Application of dynamical error estimation for statistical optimization of radio occultation bending angles

Martin S. Lohmann

Joint Center for Satellite Data Assimilation, National Environmental Satellite Data and Information Service, NOAA, Camp Springs, Maryland, USA

Received 9 July 2004; revised 19 January 2005; accepted 1 March 2005; published 8 June 2005.

[1] Errors in stratospheric bending angles retrieved from GPS radio occultations can be significantly reduced by using statistical optimization. In order for this technique to work optimally, the error covariance of the observations and the error covariance of the first guess must be known. This is generally not the case, and it is therefore common practice to assume that each of these errors is uncorrelated. In this study, it is shown that when this assumption is applied together with dynamical error estimation, it is important to account for the fact that first guess and the observation bending angle errors are not damped equally when refractivity profiles are computed through the Abel transform. It is demonstrated that this difference in damping can be accounted for by scaling the ratio of observation to first-guess bending angle error variances. It is shown that the scaling factor can be related to the ratio between the error correlation lengths of the observation errors and the first-guess errors. We present a simple procedure where both correlation lengths and variances are estimated dynamically and are scaled as described above. It is found that the relative errors can be reduced by up to 30% when compared to a standard statistical optimization scheme where the relative error in the first-guess bending angle profile is assumed to be 20%. However, it was also found that if the first guess is properly adjusted, the need for corrections is greatly reduced.


1. Introduction

[2] Radio occultation soundings of the Earth atmosphere allows for observations of temperature, pressure, and humidity. In the radio occultation (RO) technique these geophysical parameters are related to the Doppler shift imposed by the atmosphere on a signal emitted by a GPS satellite and received by a low Earth orbiting (LEO) satellite [see, e.g., Kursinski et al., 1997, 2000; Rocken et al., 1997]. During a RO observation the LEO satellite sets behind the Earth limb and thus provides a vertical scan of the atmosphere. Because of refractivity index gradients in the atmosphere, radio waves are bent as the radio signal traverses the atmosphere. The bending angles of the rays are directly related to the measured Doppler shifts and may be inverted into a profile of refractivity using the Abel transform [Fjeldbo et al., 1971]. As GPS signals received by a LEO satellite propagates through the ionosphere the measured Doppler shifts and bending angles include unwanted contributions from the ionosphere. For reconstruction of the refractivity profile of the neutral atmosphere this effect must be removed. At GPS frequencies the neutral atmosphere is nondispersive, whereas the ionosphere is dispersive. This allows for removal of the ionospheric contributions to first order through a linear combination of pairs, corresponding to the two frequencies in the GPS system, of either signal phase [Spilker, 1980], phase modulation [Melbourne et al., 1994], bending angle [Vorob’ev and Krasil’nikova, 1994], or Doppler shift [Ladreiter and Kirchengast, 1996]. These techniques are referred to as ionospheric calibration.

[3] Even when ionospheric calibration is applied, the retrieved bending angle profiles are dominated by noise from residual errors in the calibration at heights above 40–60 km. When refractivity profiles are computed from bending angle profiles through the Abel transform, errors propagate from higher altitudes to lower altitudes. Therefore it is essential to reduce these errors to get accurate refractivity observations in the stratosphere. A common
approach to reduce high-altitude bending angle errors is to use statistical optimization (SO) where observed bending angle profiles are combined with a priori or first-guess profiles. First-guess profiles are normally obtained from some climate model, though, in principle, any kind of a priori information about the bending angle profile can be used as the first guess, for example, a bending angle profile obtained from a weather prediction model. In this study, only first-guess profiles obtained from a climate model are considered.

If the error covariances, B_{obs}, of the measured bending angle profiles and the error covariance of the first-guess profile, B^{guess}, are known, the statistical optimal bending angle profile, \( \alpha_{opt} \), is found by minimizing the cost function

\[
J(\alpha) = (\alpha - \alpha_{obs})^T B^{-1}_{obs} (\alpha - \alpha_{obs}) + (\alpha - \alpha_{guess})^T B^{-1}_{guess} (\alpha - \alpha_{guess}) = \min,
\]

where \( \alpha_{obs} \) and \( \alpha_{guess} \) are the observation and the first-guess bending angle vectors, respectively [see, e.g., Gorbunov et al., 1996; Healy, 2001]. The solution that minimizes (1) is

\[
\alpha_{opt} = \left(B^{-1}_{obs} + B^{-1}_{guess}\right)^{-1} \left(B^{-1}_{obs} \alpha_{obs} + B^{-1}_{guess} \alpha_{guess}\right),
\]

with corresponding error covariance \( B_{opt} \) [Rodgers, 2000]:

\[
B_{opt} = \left(B^{-1}_{guess} + B^{-1}_{obs}\right)^{-1}.
\]

In practice, however, the error covariances are not known. While the observation error covariances may be estimated from the topmost points in the observed bending profiles (where the observation noise overshadows the signal from the neutral atmosphere), only rough estimates are available for the error covariances of the climate models which are normally used as the first guess in SO. The lack of knowledge of climate model error covariances constitutes the main limitation to the use of SO for RO data. Hence, for practical application of SO to RO data, some assumptions must be made about these error covariances.

