to be adjusted, resulting in reduced quality of the "enhanced IGAM scheme" using climatological background information (IGAM/MSIS scheme). Still this scheme, which has the advantage to be independent from NWP background data, shows a performance similar to other state-of-the-art retrieval schemes below 35 km and has a big potential for RO missions using further developed RO receivers yielding better data at high altitudes like GRAS/MetOp or COSMIC.

Since it is acceptable for many applications to use slightly NWP-dependent data at high altitudes, we developed a second retrieval mode using ECMWF operational analyses as background information (IGAM/ECMWF scheme). It shows relative independency from the background up to about 40 km whilst the retrieval performance, especially for temperature profiles, is significantly improved compared to the IGAM/MSIS scheme.

It is recommended to use IGAM/MSIS retrieval results on refractivity level and below 35 km if strict independence from ECMWF analyses is required. If strict independence is not

required and/or if temperature data is needed, we recommend to use results from the IGAM/ECMWF retrieval scheme up to 35 km.

After statistical optimization, the CCRv2 scheme applies the inverse Abel transform to derive refractivity and a hydrostatic integration to derive pressure and temperature without further inclusion of background information. This is a major difference to most other RO retrieval schemes. The frequently used alternative way of pressure, geopotential height, and temperature bias mitigation by a "2nd initialization" at pressure retrieval using an *a priori* model temperature is discouraged if the retrieved profiles target climatological use. The approach is problematic, since it leads to intricate error characteristics and *a priori* dependence in the stratospheric profiles down to about 20 km, which threats the crucial climate requirements of un-biasedness and of a clear understanding of the systematic and statistical error properties.

Investigation of the sensitivity of atmospheric RO profiles to atmospheric horizontal variability errors in the troposphere showed that biases and standard deviations are significantly smaller under a spherical symmetry assumption than corresponding errors in a quasi-realistic atmosphere with horizontal variability. The differences are most pronounced below ~7 km height. Temperature standard deviations, for example, remain smaller than 1 K in a spherically symmetric atmosphere, while they reach values of about 5 K in the horizontally variable atmosphere. Dry temperature profiles between 7 km and 20 km were found to be essentially bias-free in both the horizontal variability and spherical-symmetry scenarios (biases smaller than 0.1 K), which confirms the unique climate monitoring utility of GNSS occultation data.

A significant part of the total error below ~7 km can be attributed to adopting reference profiles vertically at mean tangent point locations instead of extracting them along actual 3D tangent point trajectories through the troposphere. Future work will investigate by how much the standard deviation and bias errors decrease if a tangent point trajectory deduced purely from observed data is used instead.

Below about 7 km most errors were found to increase with increasing angle of incidence. This is in line with the hypothesis that larger angles of incidence lead to more sensitivity to horizontal variability. In general, the sensitivity of bias errors to increases of the angle of incidence has been found to be relatively small, which is favorable regarding the climate monitoring utility of the data. For example, dry temperature biases between 7 km and 20 km exhibit no relevant increase with increasing angle of incidence.

The empirical error analysis for GNSS RO data, based on the quasi-realistic datasets for a METOP/GRAS system, provided both complete error covariance estimates (standard deviation profiles and correlation functions) and bias profile estimates, for all parameters from bending angle, via refractivity, pressure, geopotential height, and temperature to humidity.

The estimated bias profiles confirmed the high climate monitoring utility of RO data. In general, residual biases were found negligible below about 30 km in all data products and increasingly relevant throughout the upper stratosphere. More specifically, bending angle and refractivity biases start to become appreciable (systematically >0.25%) close to the stratopause only. Appreciable pressure, geopotential height and temperature biases are found to extend about 2 scale heights (~12–15 km) lower into the stratosphere than those of refractivity due to hydrostatic integration. This result confirms our view that a 2nd a priori information should not be introduced in the retrieval. Another encouraging result from the

climatological point of view is that biases in specific humidity are found small within 5% at all latitudes and heights, except at mid latitudes above 6 km. This finding, together with the small biases also found in temperature below 10 km, indicates that optimal estimation (1D-Var) retrieval based on ECMWF short-term forecasts as background can lead to tropospheric humidity and temperature retrievals clearly of climatological utility.

