Edinburgh Research Explorer

A nonlinear optimal, estimation inverse method for radio occultation measurements of temperature, humidity, and surface pressure

Citation for published version:

Digital Object Identifier (DOI):
10.1029/2000JD900151

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Journal of Geophysical Research

Publisher Rights Statement:
Published in Journal of Geophysical Research: Atmospheres by the American Geophysical Union (2000)

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
A nonlinear optimal estimation inverse method for radio occultation measurements of temperature, humidity, and surface pressure

Paul I. Palmer and J. J. Barnett
Department of Physics, Clarendon Laboratory, Oxford, England

J. R. Eyre and S. B. Healy
Satellite Applications Division, United Kingdom Meteorological Office, Bracknell, England

Abstract. An optimal estimation inverse method is presented which can be used to retrieve simultaneously vertical profiles of temperature and specific humidity, in addition to surface pressure, from satellite-to-satellite radio occultation observations of the Earth's atmosphere. The method is a nonlinear, maximum a posteriori technique which can accommodate most aspects of the real radio occultation problem and is found to be stable and to converge rapidly in most cases. The optimal estimation inverse method has two distinct advantages over the analytic inverse method in that it accounts for some of the effects of horizontal gradients and is able to retrieve optimally temperature and humidity simultaneously from the observations. It is also able to account for observation noise and other sources of error. Combined, these advantages ensure a realistic retrieval of atmospheric quantities. A complete error analysis emerges naturally from the optimal estimation theory, allowing a full characterization of the solution. Using this analysis, a quality control scheme is implemented which allows anomalous retrieval conditions to be recognized and removed, thus preventing gross retrieval errors. The inverse method presented in this paper has been implemented for bending angle measurements derived from GPS/MET radio occultation observations of the Earth. Preliminary results from simulated data suggest that these observations have the potential to improve numerical weather prediction model analyses significantly throughout their vertical range.

1. Introduction

Radio occultation (RO) experiments have played a prominent role in the NASA program for solar system exploration for more than two decades and have contributed to studies of the atmosphere of Mars [Fjeldbo and Eshleman, 1968], Venus [Fjeldbo and Kliore, 1971], Jupiter [Kliore et al., 1975], Saturn [Lindal et al., 1985], Uranus [Lindal et al., 1987], and Neptune [Lindal, 1992]. This method of radio occultation uses a receiver on Earth and a satellite occulted by a planetary atmosphere (which may occur from either a fly-by or by a satellite orbit of the planet). Suitably accurate atmospheric RO measurements of the Earth's atmosphere became possible with the advent of the Global Positioning System (GPS), but it was not until the late 1980s–early 1990s that the potential of RO using the GPS was widely appreciated [e.g., Gurvich and Krasil'nikov, 1990]. The radio occultation method used to sound the Earth's atmosphere is different from that used by the planetary experiments, in that both the receiver and the transmitters are orbiting the planet.

Data from the prototype GPS space-borne receiver, GPS/METeorology, launched in April 1995, confirmed the potential of obtaining accurate, global observations of the Earth's atmosphere from the radio occultation technique. Temperature comparisons between early results from the GPS/MET receiver and collocated radiosondes and numerical weather prediction (NWP) model analyses showed good agreement [e.g., Kursinski et al., 1996; Ware et al., 1996].

The analytic method of inverting radio occultation measurements to obtain meteorological parameters (i.e., the method used to sound other planetary atmospheres) involves the use of an integral transform, using the as-
sumption of a horizontally homogeneous atmosphere, to obtain a profile of refractivity (as a function of geometric height) [Fjeldbo and Eshleman, 1968]. The hydrostatic relation is used to obtain pressure and temperature from refractivity via density. For the Earth's atmosphere, where a reasonable prior knowledge of horizontal gradients is available, the analytic inversion does not represent the most suitable method since inadequate modelling of such gradients can cause large retrieval errors [e.g., Ahmad and Tyler, 1999].

Eyre [1994] addresses this issue and suggests a statistically optimal retrieval approach, using variational methods, to enable the direct assimilation of bending angle or refractivity (Healy and Eyre [2000] investigate the latter quantity). Zou et al. [1995] also looked at the impact of atmospheric radio refractivity measurements using a four-dimensional variational data assimilation approach. Their results showed that the measurements were effective in recovering the vertical profiles of water vapor, and found that the accuracy of the derived water vapor field was significantly better than that obtained through the analytic retrieval technique. The assimilation of these measurements were also shown to provide useful temperature information. There have also been several numerical experiments which have assessed simulated GPS/MET refractivity measurements to predict cyclonic disturbances [e.g., Kuo et al., 1998], and have concluded that these measurements are likely to have a significant impact on short-range operational NWP, with the caveat that the number of GPS receivers will have to be increased before the full potential impact of this measurement could be realized.

In this paper we utilize a nonlinear optimal estimation technique which is implemented and validated using an ensemble of simulated retrieval scenarios, using the bending angle quantity as the “observation”. Section 2 outlines the details of the RO methodology for the Earth's atmosphere necessary to derive the bending angle quantity, recalls the analytic inverse method, and discusses the impetus for pursuing an alternative inverse method. In section 3 we outline the theory for the nonlinear optimal estimation inverse method and give details of its implementation for GPS RO observations. Section 4 is devoted to details of the validation of the optimal estimation inverse method with reference to GPS RO observations. The sensitivity of the inverse model assumptions is also investigated, and we conclude the paper with a discussion of the results obtained.