A common approach, which has been applied in a number of studies [Gorbunov et al., 1996; Sokolovskiy and Hunt, 1996; Hocke, 1997; Rocken et al., 1997; Gorbunov and Gurvich, 1998; Steiner et al., 1999; Hajj et al., 2002; Kuo et al., 2004], is to assume that all errors are uncorrelated. In this case (2) simplifies to

\[
\alpha_{opt}(a) = \frac{a_{obs}(a) \sigma^2_{guess}(a) + a_{guess}(a) \sigma^2_{obs}(a)}{\sigma^2_{guess}(a) + \sigma^2_{obs}(a)},
\]

where \( a \) is the impact parameter and \( \sigma_{guess} \) and \( \sigma_{obs} \) are the standard deviations of the first-guess bending angle errors and the observed bending angle errors, respectively. It is worth noting that (2) also simplifies to (4) if both the observation and background error variances are constant with height and the two correlation functions defining \( B_{obs} \) and \( B_{guess} \) are identical except for a scaling factor so that \( B_{obs} \propto B_{guess} \).

When applying this method, the observation errors are normally estimated dynamically from the upper part of the occultation and the standard deviation of the first-guess errors is assumed to correspond to a fixed fraction of the first-guess bending angle, typically 10% to 20%, as originally suggested by Sokolovskiy and Hunt [1996]. The popularity of this approach is mainly due to its simplicity and the fact that this technique has been demonstrated to work well in a number of studies (see the references above). In the following we will refer to this approach as the standard method. However, this method is not optimal mainly for two reasons. First, it is assumed that the first-guess error standard deviation corresponds to a fixed ratio of the first-guess bending angle though the accuracy of climate models vary with both season and geographical location. Second, this approach does not account for vertical correlation of the observation and first-guess errors.

Healy [2001] applied the full matrix approach given by (1) assuming a Gaussian error correlation function with a correlation length of 6 km for the first-guess errors and uncorrelated observation errors. A fixed relative first-guess error of 20% and a fixed observation error standard deviation of 5 \( \mu \)rad were also assumed. Gobiet and Kirchengast [2004] also applied the full matrix approach, but in their study the observation error variance was estimated for each occultation from the upper part of the occultation following Sokolovskiy and Hunt [1996] and the relative error of the first guess was set to a fixed value of 15%. The error covariance for both the first-guess errors and the observation errors were assumed to be exponential with correlation lengths of 6 km and 1 km, respectively. To reduce biases between the climatology and the observations, Gobiet and Kirchengast [2004] selected their first guess from an independent search library of climate profiles.

Though the full matrix approach in principle accounts for vertical error correlations, the efficiency is limited by the accuracy of the assumed covariance...
functions. Currently, only crude estimates are available for climate model error covariances and, moreover, the model error covariances are a function of season and geographical location as mentioned above. For that reason, application of the full matrix solution may not necessarily lead to any improvements compared to the standard method.

[9] Gorbunov [2002] presented a combined algorithm of ionospheric calibration and noise reduction based on SO. In that study, dynamic estimates of the magnitude of both observation and first-guess errors were applied, while vertical correlations were neglected. Gorbunov [2002] also suggested to reduce systematic deviations between climatology and observations by scaling the climatology on the basis of a best fit between observations and climatology in the height range of 40–60 km. This technique was also used by Gobiet and Kirchengast [2004]. By estimating errors dynamically for each occultation, geographical and seasonal variations are automatically accounted for. However, when dynamic estimation of both first-guess errors and observation errors is applied, special care should be taken if vertical correlations are neglected and refractivities are computed through the Abel transform. The reason for this is that the Abel transform is a low-pass filter. Hence observation errors and first-guess errors are only damped (or amplified) equally if the corresponding error covariances have the same shape. Generally, climate model errors have more low spatial frequency components than observation errors. It is therefore expected that the observation errors will be damped more by the Abel transform than the model errors. Thus the combination of the two profiles, given by (4), which results in the smallest errors in the bending angle profile, is not necessarily identical to the combination of the two profiles which results in the smallest errors in the corresponding refractivity profile. Consequently, if (4) is applied together with dynamic error estimation, too much weight could be given to the model profile. In that case, even if the error standard deviations have been estimated perfectly, the retrieved refractivity profile may well deviate more from the “true” profile than a corresponding profile computed through the standard method.

[10] In this study we present an approach where first-guess error covariances and observation error covariances of the ionospherically corrected bending angle profile are estimated dynamically. The combined profile is computed from (4) with modified standard deviations based on the estimated error correlation lengths.

[11] This study is organized in the following way: an analysis of how noise is damped by the Abel transform is given in section 2. In section 3 it is demonstrated how first-guess and observation error covariances can be estimated dynamically. Section 4 presents a comparison between the standard SO method and the technique presented in this study. Both methods are applied to one month of Challenging Minisatellite Payload (CHAMP) radio occultations, and the retrieved refractivity profiles are compared to corresponding profiles computed from the European Centre for Medium-Range Forecasts (ECMWF) analysis. The importance of the first-guess profile is also demonstrated by applying two different types of first-guess profiles derived from the same climate model but adjusted differently to the observations.