Building on the heritage of the simulation study we performed an error analysis of CHAMP RO refractivity profiles The relative refractivity bias of CHAMP RO with respect to ECMWF was found to amount to around –0.4% at 5 –25 km height globally with the exception of larger deviations in low latitudes. The IGAM/ECMWF retrieval identifies this small bias quite clearly as being height-constant, which will help to eliminate it in future algorithm improvements.

For a separation of the observation error from the ECMWF model error we performed an error propagation and found the global estimate of the ECMWF refractivity error to be about 0.5% at 8–15 km increasing to 0.75% at 30 km.

The observation error for the relative standard deviation of CHAMP refractivity was found to be about 0.5% at 6–18 km increasing to 0.75% at 30 km and to 1.2% at 35 km. This is a more conservative estimate but still in line with the results of *Kuo et al.*, [2004]. When comparing these results for CHAMP RO to the simulation results for a MetOp/GRAS system it has to be kept in mind that the bending angle error differ by about a factor of 2.

In terms of error correlations, the ECMWF refractivity error correlation functions show negative correlation features in the vicinity of the peaks while the correlation functions for the combined error (CHAMP + ECMWF) show a flattening. These features suggest that the correlation wings are dominated by the observed data which may also be confirmed by the similar structure of the correlation functions seen in the simulation study. At stratospheric heights <35 km small correlations are seen resulting in a narrowing of the error correlation functions, while the broadening in the troposphere is mostly a result of the errors due to horizontal variations. We suggest a combination of the determined relative standard deviation with the total error correlation matrix for the construction of observation error covariance matrices for data assimilation systems.

These observation error covariance matrices can also be approximated with simple analytical formulations depending on a few sensibly selected adjustment/fitting parameters only. While we provided values for these parameters based on our empirical datasets reflecting GRAS-type performance, the formulations can be used, via re-adjusting/re-scaling/re-fitting the parameters, to approximate any other error estimation dataset; as an example, scaling to CHAMP RO data was addressed. The refractivity error covariance matrices should be useful for evaluation and proper specification of observational errors in data assimilation systems as well as in optimal estimation parts of retrieval algorithms.

The achieved results will foster a further application of the algorithms to create RO based global climatologies of refractivity, geopotential height, temperature, and humidity from CHAMP data on a large scale as planned in Part II of the CHAMPCLIM project. This will benefit not only the subsequent climate monitoring work but also the entire CHAMP/GPS data user community in different areas of applications. CHAMPCLIM Part II aims at using the complete CHAMP RO data flow for month-to-month, season-to-season, and year-to-year climate monitoring in temperature, geopotential height, and humidity fields. It is expected that these fields reveal the climate evolution of the atmosphere with an accuracy and vertical resolution never seen before. The nearly unbiased climatologies and analyses can, for example, serve as rigorous reference datasets to elucidate weaknesses in climate model

physics and forcing formulations or act as powerful observational constraints in climate change detection and attribution schemes.

Acknowledgments

The authors gratefully acknowledge the GFZ Potsdam, Germany, for the provision of the CHAMP level 2 and level 3 correlative data, for technical support, especially provided by T. Schmidt, and for scientific discussion and advice by J. Wickert. MIPAS temperature profiles were provided by IMK, Univ. of Karlsruhe, Germany. We especially acknowledge the support and discussions concerning the CHAMP-MIPAS comparison with D.-Y. Wang (IMK). We thank M. Fisher (ECMWF, Reading) for providing error specification for ECMWF analyses and A. Löscher for performing the error propagation. J. Fritzer and J. Ramsauer provided helpful assistance regarding modification and use of the EGOPS software. The authors are thankful for valuable discussions with M.E. Gorbunov (IAP, Moscow), S. Syndergaard and X.-Y. Huang (UCAR, Boulder, CO, U.S.A.), C. Marquardt (Met Office, Exeter, U.K.), and S. Healy (ECMWF, Reading, U.K.). A. Gobiet received financial support for the work from the CHAMPCLIM project funded by the Austrian Ministry for Traffic, Innovation, and Technology and managed under Contract No. ASAP-CO-004/03 of the Austrian Space Agency (ASA). Furthermore, A.K. Steiner, U. Foelsche, and A. Gobiet received financial support from the START research award of G.K. funded by the Austrian Ministry for Education, Science, and Culture and managed under Program No. Y103-N03 of the Austrian Science Fund (FWF).

References

Anderson, T.W. (1984), An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, New York.