2. Radio Occultation Measurements of Earth's Atmosphere

Kursinski et al. [1997] give a detailed description of the method used to measure the RO atmospheric observables; this section gives a summary of the theory, assuming no external encryption of the signals.

The GPS satellites transmit on two L-band radio frequencies, namely, L1 at 1.57542 GHz and L2 at 1.2272 GHz. Assuming a continuous link between the receiver and transmitter, when the receiver passes behind the atmosphere with respect to a GPS transmitter the signal travels through the atmosphere and is refracted in response to variations of refractive index along its path. This refraction causes the ray to travel over a longer path than it would in the absence of the atmosphere, in accordance with Fermat's principle of least time, which subsequently causes an atmospheric time delay in the received signal.

The Doppler shift of the signal is calculated from the additional atmospheric delay (the derivative of the phase delay). Using the geometry and notation of Figure 1, the Doppler shift $f_d$ of the carrier frequency $f_0$ measured by the receiver is given by

$$f_d = f_0 \left[ \frac{(v_T \cdot n_T + v_R \cdot n_R)}{c} \right],$$

(1)

Figure 1. Defining the radio occultation geometry used to obtain bending angle information from the time delay caused by the Earth's atmosphere. A tangential sphere is superimposed on to the oblate Earth (exaggerated) to emphasize the position of the local radius of curvature at the ray periapsis. The dashed lines indicate motion.
Thus the atmosphere.

There are also relativistic terms which need to be considered in equation (1) (due to different gravitational potentials and higher-order corrections for spacecraft velocity), but these can be eliminated on the basis of knowledge of orbital geometry and the Earth's gravity field [Kursinski et al., 1997]. Note that the relative positions and velocities of the two satellites can be calculated very accurately using available tracking data, which is independent of radio occultation data.

By specifying radial and tangential components of the velocity of satellite i in the plane coinciding with the ray trajectory as \( \mathbf{v}_R \) and \( \mathbf{v}_T \) and taking into account Snell's law, the angles \( \phi_R \) and \( \phi_T \) can be calculated from the following relations:

\[
f_d = f_0 e^{-c} (\mathbf{v}_R \cos \phi_R + \mathbf{v}_T \cos \phi_T + \mathbf{v}_R \sin \phi_R - \mathbf{v}_T \sin \phi_T). \tag{2}
\]

The cumulative effect of the atmosphere on the ray path can be expressed in terms of the total refractive bending angle, \( \varepsilon \), as a function of the impact parameter, \( a \). The impact parameter may be defined as the perpendicular distance between the local curvature of the Earth at the tangent point of the ray and the asymptotic straight line followed by the ray as it approaches the atmosphere.

From Bouguer's rule [Born and Wolf, 1993], and the geometry defined by Figure 1, \( \varepsilon(a) \) can be calculated thus

\[
\begin{align*}
\tau_R \sin \phi_R &= \tau_T \sin \phi_T = a, \quad (3) \\
\varepsilon &= \phi_R + \phi_T + \theta - \pi. \quad (4)
\end{align*}
\]

It is this rule that introduces the assumption of spherical symmetry \( (nr \sin \phi = a, \text{ where } a \text{ is a constant along the ray path}), \text{ i.e., a horizontally homogeneous atmosphere. Departures from this assumption can introduce significant errors if not properly accounted for. These errors have been studied by Ahmed and Tyler [1999] and Healy [1999] but are not addressed in the study presented here.}

Measurements of the time delay become possible for neutral atmospheric sounding when the GPS signal begins to transect the mesosphere at an altitude of \( \sim 85 \text{ km} \); at this altitude the atmospheric phase delay is \( \sim 1 \text{ mm (} \times 10^{-12} \text{ s) which can be observed by the low-Earth-orbit (LEO) GPS receiver [Ware et al., 1996]. Further information about the measurement characteristics may be obtained from Kursinski et al. [1997].}

Using an Abel integral transform (equation (5)) or making use of a similar integral transform when applying Fresnel diffraction theory [Mortensen and Høeg, 1998], these bending angle measurements can be inverted to obtain a profile of refractivity. For the sake of completeness the "forward" Abel integral transform (equation (6)) is presented alongside the "inverse" Abel integral transform:

\[
\ln n(x) = \frac{1}{\pi} \int_{\sigma}^{+\infty} \varepsilon(a)(a^2 - x^2)^{-1/2} \, da \tag{5}
\]

\[
\varepsilon(a) = -2a \int_{\sigma}^{+\infty} \left( \frac{\partial \ln n(x)}{\partial x} \right) (x^2 - a^2)^{-1/2} \, dx, \tag{6}
\]

where \( n \) is the refractive index and \( x \) is the refractive radius (i.e., \( x = \pi n \)).

Geometric height levels \( z \) can be obtained from the refractive index profile, as a function of impact parameter, and the local radius of curvature \( R_c \), thus

\[
z = \frac{x}{n} - R_c. \tag{7}
\]

Because refractive index is near unity, refractivity \( N \) is used to describe the refractive medium, which is given by \( N = (n - 1) \times 10^6 \).

Refractivity is affected primarily by air density (dependent on pressure and temperature) and water vapor density; thus the measurement contains information about both. Equation (8) describes this relationship,

\[
N = \frac{c_1 P_a}{T} + \frac{c_2 P_W}{T^2}, \tag{8}
\]

where \( P_a \) is the total atmospheric pressure (hPa), \( P_W \) is the partial pressure of water vapor, \( T \) is the temperature (K), and \( c_1 \) and \( c_2 \) represent constants of proportionality, whose values are 77.6 (KhPa\(^{-1}\)) and 3.73 \times 10^5 (K^2 hPa\(^{-1}\)), respectively. The form of the dry and moist terms in equation (8) is from Smith and Weintraub [1953].