2. Analysis of the Abel Transform

Frequency Response

[12] When a bending angle profile, which includes additive noise, is Abel transformed, the characteristics of the noise are changed. In this section we investigate how noise with different correlation lengths/bandwidths are damped by the Abel transform by analyzing the frequency response of the Abel transform. Similar analysis of error propagation through the Abel transform based on the impulse response of the Abel transform and on error covariance analysis can be found in work by Kursinski et al. [1997] and Syndergaard [1999], respectively.

[13] A profile of the refractivity index, n, can be computed from a profile of bending angles, \( \alpha \), under the assumption of spherical symmetry, through the following Abel transform [Fjeldbo et al., 1971]:

\[
n(r_0) = \exp \left( \frac{1}{\pi} \int_{a_0}^{a_1} \frac{\alpha(a')}{\sqrt{a'^2 - a_0^2}} da' \right),
\]

where the impact parameter, \( a_0 \), is related to the geometrical height of the tangent point, \( r_0 \), through \( a_0 = n(r_0)r_0 \), and \( a_1 \) is the impact parameter corresponding to the topmost point in the bending angle profile. To analyze how noise propagates through this Abel transform we linearize (5). In this case the refractivity, \( N = 10^6(n - 1) \), can be expressed as

\[
N(r_0) \approx 10^6 \left( \frac{1}{\pi} \int_{a_0}^{a_1} \frac{\alpha(a')}{\sqrt{a'^2 - a_0^2}} da' \right).
\]

The frequency response, \( F \), of (6) is

\[
F(k) = 10^6 \left( \frac{1}{\pi} \int_{a_0}^{a_1} \frac{\exp(ika')}{\sqrt{a'^2 - a_0^2}} da' \right),
\]
where \( k \) is the wave number. It can be shown that when \( a_1 \ll 3a_0 \) or \( a_0k \gg 2 \) the following expression yields an excellent approximation to (6):

\[
\begin{align*}
F(k) &= \exp(ika_0) \frac{10^6}{\sqrt{\pi ka_0}} \left[ C\left(\frac{2}{\pi} k(a_1 - a_0)\right) + iS\left(\frac{2}{\pi} k(a_1 - a_0)\right)\right]. 
\end{align*}
\]

In (8) \( C(x) \) and \( S(x) \) are the Fresnel cosine integral and the Fresnel sine integral, respectively, defined as

\[
\begin{align*}
C(x) &= \int_0^x \cos\left(\frac{\pi v^2}{2}\right) dv, \\
S(x) &= \int_0^x \sin\left(\frac{\pi v^2}{2}\right) dv.
\end{align*}
\]

From (8) it follows that for an infinite integration interval the magnitude response, \(|F(k)|\), of the Abel transform is simply \(10^6/(2\pi ka_0)^{1/2}\).

[14] Figure 1 shows the magnitude, \(|F(k)|\), of the frequency response of the Abel transform for \( a_0 = 6400 \) km and \( a_1 - a_0 = \Delta a = 10 \) km, 30 km, 100 km, computed from (7), and \( \infty \), computed from (8). As follows from Figure 1, the Abel transform is a low-pass filter with a cutoff frequency corresponding approximately to the inverse length of the integration interval and an infinite transition band described approximately by \(10^6/(2\pi ka_0)^{1/2}\). In RO statistical optimization, climate model errors contain more low spatial frequency components than observation errors. Consequently, it is anticipated that the observation errors will be damped more by the Abel transform than the first-guess errors.

[15] In order to investigate how first-guess errors and observation errors propagate through the Abel transform, we assume that both types of errors have a Gaussian covariance function, \( r(\tau) \), and hence a Gaussian power spectrum \( R(k) \):

\[
\begin{align*}
r(\tau) &= r(0) \exp\left(-\frac{\tau^2}{2}\right), \\
R(k) &= r(0) \frac{l}{\sqrt{\pi}} \exp\left(-\frac{(k\ell)^2}{2}\right),
\end{align*}
\]

where \( l \) is the correlation length and \( r(0) \) is the noise power density. The damping, \( d \), as a function of the correlation length can now be computed as

\[
d(l) = \frac{\int R(k)|F(k)|^2 dk}{r(0)}.
\]

From (11), the damping of the observation errors, \( d_{\text{obs}} \), and the damping of the first-guess errors, \( d_{\text{guess}} \), can be computed.

[16] Correlation lengths of observation errors are approximately in the range from 0 to 2 km depending on the filter applied for smoothing the observation phase. Correlation lengths of climate model errors are generally not known and may be in the range from a few kilometers to more than 20 km depending on the climate model used as the first guess. Figure 2 depicts the ratio

**Figure 1.** Magnitude response of Abel transform for different lengths of the integration interval, \( \Delta a \).

**Figure 2.** Ratio of damping factors for the observation noise and first-guess noise as a function of the first-guess error correlation length. The three different curves correspond to different correlation lengths of the observation noise, namely, 0, 1, and 2 km.
between the \(d_{\text{obs}}\) and \(d_{\text{guess}}\) for different correlation lengths when \(a_0 = 6400\) km and \(\Delta a = 100\) km.