Anthes, R.A., C. Rocken, and Y. Kuo (2000), Applications of COSMIC to meteorology and climate, *Terr. Atmos. and Oceanic Sci.*, *11*(1), 115-156.

Ao, C.O., T.K. Meehan, G.A. Hajj, A.J. Mannucci, and G. Beyerle (2003), Lower troposphere refractivity bias in GPS occultation retrievals, *J. Geophys. Res.*, *108*(D18), 4577, doi:10.1029/2002JD003216.

Barnett, J.J., and M. Corney (1985), Middle atmosphere reference model from satellite data, *Middle Atmosphere Program, Handbook for MAP*, Vol. 16, ed. K. Labitzke, J. J. Barnett and B. Edwards, 47-85.

Bassiri, S., and G.A. Hajj (1993), Higher-order ionospheric effects on the global positioning observables and means of modeling them, *Manuscr. Geod.*, *16*, 205-214.

Bevis, M., S. Businger, T.A. Herring, C. Rocken, R.A. Anthes, and R.H. Ware (1994), GPS meteorology: Mapping zenith wet delays onto precipitable water, *J. Appl. Met.*, *33*, 379-386.

Beyerle, G., J. Wickert, T. Schmidt, and C. Reigber (2004), Atmospheric sounding by GNSS radio occultation: An analysis of the negative refractivity bias using CHAMP observations, *J. Geophys. Res.*, *109*, D01106, doi:10.1029/2003JD003922.

Bronshtein, I.N., K.A. Semendyayev, G. Musiol, and H. Mühlig (2004), *Handbook of Mathematics*, 4th ed., 1157 pp., Springer-Verlag.

Collard, A.D., and S.B. Healy (2003), The combined impact of future space-based atmospheric sounding instruments on numerical weather-prediction analysis fields: A simulation study, *Q. J. R. Meteorol. Soc.*, *129*, 2741–2760.

Edwards, P.G., and D. Pawlak (2000), MetOp: The space segment for EUMETSAT's polar system, *ESA Bulletin*, *102*, 6-18.

ESA/EUMETSAT (1998), The GRAS instrument on METOP*, ESA/EUMET-SAT Rep. (ESA VR/3021/PI, EUMETSAT EPS/MIS/IN/9)*, 38 pp., Noordwijk, Netherlands.

Eyre, J.R. (1994), Assimilation of radio occultation measurements into a numerical weather prediction system, *ECMWF Tech. Memo. 199*, Eur. Cent. for Med. Range Weather Forecasts, Reading, UK.

Fishbach, F.F. (1965), A satellite method for temperature and pressure below 24 km, *Bull. Am. Meteorol. Soc.*, *9*, 528-532.

Fjeldbo, G., and V.R. Eshleman (1965), The bistatic radar-occultation method for the study of planetary atmospheres, *J. Geophys. Res.*, *70*, 3217–3225.

Fjeldbo, G., and V.R. Eshleman (1969), The atmosphere of Venus as studied with the Mariner 5 dual radio-frequency occultation experiment, *Radio Sci.*, *4*, 879–897.

Fjeldbo, G.F., V.R. Eshleman, and A.J. Kliore (1971), The neutral atmosphere of Venus as studied with the Mariner V radio occultation experiments, *Astron. J.*, *76*, 123–140, 1971.

Fleming, E.L., S. Chandra, M.R. Schoeberl, and J.J. Barnett (1988), Monthly mean global climatology of temperature, wind, geopotential height and pressure for 0–120 km, *NASA Tech. Memorandum*, *100697*, 85pp., National Aeronautics and Space Administration, Washington, D.C.

Foelsche, U. (1999), Tropospheric water vapor imaging by combination of spaceborne and ground-based GNSS sounding data, Ph.D thesis, Inst. for Geophys., Astrophys., and Meteorol., Univ. of Graz, Austria.

Foelsche, U., G. Kirchengast, and A.K. Steiner (2003), Global climate monitoring based on CHAMP/GPS radio occultation data, in *First CHAMP Mission Results for Gravity, Magnetic and Atmospheric Studies*, edited by C. Reigber et al., pp. 397–407, Springer-Verlag, Berlin Heidelberg.

Foelsche, U., and G. Kirchengast (2004a), Sensitivity of GNSS radio occultation data to horizontal variability in the troposphere, *Phys. Chem. Earth*, *29*, 225–240.

