Empirical analysis and modeling of errors of atmospheric profiles from GPS radio occultation

B. Scherllin-Pirscher1,2,3, A. K. Steiner3, G. Kirchengast3, Y.-H. Kuo2, and U. Foelsche3

1 Advanced Study Program, National Center for Atmospheric Research (NCAR), Boulder, Colorado, USA
2 COSMIC Project Office, University Corporation for Atmospheric Research (UCAR), Boulder, Colorado, USA
3 Wegener Center for Climate and Global Change (WEGC) and Institute for Geophysics, Astrophysics, and Meteorology/Institute of Physics (IGAM/IP), University of Graz, Graz, Austria

Abstract. The utilization of radio occultation (RO) data in atmospheric studies requires precise knowledge of error characteristics. We present results of an empirical error analysis of GPS RO bending angle, refractivity, dry pressure, dry geopotential height, and dry temperature. We find very good agreement between data characteristics of different missions (CHAMP, GRACE-A, and Formosat-3/COSMIC (F3C)). In the global mean, observational errors (standard deviation from “true” profiles at mean tangent point location) agree within 0.3 % in bending angle, 0.1 % in refractivity, and 0.2 K in dry temperature at all altitude levels between 4 km and 35 km. Above 35 km the increase of the CHAMP raw bending angle observational error is more pronounced than that of GRACE-A and F3C leading to a larger observational error of about 1 % at 42 km. Above ≈20 km, the observational errors show a strong seasonal dependence at high latitudes. Larger errors occur in hemispheric wintertime and are associated mainly with background data used in the retrieval process particularly under conditions when ionospheric residual is large. The comparison between UCAR and WEGC results (both data centers have independent inversion processing chains) reveals different magnitudes of observational errors in atmospheric parameters, which are attributable to different background fields used. Based on the empirical error estimates, we provide a simple analytical error model for GPS RO atmospheric parameters for the altitude range of 4 km to 35 km and up to 50 km for UCAR raw bending angle and refractivity. In the model, which accounts for vertical, latitudinal, and seasonal variations, a constant error is adopted around the tropopause region amounting to 0.8 % for bending angle, 0.35 % for refractivity, 0.15 % for dry pressure, 10 m for dry geopotential height, and 0.7 K for dry temperature. Below this region the observational error increases following an inverse height power-law and above it increases exponentially. For bending angle and refractivity we also include formulations for error correlations in order to enable modeling of full error covariance matrices for these primary data assimilation variables. The observational error model is the same for UCAR and WEGC data but due to somewhat different error characteristics below about 10 km and above about 20 km some parameters have to be adjusted. Overall, the observational error model is easily applicable and adjustable to individual error characteristics.

1 Introduction

The radio occultation (RO) technique (Melbourne et al., 1994; Kursinski et al., 1997; Hajj et al., 2002) provides accurate profiles of atmospheric parameters in the troposphere and lower stratosphere. These profiles are used in atmospheric and climate research (e.g. Leroy et al., 2006; Borsche et al., 2007; Ho et al., 2007; Foelsche et al., 2008; Zeng et al., 2008; Steiner et al., 2009; Pirscher et al., 2010) as well as for data assimilation in numerical weather prediction (e.g. Kuo et al., 2000; Healy et al., 2005). As shown for example by Healy and Thépaut (2006); Aparicio and Deblonde (2008); Cucurull and Derber (2008); Rennie (2010), the assimilation of RO data significantly improved operational analyses at stratospheric levels. Current data assimilation centers assimilate either RO raw bending angle or RO refractivity as provided by the German Research Centre for Geosciences (GFZ) (CHAMP, GRACE-A, and TerraSAR-X data), the University Corporation for Atmospheric Research (UCAR) (FORMOSAT-3/COSMIC, F3C data), the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) (MetOp/GRAS bending angle), and

The RO method is based on measurements of the path of radio waves (radio signals continuously transmitted by Global Positioning System, GPS, satellites) passing through the atmosphere. An occultation event occurs when a GPS satellite rises or sets across the limb with respect to a Low Earth Orbit (LEO) satellite. On the way through the atmosphere these signals are refracted due to vertical density gradients and received on a LEO satellite. The refraction causes an accumulation of the excess phase path in the GPS phase measurements. Due to the relative motion of the GPS and the LEO satellite, the radio signals penetrate the atmosphere at different tangent heights and the atmosphere is scanned from top downwards (setting event) or from bottom up (rising event), yielding information with high vertical resolution. The result is a near vertical profile of phase measurements as a function of time. An RO event lasts approximately one to two minutes. Given that the signal reception is stable within this measurement time, the measurement is self-calibrating and therefore also long-term stable because precise clocks are traceable to the international time standard (SI second).

The RO processing starts with precise orbit determination and atmospheric excess phase processing at both GPS frequencies. Knowledge of atmospheric excess phase profiles as a function of time allows the calculation of Doppler shift. Involvement of precise orbit information yields bending angle profiles as a function of impact parameter. The main ionospheric contribution of the measurement is removed by a linear combination of two bending angles, derived for both GPS frequencies separately. In the upper stratosphere and above, the signal-to-noise ratio of the measurement declines resulting in large bending angle noise at these altitudes. However, the inverse Abel transform, which is used to calculate refractivity from bending angle, requires data knowledge at highest altitudes. Noisy and erroneous bending angles cause non-negligible errors in the refractivity profile, which are then further transported downward by the hydrostatic integral in the pressure and temperature profiles. For noise reduction at high altitudes, raw bending angle profiles are initialized with background data leading to optimized bending angle profiles (note that raw bending angles are assimilated in numerical weather prediction systems). An Abel inversion of the ionosphere-corrected, initialized bending angle profile gives a refractivity profile as a function of altitude. Neglecting the moist contribution of refractivity in a so-called “dry air retrieval” yields atmospheric dry density profiles. Dry pressure profiles as a function of altitude, geopotential height profiles as a function of dry pressure altitude, and dry temperature profiles as a function of altitude are then calculated using the hydrostatic equation and the equation of state.