Figure 2 shows that a few kilometers difference between the correlation lengths of the first-guess errors and the correlation length of the observation errors can lead to significantly different damping of the two noise sources when the Abel transform is applied. For instance, the observation noise power will be damped four times more than the noise power of the first guess if the correlation length of the observation noise is 1 km and the correlation length of the first-guess errors is 5 km. This clearly illustrates the necessity for accounting for the difference in correlation lengths when dynamic error estimation is applied. Without proper account, the first guess may be overweighted after applying Abel inversion of the optimized bending angle. In the standard SO technique this is not necessarily a problem as a fixed relative error of the first guess is assumed, and if this ratio is chosen to be reasonably large the damping effect of the Abel transform is implicitly accounted for.

### 3. Statistical Optimization With Dynamic Estimation of Error Covariances

In section 2 it was demonstrated that when dynamic error estimation is applied it may be necessary to account for the difference between the noise power spectrum of the observation errors and the noise power spectrum of the first-guess errors. In the following, we therefore present a simple approach where the first-guess profile is combined with the observed profile, using (4), together with dynamic error estimation. A distinctive feature of this approach is that the estimated errors are adjusted on the basis of estimates of the correlation lengths of these errors according to the damping caused by the Abel transform.

The first step in this approach is to estimate variances and correlation lengths for the two types of errors. For the observation errors, the variance and the correlation length can be estimated from the differences, \(\Delta a\), between the measured ionosonde free bending angles and the first-guess bending angles in the height range from 60 to 80 km where the observation noise overshadows the signal from the neutral atmosphere [Sokolovsky and Hunt, 1996; Kuo et al., 2004]. These estimates can be applied also at heights below 60 km as the structure and magnitude of the observation errors are rather uniform below the \(E\) layer [Kuo et al., 2004]. This variance is simply estimated as

\[
\sigma_{\text{obs}}^2 = \left\langle (\Delta a(a))^2 \right\rangle \approx \frac{1}{M} \sum_i (\Delta a_i)^2,
\]

where angle brackets represents ensemble average, \(a_i\) is the impact parameter corresponding to sample number \(i\), and \(M\) is the number of samples at altitudes from 60 to 80 km. The correlation length, \(l_{\text{obs}}\), can be estimated by applying a Gaussian fit to the covariance function, \(r_{\text{obs}}(\tau)\), of \(\Delta a\) where \(r_{\text{obs}}(\tau)\) is computed using the following estimator [Proakis and Manolakis, 1996]:

\[
r_{\text{obs}}(\tau) = \left\langle \Delta a(a)\Delta a(a + \tau) \right\rangle \\
\approx \frac{1}{M_f} \sum_i \Delta a_i \Delta a_i + \tau),
\]

where \(M_f\) is the number of samples in the height range from 60 to \((80 - \tau)\) km.

Estimation of the first-guess error variance and correlation length is more complicated because the magnitude of the first-guess errors varies with height. Thus some assumption about the structure of these errors must be made. Here we apply the assumption that the magnitude of the first-guess error standard deviation is equal to a fixed fraction, \(K\), of the first-guess bending angles. Higher accuracy may be achieved by applying a more advanced error model which could take into account the fact that the relative errors generally increase with height. Now, as the observation errors and first-guess errors are expected to be uncorrelated, the following expression must be valid at any \(a\):

\[
\left\langle \Delta a(a)\Delta a(a + \tau) \right\rangle = r_{\text{obs}}(\tau) \\
\quad + \frac{r_{\text{guess}}(\tau)}{r_{\text{guess}}(0)} K^2 \alpha_{\text{guess}}(a) \alpha_{\text{guess}}(a + \tau),
\]

At high altitudes, the last term in (14) is negligible compared to \(r_{\text{obs}}(\tau)\), and (14) simplifies to (13).

The correlation function can be estimated at zero lag in the height range from 20 to 60 km through

\[
K^2 \approx \frac{1}{N} \sum_i (\Delta a_i)^2 - \frac{1}{N} \sum_i \sigma_{\text{obs}}^2(a_i),
\]

where \(N\) is the number of samples in the height range from 20 to 60 km. Similarly, the correlation function can be estimated as

\[
\frac{r_{\text{guess}}(\tau)}{r_{\text{guess}}(0)} \approx \frac{1}{N_f} \sum_i \Delta a_i \Delta a_i + \tau - \frac{1}{N_f} \sum_i \alpha_{\text{guess}}(a_i) \alpha_{\text{guess}}(a_i + \tau)
\]

where \(N_f\) is the number of samples in the height range from 20 to \((60 - \tau)\) km.
can be computed by (11). For simplicity the following approximation can be used:

\[
\frac{d(l_{\text{obs}})}{d(l_{\text{guess}})} \approx \left( \frac{l_{\text{obs}}}{l_{\text{guess}}} \right)^{0.82},
\]

(17)

which allows for adjustment of the ratio between \(\sigma_{\text{obs}}\) and \(\sigma_{\text{guess}}\) before (4) is applied. The factor 0.82 was determined by a best fit for \(\Delta a > 20\) km and 0.1 km < \(l\) < 10 km and offers a reasonable approximation in that interval, with increasing accuracy with increasing \(\Delta a\). It should be noted that (17) is based on Gaussian error covariance functions, which have been assumed for simplicity.