Foelsche, U., and G. Kirchengast (2004b), Sensitivity of GNSS radio occultation profiles to horizontal variability in the troposphere: A simulation study, in *Occultations for Probing Atmosphere and Climate*, edited by G. Kirchengast et al., pp. 127–136, Springer-Verlag, Berlin Heidelberg.

Gaspari, G., and S.E. Cohn (1999), Construction of correlation functions in two and three dimensions, *Q. J. R. Meteorol. Soc.*, *125*, 723–757.

GCOS (2003), The second report on the adequacy of the global observing systems for climate in support of the UNFCCC, *WMO/TD No. 1143*, 74 pp., WMO, Geneva, Switzerland.

Gobiet, A., and G. Kirchengast (2002), Sensitivity of atmospheric profiles retrieved from GNSS occultation data to ionospheric residual and high-altitude initialization errors, *Tech. Rep. ESA/ESTEC 1/2002*, 56 pp., Inst. for Geophys., Astrophys., and Meteorology, Univ. of Graz, Austria.

Gobiet, A., and G. Kirchengast (2004a), Advancement of GNSS radio occultation retrieval in the upper stratosphere, in *Occultations for Probing Atmosphere and Climate*, edited by G. Kirchengast et al., pp. 137–148, Springer-Verlag, Berlin Heidelberg.

Gobiet, A., and G. Kirchengast (2004b), Advancements of GNSS radio occultation retrieval in the upper stratosphere for optimal climate monitoring utility, *J. Geophys. Res.*, submitted.

Gobiet, A., A.K. Steiner, C. Retscher, U. Foelsche, and G. Kirchengast (2004a), Radio Occultation Data and Algorithms Validation Based on CHAMP/GPS Data, *Tech. Rep. ASA No. 1/2004*, Inst. for Geophys., Astrophys., and Meteorol., Univ. of Graz, Austria.

Gobiet, A., G. Kirchengast, J. Wickert, C. Retscher, D.-Y. Wang, and A. Hauchecorne (2004b), Evaluation of stratospheric radio occultation retrieval using data from CHAMP, MIPAS, GOMOS, and ECMWF analysis fields, *Proc. 2nd CHAMP science meeting,* Potsdam, Germany, Sept. 2003, in press.

Gorbunov, M.E. (2002a), Canonical transform method for processing radio occultation data in the lower troposphere, *Radio Sci*., *37*, 1076, doi: 10.1029/2000RS002592.

Gorbunov, M.E. (2002b), Radio-holographic analysis of Microlab-1 radio occultation data in the lower troposphere, *J. Geophys. Res*., *107*(D12), 4156, doi: 10.1029/2001JD000889.

Gorbunov, M.E. (2002c), Ionospheric correction and statistical optimization of radio occultation data, *Radio Sci*., *37*, 1084, doi: 10.1029/2000RS002370.

Gorbunov, M.E., and A.S. Gurvich (1998), Algorithms of inversion of Microlab-1 satellite data including the effects of multipath propagation, *Int. J. Remote Sensing*, *19*, 2283-2300.

Gorbunov, M.E., S.V. Sokolovskiy, and L. Bengtsson (1996), Space refractive tomography of the atmosphere: modeling of direct and inverse problems, *Tech. Rep. No. 210*, Max-Planck Institute for Meteorology, Hamburg, Germany.

Gorbunov, M.E., S.V. Sokolovskiy, and L. Bengtsson (1996), Advanced algorithms of inversion of GPS/MET satellite data and their application to reconstruction of temperature and humidity, *Tech. Rep. No. 211*, Max-Planck Institute for Meteorology, Hamburg, Germany.

GRAS-SAG (1998), The GRAS instrument on MetOp, *ESA/EUMETSAT Rep.* (ESA No. VR/3021/PI, EUMETSAT No. EPS/MIS/IN/9), ESA/ESTEC, Noordwijk, Netherlands.

Gurvich, A.S. and T.G. Krasil'nikova (1987), Navigation satellites for radio sensing of the earth's atmosphere. *Sov. J. Remote Sensing*, *6*, 89-93 (in Russian), 6, 1124-1131 (in English).

Hajj, G.A., E.R. Kursinski, L.J. Romans, W.I. Bertiger, S.S. Leroy (2002), A technical description of atmospheric sounding by GPS occultation, *J. Atmos. Sol. Terr. Phys*., *64*, 451-469.