The assumption of dry air is valid where water vapor density is low, i.e., above ≈14 km at low latitudes and ≈9 km at high latitudes (Foelsche et al., 2008; Scherllin-Pirscher et al., 2011). In this case dry air parameters are essentially physical parameters. Below, however, the proportion of water vapor becomes significant and the difference between physical and dry temperature can reach several tens of kelvins in the lower troposphere (Foelsche et al., 2008; Scherllin-Pirscher et al., 2011). In this case auxiliary information is required for the retrieval of physical atmospheric parameters, e.g., through a one-dimensional variational (1DVar) method. Temperature, pressure, and water vapor are then derived using auxiliary temperature and/or humidity data as provided, e.g., by operational weather analysis centers (Healy and Eyre, 2000).

Several studies showed that RO data are of highest quality in the upper troposphere and lower stratosphere (UTLS) region. The theoretical study of Kursinski et al. (1997) provides a good overview on various error sources in RO data. It showed that the refractivity retrieval error is about 0.2 % between 10 km and 20 km altitude, increasing to 1 % at the surface and to 1 % at 40 km (daytime, solar maximum ionosphere). Using GPS/Met data from 1995 to 1997 Rocken et al. (1997) showed that the accuracy of RO derived temperature is better than 1 K in the UTLS region, confirmed by Steiner et al. (1999). Gobiet et al. (2007) found CHAMP accuracy to be better than 0.5 K between 10 km and 30 km. The precision of RO data was investigated by Hajj et al. (2004), who compared co-located CHAMP and SAC-C profiles, and Schreiner et al. (2009), who compared co-located F3C profiles. The latter study showed that the root-mean-square (RMS) difference of co-located F3C refractivity profiles between 10 km and 20 km altitude is less than 0.2 %.

So far a range of dedicated studies was carried out for the provision of RO observation error estimates. Syndergaard (1999) performed a theoretical analysis on the step-by-step covariance propagation from phase level via bending angle to refractivity, pressure, and temperature profiles. Based on optimal estimation methodology, applied to RO retrievals (Palmer et al., 2000; Palmer and Barnett, 2001), a complete theoretical error characterization was carried out by Rieder and Kirchengast (2001a) accounting also for error correlations. In an empirical analysis using simulated RO profiles, Steiner and Kirchengast (2005) investigated systematic differences to ECMWF analysis fields, standard deviation profiles, and correlation functions for the full set of RO parameters. Kuo et al. (2004) separated the RO error from the background error and provided observational error estimates for refractivity based on CHAMP and SAC-C data. Following a different approach regarding ECMWF error estimation, Steiner et al. (2006) found consistent results and provided simple analytical error formulations for CHAMP refractivity. Chen et al. (2011) used data from F3C and estimated observational errors of linear excess phase and refractivity for the troposphere region over East Asia and the Western Pacific.
This study extends former work to the investigation of error characteristics for current RO data sets and aims at providing observational error estimates for RO atmospheric profiles from bending angle, refractivity, pressure, and geopotential height to temperature. We note that all refractivity error characteristics, given in percent, apply to dry density errors (in %) as well (Rieder and Kirchengast, 2001a). The error analysis is based on UCAR and WEGC (Wegener Center for Climate and Global Change/Univ. of Graz) RO data, which are investigated for different latitude regions in the altitude range between 4 km and 35 km. UCAR bending angle and refractivity are analyzed up to 50 km in view of their use in numerical weather prediction (NWP) assimilation systems.

Section 2 gives a description of all data sets used. In Sect. 3 we present the estimation of the GPS RO observational error, defined as standard deviation of RO profiles from the “true” profiles at mean tangent point location. We base this estimation on the analysis of individual RO profiles from the satellite missions CHAMP, GRACE-A, and F3C. Furthermore, we provide a simple general model for the observational error by following and further advancing the approach of Steiner and Kirchengast (2005) and Steiner et al. (2006). Conclusions are drawn in Sect. 4.

2 Data

We analyze raw and statistically optimized bending angles as a function of impact altitude (impact parameter with local radius of curvature and geoid undulation subtracted), refractivity, dry pressure, and dry temperature as a function of mean sea level (m.s.l.) altitude, and geopotential height as a function of dry pressure altitude. We use dry atmospheric profiles delivered by both the UCAR and the WEGC processing (raw bending angle profiles are only available at UCAR). We use data from CHAMP, GRACE-A, and F3C. CHAMP, which near-continuously delivered data from 2001 to 2008, and GRACE-A are single satellites (GRACE-B does not provide RO measurements operationally), which deliver(ed) on average 140 and 120 fully processed RO events per day, respectively (Wickert et al., 2009). Each single F3C satellite samples setting and rising RO events, which doubles the number of RO events per day.

2.1 RO data

UCAR and WEGC established their own processing chains with independent implementations and different assumptions. Ho et al. (2009) described and compared RO retrievals of different data centers at refractivity level and provided a detailed description of UCAR and WEGC RO processing schemes. Here, we briefly summarize the main differences in both processing chains.