Refractivity is also affected by charged particles in the ionosphere and the scattering by water droplets suspended in the atmosphere. The first-order ionospheric contribution to refractivity can be removed by combining the two GPS signals (described by Vorob'ev and Krasil'nikova [1994]), leaving higher-order terms, and the scattering contribution is found to be negligible compared to the contribution due to air and water vapor density [Kursinski et al., 1997].

There is no measurement information to allow the separation of the effects of temperature and water vapor, and therefore these quantities can be retrieved only using prior information. If the absolute humidity is judged to be small (e.g., in the coldest regions of the troposphere and stratosphere, with temperatures < 250 K), it may be neglected, and density calculated from refractivity. The hydrostatic relation can be used to calculate values of pressure and hence temperature. However, if humidity is judged to be significant, then an iterative process may be used to calculate temperature/humidity if an a priori profile of humidity/temperature is used. This prior information can be taken from various sources, such as collocated NWP model output.
The inability of this inverse method to account for horizontal refractivity inhomogeneities and the suboptimal way this method retrieves temperature and humidity with prior values represent two key disadvantages of this method and form part of the impetus to develop a new inverse method.

The hydrostatic relation, used to compute values of pressure on the retrieved height levels, requires an assumed pressure value at a particular geometric height level. A variety of methods have been implemented to tackle this initial value problem. Kursinski et al. [1996] assumed a temperature of 260 K at 50 km. This method has the problem that if the assumed temperature is inconsistent with the measurements an error is introduced, which decreases exponentially with depth. More elaborate methods (e.g., Rocken et al., 1997; Steiner et al., 1999) initialize the GPS/MET retrieval at some high altitude (e.g., 100 km) using climate model data and combine the measurements and model data to minimize downward propagation of errors. In principle, the method presented in this paper also combines the model data and the observations but achieves this in an optimal way. The method presented also has the advantage of a straightforward error characterization. Since GPS RO observations sometimes reach near-surface altitudes (i.e., < 1 km from the surface), surface pressure is also retrieved using the optimal estimation inverse method.

3. Optimal Estimation Inverse Method

The method outlined here is also known as one-dimensional variational data analysis. The main advantage of this method is that it provides simultaneous estimates of temperature and humidity profiles that are statistically optimal, given prior estimates from an NWP model (together with their error covariances). It also provides a framework for assessing the error characteristics of the estimates.

In this study, only the 1-D problem has been studied. However, the errors introduced by the neglect of the horizontal gradients have been estimated and allowed for as part of the error budget (see section 3.2).

3.1. Theory

A brief description of the theory used in optimal estimation in presented here; a more detailed description is given by Rodgers [1976, 1990]. For brevity, the observation noise, the error associated with any forward modeling parameters, and the forward model error (which includes the representativeness error [Lorenc, 1986]) will be accounted for in one vector, which will be denoted by $e$, and its ensemble characteristics described by the covariance matrix $E$.

The rationale behind optimal estimation is to minimize a cost functional $J(x)$ (or to solve $\nabla_x J(x) = 0$), which measures the degree of fit of estimates of the atmospheric state to the measurements and to some prior information and possibly to some other physical or dynamical constraints. In this case, $J(x)$ is given by

$$J(x) = \left[ y^o - y(x) \right]^T E^{-1} \left[ y^o - y(x) \right] + (x - x^b)^T C^{-1} (x - x^b) ,$$

where $x^b$ and $x$ represent the background and updated state vectors, respectively; $y^o$ and $y(x)$ represent the observation vector and the estimated observation vector calculated from the state vector [Eyre, 1994], respectively; and $C$ represents the background error covariance matrix.

There are a number of methods available to minimize $J(x)$. The scheme described here uses the Levenberg-Marquardt iterative method [e.g., Press et al., 1992]:

$$x_{i+1} = x^b + \left[ (1 + \gamma) C^{-1} + K^T E^{-1} K \right]^{-1} \left[ (K^T E^{-1} (y^o - y(x_i))) + (\gamma C^{-1} + K^T E^{-1} K)(x_i - x^b) \right],$$

where $K$ is $\nabla x_i y(x_i)$, $\gamma$ is a nondimensional weighting factor (for increasing values of $\gamma$ this minimization method degenerates into the method of steepest descent), and all other variables are as before.

Using the optimal estimation theory, it is possible to obtain an error covariance for the retrieved products. Indeed, it can be argued that the retrieved products are of limited value without an estimate of their uncertainty. The solution error covariance $\tilde{S}$ is given approximately (i.e., at the linear limit) by

$$\tilde{S} = (C^{-1} + K^T E^{-1} K)^{-1}.$$

The solution error covariance can then be compared with the prior error covariance to ascertain how the retrieval has improved upon the prior knowledge of the atmospheric state.

3.2. Implementation of the Optimal Estimation Inverse Model

This subsection describes in detail the components of equation (9).

3.2.1. Background state vector and its uncertainty covariance matrix. In this case the background knowledge of the atmospheric state $x^b$ was obtained from short-range forecasts provided by the U.K. Meteorological Office unified model [Cullen, 1993]. The model from which the data are derived had 19 levels, which were expressed on hybrid-sigma pressure coordinates (surface—10 hPa). The global model had a resolution of 0.833° (360°/288) longitude and 1.25° (360°/288) latitude. The data used are 6-hour forecasts which have been interpolated to occultation event positions, using the mean latitude and longitude of each occultation (The mean latitude and longitude of an occultation corresponds typically to altitudes in the lower stratosphere/upper troposphere.).