Hajj, G.A., C.O. Ao, P. A. Iijima, D. Kuang, E.R. Kursinski, A.J. Mannucci, T.K. Meehan, L.J. Romans, M. de la Torre Juarez, and T.P. Yunck (2004), CHAMP and SAC-C atmospheric occultation results and intercomparisons, *J. Geophys. Res.*, *109*, D06109, doi:10.1029/2003JD003909.

Healy, S.B. (2001), Smoothing radio occultation bending angles above 40 km, *Ann. Geophys*., *19*, 459–468.

Healy, S.B., and J.R. Eyre (2000), Retrieving temperature, water vapour and surface pressure information from refractive index profiles derived by radio occultation: A simulation study, *Q. J. R. Meteorol. Soc.*, *126*, 1661–1683.

Healy, S.B., A. Jupp, and C. Marquardt (2003), A forecast impact trial with CHAMP radio occultation measurements, presentation at EGS-AGU-EUG Joint Assembly 2003, Nice, France, 7–11 April.

Hedin, A. E. (1991), Extension of the MSIS thermosphere model into the middle and lower atmosphere, *J. Geophys. Res.*, *96*, 1159–1172.

Hocke, K. (1997), Inversion of GPS meteorology data, *Ann. Geophysicae*, *15*, 443-450.

Hocke, K., G. Kirchengast, and A.K. Steiner (1997), Ionospheric correction and inversion of GNSS occultation data: problems and solutions, *Tech. Rep. ESA/ESTEC 2/1997*, 35 pp., Inst. for Geophys., Astrophys., and Meteorology, Univ. of Graz, Austria.

Hocke, K., K. Igarashi, and T. Tsuda (2003), High-resolution profiling of layered structures in the lower stratosphere by GPS occultation, *Geophys. Res. Lett*., *30*(8), 1426, doi:10.1029/2002GL016566.

Hoeg, P., and G. Kirchengast (2002), ACE+ - Atmospheric and Climate Explorer based on GPS, GALILEO, and LEO-LEO radio occultation (ESA Earth Explorer Opportunity Mission Proposal), *Wissenschaftl. Ber. No. 14*, 121 pp., Inst. for Geophys., Astrophys., and Meteorol., Univ. of Graz, Austria.

Hollingsworth, A., and P. Lönnberg (1986), The statistical structure of short range forecast errors determined from radiosonde data. Part I: The wind field. *Tellus*, *38*A, 11–136, 1986.

IPCC (2001), Climate change 2001 – the scientific basis. contribution of WG I to the *3rd Assessment Report of the IPCC*, edited by D.L. Albritton et al., Cambridge University Press, Cambridge.

Jensen, A.S., M.S. Lohmann, H.-H. Benzon, and A.S. Nielsen (2003), Full Spectrum Inversion of radio occultation signals, *Radio Sci*., *38*(3), 1040, doi: 10.1029/2002RS002763.

Kirchengast G (1998) End-to-end GNSS Occultation Performance Simulator overview and exemplary applications. *Wissenschaftl Ber 2/1998*, IGAM/Univ Graz, Austria, 138 pp

Kirchengast, G., and P. Hoeg (2004), The ACE+ Mission: An Atmosphere and Climate Explorer based on GPS, GALILEO, and LEO-LEO radio occultation, in *Occultations for Probing Atmosphere and Climate*, edited by G. Kirchengast et al., pp. 201–220, Springer-Verlag, Berlin Heidelberg.

Kirchengast, G., A.K. Steiner, U. Foelsche, L. Kornblueh, E. Manzini, and L. Bengtsson (2000), Spaceborne climate change monitoring by GNSS occultation sensors, Proc. 11th

Symp. on Global Change Studies, Amer. Met. Soc. Annual Meeting 2000, Long Beach/CA, U.S.A., 62-65.

Kirchengast, G., J. Fritzer, and J. Ramsauer (2002), End-to-end GNSS Occultation Performance Simulator Version 4 (EGOPS4) Software User Manual (Overview and Reference Manual). *Tech. Rep. ESA/ESTEC 3/2002*, 472 pp., Inst. for Geophys., Astrophys., and Meteorol., Univ. of Graz, Austria.

Kuo, Y.-H., S.V. Sokolovskiy, R.A. Anthes, and F. Vandenberghe (2000), Assimilation of GPS radio occultation data for numerical weather prediction, *Terr. Atmos. and Oceanic Sci.*, *11*(1), 157–186.