2.1.1 RO processing at UCAR

The UCAR data processing as described by Kuo et al. (2004) and Ho et al. (2009) starts with raw GPS amplitude and phase measurements as well as raw GPS and LEO orbit tracking data (level 0 data). Level 1 processing comprises the reconstruction of precise LEO and GPS position and velocity vectors as well as the determination of atmospheric excess phase profiles. This step includes the removal of satellite clock errors and Doppler shift due to the satellites motion. The inversion process (level 2 processing) delivers atmospheric profiles of the well known RO atmospheric parameters raw bending angle \(\alpha_{\text{raw}}\), optimized bending angle \(\alpha_{\text{opt}}\), refractivity \(N\), dry density \(\varrho_{\text{dry}}\), dry pressure \(p_{\text{dry}}\), dry geopotential height \(Z_{\text{dry}}\), and dry temperature \(T_{\text{dry}}\), respectively. To get better retrieval results in regions with high atmospheric moisture gradients, which causes atmospheric multipath of GPS signals, UCAR applies a wave optics (WO) retrieval (Gorbunov, 2002; Jensen et al., 2003) below approximately 20 km. A geometric optics (GO) retrieval is applied above. For bending angle initialization, UCAR uses a background model which is based on an NCAR climatology exponentially extrapolated to 150 km (Randel et al., 2002). Background and observational information are optimally combined (Sokolovskiy and Hunt, 1996). To calculate pressure and temperature, the hydrostatic integral is initialized at 150 km by setting pressure and temperature to zero. This initialization does not affect retrieval results below 80 km (Kuo et al., 2004). UCAR operational retrieval products are available up to 60 km.

Most UCAR profiles penetrate below 1 km (Anthes et al., 2008), however, since we focus on dry atmospheric profiles we only use them above 4 km at polar altitudes and above 8 km at tropical altitudes (Foelsche et al., 2008). In this study we use high quality CHAMP, GRACE-A, and F3C data from 2008 and 2009 (CHAMP data are only available until September 2008) with data versions 2009.2650 (CHAMP) and 2010.2640 (GRACE-A and F3C) from UCAR. Profiles are available from \(http://cosmic-io.cosmic.ucar.edu/cdaac/index.html\).

2.1.2 RO processing at WEGC

The current WEGC processing scheme (OPSv5.4) (Steiner et al., 2009; Pirscher, 2010) starts with excess phase profiles and precise orbit information (level 1 data) provided by UCAR, which implies that both processing chains are not fully independent as they share the raw processing towards level 1 data. To derive RO bending angles, the WEGC OPSv5.4 scheme applies a GO retrieval only. Furthermore, WEGC OPSv5.4 is not able to retrieve F3C data received in open loop, which causes WEGC F3C profiles to penetrate not below \(\approx 8\) km.

Since CHAMP and GRACE-A RO measurements are not received in open-loop mode (both receivers operate in
closed-loop mode, optionally with fly-wheeling in the lower to mid troposphere), profiles of these satellites reach further down. However, we note that F3C soundings in general penetrate further down than CHAMP and GRACE-A profiles if open-loop data can be utilized (Anthes et al., 2008). In this context previous work showed that lower tropospheric refractivity retrieved from fly-wheeling measurements exhibits significantly larger biases and standard deviations than retrievals from open-loop measurements (Beyerle et al., 2006). Since these larger errors are relevant at altitudes below 5 km, while this study investigates RO data from 4 km (high latitudes) to 8 km (low latitudes) upwards throughout the UTLS, we can disregard them here.

WEGC bending angles are initialized with background information derived from ECMWF short-range forecast fields (24 h to 30 h forecasts), which are extended with MSIS data (Hedin, 1991) up to 120 km. Statistical optimization is performed with inverse covariance weighting above 30 km (Gobiet and Kirchengast, 2004). The upper bound of the hydrostatic integral is set to 120 km. WEGC retrieval output is not operationally available above 35 km.

We use profiles from CHAMP, GRACE-A, and F3C, which pass WEGC quality control from 2007 to 2009 with data version OPSv5.4. Profiles can be downloaded from the global climate monitoring website, http://www.wegcenter.at/globclim. To compare the data sets at the same vertical resolution in this study, the operational output of WEGC statistically optimized bending angle profiles is smoothed with a boxcar average filter with 11-points width (i.e. approximately 0.5 km at an altitude of 5 km, 1 km at tropopause level, and 2 km at an altitude of 30 km).

Given the type of filtering involved, the vertical resolution of both UCAR and WEGC profiles is about 1 km from near 5 km up to an altitude of 30 km, above which the resolution can be approximated as increasing linearly to reach 2 km at 50 km altitude.

2.2 ECMWF data

The Integrated Forecasting System (IFS) of the ECMWF operationally generates global analysis fields for four time layers each day, centered at 00:00 UTC, 06:00 UTC, 12:00 UTC, and 18:00 UTC. We use these global analysis fields and extract co-located reference profiles for each atmospheric parameter and for each RO event. ECMWF fields are used at a horizontal resolution of T42, which corresponds to the horizontal resolution of RO profiles (about 300 km). Data are available at 91 vertical levels (L91). The co-located profile is extracted from that ECMWF field, of which the time layer is closest to the mean RO event time. Co-location is obtained from spatial interpolation to the mean geographic event location. To avoid vertical interpolation errors, co-located ECMWF profiles are generated in a similar way as RO profiles are retrieved. In a first step refractivity is calculated from the ECMWF analysis data. Bending angle profiles are then derived from Abel transformed refractivity; pressure, temperature, and geopotential height are obtained from an RO-type retrieval.

ECMWF operationally assimilates RO data since December 2006 (Healy, 2007), which means that ECMWF and RO data are not independent anymore. We are aware that this might somewhat diminish the magnitude of our observational error estimates but the related uncertainty is estimated to be within 20 % because the statistical error of ECMWF analyses did not change much due to starting RO assimilation while bias correction was significantly improved, however (Luntama et al., 2008).

3 Estimation of GPS RO observational error

3.1 GPS RO plus ECMWF combined error

RO error estimates are based on the comparison of retrieved RO profiles to co-located profiles retrieved from ECMWF analysis fields, giving the combined (RO observational plus ECMWF model) error.