[23] The correlation functions estimated from (13) and (16) could in principle be used in the full matrix approach given by (1). However, (17) is preferred as it is simple to implement and computationally efficient. Also, it should be kept in mind that the error correlation lengths computed from the procedure described above are not exact but should rather be considered as a measure of the noise bandwidths. Finally, we can summarize the approach described in this section as follows.


[27] 4. Find correlation lengths by applying Gaussian fit to the estimated error correlation functions.

[28] 5. Compute the ratio between the damping of the observation errors and the first-guess errors using (17).

[29] 6. Multiply the estimated observation variance with \(d_{\text{obs}}/d_{\text{guess}}\) or multiply the estimated first-guess variance with \(d_{\text{guess}}/d_{\text{obs}}\).


[31] 8. Use Abel transform to compute refractivity profile from \(\alpha_{\text{opt}}\).

4. Application to CHAMP Occultations

[32] We now apply the new SO technique described in section 3 and the standard technique, introduced in section 1, to 1 month of CHAMP occultations from August 2002. In the standard technique we use a fixed relative first-guess error of 20%. For both schemes the retrieved refractivity profiles are compared to a global operational analysis from ECMWF. The ionospheric free bending angle profiles used as input to the two SO schemes are based on geometrical optics above the troposphere as implemented in the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) Data Analysis and Archive Center (CDAAC) [Kuo et al., 2004].

[33] As the first guess we apply a National Center for Atmospheric and Climate Research (NCAR) climate model that is based on the results from the recent Stratospheric Processes and Their Role in Climate (SPARC) study [Randel et al., 2002]. To remove systematic deviations between the climatology and the observations, the climate model bending angle profiles are scaled to provide a least squares fit to the measured bending angles in the interval from 40 to 60 km as suggested by Gorbunov [2002]. The modified climate model bending angle profiles are used as the first guess for both the new SO scheme and the standard scheme.

[34] For both schemes we apply (4) from the top of the observed bending angle profile and down to 20 km altitude where the weight of the climatology generally is negligible. In the comparison with ECMWF we exclude occultations with very large ionospheric noise, defined as occultations for which the mean and the standard deviation of the ionosphere free bending angle from the first guess at altitudes between 60 and 80 km are larger than \(10^{-4}\) rad and \(1.5 \times 10^{-4}\) rad, respectively. Also, occultations believed to be affected by tracking errors were excluded. In total, approximately 10% of the occultations were rejected by the quality control.

[35] Furthermore, before equation (17) is applied to scale the estimated variances, error and correlation length estimates which are believed to be either unrealistically big or small are corrected, so that \(K\) estimates smaller than 1% are set to 1%. Similarly, the allowed intervals for \(l_{\text{obs}}\) and \(l_{\text{guess}}\) are 0 to 1.4 km and \(l_{\text{obs}}\) to 15 km, respectively. The latter criterion ensures that the ratio \(l_{\text{obs}}/l_{\text{guess}}\) is always less than unity. This criterion is introduced to prevent small errors in the correlation length estimates from resulting in an overestimation of the relative damping of the first-guess errors, since the ratio \(d_{\text{obs}}/d_{\text{guess}}\) becomes very sensitive to even small errors in \(l_{\text{obs}}\) and \(l_{\text{guess}}\), when \(l_{\text{obs}}/l_{\text{guess}} > 1\) (see Figure 2). The fraction of the occultations for which the correlation lengths exceeds these bounds depends on the first guess. In this study, two different types of first guess are used. The first type, described above, is based on correcting a bending profile \(\alpha_{\text{clim}}\) computed from a climatology by applying a single scaling factor, \(\beta\), that is, \(\alpha_{\text{guess}} = \beta \alpha_{\text{clim}}\). The second type, which will be described later, is derived from a climatology by applying both a scaling factor and an offset factor, \(\Delta\), so that \(\alpha_{\text{guess}} = \beta \alpha_{\text{clim}} + \Delta\). In the former case, it was found that the correlation lengths were corrected for 10% of the occultations, whereas the in the latter case this was true for 25% of the occultations. In both cases most correction was due to \(l_{\text{obs}}/l_{\text{guess}} > 1\).

[36] Figure 3 depicts the mean and the standard deviation of the fractional refractivity differences between retrieved refractivities and corresponding refractivities derived from the ECMWF analysis. Results are pre-
The fractional difference in refractivity is defined as follows:

\[
\frac{N_{\text{champ}} - N_{\text{ecmwf}}}{N_{\text{ecmwf}}} \; ;
\]

where \(N_{\text{champ}}\) is the observed refractivity and \(N_{\text{ecmwf}}\) is the refractivity from the ECMWF analysis interpolated to the time and location of the corresponding CHAMP occultations.