These 19 levels are linearly interpolated (in pressure) onto the state vector levels used for TOVS retrievals [Eyre, 1989]. CIRA 1986 climatology is assumed above the UKMO model, accounting for the
latitudinal and seasonal variation of the profile. This climatology provides a reasonable prior and first guess information in the upper stratosphere. For the forecast error covariance matrix $C$ (described by Eyre [1989]), lower atmospheric values (surface–50 hPa) were generated from radiosonde–forecast difference statistics, and upper stratosphere values were found by regression from the levels provided [Eyre, 1989].

The radio occultation retrieval uses 40 temperature elements, 15 $\ln$(specific humidity) elements and a surface pressure element from this forecast error covariance matrix. Specific humidity is expressed as the natural logarithm of specific humidity since forecast errors in this quantity are more constant than in specific humidity. The temperature and $\ln$(specific humidity) interquantity covariance values have been set to zero. These interquantity covariances are not well known, and assuming zero covariance between them is more conservative than an erroneous covariance. The surface pressure element is uncorrelated with both temperature and $\ln$(specific humidity).

Since CIRA climatology is used to form the a priori (and the first guess), it is necessary to consider the errors that may be attached to such information. In general, if the standard deviation values from the diagonal elements of the forecast error covariance matrix are smaller than the uncertainties assumed for the climatology, then the climatological errors are used at the levels in the upper atmosphere described by the CIRA climatology (off-diagonal elements remain the same): at latitude $\theta$, for $|\theta| > 45^\circ$, $\sigma=15$ K (winter) and 5 K (summer) and for $|\theta| \leq 45^\circ$, $\sigma=5$ K.

The values for the diagonal elements of the UKMO forecast error covariance matrix are shown by Figure 2. As expected for temperature, the lower atmosphere forecast errors are reasonably small (of the order of 1.5 K) and increase as a function of altitude. The actual values have been developed over recent years at the UKMO and reflect the average error in 6-hour forecasts.

3.2.2. Observation vector and its error covariance matrix. The observation vector $y^o$ contains bending angle measurements as a function of impact parameter (section 2). In practice, atmospheric phase delay measurements from the GPS/MET receiver are low-pass filtered to reduce noise. The cutoff frequency of the filter is tuned to pass phase variations corresponding to vertical scales of 2–3 km in the stratosphere and $\sim$200 m in the lower troposphere [Rocken et al, 1997]. From the phase observable, Doppler shifts (and subsequently bending angle profiles) for the two GPS signals are computed. First-order ionospheric effects are removed from the data by combining the two signals to form a single corrected profile [Vorob'ev and Krasil'nikova, 1994]. Typically, after filtering, there are 100–200 neutral atmosphere bending angle measurements, which span a vertical range of $\sim$0.5–60 km. In addition to the observation error covariance matrix consisting of observation noise estimates, errors from the forward modeling and forward model parameters are considered.

3.2.3. Observation noise. Observation errors are created by the hardware of the measurement system and by the preprocessing of the observations. The observation noise estimates are taken from Luntama [1997]. They include thermal noise, residual errors from the ionospheric correction, local multipath (distortions when the transmitted signal is reflected from a surface near the signal propagation path), orbit determination accuracy, and clock instabilities of LEO receiver and GPS satellites and ground stations.

These observation error estimates are based on phase noise levels during the measurement or estimated from other noise sources during a radio occultation event. The effects from satellite clock errors and from the selec-
Figure 3. Measurement noise budget and total measurement error budget for the RO optimal estimation inverse method (after Luntama [1997]). The error estimates shown in Figure 3a represent normal atmospheric conditions, i.e., small multipath error (3 mm) and normal ionospheric conditions. Figure 3b shows the percentage contributions from the different error sources to the total bending angle error. Figure 3c shows typical standard deviation values from the principal diagonal from each contribution to the total measurement error covariance matrix, and Figure 3d shows the individual measurement error sources as a percentage proportion of the total error.

The observation error estimates used in the optimal estimation technique are shown by Figure 3. These estimates have been computed by a nominal bending angle profile defined using an exponential curve with a scale height of ~7 km (J. P. Luntama, personal communication, 1998). These estimates represent normal atmospheric conditions with a relatively small multipath error (3 mm) and normal ionospheric conditions. Figure 3a shows that there are a number of noise contributions of comparable size in the lower atmosphere. At altitudes above 30 km the residual ionospheric correction error begins to dominate the total bending angle error curve (Figure 3b).

These noise estimates are assumed to be fully independent, i.e., their interlevel (and interquantity) covariance is zero. However, the bending angle measurements do contain a small, local correlation between successive levels due to filtering of the phase measurements. Because this correlation is small, the diagonal form is a good approximation to the full matrix (J. P. Luntama, personal communication, 1998).
Real observations from the GPS/MET data used have been found to be noisier than theoretical estimates [Luntama, 1997]. Bending angle fluctuations in the upper stratosphere (of the order of $10^{-5}$ radians) are present and are thought to be due to residual errors from the LEO satellite clock calibration in the differencing decryption technique (see "Observation noise") [Syndergaard, 1999]. As such, a suitable error is attached to reflect the upper atmosphere measurements.