The calculation of error statistics follows Steiner and Kirchengast (2004). For each single RO profile $x_{RO}$, a co-located ECMWF profile $x_{ECMWF}$ is extracted to compute the difference profile $\Delta x$:

$$\Delta x \left( z_j \right) = x_{RO} \left( z_j \right) - x_{ECMWF} \left( z_j \right),$$

(1)

where all profiles are interpolated to the same altitude levels $z_j$. We interpolate RO data to the ECMWF vertical grid with the top level at about 34 km. The difference profile of atmospheric parameters which decrease exponentially with height (i.e. bending angle, refractivity, and dry pressure) is obtained as a percentage difference by

$$\frac{\Delta x}{x} \left( z_j \right) = 100 \cdot \frac{x_{RO} \left( z_j \right) - x_{ECMWF} \left( z_j \right)}{x_{ECMWF} \left( z_j \right)}.$$  

(2)

The mean systematic differences $m_{\Delta x} \left( z_j \right)$ and $m_{\Delta x/x} \left( z_j \right)$ are then calculated by

$$m_{\Delta x} \left( z_j \right) = \frac{1}{n \left( z_j \right)} \sum_{i=1}^{n(z_j)} \Delta x_i \left( z_j \right),$$  

(3)

and

$$m_{\Delta x/x} \left( z_j \right) = \frac{1}{n \left( z_j \right)} \sum_{i=1}^{n(z_j)} \frac{\Delta x_i \left( z_j \right)}{x_i \left( z_j \right)}.$$  

(4)

for non-exponential and exponential parameters, respectively, with $n \left( z_j \right)$ being the number of profiles at altitude level $z_j$. The number of profiles decreases with decreasing altitude because increasing humidity leads to atmospheric multipath and signal degradation. The standard deviations of the difference profiles, $s_{\Delta x} \left( z_j \right)$ and $s_{\Delta x/x} \left( z_j \right)$, further denoted as combined error $s_{combined}$, are finally obtained as well via the usual textbook formulae for these standard error estimators.
The statistics is calculated for different latitude regions (low, mid, and high latitudes, 30° S to 30° N, 30° S/N to 60° S/N, 60° S/N to 90° S/N, respectively, as well as for the global, Northern Hemisphere (NH), and Southern Hemisphere (SH) regions).

As an introductory example, Fig. 1 shows global mean UCAR raw bending angle (left panel) and WEGC dry temperature (right panel) systematic difference and standard deviation relative to ECMWF in July 2008. As can be noticed from Fig. 1, data characteristics are very similar for all satellites. UCAR bending angle systematic difference is very close to zero between 9 km and 14 km, slightly negative up to about 37 km and positive above. Bending angle standard deviation is within 2% in the altitude range between 9 km and 32 km. Due to the smaller number of CHAMP and GRACE-A occultation events, standard deviation is slightly larger than for F3C. Maximum differences between CHAMP, GRACE-A, and F3C are observed above 35 km where deviations become larger than 0.5%. Above about 42 km CHAMP raw bending angle observational error exceeds that of the other satellites by 1%. This mainly derives from the larger bending angle noise observed in CHAMP data (Pirscher, 2010; Foelsche et al., 2011a).

WEGC dry temperature systematic difference is close to zero up to 24 km and positive above. Standard deviation does not exceed 2 K within 5 km and 30 km. CHAMP and GRACE-A standard deviations are up to 0.2 K larger than those of F3C.

Figure 2 illustrates the temporal variation of the monthly statistics calculated for WEGC optimized bending angle (top panels) and dry temperature (bottom panels) for one single satellite (F3C/FM-1).

The systematic difference between RO and ECMWF bending angle is slightly negative during the whole time period, with differences being, in general, smaller than −0.5%. Differences are slightly larger at high southern latitudes than at high northern latitudes. The temperature systematic difference is very close to zero up to 20 km, increasing to about +1 K at 34 km.

Bending angle and temperature standard deviations show a clear seasonal variation above 20 km, where standard deviation always increases in the winter hemisphere. This seasonal variation is somewhat more pronounced in the Southern Hemisphere: bending angle standard deviations reach almost 3% at high southern latitudes at 34 km in June, July, and August, temperature standard deviations become even larger than 4 K above 32 km in these particular winter months. These large errors might result from both RO data and ECMWF analysis profiles in roughly equal measure. RO retrieved data are affected by background data used for bending angle initialization (for the reduction of bending angle noise). Since the background does often not properly capture high latitude winter situations this may result in initialization errors in such cases. However, ECMWF analysis fields also feature worst quality at high latitudes in the winter hemisphere, because there are only rather few good observational data available at high latitudes although a large number of observations is crucial to capture high atmospheric variability.

Figure 3 depicts differences between RO data retrieved at UCAR and at WEGC for different atmospheric parameters. At bending angle level (top panels), data are in very good agreement. The slightly larger standard deviation of UCAR data below 20 km results from the UCAR WO retrieval, which captures more (true) small-scale atmospheric variability than both, the WEGC GO retrieval and the ECMWF analyses.

Up to 30 km UCAR and WEGC refractivities show a similar error characteristics relative to ECMWF. Above ≈20 km, dry pressure, dry geopotential height, and dry temperature error characteristics differ in terms of systematic difference and standard deviation. In this case, the differences are connected with a different approach in bending angle initialization at high altitudes. The monthly NCAR background climatology, which is used at UCAR significantly

---

**Fig. 1.** RO minus ECMWF systematic difference (solid) and standard deviation (dashed) in July 2008 for UCAR raw bending angle (left panel) and WEGC dry temperature (right panel). Note different x- and y-axes of the two panels.
underestimates small-scale atmospheric variability, which explains the larger UCAR standard deviation relative to ECMWF. Since most of the background information enters the retrieval at high altitudes (>45 km), this difference cannot be observed in bending angle and refractivity below 35 km. However, it propagates down in the retrieval process through the hydrostatic integral (Rieder and Kirchengast, 2001a; Steiner and Kirchengast, 2005). Near 34 km, the global UCAR dry pressure standard deviation is larger than the corresponding WEGC error by 0.8 %. UCAR dry geopotential height and dry temperature standard deviations are larger by about 50 m, and 0.7 K, respectively.

3.2 ECMWF model error

We use the global mean ECMWF temperature error as provided by the ECMWF and made available from the ROPP 4.0 package (GRAS SAF, 2009). It is available at 91 vertical hybrid levels up to a height of 80 km and is given in terms of standard deviation.