From Figure 3 it is seen that in the height range from 25 to 36 km the standard deviation of the new SO scheme is consistently smaller than the standard deviation of the standard scheme, for all latitudes, with the largest differences at the topmost end of that interval. The main improvement of approximately one percentage point, corresponding approximately to an improvement of 30\% in the absolute error, is observed in the Northern and Southern Hemispheres, while the improvements in the tropics are somewhat smaller. This can be explained by the lower ionospheric noise in the tropics which reduces the weight of the first guess relative to the observations and thus reduces the influence of SO. This is clearly seen in Figure 4, which shows the estimated standard deviations as function of latitude for the estimated ionospheric noise, which corresponds to the observation noise.
The differences in bias between the results from the two schemes are considerably smaller than the differences between the standard deviations, though the new scheme generally has a slightly smaller bias relative to ECWMF than the standard scheme. From Figure 3 it is seen that fractional refractivity differences relative to ECMWF are, in general, in the interval from 1 to 4%, whereas biases are in the interval from −1 to 1 percentage points for the considered height interval. The performance of both schemes is found to be somewhat poorer in the Southern Hemisphere than in the tropics and the Northern Hemisphere. The reason for this is likely to be significant day to day variations which are common for the stratospheric polar winter [Kuo et al., 2004]. Such variations increase the error of the first guess and result in larger errors in the retrieved refractivities. This is also seen in Figure 5, which shows the estimated relative errors of the first guess as a function of latitude. In the Southern Hemisphere the estimated relative errors are within range from 1 to 50%, while in the tropics and in the Northern Hemisphere most error estimates are within the range from 0.5 to 10%. The slight increase in the relative error observed at the equator is believed to be related to slightly higher interdiurnal variations in that region.

It is also informative to study the distribution of the estimated error correlation lengths. Figures 6 and 7 depict the estimated correlation lengths as a function of latitude for first-guess noise and observation noise, respectively. Figure 6 shows, as expected, that the correlation length of the observation noise is independent of latitude. The mean value is 0.8 km, which is in agreement with the fact that a low-pass filter with a bandwidth of 2 Hz has been applied to the measured L1 and L2 phases in the processing chain. The estimates of first-guess error correlation lengths show far more variability than the observation error correlation lengths. The largest variations are found in the Southern Hemisphere, where the majority of the estimated correlation lengths lie within the interval from 0.5 to 25 km. Smaller variations are found in the latitude band from 20°S to 90°N, where a large part of the observed correlation lengths are concentrated in the interval 0.25 km and 2 km. The larger correlation lengths in the Southern Hemi-
sphere are believed to be related to the higher interdiurnal variations of the stratospheric polar winter, which also resulted in higher relative first-guess errors.

In order to assess the importance of the accuracy of the first guess, we repeat the calculations, applying an alternative fitting technique to the climate model profile. In SO it is assumed that errors are unbiased and the technique is therefore sensitive to biases between first-guess profiles and observations [Gobiet et al., 2002; Gobiet and Kirchengast, 2003, 2004]. To reduce the first-guess bias, the climatology is now adjusted by both a scaling factor and an offset factor in the height range from the surface to 50 km. A smooth transition between the modified and the original profile is performed in the height range 50–60 km. Offset and scaling factors are computed for each occultation using a linear fit between pairs of observation and model bending angles at heights from 25 to 60 km, which is equivalent to minimizing

$$\sum [\alpha_{\text{obs}}(a_i) - (\beta \alpha_{\text{clim}}(a_i) + \Delta)]^2$$

with respect to $\beta$ and $\Delta$ in that height range. The height range used in the fitting approach was chosen on the basis of experience.

Figure 8 shows the resulting fractional errors as compared to ECWMF. Comparison between the results in Figures 8 and 3 shows that the modification to the fitting technique has a larger impact on the performance of the standard SO scheme than on the performance of the new scheme. The reduction in the standard deviations for the standard scheme is more than 0.5 and one percentage point (around 15 and 30% in absolute values) in the Northern and Southern Hemispheres, respectively, whereas only minor improvements are found in the tropics. For the new SO scheme the standard deviations are reduced by approximately 0.5 percentage points in the Southern Hemisphere and Northern Hemispheres, whereas the improvements at other latitudes are small. For both SO schemes the bias is nearly unaffected except in the Southern Hemisphere, where the bias relative to ECMWF is slightly increased, most notably for the new SO scheme. From Figure 8 it is also seen that when scaling and offset correction is applied to adjust the climate model bending angle profiles, the difference in performance between the two schemes is reduced. For heights with a significant number of matches with ECMWF, the largest differences of up to 0.25 percentage point are found in the Northern Hemisphere, which is somewhat smaller than the differences of one percentage point that occurred when the climate model was adjusted by scaling.

The fact that the change in fitting technique has the largest impact on the standard SO scheme indicates that the standard scheme is more sensitive to the choice of first-guess profile. This is not surprising as the standard technique requires a priori knowledge of the relative error of the first-guess bending angles, which is basically a tuning parameter. In the new SO approach the relative error is estimated dynamically, which makes this technique less sensitive to the choice of the first-guess profile. However, from Figure 8 it is also seen that in the Southern Hemisphere the standard deviations of the new scheme is only marginally smaller than the standard deviations of the standard scheme. This implies that with the right choice of relative error the performance of the standard scheme might be equal to or, perhaps, even better than the new scheme presented in this study. However, it should be noted that such an optimal value of the relative error would be a function of both season and latitude.