### 3.2.4. Forward Modeling Errors

In this work the forward model used to map from state space to observation space is described by Eyre [1994] but applied to an atmosphere approximated as spherical symmetric about the given profile at the tangent point. Essentially, the geophysical parameters are converted to refractivity as a function of height and subsequently impact parameter using the local radius of curvature. The resulting profile is mapped into observation space using the "forward" Abel integral transform (equation (6)).

The two main forward modeling errors are due to the assumption of a horizontally homogeneous atmosphere and a representativeness error. Estimates for the first of these errors are obtained using a version of the forward model which can account for horizontal inhomogeneities in the plane of the ray path (described by Eyre [1994]) and comparing observation vectors with the version of the forward model which assumes a horizontal homogeneous atmosphere [Palmer, 1998]. Midlatitude two-dimensional NWP fields (0°–360°) were used to simulate typical horizontal gradients in temperature and humidity. By considering small sections of the field at a time (typical of the horizontal resolution of radio occultation measurements which is of the order of 300 km) the complete field was traversed. Computing the ensemble mean from the difference between the two versions of the forward models allowed a reasonable estimate of the forward modeling error incurred by the spherical symmetry assumption to be computed. It is noted that this error estimate does not represent the true error in observation space since the forward model does not simulate the full error characterization. Both Ahmad and Tyler [1999] and Healy [1999] consider bending angle errors from horizontal gradients for specific cases. However, there is no published material that quantifies this error statistically. Simulation with a full 3-D ray tracer through the UKMO mesoscale model fields suggest that the errors are ~3% for ray paths near the surface, which is consistent with the value used in this work.

An error arises from representing an intrinsically high resolution problem with a crude resolution. This type of error is often called a representativeness error and describes the error from the inability of NWP model vertical grids to represent small-scale atmospheric structure, which are evident in GPS/MET RO measurements [Kursinski et al., 1997]. The method used to estimate this quantity is described by Healy [1998] and is found to be of the order of 2% of the bending angle measurement in the troposphere and upper stratosphere, decreasing slightly in the middle stratosphere. This variation in the error is associated with the temperature variations in this region.

The forward model used is based on geometric optics and therefore does not account for atmospheric diffraction. However, Kursinski et al. [1997] have shown that the geometric optics assumption is successful in describing propagation characteristics above a certain diffraction limit, and diffraction is therefore not considered here.

### 3.2.5. Forward model parameter errors

Uncertainties associated with the physical constants used to model the physical system also cause modeling errors. The major physical constants used in the forward model are the refractivity coefficients ($c_1$ and $c_2$ in equation (8)) and the local radius of curvature.

The uncertainty of the refractivity coefficients does not represent the error associated with their values but the uncertainty of the measured quantity; for this reason the information will not be included in the total observation error budget since it will result in a bias in the retrieval (Optimal estimation theory assumes that all the errors are unbiased.). The local radius of curvature assumed for zero altitude is a parameter that is used to compute the bending angle observations from atmospheric phase delay, therefore any error associated with this parameter will be present in the bending angle observations. This parameter is also used to compute geometric height levels from impact parameter levels. The uncertainty of this value is estimated to be ~100 m [Kursinski et al., 1997].

### 3.2.6. Total observation error

The total observation error covariance $E$ is constructed by adding the covariance matrices from observation noise, forward modeling and forward model parameters. The diagonal terms have also been constrained not to fall below a minimum value to account for the noisy upper stratosphere measurements. Figure 3 shows how the standard deviation values of the principal diagonal from each error contribute to the observation error covariance matrix.

The dominant source of error for the majority of the vertical range considered is the forward modeling error, i.e., representativeness error and horizontal inhomogeneity error, with the upper stratospheric noise limit providing the second largest contribution to the total error. In the upper stratosphere the total bending angle reverts to the upper level minimum noise used. At near-surface altitudes the forward model parameter error, i.e., local radius of curvature, is significant. The overall effect from the local radius of curvature decreases exponentially due to the hydrostatic relation.

### 3.3. Convergence and Quality Control

The method used to judge convergence relies upon values of the cost function, that is, if the relative change is smaller than a specified value (0.5%), then the so-
lution is determined to have converged. This method alone is found to be a good indicator of convergence in this case. For the work presented in this paper, the maximum number of iterations considered is 10; if the solution has not converged (determined by the method presented above within 10 iterations), then the calculation is halted and a numerical "flag" set.

If, after convergence has been determined, the $J(x)$ value is greater than the $\chi^2$ value given the number of degrees of freedom at a set confidence level (in this case 99.9%), then a numerical "flag" is set. Retrievals with flags set are omitted from any statistics. Furthermore, to ensure the solution computed at each iteration is physical, the $\ln$(specific humidity) elements are checked for supersaturation and corrected if necessary.

4. Results

In this section the performance of the optimal estimation retrieval scheme is examined using simulated profiles and realistic error estimates. For each simulated profile a "true" profile is established by taking one of a set of profiles of UKMO model analyses from which to compute the "true" observation vector. The associated background atmospheric profile is calculated by perturbing the "true" profile thus

$$x^b = x^t + \sum_{i=1}^{N} \epsilon_i \lambda_i^{1/2} P_i,$$

where the superscript $t$ denotes the "truth", $\lambda_i$ and $P_i$ are the $i$th eigenvalues and eigenvectors of the forecast error covariance matrix, and $\epsilon_i$ represents the $i$th number drawn from a normal distribution of random numbers. Observation noise is modeled and superimposed onto the observation vector using the method analogous to equation (12), utilizing the eigenvector and eigenvalues from the total observation error covariance matrix.