Kuo, Y.-H., T.-K. Wee, S. Sokolovskiy, C. Rocken, W. Schreiner, D. Hunt, and A. Anthes (2004), Inversion and error estimation of GPS radio occultation data*, J. Meteorol. Soc. Japan*, *82*, 507–531.

Kursinski, E.R., and G.A. Hajj (2001), A comparison of water vapor derived from GPS occultations and global weather analyses, *J. Geophys. Res.*, *106*(D1), 1113–1138.

Kursinski, E.R., G.A. Hajj, W.I. Bertiger, S.S. Leroy, T.K.Meehan, L.J. Romans, J.T. Schofield, D.J. McCleese, W.G. Melbourne, C.L. Thornton, T.P. Yunck, J.R. Eyre, and R.N. Nagatani (1996), Initial results of radio occultation of earth's atmosphere using the global positioning system, *Science*, *271*, 1107-1110.

Kursinski, E.R., G.A. Hajj, J.T. Schofield, R.P. Linfield, and K.R. Hardy (1997), Observing earth's atmosphere with radio occultation measurements using the global positioning system, *J. Geophys. Res*., *102*, 23429-23465.

Ladreiter, H.P., and G. Kirchengast (1996), GPS/GLONASS sensing of the neutral atmosphere: Model independent correction of ionospheric influences, *Radio Sci*., *31*, 877-891.

Landolt-Börnstein (1984), *Geophysics of the Solid Earth, the Moon and the Planets – Numerical data and functional relationships in science and technology*, vol. V/2a (K Fuchs and H Stoffel, Eds), Springer Verlag, Berlin–Heidelberg, 332–336

Lee, L-C., C. Rocken, and R. Kursinski (Eds.) (2001), *Applications of Constellation Observing System for Meteorology, Ionosphere, and Climate*, Springer-Verlag, Hong Kong.

Leitinger, R., and G. Kirchengast (1997a), Easy to use global and regional models – a report on approaches used in Graz, *Acta Geod. Geophy. Hung*., *32*, 329-342.

Leitinger, R., and G. Kirchengast (1997b), Inversion of the plasma signal in GNSS occultations – simulation studies and sample results, *Acta Geod. Geophy. Hung*., *32*, 379-394.

Leitinger, R, J. E. Titheridge, G. Kirchengast, W. Rothleitner (1996), A "simple" global empirical model for the F layer of the ionosphere (in German; English version avail. from the authors). *Kleinheubacher Ber. 39*, 697–704.

Leroy, S. (1997), The measurement of geopotential heights by GPS radio occultation, *J. Geophys. Res.*, *102*(D6), 6971–6986.

Leroy, S.S. and G.R. North (2000), The application of COSMIC to global change research, *TAO*, *11*(1), 187-210.

Loiselet, M., N. Stricker, Y. Menard, and J.-P. Luntama (2000), GRAS –MetOp's GPS-based atmospheric sounder, *ESA Bulletin*, *102*, 38-44.

Lusignan, B., G. Modrell, A. Morrision, J. Pomalaza, and S.G. Ungar (1969), Sensing the Earth's atmosphere with occultation satellites, *Proc. IEEE*, *4*, 458-467.

Mardia, K.V., J.T. Kent, and J.M. Bibby (1979), *Multivariate Analysis*, Academic Press, London.

Melbourne, W.G., E.S. Davis, C.B. Duncan, G.A. Hajj, K.R. Hardy, E.R. Kursinski, T.K. Meehan, L.E. Young, and T.P. Yunck (1994), The application of spaceborne GPS to atmospheric limb sounding and global change monitoring, *JPL Publ.*, *94-18*, 147 pp., Jet Prop. Lab., Pasadena, CA.

Palmer, P.I., J.J. Barnett (2001), Application of an optimal estimation inverse method to GPS/MET bending angle observations, *J. Geophys. Res*., *106*(D15), 17147-17160.

Palmer, P.I., J.J. Barnett, J.R. Eyre, and S.B. Healy (2000), A non-linear optimal estimation inverse method for radio occultation measurements of temperature, humidity, and surface pressure, *J. Geophys. Res.*, *105*(D13), 17513–17526, doi:10.1029/2000JD900151.

Ramsauer, J., and G. Kirchengast (2001), Sensitivity of atmospheric profiles retrieved from GNSS radio occultation data to instrumental errors, *Tech. Rep. ESA/ESTEC 6/2001*, 65 pp., Inst. for Geophys., Astrophys., and Meteorology, Univ. of Graz, Austria.