To estimate the error for atmospheric parameters other than temperature, we empirically derive conversion factors between the relevant standard deviations:

\[
\Delta N/N \quad [\%] = \Delta T_{2N} \quad s_{\Delta T_{2N}} \quad [K] \\
\Delta \alpha/\alpha \quad [\%] = \Delta N_{2\alpha} \quad s_{\Delta N_{2\alpha}} \quad [\%] \\
\Delta p_{\text{dry}}/p_{\text{dry}} \quad [\%] = \Delta N_{2p} \quad s_{\Delta N_{2p}} \quad [\%] \\
\Delta Z_{\text{dry}} \quad [m] = \Delta p_{2Z} \quad s_{\Delta p_{2Z}} \quad [\%].
\]

We find \(\Delta T_{2N} = 0.5 \%/K, \quad \Delta N_{2\alpha} = 2.4 \%/\%, \quad \Delta N_{2p} = 0.45 \%/\%,\) and \(\Delta p_{2Z} = 65 \text{ m}/\%,\) where likewise the factors in reverse direction are \(\Delta N_{2T} = 1/c_{T_{2N}} = 2.0 \text{ K}/\%, \quad \Delta p_{2N} = 1/c_{N_{2p}} = 0.42 \%/\%, \quad \Delta p_{2\alpha} = 1/c_{N_{2\alpha}} = 2.2 \%/\%,\) and \(\Delta p_{2N} = 1/c_{N_{2p}} = 0.015 \%/m,\) respectively. The latter are used in Fig. 4, which illustrates their validity. It emphasizes the consistency of all parameters, which are scaled to a common temperature scale, especially within the core region of about 8 km to 20 km. These conversion factors from empirically comparing real-data error estimates after different retrieval steps are – as they should which is reassuring – consistent.
Fig. 3. UCAR (blue) and WEGC (green) CHAMP, GRACE-A, and F3C systematic difference (solid) and standard deviation (dashed) relative to ECMWF. Optimized bending angle, refractivity, dry pressure, dry geopotential height, and dry temperature statistics are shown for January 2008 (left panels) and July 2008 (right panels).
with theoretical and simulation studies of how RO errors propagate between retrieval steps from bending angle to dry temperature (Kursinski et al., 1997; Syndergaard, 1999; Rieder and Kirchengast, 2001a; Sofieva and Kyrölä, 2004).

Steiner et al. (2006) tested the sensitivity of the estimated L60 ECMWF refractivity error with respect to the temperature input for four different cases, where the temperature standard deviation was multiplied by 1.0, 1.5, 2.0, and 2.5. They found that the results judged most reasonable for the case of doubling the ECMWF temperature standard deviation (2.0 × T). However, our sensitivity tests yield the 1.5 case to be the most reasonable case. This reduction is attributable to the increase of the number of vertical levels from 60 to 91 and to the improved quality of ECMWF analysis fields since 2006. Thus using data as of 2007 here, we estimate the ECMWF error to be 1.5 × ECMWF standard deviation as given by the ROPP 4.0 package (yielding 0.6 K at 5 km to 15 km, increasing to about 1 K at 35 km).

The ECMWF refractivity error is of the order of 0.3 % at 4 km to 15 km increasing to 0.5 % at 30 km, to 0.55 % at 35 km, and to 1.1 % at 50 km, which is slightly smaller than the findings of Kuo et al. (2004), who performed an estimation of short-range forecast errors using the Hollingsworth-Lönneberg method.

We note that this ECMWF error is only a rough estimate, which does not reflect the true ECMWF error in regions with high atmospheric variability. Since the error does not vary with latitude and month, it might be too small at high latitudes during the winter season.

3.3 GPS RO observational error

The combined error $s_{\text{combined}}$ (see Sect. 3.1) contains the RO observational error $s_{\text{obs}}$ and the ECMWF model error $s_{\text{ECMW}}$. The separation of these errors is performed by subtracting the estimated ECMWF error from the combined error in terms of variances

$$s_{\text{obs}} = \sqrt{s_{\text{combined}}^2 - s_{\text{ECMW}}^2},$$

which gives the observational error. It turns out that the estimated ECMWF error and the observational error are approximately of the same order of magnitude. A simple estimate of the observational error $s_{\text{obs}}$ can thus be given by scaling the combined error with $1/\sqrt{2}$. This is a reasonable approximation as can be seen in Fig. 5 and as was also found by Kuo et al. (2004) and Steiner et al. (2006).

Figure 5 shows UCAR (top two rows) and WEGC (bottom two rows) refractivity error estimates in different geographical regions in July 2008 including all four error types defined above (and in addition the error according to the simple analytical model, $s_{\text{model}}$, discussed further below).

In the region where RO data are known to feature best quality (between 9 km and 25 km), the observational error sometimes becomes smaller than zero because the adopted (global static) ECMWF error is larger than the combined error. In this region the estimated observational error amounts approximately to 0.3 %, which is in agreement with observational errors found by Kuo et al. (2004). Below 9 km and above 25 km the observational error agrees well with the estimated observational error. The global mean observational error in the lower stratosphere increases from $\approx 0.3 %$ at 25 km to 0.9 % at 34 km. The increase is less in the (summer) Northern Hemisphere than in the (winter) Southern Hemisphere. Larger errors above 25 km are mainly attributable to the use of background data in bending angle retrieval as well as observational noise, which remains from uncalibrated ionospheric effects (Kursinski et al., 1997; Kuo et al., 2004). Below about 8 km the observational errors increase towards the surface mainly due to the complicated structure of humidity, superrefraction, oblique tangent point trajectories, and receiver tracking errors (Foelsche and Kirchengast, 2004; Kuo et al., 2004; Foelsche et al., 2011b).