4.1. Ensemble of Numerical Simulations

Using the method described by equation (12), simulated observations and realistic background profiles were produced. These were used as inputs to the inversion scheme to obtain retrieved profiles which were subsequently compared with the "true" profiles to assess the impact of the observations on the background information.

The observation level values (i.e., impact parameter, local radius of curvature and geographical position) have been taken from data during "prime-times" (Pe-

riods of time when the received signals are free from antispoofing military encryption.) 1 and 2 [Rocken et al., 1997] in an effort to simulate realistic retrieval scenarios. The latitudinal and longitudinal distributions of these occultation events (determined by GPS sampling) are varied, thus providing a mixed ensemble of polar, tropical, and midlatitudinal occultation events.

Five hundred profiles with random temperature, humidity, and surface pressure conditions have been tested, and successful retrievals (i.e., which pass quality control) were obtained in all but eight cases. In most cases, convergence is obtained within three or four iterations. In general, the $J(x)$ values at convergence were comparable to the number of degrees of freedom considered, as expected [Marks and Rodgers, 1993]. This suggests that although the $\chi^2$ quantity is only strictly valid for linear problems, it can be used reliably as a quality control for the retrieval.

The eight cases which do not pass the specified quality control have been examined. They are found to be cases in which the cost function has been minimized successfully but the converged value is too large compared with the $\chi^2$ distribution. These spurious converged profiles represent artificial outliers, which are generated when the increment described by equation (12) is large enough to make the inverse problem grossly nonlinear. A small number of profiles are expected to have this problem due to the normal distribution of random numbers used in the method of simulating atmospheric profiles. The profiles which failed the quality control have not been included in the statistics shown.

For each successful retrieval the retrieval error and the background error (first guess error) have been calculated, and the mean and standard deviation values of these data have been computed. The standard deviation values represent the errors ascribed to each element of the solution and background state vector and can be compared directly with the square root values of the principal diagonal of the background error covariance matrix assumed in the retrieval. The ratio of the retrieval error estimates to the forecast error estimates is related to the amount of information the measurements supply to the NWP system.

An improvement vector is defined which indicates how the retrieval has improved the knowledge of the background state throughout the atmospheric profile and will be used to complement the rms statistics presented. The improvement vector $\eta$ is given by

$$\eta_j = 100 \left[ 1 - \left( \frac{S_{ij}}{C_{ij}} \right)^{\frac{1}{2}} \right],$$

where $j$ is the matrix and improvement vector element index and all other variables are as before.

The results from the ensemble of simulated retrievals are summarized by Figure 4. Figures 4a and 4b show the computed rms errors from the the simulated retrievals. The forecast errors resemble those shown by Figure 2, as expected, modified slightly by the modeled observation noise.

The temperature improvement vector suggests that optimal estimation considerably improves upon the prior knowledge of the atmospheric temperature, from the lower-troposphere to the middle stratosphere. The temperature improvement vector declines in the upper stratosphere partly because of the upper noise limit used, which is comparable to forecast errors in
Figure 4. Theoretical error estimates from the ensemble of simulated retrievals. Figures 4a and 4b show the theoretical rms errors, where the solid and dotted lines represent the retrieval and background errors, respectively. Figure 4a shows the errors from the temperature solution, and the Figure 4b shows the errors from the ln(specific humidity) solution with the surface pressure solution error inset. The lower panel show the corresponding improvement vector. Figures 4c and 4d show the temperature and ln(specific humidity) elements with the surface pressure improvement inset in Figure 4d.

observation space, and partly because it represents the lower limit of the observation vectors used in the ensemble of simulations. The temperature improvement vector declines in the lower troposphere where humidity becomes more significant, and both quantities are retrieved simultaneously; thus the emphasis is shifted from temperature to ln(specific humidity). It is clear from the plot that there is a gradual decline in the temperature retrieval quality from 300 to 1000 hPa as the humidity retrieval quality improves, where the background uncertainty information is being used to resolve the temperature-specific humidity ambiguity in the refractivity.

Both the temperature and humidity knowledge decline near the surface because the majority of occultation events presented here terminate typically above 1 km.

The results shown by Figure 4 confirm that the method is suitable for the purpose of nonlinear optimal estimation using RO measurements. Together the theoretical rms errors and the computed improvement vectors suggest that there are improvements in the prior knowledge of the atmospheric state from near surface to the upper stratosphere. In particular, the level of surface pressure improvement suggests that the RO observations can improve the prior knowledge of the surface pressure.

4.2. Solution Error Characterization

Using the optimal estimation inverse theory outlined in section 3, an error analysis can be obtained which allows a full characterization of the solution (for a detailed account, see Rodgers [1990]). The error associated with the solution vector can be split into its constituent parts, namely, error from the background error estimates, forward modeling, forward model parameters, and observation noise. A mid latitude retrieval (henceforth referred to as the example occultation profile), which spans 0.7–60 km, has been used as an example to illustrate the method (Figure 5).

Figures 5a and 5b show that the a priori provides almost all the information to the temperature and ln(specific humidity) retrieval above 10 hPa and 500 hPa, respectively. Figures 5c and 5d show the observation noise contribution to the retrieval error which is comparatively small. Figures 5e and 5f show
the forward modeling contribution to the retrieval error. The structure of this contribution is very similar to that of the observation noise but has a larger associated error. In general, forward modeling represents the second largest contribution to retrieval error.