By using scaling and offset correction instead of scaling, both the relative errors and the error correlation lengths are greatly reduced. This is evident from Figures 9 and 10, which depict the estimated relative errors and the error correlation lengths corresponding to the climate model profiles modified by using scaling and offset correction. For this type of first guess the majority of the relative errors are within the interval from 0.5 to 10% in the Southern Hemisphere and within the intervals 1–4% and 0.5–3% in the tropics and in the Northern Hemisphere, respectively.

It is worth noting that the reduction in the deviations between the first guess and the observations, which are reflected in Figure 9, is due to two factors, the first being a genuine improvement of the first guess and the second being a result of using a fit with one more degree of freedom. For that reason, the relative error
estimates in Figure 9 represent lower bounds of the true errors. However, as the ionospheric correction is expected to work well for large scales, the large-scale deviations are dominated by large-scale errors in the climatology. The use of scaling and offset correction of the background profiles is basically a large-scale correction scheme of the climatology based on the observed large-scale variations. For that reason, it is believed that the reductions in the relative errors are due mainly to genuine improvements of the first guess, and the use of these error estimates is only expected to result in a slight overweighing of the first guess.

As for the correlations lengths, the largest variations are found in the Southern Hemisphere, where the majority of the estimated correlation lengths are within the interval from 0.25 to 7 km. The smallest variations are found in the latitude band from 20°S to 20°N, where a vast majority of the observed correlation lengths are between 0.25 and 1 km. In the Northern Hemisphere, most correlation lengths are within the interval from 0.25 to 2 km, though there are also a number of correlation lengths in the interval from 2 to 6 km. These results show that by both correcting the climate model for an offset and by scaling the performance of the SO, schemes are improved in two ways as compared to using simple scaling. First, the overall quality of the first guess is improved as the relative error is reduced. Second, the correlation lengths of the first-guess errors are reduced, which reduces the differences between the correlation lengths of the observation errors and the first-guess errors and thus reduces the errors introduced by neglecting vertical correlations.

The latitudinal dependence in the distributions of both relative errors and error correlation lengths is
generally unaffected by the change of fitting technique, though both mean and spread of these estimates are reduced considerably. The reason for this is that the overall trend is related to variations in the atmosphere which are not captured by the climate model nor completely accounted for by adjusting the model bending angles.

From Figures 6 and 10 it follows that the error correlation lengths of the observations are comparable to the error correlation lengths of the first guess in the latitude band from $-30^\circ$–$90^\circ$ when scaling and offset correction is applied. This suggests that it may not be necessary to correct for the differences in correlation lengths in that latitude band. To assess the merits of correlation length corrections, the calculation are repeated for the two adjustment schemes without any corrections based on the correlation lengths, which is equivalent to assuming that $l_{\text{obs}} = l_{\text{guess}}$. Figures 11 and 12 show the corresponding comparisons with ECMWF for scaling correction and scaling and offset correction of the climatology, respectively.

When the climatology is adjusted by scaling, the differences in correlation lengths are significant, and the correlation length corrections are expected to improve the agreement with ECMWF. This is seen in Figure 11, which shows that the correlation length corrections results in an improvement of up to nearly one percentage point. The largest improvements are in the Southern Hemisphere, where the biggest differences in correlation lengths are found. On the other hand, when the climatology is adjusted by both a scaling factor and an offset factor, the importance of correlation length corrections is significantly reduced, as expected.

5. Discussion

The results presented in section 4 show that above approximately 25 km the accuracy of radio occultation measurements is sensitive to the type of first-guess profile applied in the statistical optimization scheme. This is in agreement with Healy [2001] and Gobiet and Kirchengast [2004], who also found that SO is sensitive to the choice of first guess. It would therefore be desirable to use a more accurate first guess than an adjusted climate model profile. With the launch of the joint U.S.-Taiwan ROCSAT-3/COSMIC mission in late 2005 approximately 3000 RO soundings will be collected each day. With such a large number of occultations it might be possible to replace the climate model, as the first guess, with an averaged bending angle profile computed from earlier nearby occultations. For example, a 14 day average based on the 5% nearest occultations would give a first-guess profile with an error standard deviation due to ionospheric noise which is approximately 50 times smaller than the corresponding standard deviation of a single occultation. By doing so,
the RO measurements would also become independent of external data. This is particularly important if the data are to be used for climate monitoring or assimilation into numerical weather prediction models. This approach is generally expected to work well as the variability in the stratosphere is usually on timescales ranging from a week to months. However, the polar winters may require a more sophisticated approach because of larger diurnal variation.

It is also important to note that the influence of the first guess is smallest in the tropics, which makes occultations in that latitude band particularly suitable for climate research. For the same reason, the altitude to which the occultations can be considered as independent observations is also higher in this region.