In the upper troposphere and lower stratosphere CHAMP and GRACE-A observational errors are slightly larger than F3C observational errors but differences rarely exceed 0.1 % (cp. also Fig. 1). UCAR RO plus ECMWF refractivity combined and observational errors clearly feature a local maximum near the tropopause, where the difference variability between RO and ECMWF is naturally increased. At low latitudes it occurs at an altitude near 17 km, at high latitudes it decreases to approximately 10 km. This feature is likewise noticeable, though less pronounced due to GO retrieval instead of WO retrieval, in the WEGC data set. However, WEGC RO plus ECMWF combined errors as well as observational errors exhibit a distinctive local maximum at 15 km
Fig. 5. RO plus ECMWF combined error (yellow), estimated ECMWF error (green), observational error (light blue), estimated observational error (dark blue), and observational error model (red) in July 2008. UCAR (top two rows) and WEGC (bottom two rows) refractivity errors are shown for different latitude regions (global, NH, SH, low, mid, and high latitudes). The estimation includes RO data from CHAMP, GRACE-A, and F3C.
in all latitude regions. This error characteristic seems to be connected with the WEGC retrieval and related to the specific handling of the degrading quality of the L2 signal in the lower atmosphere, which necessitates the extrapolation of the signal down into the troposphere (Steiner et al., 1999). In the WEGC OPSv5.4 scheme (and earlier versions) the top height of this extrapolation is set to a fixed impact height of 15 km, statistically increasing fractional errors immediately below, while in the UCAR scheme this height is set dynamically, according to the quality of the L2 signal. We note that this characteristic in the WEGC retrieval does not cause systematic errors (see e.g. Fig. 2, where systematic differences of WEGC and UCAR to ECMWF do not differ), because the increased error behaves randomly from event to event.

### 3.4 GPS RO observational error model

Based on the above estimated observational errors we formulate a simple observational error model (indicated by the red lines in the panels of Fig. 5). This error model is derived from fitting simple analytical functions to the GPS RO observational error. The vertical structure of the model follows Steiner and Kirchengast (2005) and Steiner et al. (2006). Their error model $s_{\text{model}}$ as function of altitude $z$ is formulated as:

$$s_{\text{model}} = \begin{cases} s_0 + q_0 \left( \frac{1}{z} - \frac{1}{z_{\text{Top}}} \right) & \text{for } 4 \text{ km} < z \leq z_{\text{Top}} \\ s_0 & \text{for } z_{\text{Top}} < z < z_{\text{Bot}} \\ s_0 \times \exp \left( -\frac{z - z_{\text{Bot}}}{H_S} \right) & \text{for } z_{\text{Bot}} \leq z < 50 \text{ km} \end{cases}$$

where $z_{\text{Top}}$ is the top level of the troposphere domain, $z_{\text{Bot}}$ the bottom level of the stratosphere domain, $s_0$ is the error in the upper troposphere/lower stratosphere domain, $q_0$ the best fit parameter for the tropospheric model, $b$ the exponent, and $H_S$ the stratospheric error scale height. Note that the unit of $q_0$ depends on the value of $b$.

In addition to this vertical error model, we allocate a latitudinal and seasonal dependence in the form that any parameter $x$ can be modeled as a function of latitude $\varphi$ and time $\tau$:

$$x(\varphi, \tau) = x_0 + \Delta x f(\varphi) \cdot g(\tau, \varphi),$$

where $x_0$ is the basic mean magnitude of the parameter and $\Delta x$ is the maximum amplitude of seasonal variations.

The function $f(\varphi)$ accounts for latitudinal variations, the function $g(\tau, \varphi)$ combines latitudinal and seasonal changes of atmospheric conditions:

$$f(\varphi) = \max \left\{ 0, \min \left[ \frac{1}{\varphi - \varphi_{\Delta \chi_{\text{lo}}}}, 1 \right] \right\},$$

$$g(\tau, \varphi) = \text{sign}(\varphi) \cos(2\pi\tau).$$

Fig. 6. Temperature stratospheric error scale height as a function of latitude for different months (different colors) as modeled by Eq. (15).

with

$$\tau = \begin{cases} \frac{(m - 1) - m_{\text{lag}}}{12} & \text{for } m \in \{1, \ldots, 12\} \\ 3 \cdot s - m_{\text{lag}} \cdot \frac{1}{12} & \text{for } s \in \{1, \ldots, 4\} \\ (d - 15) - 30.5 \cdot m_{\text{lag}} \cdot \frac{1}{366} & \text{for } d \in \{1, \ldots, 366\} \end{cases}$$

The function $f(\varphi)$ is zero at low latitudes (equatorwards of $\varphi_{\Delta \chi_{\text{lo}}}$). Between $\varphi_{\Delta \chi_{\text{lo}}}$ and $\varphi_{\Delta \chi_{\text{hi}}}$ it linearly increases to +1, polewards of $\varphi_{\Delta \chi_{\text{hi}}}$ it is constant (+1) again. The function $g(\tau, \varphi)$ yields always positive values in the winter hemisphere and negative values in the summer hemisphere. The model can be applied on a daily base with $d$ being days of year, on a monthly base with $m = 1$ representing January and $m = 12$ being December, or on a seasonal base starting with $s = 1$ in March-April-May (MAM).

We apply Eq. (11) to the stratospheric error scale height $H_S$ used in Eq. (10), i.e. we set $x_0 = H_{S0}$ and $\Delta x = -H_S$ and account by the latter parameter for larger/smaller observational errors in the stratosphere at high latitudes during wintertime/summertime, where the error scale height thus needs to be smaller/larger. This results in

$$H_S(\varphi, \tau) = H_{S0} - \Delta H_S f(\varphi) \cdot g(\tau, \varphi).$$

For $f(\varphi)$ we chose $\varphi_{\Delta H_{S0}} = 30^\circ$ and $\varphi_{\Delta H_{S0}} = 60^\circ$. Since the annual cycle of atmospheric parameters does not show a temporal lag at high latitudes, we chose $m_{\text{lag}} = 0$.