The temperature solution error contributions shown by Figures 5c and 5e peak at 3 hPa, at the point where the stratospheric noise limit contribution to the total observation error budget peaks (Figure 3). The observation contribution to the solution error begins to increase above \( \sim 0.3 \) hPa. Above this pressure level the a priori increases more rapidly, and so the observation is given more weight, resulting in a larger contribution to the solution error.

Figures 5g and 5h show that forward model parameter error represents the smallest contribution to the total retrieval error. This contribution to the temperature and ln(specific humidity) solutions peaks near 1000 hPa owing to the local radius of curvature error. The contribution to the surface pressure solution represents a small fraction of the total retrieval error.

Figure 5 indicates that the dominant contributions to the solution error are from the a priori and forward modeling, suggesting that particular efforts should be made to improve their accuracies. It should be noted that the solution error depends on the assumptions made about the uncertainty statistics.

4.3. Quality of Surface Pressure Retrievals

It is found that the quality of the surface pressure retrievals is dependent on the vertical extent of the occultation, i.e., how closely it approaches the surface. To illustrate this point, the example occultation profile is used. By systematically removing observations near the surface and re-retrieving temperature, ln(specific humidity), and surface pressure it can be shown how the vertical range of the occultation is important to the quality of the surface pressure retrieval (Figure 6).

Figures 6a–6c show the temperature, ln(specific humidity), and surface pressure improvement vectors get smaller in the lower atmosphere with increasing values for the lowest geometric height level, as expected. Figure 6d shows that the retrieval, which includes all the observations, provides enough information to retrieve accurately the true value for surface pressure; as the number of near-surface observations decreases, the retrieval becomes smaller and smaller, approaching the prior value.

It is interesting to note that the surface pressure information does not decrease as quickly as expected. Indeed, there is still a considerable amount of surface pressure information toward the upper troposphere. This variation in surface pressure information is due to the link between height and pressure through the hydrostatic relation.
It can be concluded from this experiment that using optimal estimation, radio occultation measurements of the Earth possess surface pressure information even if the occultation has been completed in the mid troposphere (e.g., due to atmospheric multipath interrupting the transmitted signal). This is important to appreciate when validating surface pressure retrievals using real data.

4.4. Sensitivity to Inverse Model Statistics

In this section we present the results from a sensitivity study in which the statistics used to compute the optimal estimate are changed in order to investigate the retrieval sensitivity to such alterations. The three statistics that are altered are the observation noise, the forward modeling error and the forecast error since these represent the largest contributions to the solution error, as shown by Figure 5.

Altering the observation noise may simulate the possible changes in observation noise sources, e.g., improved high-order ionospheric or poor quality clocks aboard the GPS receivers. Changing the forward modeling error is a crude method of simulating the possibility of assimilating intrinsically high resolution GPS radio occultation measurements with lower- or higher-resolution NWP model fields and/or an occultation through a frontal system or a relatively horizontally inhomogeneous atmosphere.

Altering the prior error covariance matrix represents the effect of changing the prior knowledge of the atmosphere. Increasing these error estimates can represent the extent of the knowledge of the atmosphere in the Southern Hemisphere where other atmospheric information is sparse. Decreasing this error may represent a more realistic model dynamics/climatology and/or greater confidence in other similar atmospheric observations used to initialize the model. For this particular test the variance values are changed while retaining the existing correlations.

Figure 7 shows the improvement vectors from the retrieval sensitivities outlined above. Improvement vectors are used to present the results of this study because
it is the relative improvement on the prior estimate of the atmospheric state that we are interested in.

Figure 7a and 7b shows that doubling or halving the observation noise has little effect on the overall improvement, reflecting the contribution from this error source to the total observation error covariance matrix (Figure 3). The changes due to ln(specific humidity) are very slight. The surface pressure improvement changes typically by a few percent.

Figure 7c and 7d correspond to increasing or decreasing the forward modeling error (by 50%). This error source provides the largest contribution to the total observation error and, as such, has a large influence on the degree of improvement. The difference in the temperature improvement is of the order of 15% in the upper troposphere and lower stratosphere, above which other error sources are more important and below which (the lower troposphere) the emphasis is shifted from the temperature retrieval to the ln(specific humidity) and surface pressure retrieval. The improvement response for ln(specific humidity) is less pronounced than for temperature peaking at ±8%. The surface pressure improvement variation is approximately ±15%.

Figure 7e and 7f shows that increasing or decreasing the standard deviation of the prior error increases or decreases the improvement of temperature and ln(specific humidity) throughout the vertical range of the observations as expected. Decreasing the a priori error means

Figure 7. Improvement vectors from the retrieval sensitivity studies using the example occultation profile. In general, left plots show the temperature improvement vector elements, and the right plots show the ln(specific humidity) improvement vector elements, with the surface pressure improvement vector element inset. For Figures 7a–7f, dashed lines represent a reduction in the quantity altered, dotted lines represent an increase in the quantity altered and solid lines represent the control case. Figure 7a and Figure 7b show the results from doubling/halving the measurement noise, Figure 7c and Figure 7d show the results from increasing and decreasing the representative error by ±50%, Figure 7e and Figure 7f show the results from increasing and decreasing the principal diagonal standard deviation errors from the a priori error covariance matrix by ±50%.
better background knowledge; consequently, the weight-
ing of the a priori/observation is increased/decreased.
This corresponds to a small improvement relative to the
background atmospheric knowledge. Positive and neg-
ative temperature improvement differences are of the
order of 20% throughout the range described by the
observation vector; the ln(specific humidity) improve-
ment is of the order of ±10%; and the surface pressure
improvement is of the order of ±10%.