Randel, W., E. Fleming, N. Geller, M. Gelman, K. Hamliton, D. Karoly, D. Ortland, S. Pawson, R. Swinbank, P. Udelhofen, F. Wu, M. Baldwin, M.-L. Chanin, P. Keckhut, A. Simmons, D. Wu (2003), The SPARC intercomparison of middle atmosphere climatologies, *SPARC Newsletter*, *20*, 10-15.

Reigber, Ch., H. Lühr, and P. Schwintzer (2000), CHAMP mission status and perspectives, *Suppl. to EOS, Transactions, AGU, 81*, 48, F307.

Rieder , M. J., and G. Kirchengast (2001a), Error analysis for mesospheric temperature profiling by absorptive occultation sensors, *Ann. Geophys.*, *19*, 71–81.

Rieder, M. J., and G. Kirchengast (2001b), Error analysis and characterization of atmospheric profiles retrieved from GNSS occultation data, *J. Geophys. Res.*, *106*(D23), 31755–31770.

Rocken C., R. Anthes, M. Exner, D. Hunt, S. Sokolovskiy, R. Ware, M. Gorbunov, W. Schreiner, D. Feng, B. Herman, Y.H. Kuo, and X. Zou (1997), Analysis and validation of GPS/MET data in the neutral atmosphere, *J. Geophys. Res*., *102*(D25), 29849-29866.

Rocken, C., Y.-H. Kuo, W.S. Schreiner, D. Hunt, S. Sokolovskiy, and C. McCormick (2000), COSMIC system description, *Terr. Atmos. and Oceanic Sci.*, *11*(1), 21–52.

Rodgers, C. D. (2000), *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific Publ., Singapore.

Roeckner E, L. Bengtsson, J. Feichter, J. Lelieveld, H. Rodhe (1999), Transient climate change simulations with a coupled atmosphere-ocean gcm including the tropospheric sulfur cycle, *J. Climate*, *12*, 3004-3032.

Salby, M.L. (1996), *Fundamentals of Atmospheric Physics*, Academic Press, San Diego

Schmidt, T., J. Wickert, G. Beyerle, and C. Reigber (2004), Tropical tropopause parameters derived from GPS radio occultation measurements with CHAMP, *J. Geophys. Res.*, in press.

Schroeder, T., S. Leroy, M. Stendel, and E. Kaas (2003), Stratospheric temperatures probed by Microwave Sounding Units or by occultation of the Global Positioning System, *Geophys. Res. Lett.*, *30*, doi:10.1029/2003GL017588.

Silvestrin, P., R. Bagge, M. Bonnedal, A. Carlström., J. Christensen, M. Hägg, T. Lindgren, F. Zangerl (2000), Spaceborne GNSS radio occultation instrumentation for operational applications, *Proc 13th ION-GPS Meeting 2000*, Salt Lake City, UT, USA.

Smith, E.K., and S. Weintraub (1953), The constants in the equation for atmospheric refractive index at radio frequencies, *Proc. of the I.R.E.*, *41*, 1035–1037.

Sokolovskiy, S., and D. Hunt (1996), Statistical optimization approach for GPS/Met data inversions, paper presented at URSI GPS/Met Workshop, Tucson, Arizona.

Sokolovskiy S. (2003), Effect of superrefraction on inversions of radio occultation signals in the lower troposphere, *Radio Sci*., *38*(3), 1058, doi:10.1029/2002RS002728.

Spilker, J.J. (1980), GPS signal structure and performance characteristics, in *Global Positioning System*, edited by P.M. Janiczek, Institute of Navigation, Washington, DC., 29-54.

Steiner, A. K. (2004), Error analyses of refractivity profiles retrieved from CHAMP radio occultation data, *DMI Sci. Rep. 04-02*, 19 pp., Danish Meteorol. Inst., Copenhagen, Denmark.

Steiner A.K., and G. Kirchengast (2004a), Ensemble-based analysis of errors in atmospheric profiles retrieved from GNSS radio occultation data, in *Occultations for Probing Atmosphere and Climate*, edited by G. Kirchengast et al., pp. 149–160, Springer-Verlag, Berlin Heidelberg.

Steiner A.K., and G. Kirchengast (2004b), Empirical error analysis for GNSS radio occultation data based on end-to-end simulations, *J. Geophys. Res.*, submitted.