In order to illustrate the modeled seasonal dependence, Fig. 6 shows the WEGC temperature stratospheric error scale height as function of latitude and as it varies over the months of the year. The error scale height is constant (15 km) between 30$^\circ$ S and 30$^\circ$ N. At higher latitudes it increases or decreases with season. In January, for example, the error scale height is 23 km at high southern latitudes and only 7 km at high northern latitudes. Note, that a smaller error
scale height yields a stronger upward increase of observational errors. Beside the extrema, two months always feature the same error scale height according to the cosine function (e.g. February and December, March and November...).

The error model given by Eqs. (10) and (15) depends on a few parameters only. Expressing the model in general parlance, it exhibits the following simple height structure: the error is assumed to be constant around the tropopause region (a few kilometers below and above). Above this region it increases exponentially with the scale height of the error increase in the stratosphere $H_S$. Below this region the error increases closely proportional to an inverse height power-law. This functional behavior makes good physical sense as discussed by Steiner and Kirchengast (2005). We fit observational errors of all RO atmospheric parameters by the model of Eqs. (10) and (15). The most suitable parameter values for UCAR and WEGC data are summarized in Table 1.

Returning to Fig. 5 and comparing UCAR and WEGC refractivity observational errors it attracts attention that UCAR errors are distinctly larger in the lower troposphere. While UCAR errors increase from 0.3 % at 10 km to 1.5 % at 4 km, the increase in the WEGC data set is smaller: refractivity observational errors at 4 km generally remain smaller than 1.0 %. This is because the UCAR WO retrieval is able to capture more small-scale atmospheric variability than ECMWF and WEGC (GO retrieval). The parameters $q_0$ and $b$ are adjusted to account for this fact.

The error model is applicable up to 50 km for raw bending angle, optimized bending angle, and refractivity (up to 35 km for the other parameters). In addition, information on the error correlation structure is provided for these variables, for construction of error covariance matrices and their use, e.g. in NWP assimilation systems. Figure 7 shows global error estimates, error correlation functions and their models for UCAR raw and optimized bending angles, as well as refractivity up to 50 km.

The impact of statistical optimization can clearly be seen comparing raw and optimized bending angle (Fig. 7, top panels). While the raw bending angle observational error exceeds 8 % near 45 km, the observational errors of the statistically optimized bending angle and of refractivity remain smaller than 4 % up to 50 km (larger errors can occur in the lower troposphere). To account for the impact of background information at high altitudes, the observational error model for the optimized bending angle uses a lower $z_{S\text{bot}}$ and a larger $H_{S0}$ and $\Delta H_S$.

The bottom panels of Fig. 7 show the empirical and modeled error correlation functions. Very good agreement with the results obtained by Steiner and Kirchengast (2005) is given and we follow their approach to model error correlations for raw bending angle by using a Mexican Hat function of the form.

$$ S = S_{ij} = s_i s_j \cdot \left( 1 - \frac{(z_i - z_j)^2}{c \cdot L^2} \right) \cdot \tilde{f}, \quad (16) $$

where $S$ is the error covariance matrix with elements $S_{ij}$, $s_i$ and $s_j$ are the standard deviations at height levels $z_i$ and $z_j$, respectively, $c$ is a stretching factor, and $L = L(z)$ the height dependent correlation length. The function $\tilde{f}$, representing the Gaussian exponential factor in the Mexican Hat function, is modeled for robust invertability of $S$ after Gaspari and Cohn (1999) as formulated in detail by Steiner and Kirchengast (2005). We choose $c$ to be $c = 1.0$ below 14 km (hardly

### Table 1. UCAR and WEGC model parameter values for the simple analytical error model for different RO atmospheric parameters, applicable to UCAR $\alpha_{\text{raw}}$, $\alpha_{\text{opt}}$, and $N$ up to 50 km, to other UCAR and all WEGC atmospheric parameters up to 35 km. Note that the unit of $q_0$ depends on the parameter $b$, which is given in the column to the right.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>UCAR</th>
<th>WEGC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_{\text{Top}}$</td>
<td>14.0 km</td>
<td>14.0 km</td>
</tr>
<tr>
<td>$z_{S\text{bot}}$</td>
<td>28.0 km</td>
<td>25.0 km</td>
</tr>
<tr>
<td>$s_0$</td>
<td>0.8 %</td>
<td>0.8 %</td>
</tr>
<tr>
<td>$q_0$</td>
<td>$20.0% km^b$</td>
<td>$10.0% km^b$</td>
</tr>
<tr>
<td>$b$</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>$H_{S0}$</td>
<td>9.0 km</td>
<td>16.0 km</td>
</tr>
<tr>
<td>$\Delta H_S$</td>
<td>2.0 km</td>
<td>2.0 km</td>
</tr>
</tbody>
</table>
any overshooting of error correlations that allows error correlation values to be negative), linearly decreasing to $c = 0.8$ at 50 km (stronger overshooting). The correlation length $L$ was estimated to be 1.5 km at 50 km, linearly decreasing to 0.7 km at 14 km and kept at this value downwards. The error correlation model obtained from raw bending angle data is also plotted for cross-check for the statistically optimized bending angle, which is generally not used for assimilation. The model almost perfectly fits also these data up to the 30 km level below which background information is negligible, but above the influence of background information becomes noticeable through a much larger error correlation length and no overshooting. Error correlations at these altitude levels would require a different model such as, e.g. applied for refractivity.

Refractivity error correlations are modeled after Steiner et al. (2006) using an exponential function of the form

$$S = S_{ij} = s_i s_j \cdot \exp \left( - \frac{|z_i - z_j|}{L(z)} \right), \quad (17)$$

where the correlation length $L(z)$ was estimated to be 1 km up to 30 km altitude and then linearly increasing to $L = 10$ km at 50 km. This increase accounts for an increasingly larger amount of background information at higher altitudes.