This sensitivity study has looked at some of the ex-
treme case scenarios in which the statistics assumed
for the optimal estimation inverse method have been
changed. It has been shown that the retrieval method is
most sensitive to the background errors and forward
modeling errors; the latter being related to errors due to
the spherical symmetry assumption. The results from
increasing the background error estimates are especially
important to note since they represent a real possibility
when dealing with any reasonable measurements in the
data-sparse Southern Hemisphere.

5. Conclusions

In this paper we have demonstrated a prototype op-
timal estimation inverse method for GPS radio oc-
cultation observations. The method is a nonlinear,
maximum a posteriori technique which can accommodate
most aspects of the real radio occultation problem.
In particular, it is able to account for some of the
error incurred from assuming local spherical sym-
metry which is not possible using the analytic inverse
method. The optimal estimation technique handles the
temperature–water vapor ambiguity in a more rigorous
way, rather than the suboptimal manner inherent with
the analytic inverse method.

The optimal estimation inverse method is used here
as an iterative method but is found to be stable and to
converge rapidly in most cases. The value of the cost
function at each iteration can be used reliably to judge
convergence and as an indicator of sensible results, al-
lowing anomalous retrieval conditions to be recognized
and omitted, thus preventing gross retrieval errors.

The method is shown to be suitable for retrieving val-
ues for surface pressure. Hence, this method of utilizing
radio occultation observations of the Earth’s atmo-
sphere has the potential to improve both atmospheric
and oceanographic models, which may lead to improved
predictions of the weather and climate. The quality of
the surface pressure retrieval is shown to depend on the
vertical extent of the occultation, i.e., higher-quality ret-
rievals are attainable with occultations that reach low
altitudes.

It should be noted that the background statistics as-
sumed in the paper represent global statistics. How-
ever, for purposes of demonstrating a prototype re-
treeval scheme for radio occultation observations they
are found to be adequate. It has been shown that the
retrieval accuracies and hence the weight which should
be given to the data in the subsequent model analy-
sis are sensitive to both forecast error uncertainties and
values used to describe the forward modeling error.

Acknowledgments. The authors would like to thank
the Jet Propulsion Laboratory for providing the GPS/MET
radio occultation data set, J. P. Luntama for the bending
angle error estimates, and C. D. Rodgers, A. Dudhia, and
H. Roscoe for comments on earlier drafts. We gratefully
acknowledge the University Corporation for Atmospheric Re-
search for operating the GPS/MET program. Three anony-
umous reviewers provided thorough and thoughtful comments
which helped to improve the paper significantly.

References

Ahmad, B. and G. L. Tyler, Systematic errors in atmo-
spheric profiles obtained from Abel inversion of radio
occultation data: Effects of large-scale horizontal gradi-
Born, M. and E. Wolf, Principles of Optics, Pergamon, New
Eyre, J. R., Inversion of cloudy satellite sounding radiances
by nonlinear optimal estimation. I: Theory and simula-
tion for TOVS, Q. J. R. Meteorol. Soc., 116, 401-434,
1989.
Eyre, J. R., Assimilation of radio occultation measurements
into a numerical prediction system, Tech. Memo. 199,
Eur. Cent. for Medium-Range Weather Forecasts, Reading,
Fjeldbo, G., and V. R. Eshleman, The atmosphere of Mars
analyzed by integral inversion of the Mariner IV occulta-
Fjeldbo, G., and A. J. Kliore, The neutral atmosphere of
Venus as studied with the Mariner V radio occultation
Gurvich, A. S., and T. G. Krasil’nikova, Navigation satel-
lites for radio sensing of the Earth’s atmosphere, Sov. J.
Healy, S. B., A statistical comparison of GPS/MET radio
occultation data with numerical weather prediction ana-
Healy, S. B., Radio occultation bending angle errors caused
by horizontal gradients: a simulation study, Tech. Memo.
Healy, S. B., and J. R. Eyre, Retrieving temperature, water
vapour and surface pressure information from refractive
index profiles derived by radio occultation: a simulation
Kliore, A. J., G. Fjeldbo, B. L. Seidel, D. N. Sweeney,
T. T. Sesplaukis, P. M. Woiceshy, and S.I. Rasool, The
atmosphere of Io from Pioneer 10 radio occultation mea-
Kuo, Y. H., X. Zou, and W. Huang, The impact of Global
Positioning System data on the prediction of an extratrop-
ical cyclone: An observing system simulation experiment,
Kursinski, E. R., et al., Initial results of radio occultation of
Earth’s atmosphere using GPS, Science, 271, 1107-1110,
1996.
Kursinski, E. R., G. A. Hajj, K. R. Hardy, J. T Schofield,
and R. Linfield, Observing the Earth’s atmosphere with
radio occultation measurements using GPS, J. Geophys.


J. J. Barnett, Department of Physics, Clarendon Laboratory, Oxford, OX1 3PU, England, U.K. (j.barnett@physics.ox.ac.uk)

J. R. Eyre and S. B. Healy, Satellite Application Division, United Kingdom Meteorological Office, Bracknell, RG12 2SZ, England, U.K. (jreyre@meto.gov.uk; sb-healy@meto.gov.uk)

P. I. Palmer, Division of Engineering and Applied Science, Pierce Hall, Harvard University, Cambridge, MA 02138 (pip@io.harvard.edu)

(Received October 15, 1999; revised February 18, 2000; accepted February 24, 2000.)