Steiner, A.K., G. Kirchengast, and H.P. Ladreiter (1999), Inversion, error analysis, and validation of GPS/MET occultation data, *Ann. Geophys.*, *17*, 122–138.

Steiner, A. K., G. Kirchengast, U. Foelsche, L. Kornblueh, E. Manzini, and L. Bengtsson (2001), GNSS occultation sounding for climate monitoring, *Phys. Chem. Earth (A)*, *26*, 113–124.

Steiner, A. K., U. Foelsche, A. Gobiet, G. Kirchengast, J. Wickert, and T. Schmidt (2004), CHAMPCLIM: Climatologies based on radio occultation observations, poster presented at 1st International CLIVAR Science Conference, Baltimore, U.S.A., 21–25 June.

Syndergaard, S. (1999), Retrieval analysis and methodologies in atmospheric limb sounding using the GNSS radio occultation technique, *DMI Sci. Rep. 99-6*, 131 pp., Danish Meteorol. Inst., Copenhagen, Denmark.

Syndergaard, S. (2000), On the ionosphere calibration in GPS radio occultation measurements, *Radio Sci*., *35*, 865-883.

Syndergaard, S., D. Flittner, R. Kursinski, D. Feng, B. Herman, and D. Ward (2004), Simulating the influence of horizontal gradients on retrieved profiles from ATOMS occultation measurements – a promising approach for data assimilation, in *Occultations for Probing Atmosphere and Climate*, edited by G. Kirchengast et al., pp. 221–232, Springer-Verlag, Berlin Heidelberg.

Turchin, V.F., and V.Z. Nozik (1969), Statistical regularization of the solution of incorrectly posed problems, *Phys. Atmos. Ocean*, *5*(1), 14-18.

von Engeln A., G. Nedoluha, G. Kirchengast, and S. Bühler (2003), One-dimensional variational (1-D Var) retrieval of temperature, water vapor, and a reference pressure from radio occultation measurements: A sensitivity analysis, *J. Geophys. Res*., *108*(D11), 4337, doi:10.1029/2002JD002908.

von Storch, H., and F.W. Zwiers (1999), *Statistical analysis in climate research*, Cambridge University Press, Cambridge, UK.

Vorob'ev, V.V. and T.G. Krasnil'nikova (1994), Estimation of the accuracy of the atmospheric refractive index recovery from Doppler shift measurements at frequencies used in the NAVSTAR system, *Phys. Atmos. Ocean*, *29*, 602-609.

Ware, R., M. Exner, D. Feng, M. Gorbunov, K. Hardy, B Herman, Y. Kuo, T. Meehan, W. Melbourne, C. Rocken, W. Schreiner, S. Sokolovskiy, F. Solheim, X. Zou, R. Anthes, S. Businger, and K. Trenberth (1996), GPS sounding of the atmosphere from low earth orbit: preliminary results, *Bull. Amer. Met. Soc*., *77*, 19-40.

Wickert, J., C. Reigber, G. Beyerle, R. Konig, C. Marquardt, T. Schmidt, L. Grunwaldt, R. Galas, T.K. Meehan, W.G. Melbourne, and K. Hocke (2001), Atmosphere sounding by GPS radio occultation: first results from CHAMP, *Geophys. Res. Lett*., *28*, 3263-3266.

Wickert, J., T. Schmidt, G. Beyerle, R. König, C. Reigber, and N. Jakowski (2004), The radio occultation experiment aboard CHAMP: Operational data analysis and validation of vertical atmospheric profiles, *J. Meteorol. Soc. Japan*, *82*, 381-395.

Yunck, T.P., G.F. Lindal, and C.H. Liu (1988), The role of GPS in precise earth observation. *Proc. IEEE Position*, Location and Navigation Symposium (PLANS 88), Orlando, FL., Nov. 29-Dec. 2.

Yunck, T.P., C. Liu, and R. Ware (2000), A history of GPS sounding, *Terr. Atmos. and Oceanic Sci.*, *11*(1), 1-20.

Zou, X., B. Wang, H. Liu, R.A. Anthes, T. Matsumura, and Y.-J. Zhu (2000), Use of GPS/MET refraction angles in three-dimensional variational analysis, *Q. J. R. Meteorol. Soc.*, *126*, 3013–3040.

Ω end of document Ω