The results presented in this study indicate that a fixed value of the relative error might be a better choice than using dynamical estimation if this value is chosen correctly. The methods presented in this study could be used to generate an “atlas” of the relative errors by averaging relative errors adjusted according to (17). This would give a “look-up” table of the relative error as a function of season and latitude which could be used to determine the relative error for individual occultations. Finally, it is worth noting that by estimating errors and error correlation lengths it becomes straightforward
to compute the stratospheric error covariance for individual refractivity and temperature profiles [see, e.g., Syndergaard, 1999].

6. Summary and Conclusion

[52] This study demonstrates that the Abel inversion of the bending angles, statistically optimized under the assumption of vertically uncorrelated errors, requires adjustment (scaling) of the estimated error magnitudes according to their estimated vertical correlation lengths. The adjustment is required because the Abel transform is a low-pass filter, which generally damps observation errors more than the errors in the climate models which are normally used as the first guess.

[53] The degree of the adjustment depends on the ratio between the correlation length of the observation errors and the first-guess errors. The more this ratio deviates from unity, the larger is the needed adjustment.

[54] These findings were applied in a new statistical optimization scheme in which the error variance and the error correlation length of both the ionospherically corrected bending angle profile and of the first guess are estimated for individual occultations. The observation error magnitude and correlation length were estimated from differences between first-guess bending angles and measured bending angles at 60 to 80 km heights, whereas the first-guess relative error and error correlation length were estimated from differences between first guess and observations at 20 to 60 km heights.

[55] This new statistical optimization (SO) scheme was applied to one month of CHAMP occultations from August 2002 and the retrieved refractivities were compared to corresponding refractivities derived from

Figure 12. Same as Figure 11 but with scaling and offset correction of climate model bending angle profile.
the ECMWF analysis. For comparison, similar calculations were performed by the commonly used “standard” statistical optimization scheme. In the standard scheme, the observation noise is estimated dynamically, whereas the standard deviation of the first-guess errors is assumed to be equal to 20% of the first-guess bending angles. Bending angle profiles derived from a NCAR stratospheric climate model were applied as the first guess. For each occultation this model profile was fitted to the observed bending angles in order to reduce systematic deviations between the first guess and the observations. Two different fitting techniques were tested for both SO schemes: (1) scaling of the climate model profile and (2) scaling and offset correction.

[56] It was found that, in general, the fractional refractivity differences relative to ECMWF were in the interval from 1 to 4% whereas biases were in the interval from –1 to 2% for the considered height interval, 20 to 40 km. The largest deviations from ECMWF occurred at heights above 35 km. Fractional errors were found to be significantly higher in the Southern Hemisphere than in the tropics and the Northern Hemisphere. Also, it was found that the fractional errors were smaller for the new scheme than for the standard scheme. However, the difference in performance was found to depend on how the climate model profile is fitted to the measured bending angles. When scaling of the climate model profile is applied, significant differences between the two SO schemes were observed. The largest differences were found in the Northern and Southern Hemispheres, where the fractional refractivity differences were up to one percentage point, corresponding to an improvement of up to 30% in the absolute error. When scaling and offset correction were applied to the climate model profile, the difference in performance between the two SO schemes was found to be smaller, with a maximum difference of approximately 0.25 percentage point or nearly 10% in absolute values. For both schemes the application of scaling and offset correction of the climate model profile led to improvements in performance as compared to using scaling, most notably for the standard SO scheme.

[57] To assess the importance of correcting for correlation length differences, the occultation data were also processed without correlation length corrections. Comparison with ECMWF showed that the correlation length corrections led to significant reduction in the deviations between observations and the ECMWF analysis when the climatology was adjusted by scaling, whereas only small improvements were achieved when the climatology was adjusted using scaling and offset correction.

[58] The comparisons with ECMWF showed that at heights above 25 km, the accuracy of the retrieved refractivities becomes sensitive to how the model profile is adjusted to the observations. In that region, it was also found that the new SO scheme is considerably less sensitive to the choice of first guess than the standard scheme.

[59] When scaling is used to fit the climate model bending angles to the observations, the estimated relative errors and error correlation lengths of the first-guess profile were found to be mainly in the ranges from 1 to 50% and 0.5 to 25 km, respectively. The largest relative errors and error correlation lengths were found in the Southern Hemisphere. For the climate model profiles modified by scaling and offset correction, the estimated relative errors and error correlation lengths were generally smaller and were found to be mainly within the intervals from 0.5 to 10% and 0.25 to 7 km, respectively. Also, in this case, the largest relative errors and error correlation lengths were found in the Southern Hemisphere. The larger relative errors of the first guess and larger fractional errors compared to ECMWF observed in the Southern Hemisphere are believed to be caused by significant day to day variations which are common for the stratospheric polar winter.

[60] Acknowledgments. The author is grateful to Sergey Sokolovskiy for valuable discussions on the topic, which have helped to improve this paper, and to Doug Hunt; without his help with the CDAAC database this work would not have been possible. This work was supported by National Environmental, Satellite, Data, and Information Service (NESDIS).

References


Gorbunov, M. E., and A. S. Gurvich (1998), Algorithms of inversion of Microlab-1 satellite data including effects of