The error correlation functions are also usable for the WEGC data over their domain of availability up to 35 km.

Figure 8 shows UCAR and WEGC observational errors and the corresponding simple error model for different atmospheric parameters. Apart from the troposphere below 10 km, bending angle and refractivity observational errors are similar for both data sets; the observational error model is the same. However, pressure, geopotential height, and temperature observational errors are noticeably larger in the UCAR data set in the lower stratosphere. These larger UCAR observational errors are attributable to the NCAR climatology, which is used for bending angle initialization. This monthly mean climatology captures less atmospheric variability than both ECMWF and WEGC, which results in larger observational errors due to downward propagation of initialization errors in the retrieval (hydrostatic integration) of pressure and further parameters. To account for these errors, $H_S$, which is modeled utilizing $H_S$ and $\Delta H_S$, is smaller for UCAR than for WEGC.
Fig. 8. Global UCAR (left panels) and WEGC (right panels) error estimates for optimized bending angle, refractivity, dry pressure, dry geopotential height, and dry temperature in July 2008.
4 Summary and conclusions

In this study, we presented results of an empirical error analysis of GPS radio occultation (RO) data. We analyzed RO profiles of bending angle, refractivity, dry pressure, dry geopotential height, and dry temperature using data from CHAMP, GRACE-A, and Formosat-3/COSMIC (F3C). Comparing data processed at UCAR and at WEGC, we focused on the time period from January 2007 to December 2009.

In general, CHAMP, GRACE-A, and F3C data characteristics are in very good agreement. Global mean CHAMP and GRACE-A observational errors are only slightly larger than those of F3C, the differences being less than 0.3 % in bending angle, 0.1 % in refractivity, and 0.2 K in dry temperature at all altitude levels between 4 km and 35 km. Above about 35 km the CHAMP raw bending angle observational error increases stronger than that of GRACE-A and F3C. The difference exceeds 1 % above 42 km. Due to bending angle initialization at high altitudes, differences between the satellites above 35 km are smaller for other atmospheric parameters. Observational errors for the different F3C satellites exhibit the same characteristics. Above ≈20 km, the observational error shows a distinct seasonal dependence at high latitudes. Largest stratospheric errors are observed at high latitudes during wintertime, when atmospheric variability is most pronounced. These errors are mainly associated with background data used in the initialization for the reduction of bending angle noise. The impact of background data used in the processing chain is also noticeable when comparing retrieval results from UCAR and WEGC. While UCAR uses a monthly NCAR climatology, WEGC utilizes co-located ECMWF forecast profiles for initialization. Since the monthly climatology significantly underestimates true atmospheric variability, UCAR observational errors are larger than WEGC errors above ≈20 km in dry pressure, dry geopotential height, and dry temperature. However, below 35 km these differences are not seen in bending angle and refractivity. Larger UCAR bending angle observational errors below 20 km are attributable to the UCAR WO retrieval, which captures more (true) small-scale atmospheric variability than comparison data sets.

Based on the empirically derived error estimates we provided a simple analytical model of the observational error for GPS RO atmospheric profiles of bending angle, refractivity, dry pressure, dry geopotential height, and dry temperature. Using the formulations introduced by Steiner and Kirchengast (2005) and Steiner et al. (2006) for the vertical error structure, we extended the model to also represent the latitudinal and seasonal dependencies of the observational error. In the model a constant error is adopted around the tropopause region (a few kilometers above and below), which amounts to 0.8 % for bending angle, 0.35 % for refractivity, 0.15 % for dry pressure, 10 m for dry geopotential height, and to 0.7 K for dry temperature. Below this region the observational error increases following an inverse height power-law. Above this region the error increases exponentially with the latitudinal and seasonal behavior modeled through variation of the stratospheric error scale height.

The error correlation structure of bending angles follows a Mexican hat function including negative correlations while the error correlations of refractivity can be modeled by using a simple exponential function.

The observational error model is the same for UCAR and WEGC data but due to slightly different error characteristics below about 10 km and above about 20 km some parameter values had to be adjusted. Basically, the model is easily applicable in fields like operational meteorology or climate monitoring (see also Scherllin-Pirscher et al., 2011) and all parameter values in the model can be readily adjusted for individual error characteristics of other RO data sets, e.g. from other processing chains than UCAR and WEGC.

The empirical estimates of the observational error from real data in this study were found somewhat larger, especially in the core domain from about 10 km to 20 km, than earlier error estimates based on theoretical and end-to-end simulation studies (Kursinski et al., 1997; Rieder and Kirchengast, 2001a; Steiner and Kirchengast, 2005). However, the empirical estimates are considered tentatively conservative, which is viewed as a good practice; more evidence and new validation results in future may further refine the model. Overall the results underpin the high quality of RO data in the upper troposphere and lower stratosphere region.

Acknowledgements. We thank ECMWF (Reading, UK) for providing analysis data. Furthermore we are grateful to the UCAR/CDAAC and WEGC GPS RO operational team members, at WEGC especially to J. Fritzter and F. Ladstädter, for their contributions in OPS system development and operations. We also thank S. Sokolovskiy for helpful discussions. AKS was funded by the Austrian Science Fund (FWF) under grant P21642-N21 (TRENDEVAL project) and the work was also co-funded by grant P22293-N21 (BENCHCLIM project). WEGC OPS development was co-funded by ESA/ESTEC Noordwijk, ESA/ESRIN Frascati, and FFG/ALR Austria. At UCAR the work was supported by the National Science Foundation under Cooperative Agreement No. AGS-0918398/CSA No. AGS-0939962. The National Center for Atmospheric Research is also supported by the National Science Foundation.

Edited by: K. B. Lauritsen

References


www.atmos-meas-tech.net/4/1875/2011/
Syndergaard, S.: Retrieval analysis and methodologies in atmospheric limb sounding using the GNSS radio occultation technique, Danish Meteorological Institute, Copenhagen, Denmark, 1999.